Logic and Argumentation

Proceedings of the colloquium,
'Logic and Argumentation',
Amsterdam, 14-17 June 1994

Logic and Argumentation

Edited by Johan van Benthem, Frans H. van Eemeren,
Rob Grootendorst and Frank Veltman

Contents

Preface

This volume finds its origin in a colloquium on Logic and Argumentation, held in June 1994 in Amsterdam and sponsored by the Royal Netherlands Academy of Arts and Sciences. The papers included have been selected for the role they can play in illuminating current thinking about the various kinds of relations between logic and argumentation. In order to complete the picture, the editors requested a few of their colleagues who hold views that were not represented at the colloquium to add to the volume by contributing a chapter.

Logic arose from argumentation theory as it was developed in antiquity. Gradually, a difference emerged between formal logic as the investigation of mechanical reasoning patterns and argumentation theory as the study of argumentative discourse in a more general sense. The latter has been mainly fed by argumentation theorists stemming from the humanities who often call themselves 'rhetoricians' or 'informal logicians'. Some of them, Toulmin and Perelman being the most prominent, even strongly argued against modern formal logic.

Although there is nothing against an academic division of labour, we feel that the opposition between logic and argumentation theory is artificial and should be overcome. Building on insights provided by Evert Beth, Paul Lorenzen, Charles Hamblin and many others, broader applications of logic can be pursued than so far have been recognized. A more clearly defined idea of the direction that a development towards a 'logical argumentation theory' may possibly take can already be gained by looking at the game-theoretical dialogical accounts of rational communication that have recently been given. Another helpful starting point can be drawn from the communication-oriented theories of argumentation that have been developed in the humanities and are, as it were, waiting to be formalized. Artificial intelligence, where many researchers have become interested in the role of argumentation-theoretical structures in programming languages, can be a further source of inspiration.

This volume aims at providing some background to the academic endeavour of exploring the connections between logic and argumentation. It offers the reader some respresentative specimina of current thinking about this subject. The volume starts with two introductory chapters. First, Frans H. van Eemeren and Rob Grootendorst give a survey of the state of the art in argumentation theory. Johan van Benthem then discusses some interfaces between current developments in logic and argumentation theory.

In several chapters, the links between argumentation and logic are immediately at issue. Robert C. Pinto attempts to clarify the relations between arguments and inferences, between the normative study of arguments and inference, and between logic as the normative study of inference and the study of argumentation. Diderik Batens devotes his chapter to the challenging task of bridging the gap between logic

and argumentation. Richard J.C.M. Starmans discusses the relation between modern argumentation theory and formal logical theories of commonsense reasoning in Artificial Intelligence.

Traditionally, intriguing ideas ensue from the study of the distinctions between validity and invalidity and the related problem of coming to grasps with the fallacies. Maurice A. Finocchiaro argues that the Oliver-Massey asymmetry between showing that a given argument is formally valid and showing that it is formally invalid does not hold. Sally Jackson proposes an explanation for the persuasiveness of fallacies. Erik C.W. Krabbe discusses some circumstances in which a formal fallacy can be tracked down. Douglas Walton gives an analysis of the straw man fallacy as a misrepresentation of someone's commitments in order to refute that person's argument. John Woods points out that the logical and semantic paradoxes push theorists, unannounced and often unaware, into idealism.

In the study of communication, argumentation has been a focus of attention from several angles. David Zarefsky distinguishes between four forces that have shaped argumentation studies in the speech communication discipline: the evolution of competitive debate, the infusion of empirical perspectives and methods from the social sciences, the recovery of practical philosophy, and the growing interest in social and cultural critique. From a psychological angle, Daniel J. O'Keefe discusses some interconnections between argumentation studies and persuasion effect research.

The volume closes with three chapters concentrating on linguistic aspects. Keith Stenning proposes a fresh approach to the tension between language as a formal structure and language as a social practice. M. Agnès van Rees makes an argument for taking into account the social interactional aspect when reconstructing discourse as a critical discussion. Francisca Snoeck Henkemans establishes a connection between the semantical descriptions of 'anyway' and 'even' and the characterization of independent and interdependent arguments.

The editors regard this volume primarily as a gambit. They hope that it will provoke the reader to follow up on it. Only if this happens there is a real chance that the various kinds of relations between logic and argumentation tentatively indicated in this volume can develop into a real bond.

Developments in Argumentation Theory

Frans H. van Eemeren and Rob Grootendorst,
University of Amsterdam

Abstract

In this paper, a survey is provided of the state of the art in argumentation theory. Some of the most significant approaches of the past two decades are discussed: Informal Logic, the formal theory of fallacies, formal dialectics, pragma-dialectics, Radical Argumentativism, and the modern revival of rhetoric. The survey is based not only on books, but also on papers published in professional journals or included in conference proceedings.

1. Introduction

Argumentation is a speech act complex aimed at resolving a difference of opinion. According to a prominent handbook definition, it is a verbal and social activity of reason carried out by a speaker or writer concerned with increasing (or decreasing) the acceptability of a controversial standpoint for a listener or reader; the constellation of propositions brought to bear in this endeavour is intended to justify (or refute) the standpoint before a rational judge.[1] *Argumentation theory* is the name given to the (systematic results of the) study of this discourse phenomenon. Argumentation theory studies the production, analysis and evaluation of argumentation with a view of developing adequate criteria for determining the validity of the point of departure and presentational layout of argumentative discourse.

The constellation of propositions advanced in argumentation is often referred to by the term *argument*, particularly by logicians and philosophers. This may lead to confusion because (in English) the word 'argument' has various meanings. Apart from (a) a reason and (b) a logical inference of a conclusion from one or more premises, 'argument' can also denote (c) a discussion and (d) a quarrel. In order to avoid ambiguity, O'Keefe (1977) distinguishes between arguments in sense (a), (c) and (d), but for the purposes of argumentation theory it is the obscuring of (a) and (b) that causes most confusion. It blurs the distinction between the logical and the pragmatic aspects of argumentative discourse.[2]

---

[1] See van Eemeren et al. (1996: 5).

[2] This distinction plays a vital part in the reconstruction of unexpressed premisses, the classification of argumentation schemes and the analysis of argumentation structures. See van Eemeren and Grootendorst (1992: 60-62).

Argumentation theory has a descriptive as well as a normative dimension. It is descriptive because it investigates the practice of argumentative discourse empirically; it is normative because it reflects critically on the reasonableness of that discourse. Normative theorists, such as those inspired by logic and philosophy, concentrate on the criteria that need to be satisfied in reasonable argumentation. Descriptive theorists, who often have a background in discourse analysis or social psychology, examine how argumentation is used to convince or persuade the interlocutors or readers. It is the divergence of normative and descriptive approaches to argumentative discourse--and the ensuing controversies[3]--that creates another source of confusion in argumentation theory. For the purposes of argumentation theory, both descriptive and normative insights are indispensable. A fully-fledged argumentation theory therefore requires a comprehensive research programme that integrates the descriptive dimension and the normative dimension.[4]

This chapter provides a survey of the state of the art in argumentation theory by describing some of the major developments that have taken place in the past two decades.[5] It is based not only on books, but also on papers published in professional journals or included in conference proceedings. Starting in the late seventies argumentation has become a subject of interest to scholars in a growing number of disciplines, and the number of publications on argumentation has risen accordingly. There has also been a considerable increase of professional journals, argumentation conferences and organizations devoted to the study of argumentation.

The most important argumentation journals published in English are the *Journal of the American Forensic Association* (since 1954), continued as *Argumentation and Advocacy* (since 1988), *Informal Logic* (since 1978), and *Argumentation* (since 1987). Argumentation conferences are generally held under the auspices of the International Society for the Study of Argumentation (*ISSA*), the Association for Informal Logic and Critical Thinking (*AILACT*), the Ontario Society for the Study of Argumentation (*OSSA*), and the American Forensic Association (*AFA*), which is part of the Speech Communication Association (*SCA*). Conferences are also often organized by specific universities or included in broader conferences on philosophy, linguistics, speech communication, law, or some other discipline.

---

[3] Extreme normativists frequently combine their position with a rejection of a relativist stand, whereas extreme descriptivists defend it. Among the former are Siegel (1982) and Biro and Siegel (1992), among the latter is Willard (1989).

[4] For a research programme that encompasses both the descriptive and the normative dimension of argumentation theory, see van Eemeren and Grootendorst (1992) and van Eemeren et al. (1993).

[5] For a more comprehensive survey, see van Eemeren et al. (1996).

2. Toulmin's model and Perelman's new rhetoric

The study of argumentation was for a long time dominated by the--still influential--contributions of Toulmin and Perelman. Their approaches are in both cases characterized by the attempt to provide an alternative to formal logic that is more suitable for dealing with everyday argumentation in ordinary language. Toulmin's *The uses of argument* (1958) presents a model of the various elements constituting an argumentation ('claim', 'data', 'warrant', etc.).[6] Perelman and Olbrechts-Tyteca's *La nouvelle rhétorique* (1958/1969) provides an inventory of effective argumentation techniques ('quasi-logical argumentation', 'argumentation based on the structure of reality', 'argumentation estabishing the structure of reality', etc.).[7]

In Toulmin's model as well as in Perelman's new rhetoric the rational procedures of judicial reasoning are taken as the starting point. In neither case, however, the aim the authors set out to tackle has been truly achieved. This may, at least partly, be due to Toulmin's and Perelman's limited views of logic. They conveniently identify logic with traditional syllogistic logic. Modern developments are largely ignored, or could--as in the case of dialogue logic--not yet be taken into account. Another inadequacy is that no justice is done to the fact that argumentation is primarily a discourse phenomenon, which is always embedded in a specific contextual and social environment. In order to study argumentation adequately, it must be viewed as a form of linguistic action that is to be approached pragmatically. The neglect of insights from both (dialogue) logic and (pragma-) linguistics has been an impediment to the development of a sound theory of argumentation. It may explain why neither Toulmin's model nor Perelman's new rhetoric offers a satisfactory alternative to formal logic, however inspiring these contributions to the study of argumentation may be.

[6]   For an elaborate discussion of Toulmin's model, see van Eemeren et al. (1996: 129-160). See also Hample (1977) and Healy (1987). Less critical are Burleson (1979) and Reinard (1984).

[7]   For an elaborate discussion of Perelman and Olbrechts-Tyteca's new rhetoric, see van Eemeren et al. (1996: 93-128). See also Ray (1978), *JAFA*'s special issue edited by Dearin (1985), Golden and Pilotta (ed., 1986), and Corgan (1987).

3. Informal Logic

Informal Logic is a movement, originating in North America in the early seventies, which grew out of dissatisfaction with the usual treatment of argumentation in introductory logic textbooks. It is inspired by the works of Toulmin and Perelman and by the ideas of some other dissenting philosophers. Since 1978, the voice of this movement has been the journal *Informal Logic*, edited by Blair and Johnson.[8]

Although the name suggests otherwise, Informal Logic is not a new kind of logic. It is rather a normative approach to argumentation in everyday language that is broader than formal logic. According to the informal logicians, the validity and cogency of argumentation is not identical to formal validity in deductive logic. Blair and Johnson (1987a) argue that the premises for a conclusion must satisfy three criteria: (1) 'relevance', (2) 'sufficiency', and (3) 'acceptability'.[9] With relevance, the question is whether the contents of the premises and the conclusion are adequately related; with sufficiency, whether the premises provide enough evidence for the conclusion; with acceptability, whether the premises are true, probable or otherwise reliable.

The informal logicians' objective is to develop norms, criteria and procedures for the interpretation, evaluation and construction of argumentation. The problems for which solutions are sought are largely the same as in other approaches: how to analyze argumentation structures, how to classify argumentation schemes, how to assess argumentation, how to identify fallacies, how to conduct a discussion, et cetera. In its present state, Informal Logic is a comprehensive research programme rather than an elaborated theory of argumentation.

There is a striking overlap between the aims and scope of Informal Logic and those of pragma-dialectics (see section 6). An important difference is that Informal Logic concentrates primarily on the relation between premises and conclusions, while pragma-dialectics pretends to cover all aspects and stages of a critical discussion. Another difference is that Informal Logic studies various uses of reasoning whereas pragma-dialectics focuses on reasoning that is directed at re-solving differences of opinion. A third difference stems from the distinct theoretical backgrounds: in analysing argumentation, Informal Logic is geared to disclosing the logical qualities of argumentative discourse, and pragma-dialectics to examining the pragma-linguistic properties that can be taken into account in its reconstruction.

---

[8]   For the object and research programme of Informal Logic, see Blair and Johnson (1987b). A brief historical survey is provided in Johnson and Blair (1980). See also van Eemeren et al. (1996: 163-188).

[9]   These criteria were first discussed in Johnson and Blair (1977). Although the labels are not always the same, the three criteria have been adopted by Govier (1985), Damer (1987), Freeman (1988), and others.

4. Formal theory of fallacies

A substantial contribution to the study of argumentation, concentrating on the fallacies, has been made by the Canadian logicians Woods and Walton (1989). Their formal approach is exhibited in a series of jointly and independently authored articles and books. Many of their co-authored papers are collected in *Fallacies: Selected Papers, 1972–1982* (1989). The basic principles of their approach of the fallacies are explained in their textbook *Argument: The Logic of the Fallacies* (1982).[10]

The Woods-Walton approach to the fallacies is *pluralistic*: in their opinion, it makes no sense to suppose that all fallacies must be given the same kind of analysis. Their general methodological view is that fallacies are usefully analyzed with the help of the structures and the theoretical vocabulary of various logical systems, including systems of dialectical logic. This does not mean that they take a fixed position on whether fallacies are inherently logical. In Woods and Walton's view, this will vary depending on the fallacy in question. It is their claim that at least a great many fallacies can best be analysed in a way that can in some sense be qualified as *formal*.

In their analysis of fallacies, Woods and Walton draw upon Hamblin's dialectical concepts of 'commitment set' and 'retraction' as methodological tools. Thus their analysis of the fallacies is formally oriented, but also dialectical. They tend to organize the many fallacies they have recognized in their writings into three grades of 'formality'. First, there are those fallacies (such as the fallacy of four terms) which are formal in the strict sense. At the next grade of formality come those fallacies (such as the fallacies of ambiguity) that are not formal in the strict sense, but whose commission is at least partly made explicable by reference to logical forms. Much more prominently realized in Woods and Walton's work is a third grade of formality that applies to theories whose key concepts are analyzable using the vocabulary and concepts of a system of logic or some other formal system.

Woods and Walton (1982) emphasize the theoretical importance of characterizing fallacies as features of arguments in actual use. A pragmatic feature of their approach is that it admits many different contexts or frameworks in which argumentation could be used. In theory, all these contexts or frameworks should be definable under the general rubric of a structure of dialogue where the participants, moves, locations, commitments, and other factors that define the dialogue exchange, are clearly and precisely defined.

---

[10]   See also Walton (1982). For a more elaborate discussion of Woods and Walton's theoretical position, see van Eemeren et al. (1996: 213-245).

5. Formal dialectics

Formal logic was given an important dialectical turn by Barth and Krabbe in *From axiom to dialogue* (1982). Building upon Lorenzen's dialogue logic, they described a formal procedure to check whether a given thesis can be logically maintained in light of certain assumptions. This dialectical interpretation of logic is known as 'formal dialectics'.[11]

In formal dialectics reasoning is viewed as a dialogue between a proponent and an opponent of a certain thesis. Together the proponent and the opponent attempt to find out whether this thesis can be successfully defended against critical attacks. In his defence, the proponent of the thesis can make use of the opponent's 'concessions': statements that the opponent is prepared to take responsibility for. The proponent must parry any attack on one of his own statements. In this endeavour, he can either give a direct defence or undertake a counterattack on a concession by the opponent.

The opponent is obliged to defend any concession that comes under attack. If this would result in him being unable to do anything other than assert something that he had attacked earlier in the dialogue, this would benefit the proponent. Therefore, the proponent attempts to manoeuvre the opponent into this position by cleverly using the opponent's concessions. If he is able to do so, according to the rules of the game, the proponent has successfully defended his position, thanks to the opponent's concessions, hence *ex concessis*.

The discussion envisioned in formal dialectics differs fundamentally from ordinary argumentative practice. The assumed starting-point can only occur after a party in a discussion has already presented his argumentation in defence of a standpoint. It arises if he and the other party then decide to discover whether this standpoint can be maintained in the light of the argumentation. The parties then initiate a procedure to check whether the standpoint can be logically concluded from the premises that have been presented in the argumentation. If the other party is indeed prepared to take on the role of opponent, he needs to add the proponent's argumentation as a set of concessions to his own commitments.

6. Pragma-dialectics

There are a number of direct links between formal dialectics and 'pragma-dialectics', the theory of argumentation developed by van Eemeren and Grootendorst (1984, 1992). As indicated by the joint use of the term *dialectic*, the general objective is in both cases the same. The theoretical orientation of pragma-dialectics, however, is different from that of formal dialectics. This difference is clearly

---

[11]   For an explanation of Lorenzen's dialogue logic and Barth and Krabbe's formal dialectics, see van Eemeren et al. (1996: 246-273).

expressed in the choice of the prefix *pragma(tic)* rather than *formal*. Pragma-dialectics is primarily a theory of argumentative discourse, not a theory of logic.[12]

Unlike the formal dialectical rules for generating rational arguments, the pragma-dialectical rules for resolving a difference of opinion are envisaged as representing necessary conditions for carrying out a critical discussion in argumentative discourse. In *Speech acts in argumentative discussions* (1984), van Eemeren and Grootendorst introduced a code of conduct for resolving differences of opinion in a reasonable way. In *Argumentation, communication, and fallacies*, they summarized this discussion procedure in ten basic rules: the "Ten Commandments" of a critical discussion (1992: 208-209).[13]

In a critical discussion, one language user (the 'protagonist') expresses a standpoint and another language user (the 'antagonist') expresses doubt with respect to this standpoint or advances a contradictory standpoint. The protagonist defends his standpoint by putting forward argumentation, and if confronted with critical reactions, further argumentation to support his prior argumentation. The difference of opinion is resolved when either the antagonist is convinced by the protagonist's argumentation and accepts the defended standpoint or the protagonist withdraws his standpoint as a result of the antagonist's critical reactions.

Analytically, four stages are distinguished in the conduct of a critical discussion: defining the difference of opinion ('confrontation' stage), establishing the starting-point of the discussion ('opening' stage), exchanging arguments and critical reactions in order to resolve the difference ('argumentation' stage), and determining the result of the discussion ('concluding' stage).[14] At every stage of a discourse aimed at bringing about a critical discussion, specific obstacles may arise that can impede the resolution of the difference of opinion. The pragma-dialectical rules are designed to prevent such obstacles from arising; they provide a definition of the general principles of constructive argumentative discourse. Supposedly, obeying all the rules obviates the obstacles which are traditionally known as 'fallacies'.[15]

---

[12] As for its dialectical starting-point, pragma-dialectics has been inspired by insights from Crawshay-Williams (1957), Popper (1972, 1974) and Barth and Krabbe (1982); its pragmatic theoretical orientation is based on insights from Austin (1962), Searle (1969, 1979) and Grice (1975). For a more elaborate discussion of pragma-dialectics, see van Eemeren et al. (1996: 274-311).

[13] A precondition for resolving a difference by means of a critical discussion is that the appropriate 'higher order' conditions have already been met. See van Eemeren and Grootendorst 1988: 287-288) and van Eemeren et al. (1993: 30-34).

[14] See van Eemeren and Grootendorst (1992: 35).

[15] This is why the rules are purported to be 'problem-valid'; see Barth and Krabbe (1982: 21-22). For a discussion of the fallacies as violations of pragma-dialectical discussion rules, see van Eemeren and Grootendorst (1992: 102-207).

A crucial difference between the pragma-dialectical rules and the rules of formal dialectics is that the former are linked to ordinary discussions in everyday language. Their scope extends over all aspects of a critical discussion, inclusive of the logical inference relations between premisses and conclusions. The rules cover all speech acts performed in all stages of a discourse aimed at resolving a difference of opinion. In *Reconstructing argumentative discourse*, van Eemeren, Grootendorst, Jackson and Jacobs (1993) have shown how the model of a critical discussion can be applied to the analysis of argumentative discourse as it occurs in various kinds of practices.[16]

7. Radical Argumentativism

In the seventies, the French linguists Ducrot and Anscombre started to develop a linguistically-oriented approach to argumentative discourse. They label this approach "Radical Argumentativism", because in their view every form of language use has an argumentative aspect (Anscombre and Ducrot 1986). The outlines of Radical Argumentativism have been presented in *Les échelles argumentatives* by Ducrot (1980), *L'argumentation dans la langue* by Anscombre and Ducrot (1983), and *Le dire et le dit* by Ducrot (1984).[17]

Ducrot and Anscombre's basic idea is that every piece of discourse contains an explicit or implicit dialogue. They describe how 'argumentative connectors' (such as *but*, *even*, and *at least*) and 'argumentative operators' (such as *only*, *no less than*, and *very*) give specific 'argumentative power' and 'argumentative direction' to the discourse by activating a certain *topos*.[18] According to Ducrot and Anscombre's theory of 'many-voicedness' or 'polyphony', argumentative connectors such as *but* can be responsible for a conflicting argumentative direction since they create a silent second voice which reveals the structural presence of two incompatible conclusions.

In 'That book is fantastic, but it is hard to understand', for example, the listener may conclude on the basis of the first part of the sentence that it would be wise to read the book; on the basis of the second part, he might conclude that this is not so wise. The opposing conclusions suggest different 'argumentative principles' or

---

[16]   In this endeavour, van Eemeren et al. explain which transformations need to be performed in order to deal with digressions and repetitions, to do justice to implicit and indirect speech acts, et cetera. Jackson and Jacobs have also made an important contribution to the study of conversational argument in it's own right. Their publications include, for example, Jackson and Jacobs (1980, 1989).

[17]   Only lately Ducrot and Anscombre's theory has become somewhat better known outside the French-speaking world, due to articles in English such as Lundquist (1987), Verbiest (1991) and Nølke (1992). For a more elaborate discussion of Radical Argumentativism, see van Eemeren et al. (1996: 312-321).

[18]   For a brief explanation of the meaning of the term *topos* in classical dialectic and rhetoric, see van Eemeren et al. (1996: 37-50).

topoi: 'The more fantastic a book is, the more reason there is to read it' and 'The less understandable a book is, the more reason there is not to read it'. The use of argumentative operators can have the same effect. Compare the sentence 'The ring costs only one hundred dollars' with the sentence 'The ring costs no less than one hundred dollars'. In a certain context, the first sentence can point to the conclusion 'Buy the ring', the second to the conclusion 'Do not buy the ring'. In the first sentence, the argumentative operator *only* activates the *topos* 'The cheaper the ring is, the more reason there is to buy it'; in the second sentence, the argumentative operator *no less than* activates the *topos* 'The more expensive a ring is, the more reason there is not to buy it'.

In the field of argumentation theory, Ducrot and Anscombre's view that 'argumentativity' is a feature of all language use is not generally accepted: argumentation is usually seen as a special form of discourse with a specific communicative and interactional function. Another distinctive feature of Ducrot and Anscombre's Radical Argumentativity is that it is not aimed at developing norms and criteria for the evaluation of argumentation. Its aim is exclusively descriptive: providing a description of the syntactic and semantic elements that play a role in the argumentative interpretation of sentences.[19]

8. Modern revival of rhetoric

Over the past few years, a powerful revaluation of classical rhetoric has been in progress. It has become accepted in the professional literature that the a-rational--sometimes anti-rational--image of rhetoric must be revised. More or less as a consequence, the sharp opposition to dialectics should be moderated too. A number of authors claim that rhetoric as the study of effective techniques of persuasion is not incompatible with the critical ideal of reasonableness upheld in dialectics. Others maintain that there are fundamental differences between a rhetorical and a dialectical conception of reasonableness, but see no reason to regard the rhetorical conception as inferior to the dialectical conception.

The rehabilitation of rhetoric goes together with a general acknowledgement that the non-rhetorically oriented theories of argumentation are saturated with insights from classical rhetoric. It is striking that the rise of rhetoric has progressed almost simultaneously in different countries.[20] The survey *Contemporary perspectives on rhetoric* by Foss, Foss and Trapp (1985) discusses most of the works that have contributed significantly to the resurgence of rhetoric in the United States. Farrell (1977) and McKerrow (1977, reprinted 1992) in particular have defended the

---

[19] Ducrot and Anscombre's theory has been the basis for carrying out empirical research regarding the interpretation of sentences. See, for example, Bassano (1991) and Bassano and Champaud (1987a, 1987b, 1987c).

[20] For a more elaborate discussion of recent developments in the rhetorical approach to argumentation, see van Eemeren et al. (1996: 189-212, 345-349).

rational qualities of rhetoric. Rhetoric is also given its due by Wenzel (1980, reprinted 1992), but emphatically in relation to logic, and primarily dialectics.

In France, it is first of all Reboul who is responsible for giving rhetoric a fully-fledged position in the study of argumentation. In 'Can there be non-rhetorical argumentation?' (1988) he discusses the rhetorical characteristics of argumentation: its formulation in ordinary language, its orientation to an audience, the probability (at best) of its premisses, the lack of logical necessity in the connection between its premisses and its conclusion. Although Reboul (1990) regards rhetoric and dialectic as two different disciplines, they do exhibit some common traits. Rhetoric is dialectic applied to discussions of social issues; at the same time, dialectic is part of rhetoric, because it provides rhetoric with its intellectual instruments.

In Germany, Kopperschmidt goes a step further. In a sequel to an exploratory article on the relation between rhetoric and argumentation theory (1977), he contends that rhetoric is *the* subject of research in argumentation theory. This agrees with the historical view of rhetoric (ed., 1990). The Austrian Kienpointner (1991b) offers an even more radical revaluation of rhetoric: he defends a relativistic conception of reasonableness and contends that rhetoric constitutes the most productive instrument for resolving social dissension.

In the Netherlands, Braet has been active in stimulating rhetoric. In *De klassieke statusleer in modern perspectief* ('Classical theory of status in modern perspective'), he emphasizes the importance of the classical theory of status for modern argumentation studies (1984). He illustrates his point by comparing this classical theory with the theory of stock issues in American academic debate.

9. Other significant approaches

Each of the approaches to argumentation discussed up to this point has been explored in a comprehensive research programme. Other recent contributions to the study of argumentation may be equally interesting, but they are less focused on developing a general theory of argumentation, more limited in scope, less elaborated, or not accessible in English. To conclude our survey, we shall mention a few.

First, there are the Swiss logician Grize and his colleagues Borel, Miéville, Apothéloz and others, who have been developing a theory of 'natural logic' at the Centre de recherches sémiologiques of the University of Neuchâtel.[21] Their main motive has been dissatisfaction with formal logic. Natural logic is designed for everyday discourse as it manifests itself in advertisements, political addresses, et cetera. Without assuming any a priori normative concepts of 'truth' and 'validity', natural logic aims to expose the 'logic' of such argumentative texts. The term *logic* here refers to the commonplaces (*topoi*) and rules used in everyday argumentation

---

[21]   See Grize (1982), Borel, Grize and Miéville (1983), Borel (1989), and Maier (1989). For a more elaborate discussion of natural logic, see van Eemeren et al. (1996: 322-328).

and reasoning, not to the formal logical systems for deductively valid reasoning. Relying on abstract 'schematizations' of persuasive forms of presentation and on discursive logical operations instrumental in the creation or elimination of contradiction or inconsistency, natural logic gives a description of argumentative language use.

Second, there is the Unité de Linguistique Française at the university of Geneva, a research group of Francophone Swiss pragma-linguists. Since the beginning of the eighties they have devoted themselves to giving pragmatic descriptions of French markers (pragmatic connectives, modal adverbs, illocutionary verbs) within a general model of discourse structure. The group includes Auchlin, Egner, Luscher, Perrin, Moeschler, (Anne) Reboul, Roulet, Schelling, and de Spengler. Their pragmatic studies are influenced by speech act theory, Ducrot and Anscombre's Radical Argumentativism, and Goffman's symbolic interactionism. An essential characteristic of their approach is that speech acts are not examined in isolation, but in their relations with other speech acts in a discourse. Making use of Sperber and Wilson's theory of relevance, they have recently added a cognitive component to the "Geneva model", which distinguishes between different levels of the discourse, describes the relations between these levels, and indicates which linguistic markers may be indicative of the various relations.[22]

Third, there is the theory of 'problematology', developed in the early eighties by the Belgian philosopher Meyer both in order to solve philosophical problems and as a model for argumentation.[23] In his skeptical attitude towards formal logic, Meyer shows himself a true disciple of Perelman, his teacher at Brussels Free University. According to Meyer, the function of argumentative discourse is, on the one hand, to provide an answer to a specific problem in a specific context. On the other hand, argumentation can also be seen as the 'problematizing' of an answer; that is, as the recognition of the question contained in a given answer. In non-formal reasoning there is no guarantee that a posed question will not remain an open question, and final answers are not to be expected: they can only be given in the formal language of a logic in which there is no room for doubt or contradictory propositions. In problematology, there is only room for a non-formal logic governing 'nonconstraining reasoning' (1986a: 130-131).

Fourth, there is the German argumentation tradition.[24] Its most prominent representative is Kopperschmidt, whose normative approach to argumentation combines insights from classical rhetoric with insights from speech act theory, text

---

[22]   See, for example, Roulet et al. (1985), Moeschler (1982, 1989a, 1989b), Anne Reboul (1988), Luscher (1989), and also van Eemeren et al. (1996: 35-351).

[23]   See Meyer (1982a, 1982b (English translation 1986a), 1986b) and van Eemeren et al. (1996: 343-344).

[24]   For a survey, see Kienpointner (1991a). See also van Eemeren et al. (1996: 341-343, 347-348, 350, 354-355).

linguistics, and Habermas's theory of communicative rationality.[25] Habermas's influence is equally apparent in the work of the linguistically and descriptively oriented German theorists who attempt to apply speech act theory and conversation analysis to spoken and written argumentative discourse.[26] Their work has been strongly influenced by Toulmin too.[27] Another German contribution to the development of argumentation theory is the dialogue logic of the Erlangen School of Lorenzen *cum suis*, which is fundamental to Barth and Krabbe's formal dialectics and has already been mentioned in section 5.[28]

Fifth, there is the richly varied American tradition in the field of speech communication, with prominent scholars such as Willard, Zarefsky, and Goodnight. Willard has developed a social-epistemological approach to argumentation based on insights from phenomenology, symbolic interaction and constructivism. In his view, argumentation is a form of conversation ensuing from differences of opinion; the interaction between arguments is a source of human knowledge.[29] By other American rhetoric and communication scholars argumentation is approached with divergent interests. A useful survey of the main contributions to the various areas is offered by Benoit, Hample and Benoit in *Readings in argumentation* (eds., 1992). This collection includes classical articles by (Pamela) Benoit, (William) Benoit, Brockriede, Burleson, Ehninger, Gouran, Gronbeck, Hample, Jackson, Jacobs, Kneupper, McKerrow, (Daniel) O'Keefe, Rowland, Trapp, Wallace, Wenzel, Willard, and Zarefsky. Some of these names have already appeared in earlier sections of this survey.

Last but not least, there are numerous authors who, from various theoretical starting points, have given special attention to specific topics: validity, unexpressed (or implicit) premises, argumentation schemes, argumentation structures, fallacies, relevance, cognitive processing of argumentative discourse, acquisition of argumentative skills, teaching of argumentative skills, conversational argument, field-dependent argumentation, and intercultural argumentation. On most of these topics vast numbers of publications have appeared.

---

[25]   See Kopperschmidt (1978, 1980, 1989). For a brief discussion in English of Kopperschmidt's ideas, see Kopperschmidt (1985, 1987) and van Eemeren et al. (1996: 342-343).

[26]   For the remarkable influence of Habermas on speech communication in the United States, see the special issue of the *Journal of the American Forensic Association* (1979), with contributions from Burleson, Farrell, and Wenzel. See also Doxtader (1991).

[27]   For Habermas's influence, see, for example, Berk (1979); for Toulmin's influence, Göttert (1978), Quasthoff (1978), Völzing (1979), Öhlschläger (1979), and Kienpointner (1983).

[28]   For an introduction to dialogue logic, see Lorenzen and Lorenz (1978) and van Eemeren et al. (1996: 253-262).

[29]   See Willard (1979a, 1979b, 1983, 1989), and also van Eemeren et al. (1996: 197-198).

10. Epilogue

Without making any claim to being exhaustive, we think that the following tendencies are worth noting in the study of argumentation during the last two decades:

(1) A growing interest in developing a fully-fledged theory of argumentation among scholars from a variety of disciplines. This interest manifests itself internationally in publications by philosophers and logicians, rhetoric and communication scholars, linguists and discourse analysts, lawyers, psychologists and other social scientists. It is accompanied by an increased awareness of the need for multidisciplinary and interdisciplinary collaboration.

(2) A keen interest in the prospects that recent developments in formal logic, especially in dialogue logic, may have to offer for the study of argumentation. This interest is usually coupled with a realistic appreciation of the limitations of a formal approach when dealing with the peculiarities of genuine argumentative discourse. It expresses itself in various kinds of dialectical approaches to argumentation that have resulted in several models for analysing argumentative discourse.

(3) A spectacular revaluation of the importance of rhetoric for the study of argumentation. This revaluation has led to the realisation that a dialectical approach to argumentation, even if it is primarily normative, and a rhetorical approach, even if it is primarily seen as descriptive, need not necessarily be at loggerheads. Since it has become clearer to many that rhetoric is not by definition related to effective though often irrational persuasion techniques, the idea has gained ground that studying rhetoric may lead to beneficial insights concerning the reasonableness of argumentation.

(4) An increased empirical interest in how argumentative discourse is conducted in various kinds of argumentative practices or fields. Such practices can be highly institutionalised, as in the case of most judicial argumentation, but they can also be more or less informal. Paying special attention to the linguistic means that are brought to bear in the execution of specific argumentative practices, detailed studies have been undertaken or are being undertaken into the characteristic features of academic discussions, mediation talks, policy making and negotiation. Besides field-dependent argumentative conventions, the intercultural differences in argumentation styles seem to have become another focus of attention.

(5) A renewed interest in "old" theoretical concepts such as 'relevance' and the 'fallacies', which are crucial to the development of an adequate theory of argumentation. Argumentation theory can only live up to its practical ambitions, if a satisfactory treatment can be given of these concepts. After Hamblin's devastating critique of the logical "Standard Treatment" of the fallacies, various new approaches have evolved. These approaches tend to be much broader in scope than the logical Standard Treatment and they are usually in a dialectical vein.

References

Anscombre, J.-C. and O. Ducrot (1983). *L'argumentation dans la langue*. Bruxelles: Mardaga.
Anscombre, J.-C. and O. Ducrot (1986). 'Argumentativité et informativité'. In: Meyer (ed.) (1986), 79-94.
Austin. J.L. (1962). *How to do things with words*. London etc.: Oxford University Press.
Barth, E.M. and E.C.W. Krabbe (1982). *From axiom to dialogue: a philosophical study of logics and argumentation*. Berlin etc.: Walter de Gruyter.
Bassano, D. (1991). 'Opérateurs et connecteurs argumentatifs: une approche psycholinguistique'. *Intellectia* 11: 149-191.
Bassano, D. and Chr. Champaud (1987a). 'Argumentative and informative functions of French intensity modifiers *presque* (almost), *à peine* (just, barely) and *à peu près* (about): an experimental study of children and adults'. *Cahiers de Psychologie Cognitive* 7 (6): 605-631.
Bassano, D. and Chr. Champaud (1987b). 'La fonction argumentative des marques de la langue'. *Argumentation* 1 (2): 175-199.
Bassano, D. and Chr. Champaud (1987c). 'Fonctions argumentatives et informatives du langage: le traitement des modificateurs d'intensité *au moins, au plus* et *bien* chez l'enfant et chez l'adulte'. *Archives de Psychologie* 55 (212): 3-30.
Benoit, W.L., D. Hample and P.J. Benoit (eds.) (1992). *Readings in argumentation*. Berlin etc.: Foris (PDA 11).
Berk, U. (1979). *Konstruktive Argumentationstheorie*. Stuttgart etc.: Frommann-Hotzboog (Problemata 72).
Biro, J. and H. Siegel (1992). 'Normativity, argumentation and an epistemic theory of fallacies'. In: van Eemeren, Grootendorst, Blair and Willard (eds.) (1992), 85-103.
Blair, J.A. and R.H. Johnson (eds.) (1980). *Informal logic. The first international symposium*. Point Reyes, CA: Edgepress.
Blair, J.A. and R.H. Johnson (1987a). 'Argumentation as dialectical'. *Argumentation* 1 (1), 41-56.
Blair, J.A. and R.H. Johnson (1987b). 'The current state of informal logic'. *Informal Logic* 9 (2/3), 147-151.
Borel, M.-J. (1989). 'Comment dit-on d'une *logique* qu'elle est *informelle*?' In: M.-J. Reichler (ed.) (1989).
Borel, M.-J., J.-B. Grize and D. Miéville (1983). *Essai de logique naturelle*. Berne etc.: Peter Lang (Sciences pour la Communication 4).
Braet, A. (1984). *De klassieke statusleer in modern perspectief, Een historisch-systematische bijdrage tot de argumentatieleer*. [Classical theory of status in modern perspective. A historical-systematic contribution to the theory of argumentation]. Groningen: Wolters-Noordhoff.
Burleson, B.R. (1979) 'On the analysis and criticism of arguments: some theoretical and methodological considerations'. *Journal of the American Forensic Association* 15: 137-147.
Cole, P. and J.L. Morgan (eds.) *Syntax and semantics 3: Speech acts*. New York: Academic Press, 45-58.
Corgan, V.C. (1987). 'Perelman's universal audience as critical tool'. *Journal of the American Forensic Association* 23 (3): 147-157.
Crawshay-Williams, R. (1957). *Methods and criteria of reasoning. An inquiry into the structure of controversy*. London: Routledge & Kegan Paul.
Damer, T.E. (1987). *Attacking faulty reasoning*. 2nd ed. Belmont, CA: Wadsworth (1st ed. 1980).
Dearin, R.D. (ed.) (1985). 'A special issue to honor the memory of Chaim Perelman'. *Journal of the American Forensic Association* 22 (2): 63-95.
Doxtader, E.W. (1991). 'The entwinement of argument and rhetoric: a dialectical reading of Habermas' theory of communicative action'. *Argumentation and Advocacy* 28 (2): 51-63.
Ducrot, O. (1980). *Les échelles argumentatives*. Paris: Les éditions de minuit.

Ducrot, O. (1984). *Le dire et le dit.* Paris: Les éditions de Minuit.

Dijk, T.A. van (1985). *Handbook of Discourse Analysis,* vol. 2 (Dimensions of Discourse). London etc.: Academic Press.

Eemeren, F.H. van and R. Grootendorst (1984). *Speech acts in argumentative discussions. A theoretical model for the analysis of discussions directed towards solving conflicts of opinion.* Dordrecht etc.: Foris (PDA 1).

Eemeren, F.H. van and R. Grootendorst (1988). Rationale for a pragma-dialectical perspective. *Argumentation* 2 (2): 271-291.

Eemeren, F.H. van and R. Grootendorst (1992). *Argumentation, communication, and fallacies.* Hillsdale, N.J.: Lawrence Erlbaum.

Eemeren, F.H. van, R. Grootendorst, S. Jackson and S. Jacobs (1993). *Reconstructing argumentative discourse.* Tuscaloosa etc.: The University of Alabama Press.

Eemeren, F.H. van, R. Grootendorst, A.F. Snoeck Henkemans, J.A. Blair, R.H. Johnson, E.C.W. Krabbe, Ch. Plantin, D.N. Walton, Ch.A. Willard, J. Woods and D. Zarefsky (1996). *Fundamentals of argumentation theory. A handbook of historical backgrounds and contemporary developments.* Mahwah, N.J.: Lawrence Erlbaum.

Eemeren, F.H. van, R. Grootendorst, J.A. Blair and Ch.A. Willard (eds.) (1987a). *Argumentation: Across the lines of discipline. Proceedings of the conference on argumentation 1986.* Dordrecht etc.: Foris (PDA 3)

Eemeren, F.H. van, R. Grootendorst, J.A. Blair and Ch.A. Willard (eds.) (1987b). *Argumentation: Perspectives and approaches. Proceedings of the conference on argumentation 1986.* Dordrecht etc.: Foris (PDA 3A).

Eemeren, F.H. van, R. Grootendorst, J.A. Blair and Ch.A. Willard (eds.) (1991). *Proceedings of the second international conference on argumentation.* Amsterdam: SICSAT/ISSA (SICSAT 1A en 1B).

Eemeren, F.H. van, R. Grootendorst, J.A. Blair and Ch.A. Willard (eds.) (1992). *Argumentation illuminated.* Amsterdam: SICSAT/ISSA (SICSAT 1).

Farrell, Th.B. (1977). 'Validity and rationality: the rhetorical constituents of argumentative form'. *Journal of the American Forensic Association* 13: 142-149.

Farrell, Th.B. (1979). Habermas on argumentation theory. Some emerging topics. *Journal of the American Forensic Association* 16: 77-82.

Foss, S.K., K.A. Foss and R. Trapp (1985). *Contemporary perspectives on rhetoric.* Prospect Heights, Illinois: Waveland Press. (2nd ed. 1991).

Freeman, J.B. (1988). *Thinking logically. Basic concepts for reasoning.* Englewood Cliffs, N.J.: Prentice-Hall.

Golden, J.L. and J.J. Pilotta (eds.) (1986). *Practical reasoning in human affairs. Studies in honor of Chaim Perelman.* Dordrecht etc.: Reidel (Synthese Library 183).

Göttert, K.-H. (1978). *Argumentation. Grundzüge ihrer Theorie im Bereich theoretischen Wissens und praktischen Handelns.* Tübingen: Niemeyer (Germanistische Arbeitshefte 23).

Govier, T. (1985). *A practical study of argument.* Belmont: Wadsworth.

Grice, H.P. (1975). 'Logic and conversation'. In: P. Cole and J.L. Morgan (eds.) (1975), 45-58.

Grize, J.-B. (1982). *De la logique à l'argumentation.* Genève: Librairie Droz (Travaux de droit, d'économie, de sciences politiques, de sociologie et d'anthropologie 34).

Hample, D. (1977). 'The Toulmin model and the syllogism'. *Journal of the American Forensic Association* 14: 1-8.

Healy, P. (1987). 'Critical reasoning and dialectical argument'. *Informal Logic* 9 (1): 1-12.

Jackson, S. and S. Jacobs (1980). 'Structure of conversational argument. Pragmatic bases for the enthymeme'. *Quarterly Journal of Speech* 66: 251-265.

Jackson, S. and S. Jacobs (1989). 'About coherence'. *Communication Yearbook* 12, ed. J.A. Anderson, 146-156. Newbury Park, CA: Sage.

Johnson, R.H. and J.A. Blair (1977). *Logical self-defense*. Toronto: McGraw-Hill Ryerson. (2nd ed. 1983, 3rd ed. 1993).

Johnson, R.H. and J.A. Blair (1980). 'The recent development of informal logic'. In: Blair and Johnson (1980), 3-28.

Kienpointer, M. (1983). *Argumentationsanalyse*. Innsbruck: Verlag des Instituts für Sprachwissenschaft der Universitäts Innsbruck (Innsbrucker Beiträge zur Kulturwissenschaft, Sonderheft 56).

Kienpointner, M. (1991a). 'Argumentation in Germany and Austria: an overview of the recent literature'. *Informal Logic* 13 (3): 129-136.

Kienpointner, M. (1991b). 'Rhetoric and argumentation-relativism and beyond'. *Philosophy and Rhetoric* 24 (1): 43-53.

Kopperschmidt, J. (1977). 'Überzeugen - Problemskizze zu den Gesprächschancen zwischen Rhetorik und Argumentationstheorie'. In: Schecker (ed.) (1977), 203-240.

Kopperschmidt, J. (1978). *Das Prinzip der vernünftigen Rede. Sprache und Vernunft, Teil 1*. Stuttgart etc.: Kohlhammer (Urban-Taschenbücher 289).

Kopperschmidt, J. (1980). *Argumentation. Sprache und Vernunft II*. Stuttgart etc.: Kohlhammer.

Kopperschmidt, J. (1985). 'An analysis of argumentation'. In: T.A. van Dijk (ed.) (1985), 159-168.

Kopperschmidt, J. (1987). 'The function of argumentation: a pragmatic approach'. In: van Eemeren, Grootendorst, Blair and Willard (eds.) (1987a), 179-188.

Kopperschmidt, J. (1989). *Methodik der Argumentationsanalyse*. Stuttgart: Fromann-Holzbog.

Kopperschmidt, J. (ed.) (1990). *Rhetorik. Bd. 1: Rhetorik als Texttheorie*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Lorenzen, P. and K. Lorenz (1978). *Dialogische Logik*. [Dialogue logic]. Darmstadt: Wissenschaftliche Buchgesellschaft.

Lundquist, L. (1987). 'Toward a procedural analysis of argumentative operators in texts'. In: van Eemeren, Grootendorst, Blair and Willard (eds.) (1987b), 60-69.

Luscher, J. (1989). 'Connecteurs et marques de pertinence: l'example de *d'ailleurs*'. [Connectives and markers of relevance: The case of *anyway*]. *Cahiers de Linguistique Française*, 10, 101–145.

Maier, R. (1989). 'Natural logic and norms in argumentation'. In: R. Maier (ed.) (1989), 49-65.

Maier, R. (ed.) (1989). *Norms in argumentation. Proceedings of the conference on norms 1988*. Dordrecht etc.: Foris (PDA 8).

McKerrow, R.E. (1977). 'Rhetorical validity: an analysis of three perspectives on the justification of rhetorical argument'. *Journal of the American Forensic Association* 13: 133-141. (Repr. in Benoit, Hample and Benoit (eds.) (1992), 297-311).

Meyer, M. (1982a). 'Argumentation in the light of a theory of questioning'. *Philosophy and Rhetoric* 15: 81-103.

Meyer, M. (1982b). *Logique, langage et argumentation*. Hachette.

Meyer, M. (1986a). *From logic to rhetoric* Amsterdam etc.: John Benjamins. (Pragmatics & Beyond VII:3). (English transl. of Meyer, 1982b)

Meyer, M. (1986b). *De la problématologie. Philosophie, science et langage*. Bruxelles: Pierre Mardaga (Philosophie et Langage).

Meyer, M. (ed.) (1986). *De la métaphysique à la rhétorique*. Brussel: Editions de l'Université de Bruxelles.

Moeschler, J. (1982). *Dire et contredire* [Asserting and countering]. Bern: Peter Lang.

Moeschler, J. (1989a). 'Pragmatic connectives, argumentative coherence and relevance'. *Argumentation*, 3, 321–339.

Moeschler, J. (1989b). *Modélisation du dialogue. Représentation de l'inference argumentative* [Modelling dialogue. The representation of argumentative inference]. Paris: Hermès.

Nølke, H. (1992). 'Semantic constraints on argumentation: from polyphonic micro-structure to argumentative macro-structure'. In: van Eemeren, Grootendorst, Blair and Willard (eds.) (1992), 189-200.

Öhlschläger, G. (1979). 'Beiträge zur Theorie und Praxis der Argumentation'. *Zeitschrift für Germanistische Linguistik* 7: 83-103.

O'Keefe, D. (1977). 'Two concepts of argument'. *Journal of the American Forensic Association* 13: 121-128.

Perelman, Ch. and L. Olbrechts-Tyteca (1958). *La nouvelle rhétorique: traité de l'argumentation.* Bruxelles: l'Université de Bruxelles.

Perelman, Ch. and L. Olbrechts-Tyteca (1969). *The new rhetoric. A treatise on argumentation.* Notre Dame etc.: University of Notre Dame Press. (English transl. of Perelman and Olbrechts-Tyteca, 1958).

Popper, K.R. (1972). *Objective knowledge: An evolutionary approach.* Oxford Clarendon Press.

Popper, K.R. (1974). *Conjectures and refutations. The growth of scientific knowledge.* 5th ed. London: Routledge & Kegan Paul.

Quasthoff, U. (1978). 'The uses of stereotype in everyday argument'. *Journal of Pragmatics* 2: 1-48.

Ray, J.W. (1978). 'Perelman's universal audience'. *Quarterly Journal of Speech* 64: 361-384.

Reboul, A. (1988). 'Les problèmes de l'attente interprétative: Topoi et hypothèses projectives'. [Problems of interpretative expectation: topoi and hypotheses]. *Cahiers de Linguistique Française,* 9, 87–114.

Reboul, O. (1988). 'Can there be non-rhetorical argumentation?' *Philosophy and Rhetoric* 21 (3): 220-223.

Reboul, O. (1990). 'Rhétorique et dialectique chez Aristote'. *Argumentation* 4 (1): 35-52.

Reichler, M.-J. (ed.) (1989) *Perspectives méthodologiques dans les sciences du langage.* [Methodological perspectives in the language sciences.] Bern: P. Lang.

Reinard, J.C. (1984). 'The role of Toulmin's categories of message development in persuasive communication. Two experimental studies on attitude change'. *Journal of the American Forensic Association* 20 (4): 206-223.

Roulet, E., Auchlin, A., Moeschler, J., Rubattel, C., & Schelling, M. (1985). *L'articulation du discours en français contemporain* [The organization of discourse in contemporary French]. Bern: Peter Lang.

Schecker, M. (ed.) (1977). *Theorie der Argumentation.* Tübingen: TBL Verlag Gunter Narr.

Searle, J.R. (1969). *Speech acts. An essay in the philosophy of language.* Cambridge: Cambridge University Press.

Searle, J.R. (1979). *Expression and meaning. Studies in the theory of speech acts.* Cambridge: Cambridge University Press.

Siegel, H. (1987). *Relativism refuted. A critique of contemporary epistemological relativism.* Dordrecht etc.: Reidel (Synthese Library 189).

Toulmin, St.E. (1958). *The uses of argument.* Cambridge: Cambridge University Press.

Verbiest, A. (1991). A new supply of argumentative indicators? In: van Eemeren, Grootendorst, Blair and Willard (eds.), (1991), 448-454.

Völzing, P.-L. (1979). *Begründen, Erklären, Argumentieren. Modelle und Materialien zu einer Theorie der Metakommunikation.* Heidelberg: Quelle & Meyer (Uni-Taschenbücher 886).

Walton, D. (1982). *Topical relevance in argumentation.* Amsterdam: John Benjamins.

Wenzel, J.W. (1979). 'Jürgen Habermas and the dialectical perspective on argumentation'. *Journal of the American Forensic Association* 16 (2): 83-94.

Wenzel, J.W. (1980). 'Perspectives on argument'. In: J. Rhodes and S. Newell (eds.), *Proceedings of the 1979 summer conference on argument.* Falls Church: SCA. (Reprinted in Benoit, Hample and Benoit (eds.) (1992), 121-143).

Willard, Ch.A. (1979a). 'The epistemic function of argument: reasoning and decision-making from a constructivist/interactionist point of view'. *Journal of the American Forensic Association* 15: 169-191.

Willard, Ch.A. (1979b). 'The epistemic function of argument: reasoning and decision-making from a constructivist/interactionist point of view (Part II)'. *Journal of the American Forensic Association* 15: 211-220.

Willard, Ch.A. (1983). *Argumentation and the social grounds of knowledge.* Alabama: The University of Alabama Press.

Willard, Ch.A. (1989). *A Theory of Argumentation.* Tuscaloosa etc.: The University of Alabama Press.

Woods, J., & Walton, D. N. (1982). *Argument: The logic of the fallacies.* Toronto: McGraw-Hill Ryerson.

Woods, J., & Walton, D. N. (1989). *Fallacies: Selected papers, 1972–1982.* Dordrecht/Providence: Foris Publications, PDA 9.

Logic and argumentation

Johan van Benthem,
University of Amsterdam

Abstract

Much has changed in modern conceptions of logic: logical tools and attitudes have matured, and the initial tension found with Perelman and Toulmin seems unproductive by now. In this paper, some interfaces are discussed between current developments in logic and argumentation theory. Both logic and argumentation share a common concern with the variety and fine-structure of reasoning.

Introduction

It is now about twenty years since pleasant contacts started, over a Chinese dinner table, between a small community of philosophical logicians and some incipient argumentation theorists in the Netherlands. At that fabled time of the early seventies, we were looking for common intellectual ground. In the intervening years, however, the two groups have largely gone their own way (and not unsuccessfully). The purpose of this invited lecture is to re-assess the situation, twenty years later. Much has changed, at least, in modern conceptions of logic, and it may be of interest to compare the agendas of both fields in their current state. In what follows, I will look at some broad features of human reasoning, viewed through the eyes of a contemporary logician.

The texture of argument

The early leaders of argumentation theory often operated in conscious opposition to what they considered the tradition of 'formal logic'. For instance, Perelman & Olbrechts-Tyteca (1958) claimed that the traditional logical metaphor for human argument is fundamentally mistaken. It views arguments as mathematical proofs, viz. on the analogy of a 'chain', which becomes worthless once a single link has been broken. This rigid foundationalist view could lead, e.g., Gottlob Frege to think that the discovery of one single contradiction would bring all of mathematics down 'like a house of cards'. Real argument, however, is more like a piece of cloth: it still functions when a few strands have broken and become ragged. Its strength rather lies in a web of interconnections. Thus, in contemporary jargon: real argument admits of 'graceful degradation'. This may seem a mere play with images, but e.g., Lakoff & Johnson (1980) have shown convincingly how deep metaphors determine both our ordinary and scientific thinking in many hidden ways, sometimes beneficial, sometimes quite insidious.

Of course, less chain-ful types of argument abound, even inside the exact sciences. For instance, Lakatos (1976) has shown convincingly how real-life mathematical argument is a complex mixture of proofs, refutations and re-definitions of concepts. And more globally, even mathematicians engage in cloth-like common sense argument when 'negotiating' the importance of results and creating common perspectives and research agendas – in what Withaar (1983) has called the 'context of persuasion' in science. But also, it seems fair to say that, even at a more standard formal level, current logical conceptions of reasoning have become broader. This is caused to a large extent by influences from Artificial Intelligence, where the analysis of so-called 'common sense reasoning' has become an urgent and respected task (Hayes 1979; Davis 1990). The subsequent repercussions for our understanding of Logic are slowly making their way into some of the more enlightened text books, but have not yet changed the 'standard image' of the discipline.

Incidentally, the chain metaphor is not all bad, and conservative. When we view reasoning from a Popperian point of view of refutation, rather than justification, having a chain-like system of reasoning which is easily attacked – without a refuge of vague forms of cloth-like 'half-functioning' – may be the preferable strategy for achieving critical progress.
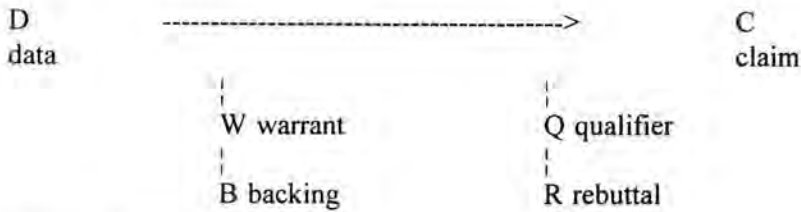
The Toulmin Schema

To demonstrate the new thinking at work, let us consider the famous 'Toulmin Schema', which has served as a rallying point for informal argumentation studies in the early seventies. In fact, Toulmin voiced three influential general criticisms of formal logic. First, reasoning is not uniform, but task-dependent: the appropriate inference mechanism may depend on the subject matter. Second, reasoning is more richly structured than the standard 'premise-conclusion' schema would have us believe. And third, what is crucial in reasoning is not the static 'form', but the dynamic 'formalities' of inferential procedure. Behind this lies a proposed paradigm shift for logic from 'mathematics' to 'law'.

Let us give away our game straightaway. By current logical lights, all three tenets in Toulmin's critical position make good sense. For instance, the dependence of human reasoning behaviour on its subject matter has been demonstrated convincingly by cognitive psychologists (cf. Wason and Johnson-Laird 1972). But there are also more internal logical reasons for appreciating the above points. Let us make this more precise, using the actual 'schema' as a convenient setting. One replaces the traditional binary view of

P   --->   C
premises   conclusion

by the following richer structure, whose various components probably speak for themselves, to a first approximation:

```
D                  -------------------------------------->          C
data                                                               claim
                   ¦                          ¦
                   ¦                          ¦
                   W  warrant                 Q  qualifier
                   ¦                          ¦
                   ¦                          ¦
                   B  backing                 R  rebuttal
```

Typology of inference

Let us first start with the role of the qualifier Q. This is the expression giving the force of the inferential transition from data to conclusion, sometimes linguistically encoded (say, by a modal adverb like "certainly" or "probably"), sometimes merely understood in context. Qualifiers can be deductive or inductive (probabilistic), or yet otherwise. This way of viewing inference is quite congenial to what has been happening in the literature on reasoning in AI. Especially, Shoham (1988) has pointed out how, in addition to classical reasoning, whose qualifier ("absolutely") says that the conclusion must hold in *all* models of the premises, there are pervasive 'preferential styles' of default reasoning. In the latter styles, the qualifier is something like "presumably", whose claim is that the conclusion holds in *all most preferred* models of the premises. Another way of describing this feature is that we are engaged in the art of reasoning 'under normal circumstances', being the most preferred cases for us to take into account. (Incidentally, this is also the art of scientific reasoning in the natural sciences!) Examples would be situations where we reason about train travel in Holland, using a mixture of logical laws in figuring out our itinerary plus default assumptions about this country, such as the absence of strikes, or the continued validity of the laws of physics. (And of course, there is always the over-riding 'mother of all defaults' in the Netherlands, prefixing every practical undertaking by the rider "assuming the dikes don't break".) Note, incidentally, that preferential reasoning is not necessarily statistical in nature: the 'most preferred' cases need not be the most frequent ones (although the two will often coincide).

Preferential reasoning differs from classical reasoning, even in its most simple domestic properties. These lie encoded in so-called 'structural rules', stateable without any reference to special logical constants. One famous structural rule which may fail now is *Monotonicity*: unlike in classical logic, preferential conclusions which follow from some set of premises need no longer follow from any extension of these premises. (Just suppose that the extension contains facts which tell us that we are in non-normal circumstances after all.) In fact, more general logics in AI are often called 'non-monotonic', a somewhat unfortunate term which emphasizes their iconoclastic character, rather than any positive virtue. (One is reminded of the now-defunct Dutch calvinist "anti-revolutionary party", which existed for one and a half century, starting from an initial program of merely opposing the principles of the French Revolution.) Another conspicuous failure of a classical structural rule is so-called non-*Transitivity*: proposition B may follow preferentially from A, and

C again from B, without C thereby being true in all most preferred models for A. (Well-known examples of transitivity failures occur in inductive logic: where exceptions may overflow any pre-set threshold in a number of steps.) As it turns out, though, preferential and classical reasoning still do agree on some familiar structural rules, such as *Permutation* of premises (their order is irrelevant to conclusions drawn) or *Contraction* (the multiplicity of occurrences of premises is irrelevant, too).

For the sake of concreteness, we list some well-known classical structural rules in their most general sequent forms:

Monotonicity $X, Y \Rightarrow A$
$$\overline{X, B, Y \Rightarrow A}$$

Transitivity $X \Rightarrow A$ $Y, A, Z \Rightarrow B$
$$\overline{Y, X, Z \Rightarrow B}$$

Permutation $X, A, B, Y \Rightarrow C$
$$\overline{X, B, A, Y \Rightarrow C}$$

Contraction $X, A, Y, A, Z \Rightarrow B$ $X, A, Y, A, Z \Rightarrow B$
$$\overline{X, A, Y, Z \Rightarrow B} \qquad\qquad \overline{X, Y, A, Z \Rightarrow B}$$

Nevertheless, there is also a basis here for a more refined positive typology of inference (cf. Makinson 1988), starting from the observation that some variants of classical structural rules do remain valid in the new setting. (Non-believers are not necessarily total rejecters.) For instance, preferential reasoning does satisfy 'Cautious Monotonicity', saying that adding already derived conclusions will not disturb inferences:

$X \Rightarrow C$ $X \Rightarrow B$
$$\overline{X, C \Rightarrow B}$$

Moreover, there is also a converse principle of 'Cautious Transitivity', telling us when indeed we can 'chain inferences':

$$X \Rightarrow C \qquad\qquad X, C \Rightarrow B$$

$$X \Rightarrow B$$

An aside. Here is a, perhaps perverse, logician's question. Can there also be structural rules that are valid for preferential reasoning, but not for classical reasoning? The answer is negative. Among all preference relations, there is the universal indifference relation, which makes all models of the premises 'most preferred'. Therefore, classical consequence amounts to preferential inference over a restricted universe of preference relations. And then, each structural rule for preferential reasoning (being a universal statement) will carry over to this subdomain, and thereby hold for classical consequence.

Another aside. A further source of refinement in the above typology is the following. Some classical structural rules may even continue to hold in their unrestricted original format, but then only for special linguistic forms of statement. For instance, in Circumscription (a popular specific system of preferential reasoning in AI; cf. McCarthy 1980), so-called 'purely universal' statements can always be added to the premises without endangering earlier conclusions.

So far, we have considered only the logical role of the qualifier Q . But there is also a modern counterpart to the rebuttal element R in Toulmin's schema. The latter is the 'rider' of the form "unless ..." which states when the qualifier admits exceptions. Similar elements have appeared in the computational literature, witness the 'abnormality predicates' in the logical formalization of circumscriptive arguments (cf. Sandewall 1992), which regulate the domain of exceptional cases.

The more general situation here suggests an agenda that can already be found, in fact, in an earlier phase of modern logic. In what may be called 'Bolzano's Program' (cf. Bolzano 1837), the aim was precisely to develop a rich typology of human styles of reasoning. Bolzano distinguished deductive and inductive varieties, as well as an especially 'strict' professional philosophical style of reasoning. Moreover, he made a sustained effort to chart the structural behaviour of these styles, including their interaction with changing vocabularies of 'fixed' and 'variable' terms – a level of refinement yet to be attained in much of the contemporary literature. Another famous logician pursuing a similar program (around 1890) is C.S. Peirce, who emphasized that humans display a variety of inferential skills, which logic should analyze and bring out (cf. the collection Peirce 1960). In particular, he distinguished both 'forward' and 'backward' styles of reasoning, which were then classified under such headings as deduction, induction and abduction. The latter is a backward process of inferring the most plausible explanation for observed facts. Another important backward reasoning process is presupposition, well-known from the linguistic and philosophical literature, which provides necessary

'preconditions' for our understanding of a sentence – whereas the usual forward reasoning rather provides 'postconditions'.

Repercussions and elaborations

The preceding general point of view has many interesting consequences. For a start, on the practical side, it will affect traditional empirical topics. Notably, it is no longer so clear what are argumentative *fallacies*. Observed inferential patterns which seem 'wrong' according to one notion of inference might just as well signal that the speaker is engaged in correct execution of another style of reasoning. E.g., take the concrete fallacy of 'affirming the consequent' (cf. Hamblin 1970). What we observe somewhere, say, is an instance of the propositional fallacy:

$$A \rightarrow B, B \text{ "and therefore" } A$$

But note that this pattern would be valid as an instance of abduction (since A is certainly the only available explanation here for B). This abductive use of 'the only available source' is also what drives logic programming in the Prolog-style (an extremely useful computational mechanism, which tends to strengthen implications to equivalences; cf. Kowalski 1979). Of course, this is not the end of the matter. The above analysis also suggests that, when confronted with 'fallacies', we extend our field of vision from observing single inferences to sequences of inferences. If the speaker is engaged in abduction, then the structural rules should not be the same as for classical logic. In particular, in this case, we do not have Monotonicity. In particular, what should not be valid, even as a specimen of abduction, is the transition:

$$C \rightarrow B, A \rightarrow B, B \text{ "and therefore" } A$$

For now, there are two possible explanations for B, and the 'best' one is rather the disjunction C-or-A. Thus, we also learn that fallacies should not be studied in isolation. Similar observations can be made about juridical reasoning (cf. Prakken 1993; Feteris 1994). For instance, there is a legal argument pattern called "a contrario", where one reasons as follows. "The law only explicitly states a penalty for male offenders. This person is a woman. Therefore, she should not be punished for this offense." Formally, we have an 'invalid' transition here 'from $A \rightarrow B$ and not-A to not-B', to which all the previous points apply.

Next, the preceding perspective also has technical consequences in logic. It is not enough to say that there exists a multitude of inferential styles in reasoning, and then rejoice. For now, the logician has acquired the task of explaining how all these styles manage to co-exist, and indeed cooperate. Thus, one needs mechanisms for combining logics (cf. Gabbay 1994), as well as 'triggers' that tell us when we are switching from one

reasoning style to another. (Here the earlier qualifiers may play a systematic role – whence we would need a more systematic logic of modal adverbs from this inferential point of view.) Here, let us just show what combination of inferential styles might involve. Assume that we have two meta-arrows ⇒ (for classical reasoning) and --> (for preferential reasoning). Then we must at least enquire into their combinations, such as:

does A --> B, B ⇒ C imply A --> C ?
            (the answer is *yes*)
does A ⇒ B, B --> C imply A --> C ?
            (the answer is *no*)
does A --> B, B ⇒ C imply A ⇒ C ?
            (the answer is *no*)
does A ⇒ B, B --> C imply A ⇒ C ?
            (the answer is *no*)

Thus, logics must now be able to manipulate and combine diverse forms of inferential information.

Parameters of inference

Another attractive feature of the Toulmin Schema is its richer structuring of the material from which conclusions are supposed to follow. This is in line with most accounts of reasoning from theories in the philosophy of science, as well as computational theories of data bases. From the binary 'premise-conclusion' pattern, one moves to a *ternary* view, where basic 'data' are distinguished from background theory:
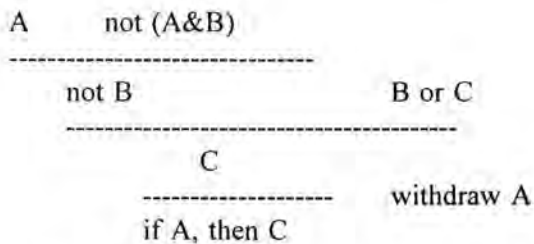
$$P \text{----------} \text{----------}> C$$
$$T$$
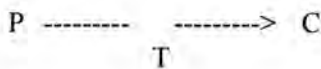
There are many examples where the third parameter T emerges naturally. For instance, in the above notion of abduction, T is indispensable for providing the available 'explanations'. This is also true more generally for scientific explanation in the Hempel-Oppenheim style, which even distinguishes further levels: 'facts', 'theoretical laws' and 'auxiliary hypotheses' (essentially, the relevant default assumptions). And the point also emerges in the linguistic study of conditionals, where the basic 'Ramsey Test' presupposes revision of some explicit 'stock of beliefs', so as to accommodate recalcitrant antecedents (Sosa 1975 collects various papers on these matters). The general situation is even more diverse, in that the third 'theory' parameter itself has inner hierarchical structure. Not all theoretical principles are equally general and important. "Structured theories" accounting for this behaviour are coming up in contemporary computer science (cf. Ryan 1992), and we may also think of the much richer structuring found in the computational literature on abstract data types (Meseguer 1989) or module algebras (Bergstra, Heering & Klint

1986). As a final thrust towards more hierarchical views of theory structure, allowing shifts in perspective, we mention Blackburn & de Rijke (1994).

Logical levels of aggregation: from propositions to proofs

The traditional field for logical analysis lies at the sentence level, where propositions are expressed. This follows standard grammatical practice in linguistics. But reasoning also involves higher levels of aggregation. Evidently, real arguments are *texts*, i.e., configurations of sentences, which shows clearly in argumentation studies (cf. the various contributions in van Eemeren & Grootendorst 1994, which mostly propose text structures). And of course, the above 'third parameter' T hints at still higher levels of organization, with configurations of texts into theories. Logic as it is does not have a well-developed theory of text structure for argumentation. Nevertheless, there is much implicit material here, once we turn to systems of logical proof and the subdiscipline of proof theory (cf. Sundholm 1986; Troelstra 1994). Let us illustrate this potential by means of a little example.

Here is a simple 'natural deduction tree' for the inference from the two premises not (A&B) and (B or C) to the conclusion (if A, then C):

```
A       not (A&B)
-----------------------------------
       not B                    B or C
       -------------------------------------------
                    C
                    --------------------  withdraw A
              if A, then C
```

The following view of argumentative texts lies behind this example. First, the structure is 'chain-like': the smallest error anywhere would invalidate the deduction. Also, explicit rule annotation is needed: we need to justify each basic step across a bar by reference to some pre-given repertoire of admissible basic steps. And finally, and very importantly, there is a dynamic pattern of changing dependencies. For instance, the intermediate conclusion C inherits the assumptions from both its ancestors 'not B' and 'B or C', three in all, but the final conclusion has lost one of these.

What of this is relevant in 'real life argument'? This is not the place to perform a detailed comparison with empirical argumentation studies, but a few things may be observed, showing the interest of such an endeavour. First, in reality, there may be a more 'cloth-like' structure, whereby one intermediate conclusion is supported by several bunches of premises. Formally, this requires AND/OR trees, rather than just AND-trees in natural deduction. Thus, we obtain 'forests', rather than trees, where conclusions can have multiple support. Such a logical system would incorporate the natural distinction

between 'subordinate' (i.e., sequential) and 'coordinate' (parallel) structures in argumentation discussed in Snoeck Henkemans (1994), with patterns like

$$
\frac{P1 \ \& \ P2 \quad \| \quad Q1 \ \& \ Q2}{C}
$$

This coordinated structure has the virtue of explaining something about our actual argumentation, namely its 'robustness'. We are seldom willing to give up a conclusion on the basis of one single problem. This need not be logical immorality or blatant self-interest. A more rational reason is again the cloth picture: that conclusion may be tied to many things supporting it. (As observed earlier on, though, a more refutation-oriented logical strategy might sometimes be preferable from a cognitive point of view – for instance, when engaging in physically dangerous endeavours.)

Next, what seems utterly lacking in reality is explicit rule annotation. The standard argumentative pattern is rather one of 'bare dependency': certain statements stand in certain 'support relations', which are seldom explicitly tagged. Thus, practical argumentative analysis uncovers a 'pre-structure' of dependencies (somewhat like 'argument structures' found in AI), operating at a level somewhat like pure grammatical constituent trees, which are still to be decorated with an appeal to specific derivation rules. (From a mathematical-logical point of view, we still have a 'type-assignment problem' ahead: cf. van Benthem 1991; Barendregt 1992.) Finally, natural deduction also makes one telling empirical point. The delicate dynamics of changing assumptions is a well-attested feature of actual human argumentation and debate. In particular, when viewed in isolation, one cannot tell what an assertion in an argument 'means': since that depends on its 'contextual burden' at the relevant stage of the argumentation. Even the same assertion may occur with different loads of assumptions in the course of one and the same argument.

Here is one more example. Consider the well-known elegant natural deduction for the propositional law $(A \leftrightarrow (A \rightarrow B)) \rightarrow A$, expressing a form of 'Löb's Paradox'. Dependencies on previous assertions are explicitly indicated at the inference bars:

$$
\begin{array}{c}
(1)\\
(A \leftrightarrow (A \to B))\\
\hline
(A \to (A \to B))
\end{array}\ (1)\qquad\qquad (2)\atop A
$$

Note the variable burdens for the occurrences of A and (A→B). Incidentally, with two more steps, this tree becomes a natural deduction for the proposition B , from just the single assumption (1). (This is the core of the mentioned 'paradox': cf. Boolos 1979.)

Discourse grammar

The preceding view of proofs as texts suggests that there is a higher level of linguistic discourse structure that may be quite relevant to logic. (Cf. Polanyi & Scha 1988, as well as the more general computational tradition of Grosz & Sidner 1986.) In particular, logical particles such as "so", "then", "unless", "although" specify various of the above-mentioned argumentative connections in texts. Moreover, there are various discourse uses of "and", signalling the earlier parallel and sequential structures. These particles will exhibit linguistic behaviour that is very similar to what happens at the sentence level. For instance, "so" is a scope-bearing operator, looking backwards from a conclusion to bring a number of previous assertions within its inferential ambit. This is why texts of the form "P1 ... Pk. Therefore C", as found in the usual discussions of argument patterns, are often ambiguous. Which of the initial assertions are in the backward scope of the operator "therefore"? Hence, one cannot draw far-reaching conclusions from untutored intuitions about such flat patterns. (This linguistic point is even relevant to discussions of potential failures of monotonicity: perhaps, in actual examples, the additional premises do not make it into the scope of the conclusion particle. See Kameyama (1993) on the topic of 'linguistic surplus information' in the analysis of puzzle solving in AI.) Similar scope behaviour is exhibited by other discourse particles, such as the 'assumer' "if", whose companion particle "then" rather functions as an anaphoric pronoun.

A more systematic study of this linguistic fine-structure may serve various purposes. For instance, one would also hope to discover explicit cues as to the 'current inferential style' being performed. Some of these cues lie in the earlier modal adverbs, but the situation can be more subtle. For instance, question-answering is often 'exhaustive' (in a Gricean sense: cf. Groenendijk & Stokhof 1984). An answer "John and Mary" to the question

"Who are dancing?" suggests that John and Mary are the only dancers. This means that we are making a preferential inference to the smallest models – in terms of individual facts– satisfying these data (cf. van Benthem 1989 on this connection with preferential reasoning). Not surprisingly again, this inference is defeasible by further premises: "And Claudia". Note that this exhaustive mode is the default, which does not need any explicit syntactic triggering. But it does seem that it can be switched off explicitly through certain linguistic (re-)formulations of our answers. For instance, the hedged reply "*at least* John and Mary", although semantically having exactly the same minimal models as the previous one, does not allow any inferences beyond the classical ones. What we learn from the latter is merely that John is dancing and Mary is – but Heaven knows who else besides.

Mechanisms of reasoning

There is another source of 'plurality' in current logical theorizing. Inference is just one of many general cognitive procedures, such as learning, updating or revising (Gärdenfors 1988). Indeed, one can discern a kind of general 'procedural turn' in recent work in artificial intelligence and linguistic semantics, emphasizing the undeniable and crucial imperative *procedural* aspects of our cognitive behaviour. And also with our present concerns, after all, much of the art in actual argumentation is sequential 'timing' and playing one's cards correctly. Now, standard logic is largely declarative, focused on static truth conditions. Thus, the new task becomes to bring these dynamic procedures within the scope of logical investigation too: focusing rather on update conditions (Groenendijk & Stokhof 1991; Kamp 1984; Stalnaker 1972; Veltman 1991; van Benthem 1991, 1994).

There are various broad technical paradigms for bringing out this dynamic structure. Traditionally, there has been the approach via games (Lorenzen & Lorenz 1979; Hintikka 1973), which continues to exist as an undercurrent in contemporary logic. It has also been the main formal face of dynamics in argumentation theory so far (cf. Barth & Krabbe 1982). But the dominant paradigm in the current logical literature comes from computer science: 'texts are programs' denoting cognitive processes that change human information states. One immediate appeal of this view to many people lies in its concrete mentalist interpretation (although its protagonists tend to be non-committal on this score). Another attractive feature is that we can now avail ourselves of the acquired expertise in computer science concerning the logical properties of procedures.

The dynamic view considerably enriches the earlier landscape of styles of inference. For instance, here is a strong contender for a notion of valid dynamic inference. 'Processing the successive premises always brings us to an information state where the conclusion should hold.' (On this view, a discourse particle "so" keys us for a change in pace from premises to conclusion: from recording to testing.) Here is a picture for this view:

Dynamic inference

• ---> • --->   ...   • ---> • ↻   C

           premises   conclusion

The loop at the end expresses that processing the conclusion will not change the information state already attained: that is, it already 'holds' there in some dynamic sense. This view of inference is quite congenial to the world of computational data base updates – which again may not be such a bad model for human reasoning either.

Like the preferential style of inference, the dynamic one loses central classical structural rules such as Monotonicity and Transitivity. This may be seen somewhat domestically by viewing the above notion as follows: the premises form a 'recipe' for achieving the conclusion. Monotonicity then says that inserting arbitrary instructions into the recipe would not change the effects previously obtained: and this is obviously implausible. But this time, there are even more dramatic divergences from classical reasoning, expressing the sequential character of imperative procedures. Permutation fails: changing the order of instructions in a recipe may produce dramatically different outcomes. And also Contraction fails. Evidently, the amount of times the same instruction is performed may matter vitally to what is produced by a recipe. Nevertheless, as before, there remains a positive typology too: dynamic inference satisfies some well-defined variants of classical Monotonicity and Transitivity, which turn out to completely determine its inferential behaviour. (There is a lot of recent work on complete proof theories for dynamic inference. Cf. Blackburn & Venema 1993; Groeneveld 1994; Kanazawa 1993.)

Again, there are many further logical repercussions of this viewpoint, which we cannot begin to enumerate here. For instance, dynamically, one has to redefine the role of the traditional 'logical constants'. These now become more like programming constructions, and can be studied using algebraic techniques from computer science, as well as from modal and so-called 'dynamic logic' (cf. van Benthem 1991, 1994). Sometimes, this makes them, say "and", "or" and "not", behave more like the above discourse particles than as the original sentence operators – but that, of course, is all to the good in a dynamic perspective on argument.

Conclusion

This brief essay by no means exhausts the potential interfaces between current developments in logic and argumentation theory. For instance, it would be of great interest to also compare actual argumentation patterns with other logical paradigms – such as the partly dynamic, partly declarative styles of reasoning formalized in logic programming (Kowalski 1979, 1989). Moreover, it might be a good idea to bring the disciplines together, not by comparing their consolidated assets, but rather by undertaking some new and challenging joint task, say the detailed exploration of juridical argumentation and procedure, using insights from both disciplines in tandem.

It seems fair to say that contemporary Logic and Argumentation Theory share a common concern with the variety and fine-structure of reasoning. Therefore, the initial tension found with Perelman and Toulmin seems unproductive by now: logical tools and attitudes have matured. Of course, such an optimistic message brings to mind commercials for detergents. The old product has totally changed, according to a 'new formula', and it is being recommended by prominent scientists and other authorities. Why should argumentation theorists buy modern logic? What is the pay-off of the new subtleties and (if the truth be told, sometimes) new complexities? I would recommend that the two communities pull their research agendas and at least begin to find out.

References

Barendregt, H. (1992). 'Lambda calculi with types'. In: S. Abramsky, D. Gabbay & T. Maibaum (eds.), *Handbook of Logic in Computer Science* (pp. 118-309), vol. 2, Oxford: Clarendon Press.

Barth, E. & E. Krabbe (1982). *From Axiom to Dialogue*. Berlin: De Gruyter.

Benthem, J. van (1989a). 'Reasoning and cognition'. In: N.O. Bernsen & H. Schnelle (eds.), *Logic and Linguistics; Research Directions in Cognitive Science* (pp. 185-208), Hove, UK: L. Erlbaum.

Benthem, J. van (1989b). 'Semantic parallels in natural language and computation. In: H-D Ebbinghaus et al. (eds.), *Logic Colloquium. Granada 1987* (pp. 331-375), Amsterdam: North-Holland.

Benthem, J. van (1991). *Language in Action. Categories, Lambdas and Dynamic Logic*. Amsterdam: Elseviers Science Publishers, (Studies in Logic 130).

Benthem, J. van (1994). *Dynamic Logic and the Flow of Information*. Institute for Logic, Language and Computation, University of Amsterdam.

Bergstra, J., J. Heering & P. Klint (1986). 'Module algebra'. Report CS-R8617, Centre for Mathematics and Computer Science, Amsterdam.

Blackburn, P. & M. de Rijke (1994). 'Zooming in - zooming out'. Centre for Mathematics and Computer Science, Amsterdam.

Blackburn, P. & Y. Venema (1993). 'Dynamic squares'. Logic Preprint 92, Department of Philosophy, University of Utrecht. To appear in *Journal of Philosophical Logic*.

Bolzano, B. (1837). *Wissenschaftslehre*. Seidelsche Buchhandlung, Sulzbach. Translated by R. George as *Theory of Science* (1972), Berkeley and Los Angeles: University of California Press.

Boolos, B. (1979). *The Unprovability of Consistency*. Cambridge: Cambridge University Press.

Davis, E. (1990). *Representation of Commonsense Knowledge*. Los Altos: Morgan-Kaufmann Publishing Co.

Eemeren, F.H. van & R. Grootendorst (eds.) (1994). *Studies in Pragma-Dialectics*. Amsterdam: SICSAT.

Feteris, E. (1994). 'Rationality in legal discussions'. In: F.H. van Eemeren & R. Grootendorst (eds.), *Studies in Pragma-Dialectics* (pp. 29-40), Amsterdam: SICSAT.

Gabbay, D. (1994). *Labeled Deductive Systems*. Department of Computing, Imperial College, London. To appear with Oxford University Press.

Gärdenfors, P. (1988). *Knowledge in Flux. Modelling the Dynamics of Epistemic States*. Cambridge, Mass: Bradford Books / MIT Press.

Groenendijk, J. & M. Stokhof (1984). *Studies on the Semantics of Questions and the Pragmatics of Answers*. Dissertation, Philosophical Institute, University of Amsterdam.

Groenendijk, J. & M. Stokhof (1991). 'Dynamic predicate logic'. *Linguistics and Philosophy* 14, pp. 39-100.

Groeneveld, W. (1994). *Logical Issues in Dynamic Semantics*. Dissertation, Institute for Logic, Language and Computation, University of Amsterdam.

Grosz, B. & C. Sidner (1986). 'Attention, intentions and the structure of discourse'. *Computational Linguistics* 12, pp. 175-204.

Hamblin, C. (1970). *Fallacies*. London: Methuen.

Hayes, P. (1979). 'The naive physics manifesto'. In: D. Michie (ed.), *Expert Systems in the Microelectronic Age*, Edinburgh: University Press.

Hintikka, J. (1973). *Logic, Language Games and Information*. Oxford: Clarendon Press.

Kameyama, M. 'The linguistic information in dynamic discourse'. Report CSLI-92- 174, Center for the Study of Language and Information, Stanford University.

Kamp, H. (1984). 'A theory of truth and semantic representation'. In: J. Groenendijk et al. (eds.), *Truth, Interpretation and Information* (pp. 1-41), Dordrecht: Foris.

Kanazawa, M. (1993). 'Completeness and decidability of the mixed style of inference with composition'. Report CSLI-93-181, Center for the Study of Language and Information, Stanford University.

Kowalski, R. (1979). *Logic for Problem Solving*. New York: North-Holland.

Kowalski, R. (1989). 'The treatment of negation in logic programs for representing legislation'. *Proceedings ICAIL-1989 Vancouver*, ACM Press, pp. 11-15.

Lakatos, I. (1976). *Proofs and Refutations*. Cambridge: Cambridge University Press.

Lakoff, G. & M. Johnson (1980). *Metaphors We Live By*. Chicago: Chicago University Press.

Lorenzen, P. & K. Lorenz (1979). *Dialogische Logik*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Makinson, D. (1988). 'General non-monotonic logic'. UNESCO, Paris. To appear in: D. Gabbay et al. (eds.), *Handbook of Logic in Artificial Intelligence and Logic Programming*, Oxford: Oxford University Press.

McCarthy, J. (1980). 'Circumscription - A form of non-monotone reasoning'. *Artificial Intelligence* 13, pp. 295-323.

Meseguer, J. (1989). 'General logics'. In: H-D Ebbinghaus et al. (eds.), *Logic Colloquium. Granada 1987* (pp. 275-329), Amsterdam: North-Holland.

Peirce, C.S. (1960). *The Collected Papers of C.S. Peirce*. Vol. IV, Ch. Hartshorne & P. Weiss (eds.), Cambridge, Mass: Harvard University Press.

Perelman, Ch. & L. Olbrechts-Tyteca (1958). *La Nouvelle Rhétorique. Traité de l'Argumentation*. Brussels: Editions de l'Université de Bruxelles.

Polanyi, L. & R. Scha (1988). 'An augmented context-free grammar of discourse'. In: *Proceedings of the 12th International Conference on Computational Linguistics* (pp. 22-27), Budapest.

Prakken, H. (1993). *Logical Tools for Modelling Legal Argument*. Dissertation, School of Law, Free University, Amsterdam.

Ryan, M. (1992). *Ordered Presentations of Theories. Default Reasoning and Belief Revision*. Ph. D. Thesis, Department of Computing, Imperial College, University of London.

Sandewall, E. (1992). *Features and Fluents*. Department of Computer and Information Science, Linköping University.

Shoham, Y. (1988). *Reasoning about Change. Time and Causation from the Standpoint of Artificial Intelligence.* Cambridge, Mass: The MIT Press.

Snoeck Henkemans, F. (1994). 'Complex argumentation in a critical discussion'. In: F.H. van Eemeren & R. Grootendorst (eds.), *Studies in Pragma-Dialectics* (pp. 69-78), Amsterdam: Sic Sat.

Sosa, E. (ed.) (1975). *Causation and Conditionals.* Oxford: Oxford University Press.

Stalnaker, R. (1972). 'Pragmatics'. In: D. Davidson & G. Harman (eds.), *Semantics of Natural Language* (pp. 380-397), Dordrecht: Reidel.

Sundholm, G. (1986). 'Proof theory and meaning'. In: D. Gabbay & F. Guenthner (eds.), *Handbook of Philosophical Logic* (pp. 471-506), vol.III, Dordrecht: Reidel.

Toulmin, S. (1958). *The Uses of Argument.* Cambridge: Cambridge University Press.

Troelstra, A. (1994). *Lectures on Proof Theory.* Institute for Logic, Language and Computation, University of Amsterdam.

Veltman, F. (1991). 'Defaults in update semantics'. Report LP-91-02, Institute for Logic, Language and Computation, University of Amsterdam. (To appear in the *Journal of Philosophical Logic.*)

Wason, P. & P. Johnson-Laird (1972). *Psychology of Reasoning.* London: Batsford.

Withaar, H. (1983). *The Context of Persuasion.* Master's Thesis, Philosophical Institute, Rijksuniversiteit Groningen.

Argumentation in the tradition
of speech communication studies

David Zarefsky,
Northwestern University

Abstract

In this paper, four broad forces are discussed that shaped the nature of argumentation studies within the speech communication discipline: the evolution of competitive debate, the infusion of empirical perspectives and methods by the social sciences, the recovery of practical philosophy, and the growing interest in social and cultural critique. Unfortunately, the growth of the discipline of argumentation is not accompanied by a clear and common sense of what is being studied. To increase coherence, a root concept for argumentation studies is proposed and explicated: argumentation as the practice of justifying decisions under conditions of uncertainty.

From debate to argumentation studies

Argumentation studies in speech communication sprang from modest roots. Late in the 19th century, as an alternative to social fraternities and athletics, American colleges and universities began competition in debate. The earliest publications were textbooks to instruct students and coaches in this new activity. The most prominent of these books was George Pierce Baker's *Principles of Argumentation*,[1] but the early books shared several common features. They were practical, how-to-do-it guides informed primarily by their authors' intuition and experience. They were unreflective, in that they treated matters of practice as neither complicated nor problematic. They paid little attention to any relationship between the species debate and the genus argumentation. And they typically did not place their instruction in a context broader than preparation for the contest activity itself.

Subsequent generations of textbooks, in the early and middle years of this century, had many of the same characteristics, but with two important qualifiers. First, they became more sophisticated in their analyses. They could rely on a growing body of experience that both codified conventional categories and permitted more textured and nuanced discussion. And, second, they began to make connections with the terms of classical rhetorical theory, particularly the concepts of common topics, issues, *stasis*, and *logos*, *ethos*, and *pathos* as modes of proof. They also revived Bishop Whately's 19th-century treatment of presumption and burden of proof. Still, they retained an emphasis on practice

---

[1]   Baker (1895).

that was fairly straightforward, without reflection on its goals, methods, and underlying assumptions. In retrospect, debate during those years has been characterized as dominated by the "stock issues" paradigm and modeled on formal logic and courtroom oratory. At the time, a term such as "stock issues" paradigm would have seemed meaningless, because that was all there was. Alternative perspectives largely escaped consideration.

The literature on debate beginning in the early 1960's represents a series of departures from this tradition. Perhaps most influential was the 1963 publication of Ehninger and Brockriede's *Decision by Debate*.[2] At least in embryonic form, this book offered a broader perspective of the debate activity. Debate was seen as a means of making decisions critically. It was described as fundamentally a cooperative rather than competitive enterprise. And it incorporated the model of argument that Stephen Toulmin had set out in *The Uses of Argument*[3] five years before. By emphasizing this model as a diagram, Ehninger and Brockriede may have reinforced a formalistic understanding of reasoning.[4] But by focusing explicitly on warrants, qualifiers, and rebuttals, they significantly undercut the analytic ideal of argument as applied formal logic. Inductive reasoning was seen not as an inferior form of logic but as the prototypical pattern of inference-making. This meant acknowledging that inferences were fallible and conclusions uncertain, and that the warrants authorizing inferences came not from logical form but from the substantive beliefs of an audience.

Subsequently, theorists of debate began to explore alternatives to the received tradition. The pages of the *Journal of the American Forensic Association* (now known as *Argumentation and Advocacy*) in the late 1960's and early 1970's are filled with articles on alternative patterns of case construction -- the comparative advantage affirmative case, the goals/criteria case, the alternative justification case -- as well as essays identifying underlying consistencies amid these seeming differences.[5] The counterplan, a negative debate strategy traditionally dismissed as weak, was revived and given theoretical anchor.[6] Writers began to focus attention on the underlying nature and goals of the process of debate itself, believing that emerging differences about theory and practice really reflected different root assumptions about debate. The late 1970's and early 1980's saw essays explicating different paradigms or models of debate -- the policy-making model, the

---

[2] Ehninger and Brockriede (1963).

[3] Toulmin (1958).

[4] Charles Willard, for example, has argued forcefully that the process of diagramming arguments fundamentally misunderstands the mix of discursive and nondiscursive elements in argument and gives too much credence to formal structure. See Willard (1976: 308-319).

[5] See, for example, Fadely (1967: 28-35); Chesebro (1968: 57-63); Chesebro (1971: 208-215); Lewinski, Metzler and Settle (1973: 458-463); Lichtman, Garvin and Corsi (1973: 59-69). An example of an essay questioning distinctions is Zarefsky (1969: 12-20).

[6] See, for example, Louis Kaplow (1981: 215-226).

hypothesis-testing model, the game-theory model, the critic-judge model, and the *tabula rasa* model, for example. The traditional perspective on debate, now renamed the stock-issues model, took its place among these alternatives.[7]

From the perspective of hindsight, this literature is not so important for its explicit content. Many of the disputes engaging debate theorists were esoteric, and many of the controversies now seem passé, not because they were solved but because they were outgrown. Rather, this phase of the debate literature is significant because it shows how conventional wisdom was rendered problematic through the imagination of alternatives. This is an important step toward developing a more reflective, self-conscious, and critical understanding of argumentation.

One of the major trends in recent writing on debate is to stress the links between debate and argumentation in general. Recognizing that debate was a specific application of more general principles, educators began to develop courses in argumentation theory and practice that were not geared specifically to debate. These courses involved larger numbers of students in the understanding of argumentation theory. To meet the needs of such courses, a new kind of textbook emerged, such as Rieke and Sillars's *Argumentation and the Decision-Making Process*, Warnick and Inch's *Critical Thinking and Communication*; and Branham's *Debate and Critical Analysis: The Harmony of Conflict.*[8] Even books oriented primarily toward debate, such as my own *Contemporary Debate,*[9] often portrayed debate as a derivative of general argumentation. This relationship was explicitly acknowledged in 1974 when the National Developmental Conference on Forensics defined forensic activities as laboratories for investigating the argumentative perspective on communication.

The linkage between debate and general argumentation has been pursued in both directions. Not only has debate drawn from an understanding of general argumentation; it also has contributed to it. To be sure, even fifty years ago one could find critical studies of legislative or political debate. Often, however, these were either simply descriptive studies or attempts to apply the principles of contest debate to situations they did not fit. Recent literature has been far more sophisticated. In 1979, my colleague Tom Goodnight delivered a paper on "the liberal and the conservative presumption," demonstrating that presumption was not just an arbitrary concept or a tie-breaking rule but a substantive concept according to which one could distinguish political positions and understand political disputes.[10] More recently, he has drawn attention to the dynamics of controversy.[11]

---

[7]   Representative articles include Lichtman and Rohrer (1980: 236-247); Zarefsky (1992: 252-262); and the special forum on "Debate Paradigms," *Journal of the American Forensic Association*, 18 (Winter, 1982), 133-160.

[8]   Rieke and Sillars (1975, in 3rd ed. 1993); Warnick and Inch (1989); Branham (1991).

[9]   Patterson and Zarefsky (1983).

[10]   Goodnight (1980: 304-337).

I do not think I am straining the concept too much to describe controversy as debate conducted over time, without *a priori* rules, boundaries, or time limits. Scholars trained in debate have employed this understanding of controversy to shed new insight on cultural and political disputes, especially related to military policy and international relations.[12]

I have dwelt at such length on the contributions of contest debate to the field of argumentation, for at least three reasons. First, debate does not get enough respect. Too often in American speech communication programs, it is seen as something of an academic stepchild rather than as an evolving intellectual tradition with far broader implications. Second, many of the leading American scholars of argumentation were introduced to the subject through contest debate, labored in the vineyards of that activity, and found it an important influence on their subsequent work. And third, the case of academic debate illustrates very well a recurrent pattern in the speech communication discipline: practice precedes theory. Rather than being driven by grand theories tested through application, the discipline has tended to construct theories as needed to explain or to solve problems encountered in practice.

I would not want to give the impression, however, that argumentation in the speech communication discipline derives directly or singly from competitive debate. It is far more complicated than that. I would like to discuss, albeit more briefly, three other contributors to our current understanding of argumentation. It is the plurality of these roots that makes the discipline both rich and diverse.

The influence of social science

One of these is the development of social-science perspectives on communication. To be sure, the discipline has always stood on the boundary between the humanities and the social sciences, drawing on the methods and research traditions of both. Even in the early years, the journals included articles whose lineage traced to classical rhetoric and others whose ancestry was traced to the 18th and 19th century beginnings of psychology. Often the tension between humanities and social sciences has led to a healthy dialectic; occasionally it has led to the academic equivalent of a holy war.

Social-science studies of communication received a significant boost from the World War II studies of persuasion and attitude change.[13] During the 1950's and 1960's they assumed greater prominence in, and sometimes came to dominate, American departments of speech communication. The social-science tradition brought at least three major

---

[11] See especially his keynote address at the 1991 Alta conference: Goodnight (1991: 1-13).

[12] See, for example, Dauber (1988: 168-180); Ivie (1987: 27-36).

[13] Carl I. Hovland, who conducted such studies during World War II, then returned to Yale University to establish the Yale Communication and Attitude Change Program. Among its research publications are Hovland, Janis and Kelley (1953); Hovland (1957); Hovland and Janis (1959); Hovland and Rosenberg (1960); Sherif and Hovland (1961).

influences to communication studies. First, it emphasized descriptive and empirical, rather than normative, studies. Instead of focusing on an ideal of what communication should be, it sought to describe communication as it actually is. Second, it sought to produce testable statements about communication in general, rather than shedding insight on particular significant cases. It was far more concerned with prediction than with retrospective explanation. Since case studies were important only as they contributed to generalizations, it was not necessary or useful to study the "great speakers." Indeed, it might be *better* to study everyday interactions among ordinary people. These interactions might be more likely to yield general theory than would the study of what by definition was an exceptional or atypical case. Third, and directly related to this last point, the social-science perspective de-emphasized formal oratory and public address in favor of studying interpersonal communication, group discussion, and bargaining and negotiation, for example.

Social-science perspectives were brought to bear on argumentation studies beginning in the 1970's, predominantly by a group of scholars then located at the University of Illinois and united by their commitment to the perspectives of constructivism. In a particularly influential essay, Daniel J. O'Keefe distinguished between two different senses of argument -- one that referred to texts and products (as in "making an argument") and the other that referred to ongoing processes (as in "having an argument"). Moreover, it challenged the assumption that the first of these senses was somehow the more foundational.[14] At about the same time, Charles Willard was beginning the work that would lead to a constructivist theory of argumentation, developed in mature form in his books *Argumentation and the Social Grounds of Knowledge* and *A Theory of Argumentation*.[15] Willard defined argumentation as an interaction in which two or more people maintain what they construe to be incompatible claims, and he urged that researchers explore what actually took place in such interactions.

Meanwhile, scholars were launching research projects to do exactly that. Sally Jackson and Scott Jacobs initiated an ongoing program of studying argumentation in informal conversations. They have tried to understand the reasoning processes individuals actually use to make inferences and resolve disputes in ordinary talk.[16] Their work has some similarities to discourse analysis in linguistics. As it has matured, it also has drawn closer

[14] O'Keefe (1977: 121-128).

[15] Willard's view was first set out in "A Reformulation of the Concept of Argument: The Constructivist/Interactionist Foundations of a Sociology of Argument," Willard (1978: 121-140), and a series of subsequent articles. The two books referred to were published by the University of Alabama Press, in 1983 and 1989 respectively.

[16] See, for example, Jackson and Jacobs (1981: 77-90); Jackson and Jacobs (1980: 251-265). Several other reports on this research program also have been published.

to the pragma-dialectical perspective of Frans van Eemeren and Rob Grootendorst, with whom they collaborated on a recent book.[17]

Another strand in the empirical literature, associated with Barbara O'Keefe and Pamela Benoit, among others, is studies of how individuals develop argumentative competence.[18] Its value is its focus on argumentation as a set of acquired skills. If we know more about how and when these skills normally are acquired, we can design more effective pedagogy and training. In a somewhat related research program, Dominic Infante has explored the distinction between argumentative competence and skills, on one hand, and argumentativeness as a personality trait, on the other.[19]

Yet another application of the empirical perspective on argument studies is the growing interest in studying argument in natural settings. Unlike the debate contest or the courtroom, these are usually informal and unstructured. School board meetings, labor-management negotiations, counseling sessions, public relations campaigns, and self-help support groups are some of the highly varied settings in which argumentation has been studied.[20] The goal of such studies is to produce what has been called "grounded theory," that is, a theory of the specific case. Of course, recurrent patterns observed in such cases also contribute to more general understanding of argumentation.

The recovery of practical philosophy

Let me now turn to a third trend affecting argumentation studies in speech communication: the recovery of practical philosophy. This theme harks back to the classical concept of *phronesis*, practical wisdom in a given case. Practical wisdom was divorced from analytic knowledge and formal logic during the 17th century. The intellectual history of the disappearance and rediscovery of practical philosophy were included in my colleague Stephen Toulmin's keynote address at this conference four years ago.[21] Toulmin himself is a major figure in the recovery of *phronesis*, especially with the 1958 publication of *The Uses of Argument* and the 1972 volume, *Human Understanding*.[22] The other major figure in this recovery is Chaim Perelman, whose *The New Rhetoric* (co-authored with

[17]   van Eemeren, Grootendorst, Jackson and Jacobs (1993).

[18]   For example, see Benoit (1983: 72-89); O'Keefe and Benoit (1982: 154-183).

[19]   Infante's bibliography is lengthy. A representative example of his research is "Trait Argumentativeness as a Predictor of Communicative Behavior in Situations Requiring Argument," *Central States Speech Journal*, 32 (Winter, 1981), 265-273.

[20]   An example of such studies is Putnam, Wilson, Waltman and Turner (1986: 63-81). The proceedings of the SCA/AFA Summer Conferences in Argumentation at Alta, Utah, often include such studies.

[21]   Toulmin (1992: 3-11).

[22]   Toulmin (1958) and (1972).

L. Olbrechts-Tyteca) also was published in 1958 in French; the English translation appeared eleven years later.[23]

Both Toulmin and Perelman were surprised to discover far more interest in their work among speech communication scholars than within their own disciplines, where they were seen as marginal. In each case, however, they offered concepts and perspectives that helped to illuminate the study of argumentation. I already have mentioned how Toulmin's model was adapted as a way to understand and systematize informal reasoning. His other concept that strongly influenced argumentation scholarship was that of "field." In *The Uses of Argument*, Toulmin said only that arguments belonged to the same field if their data and conclusions were of the same logical type,[24] without explaining what that meant. In *Human Understanding* he described fields as "rational enterprises," which he equates with intellectual disciplines, and explored how the nature of reasoning differed according to whether the discipline was compact or diffuse. This treatment led to vigorous discussion about what defined a field of argument -- subject matter, general perspective or world-view, or the arguer's purpose, to mention a few of the possibilities.[25]

The concept of fields of argument, however defined, encouraged recognition that the soundness of arguments was not universal and certain but field-specific and contingent. This belief, of course, was another step in undermining the analytic ideal and resituating argument within the rhetorical tradition. Instead of asking whether an argument was sound, the questions became "sound for whom?" and "sound in what context?" Some feared that the only alternative to formal validity was vicious relativism, according to which any argument must be deemed sound if some person could be found to accept it.[26] This concern was allayed as research on argument fields demonstrated the role of cumulative experience in shaping one's perspective and the durability and predictability of a field's standards of judgment.

The term "field," of course, was a metaphor for the location of arguments. Other metaphors have also been used. McKerrow, for example, has written of "argument communities,"[27] emphasizing that shared values, common personal bonds, and argument evaluation are mutually reinforcing. Goodnight has preferred the use of the term "spheres," emphasizing more general and all-encompassing categories. His triad of personal, technical, and public spheres stresses differences among arguments whose relevance is confined to the arguers themselves, arguments whose pertinence extends to a specialized or limited

[23] Perelman and Olbrechts-Tyteca (1969).

[24] Toulmin (1958: 14).

[25] These questions are explored in a special issue of the *Journal of the American Forensic Association*. See Zarefsky (1982: 191-203), and the essays which follow.

[26] This concern is explored and answered in Booth (1974).

[27] See, for example, McKerrow (1980: 214-227).

community, and arguments that are meaningful for people in general.[28] His project also dovetails with efforts to revitalize the "public sphere," that metaphorical place in which people transcend their personal interests and guide themselves by a sense of the common good.

Perelman and Olbrechts-Tyteca's *The New Rhetoric*, perhaps because it is more vast in scope, has proved more difficult to digest. Many of its ideas have not been plumbed by argumentation scholars, and some --such as the construct of the universal audience-- have been shown to be problematic in application. Several of Perelman and Olbrechts-Tyteca's ideas, however, have permeated argumentation scholarship. Let me briefly highlight four. First, the concept of *loci*, akin to the topics in classical rhetoric, has been used as a way to understand sources of argument. Second, the treatment of figures and tropes has made clear that they are not just ornaments applied after an argument is constructed, but that they themselves have the argumentative function of strengthening or weakening presence, that is, the salience of an idea or topic. Third, the concepts of association and dissociation -- especially the latter -- illustrate the role of definitions and stipulations in advancing or retarding arguments. And fourth, the distinction between the rational and the reasonable has, like Toulmin's work, helped to displace formal logic as the paradigm of reasoning and instead to position it as a particular, and highly limited, case.

Toulmin and Perelman probably have had more far-reaching impact on argumentation studies than other philosophers, but they are not unique in their interest or concern. For example, Henry W. Johnstone, Jr., has written provocatively about the relationship between argumentation and selfhood. To engage in argumentation, he writes, is to accept risk -- the risk of being proved wrong and of having to alter one's belief system and self-concept. But the very act of person-risking proves to be person-making, constitutive of one's sense of self.[29] Legal philosophers Gidon Gottlieb and John Rawls, as well as Perelman, have explored reasoning about the nature of justice, and by extension about other abstract values.[30] On this continent, Habermas has sketched the nature of the ideal speech situation which, though counterfactual, serves as a normative ideal for argumentation.[31] And the informal logicians, especially in Canada, have re-examined the

[28]   Goodnight (1982: 214-227).

[29]   This view is developed in Johnstone, Jr. (1959); Natanson and Johnstone, Jr. (1965); and Johnstone, Jr. (1970). For an example of his influence on argumentation scholarship, see Ehninger (1970: 101-110).

[30]   Gottlieb (1968); Rawls (1971); Perelman (1963).

[31]   The list of Habermas's works, of course, is extensive. His philosophy of communication is explicated in McCarthy (1978). The entire Fall 1979 issue of the *Journal of the American Forensic Association* is devoted to studies of the implication of Habermas's philosophy for argumentation theory.

fallacies, reinterpreting many of them as errors in argumentative practice rather than as flaws in logical form.[32]

In the late 1960's, Robert L. Scott wrote an influential essay, "On viewing rhetoric as epistemic."[33] Objecting to the view that the processes of discovering and expressing truth were distinct, he maintained that rhetorical discourse itself was a means of determining truth. His work contributed further to the emerging belief that truth is relative to argument and to audience. It stimulated studies of what sorts of knowledge are rhetorically constructed and how arguing produces knowledge. Proposed answers have included the claim that *all* knowledge is rhetorical and hence that there are no transcendent standards, to the intermediate position of my colleague Thomas Farrell, who distinguishes between technical and social knowledge and maintains that it is the latter that is achieved rhetorically, to the more limited position that there is objective knowledge but that argumentation is one means of discovering it.[34]

Although not specifically intended by Scott, one consequence of the rhetoric-as-epistemic perspective has been to foster studies of rhetoric within academic disciplines. Probably more has been written about the rhetoric of science than about other disciplinary clusters.[35] I suspect that is because the popular conception of science is that it yields certain knowledge, that it is the empirical analogue for formal logic and mathematics. Demonstrating that there is a significant rhetorical component even to what we sometimes call exact sciences, therefore, would make it easier to establish that rhetoric is a part of other ways of knowing as well. But there also have been studies of rhetoric in economics, sociology, medicine, statistics, business, history, religion, and other disciplines too numerous to list.[36] This line of inquiry received a powerful boost from the 1984 conference on The Rhetoric of the Human Sciences, held at the University of Iowa,[37] the subsequent formation of the Project on Rhetoric of Inquiry (Poroi) at that institution, and the series

[32] Here the work of John Woods and Douglas Walton has been especially significant. See, for example, Woods and Walton (1982) as well as the keynote presentation from the 1990 Amsterdam conference - Woods (1992: 23-48). Blair and others in the informal logic movement have also contributed to a re-examination of the fallacies. See Blair and Johnson (1980). The book often cited as the impetus for this effort is Hamblin's *Fallacies* (1970). Volume 1, Number 3 of *Argumentation* (1987) is devoted entirely to essays exploring the theory of the fallacies.

[33] *Central States Speech Journal*, 18 (February, 1967), 9-17.

[34] For examples of these positions, respectively, see Brummett (1976: 21-51); Farrell (1976: 1-14); Cherwitz and Hikins (1983: 249-266). The "Forum" section of the *Quarterly Journal of Speech*, 76 (February, 1990), 69-84, consists of an exchange of essays by these same writers.

[35] A recent strong example of studies of the rhetoric of science is Prelli (1989).

[36] For examples of such studies, see McCloskey (1985); Simons (1990); Hunter (1990); Kellner (1989). A series of approximately twelve books on "rhetoric of inquiry" has been published by the University of Wisconsin Press.

[37] Nelson, Megill and McCloskey (1987).

of books on rhetoric in the human sciences published by the University of Wisconsin Press.

Argument as social and cultural critique

So far I have discussed three broad forces shaping the nature of argumentation studies within speech communication: the evolution of contest debate, the infusion of empirical perspectives and methods, and the recovery of practical philosophy. I'd like to discuss one more: the growing interest in social and cultural critique.

Although it usually is not characterized this way, I believe that the work of Walter Fisher is an example of this influence.[38] Fisher began with an attempt to flesh out the meaning of "good reasons" -- what rhetoric regarded as the equivalent of deduction in formal logic. He found that good reasons often took the form of narratives, and has gone so far as to claim that story-telling is a defining aspect of the human condition. But traditionally story-telling has been excluded from the category of reasoning, because of what Fisher calls the "rational world model" of knowing. The result, he believes, is systematically to privilege certain kinds of claims over others -- in his example of the nuclear debate, it is scientific claims that are preferred over moral claims. It is not Fisher's primary purpose to do so, but his work points to the nexus between argumentation and power. It is power (whether political, social, or intellectual) that permits one to stipulate what sorts of claims "count" in any argumentative situation.

Power enables those who hold it to impose a partial perspective as if it were holistic -- the definition usually given for the term "hegemony." The most recent wave of argumentation studies seeks to explore and expose the tendency of power to foreclose discourse, and to seek emancipation by opening up alternatives. This project focuses on marginalized arguers and arguments, and is given impetus by the widespread concern throughout the academy for matters of race, gender, and class.

The intellectual underpinning of argument-as-critique is postmodernism, a pattern of thought that began in architecture and has spread through much of the arts, humanities, and social sciences.[39] There are many varieties of postmodernism, and I admittedly oversimplify, but the central core seems to me to be the denial that there are any verities or standards of judgment, and the claim that what passes for such standards really is socially constructed. In some measure, this perspective is altogether consistent with the others I've discussed, in its rejection of the analytic ideal and the location of argument in communities. But it goes on to argue that only a *part* of the relevant community has defined the standards and then hegemonically imposed them on the whole. The goal

---

[38]   See especially Fisher (1987).

[39]   For the implications of postmodernism for argumentation, see several of the essays in McKerrow (1993). The keynote address, by Joseph W. Wenzel, is titled, 'Cultivating Practical Reason: Argumentation Theory in Postmodernity.'

of critique is to shed light on this practice and to promote emancipatory potential by posing alternatives to it.

At least two different implications of the postmodern project can be suggested. The more extreme is the denial that there can be any such thing as communal norms or standards for argument. On this view, the principal goal of the project is to celebrate difference and insist that it is "difference all the way down."

The other implication is more optimistic. If communal standards have been defined by only the powerful interests in a community, then the goal of argument-as-critique is to expose this practice and to suggest alternatives, so that those who were excluded or marginalized can be brought into the process of deliberation and more inclusive and meaningful norms can be developed. This view fosters empowerment of the marginalized, not in order to tear a community apart but to bind it more closely together. The question, then, is: Should the public sphere be expanded or disbanded? I expect that the coming years will see a continuing dialectic between these two versions of the postmodern challenge.

A root concept of argumentation

As I have tried to demonstrate, the study of argumentation within the speech communication discipline is a complex and many-splendored thing, a tree growing from many roots. The most obvious common features of the four intellectual movements I've described are the dethronement of formal logic as the paradigm case of reasoning and the corollary insistence that argumentation relates to audiences and fits squarely within the rhetorical tradition. There are only so many times, however, that that basic statement needs repeating. Growth of a discipline depends more on advancing knowledge and insight than on continuing restatement of a basic premise.

Where do argumentation studies in speech communication stand in that regard? In my judgment, the record is mixed. To be sure, the literature is rich. There are two major journals, *Argumentation* and *Argumentation and Advocacy*. Several books and edited collections have been published. We now have eight volumes of proceedings of the summer conference at Alta and two sets of proceedings from this international conference. The question, though, is where this literature is going. Is it building on itself or is it fragmented? I fear the latter, largely because we are not working from a clear and common sense of what we are studying. Without that, it is hard to anchor our burgeoning literature or to see how one line of inquiry relates to another. The four-part schema that I described earlier was an idiosyncratic arrangement that I'm not sure would be shared by others. Disciplinary maturity requires a greater consensus about how we organize what we do. Having made these statements, it seems incumbent on me to sketch such a framework. To that end, I'd like to propose and explicate a root concept for argumentation studies. I believe we should regard argumentation as *the practice of justifying decisions under conditions of uncertainty*. This definition has four key elements.

First, argumentation is a *practice*. It is a social activity in which people engage. In the course of this practice they make and examine texts, but the texts should be studied as products of the practice. Unlike some subjects, however, argumentation is not a practice that can be easily isolated from other practices. It has no unique subject, and people who engage in argumentation are *also* doing other things. They may not even recognize what they are doing as argumentation. This is to say that the practice of argumentation occurs in both the natural and the critical attitude. It is something that people do, and it is also a perspective or point of view which analysts use to examine the argumentative dimension of *whatever* social actors regard as their practice.[40] Studying argumentation as a practice means that it can be studied both in general and in the specific situations in which it occurs. This view of argumentation as practice contrasts most strongly with a view of argumentation as textual or logical structure.

Second, argumentation is a practice of *justifying*. This word is critical. It stands in contrast to the word *proving*. Having dethroned the analytic ideal, we recognize that the outcomes of argument cannot be certain. On the other hand, neither are they capricious or whimsical. They are supported by what the audience would regard as good reasons warranting belief or action. To say that a claim is justified immediately raises the question, "justified to whom?" Several answers can be given, depending on the situation. Claims can be justified for oneself, for one's family or friends, for the particular audience present on the occasion, for a broader audience defined by some special interest, for the general public, or for an audience of people from diverse cultures. The questions then become whether the practical meaning of "justify" varies among these different audiences and whether the process of justification is different as well. Much of the literature on argument fields, spheres, and communities, as well as discussions of what counts as evidence for claims, could be anchored productively to this basic question.

In any case, however, the question "justified to whom?" immediately calls to our consciousness the fact that argumentation is *addressed*. It is a practice that occurs in the context of an audience, not *in vacuo*. Since it is concerned with the nexus between claims and people, it clearly is a rhetorical practice.

Third, argumentation is a practice of justifying *decisions*. Decisions involve choices, for if there were only one alternative there would be nothing to decide. But decisions also presuppose the need to choose. The alternatives are perceived as being incompatible. Taking a decision is like standing at the proverbial fork in the road. One cannot stand still; one cannot take both forks; and one cannot be sure in advance which fork will prove to be the right path.

Sometimes decisions are taken at a particular moment in time. Each of the nations in the European Union, for instance, had to decide whether to approve the Maastricht treaty, just as the United States Congress had to decide whether to ratify the North American Free Trade Agreement. These decisions took place at particular moments and were preceded by attempts to justify one decision or another. Sometimes, however, a

---

[40]   The notion that argumentation can be seen as a point of view is developed more fully in Zarefsky (1980: 228-238).

decision is taken over a long period of time, and the process of justifying the decision is likewise longitudinal. The shift from nationalism to globalism as a frame of reference is a good example. For many years now, we have witnessed an ongoing controversy about whether the national or the global economy should be the unit of analysis for policy choices. Maastricht and NAFTA might be seen, from a longer term perspective, as moments in that ongoing controversy. The practice of justifying these decisions about world-view should be examined over a long period of time, not by considering particular texts in isolation.

Decisions involve choices, but they are seldom so final that they obliterate the alternative not taken. The same forks in the road may present themselves repeatedly, if in slightly altered guise. In the United States, for example, the current controversy about how best to pay for health care is largely a re-enactment of arguments that go back sixty or eighty years, even though various specific decisions have been made along the way. The minority position is seldom vanquished completely; it may come back and win another day. Recognizing this fact, decisions should respect all of the proffered alternatives, even if only one is selected at a given time.

Fourth, argumentation is the practice of justifying decisions *under conditions of uncertainty.* It was Aristotle who wrote that no one deliberates about matters that are certain. The need to make choices when not everything can be known is the defining feature of the rhetorical situation. We might have to act in the face of incomplete information. The universe affected by the decision might be so large that only a sample possibly could be considered. Or the decision might depend upon other choices or outcomes that cannot be known. Alternatively, the situation may be uncertain because of an inferential gap between data and conclusion. Even if perfect information were available, it would not entail a conclusion. The data might be factual whereas the conclusion was a matter of belief, value, or policy. Or perhaps the information relates to present conditions whereas the decision involves predictions for the future. For whichever reason, people argue to justify decisions that cannot be taken with certainty. Hence argumentation is situated within the realm of rhetoric, not of apodeictic proof. This does not mean that outcomes are irrational but rather that they are guided by rhetorical reason. Warrants are evoked from the cumulative experience of a relevant audience, rather than from a particular structure or form.

This root conception, in my opinion, will help to organize the branches of our subject, giving greater coherence to an otherwise disparate and diffuse field. The major research traditions I've described can be grafted onto it. Its descriptive and normative dimensions are clear and it can encompass argumentation from the personal to the cultural. Likewise, I believe it can suggest the questions on which research needs to focus.

Several of these questions relate to the fundamental role of an audience or community as a validating agent: (1) Given that argumentation occurs within fields, how can it occur *across* fields? How do interfield disputes come about or how do arguers in practice

transcend field boundaries? Willard has made a beginning effort to address these questions,[41] but more attention to them is needed. (2) What is the relationship between argument fields and the public sphere? Is "the public" just another field? Or is "the public" an *alternative* to argument fields, in which case what determines its boundaries? (3) What conception of "the public" is appropriate for a 21st century world characterized increasingly by cultural diversity and globalism yet tainted by the confusion of icons, images, staged events, and spectacles with the practice of justifying decisions? As I put it a few moments ago, should the public sphere be expanded, or disbanded?

A second set of issues emanates from the concept of *justifying*: (1) What do audiences count as justification? How does this view develop, and how does it change over time? (2) How do (or should) listeners decide upon the threshold level of assent needed to justify a decision? When and how does this threshold level change? (3) When is controversy healthy for a society, so that the threshold will be high, and when is it unhealthy so that relatively little would be needed in order to count as justification? (4) How does the possession or absence of power affect what decisions need justification and what counts as justification for them? More generally, how can a commitment to the practice of justifying decisions coexist with the pursuit and attainment of power?

Other questions could be clustered around elements of the definition, but these two examples should illustrate its potential for stimulating as well as classifying inquiry. It is a view of argumentation, it should be noted, which is not without assumed values and beliefs. It does place value in the idea of a "marketplace of ideas" in which claims compete for justification. The perfect market would be found in Perelman's universal audience or in Habermas's ideal speech situation. Argumentation is also presumed to have epistemic properties, because through the practice of justifying decisions a person, group, or society determines what it regards as right. This definition values community standards as a source of validation and hence rejects the extreme postmodern view that there are no common bonds and that it is "difference all the way down." And this view places argumentation firmly within the speech communication tradition, which focuses not on discourse in the abstract but on the study of how messages affect people.

---

[41]   See especially Willard (1989) on this subject.

References

Baker, G.P. (1985). *Principles of Argumentation*. Boston: Ginn.

Benoit, P.J. (1983). 'Extended Arguments in Children's Discourse.' *Journal of the American Forensic Association*, 20 (Fall), pp. 72-89.

Blair, J.A. and R.H. Johnson (1980). *Informal Logic*. Pt. Reyes, Cal.: Edgepress.

Booth, W. (1974). *Modern Dogma and the Rhetoric of Assent*. Notre Dame: University of Notre Dame Press.

Branham, R.J. (1991). *Debate and Critical Analysis: The Harmony of Conflict*. Hillsdale, New Jersey: Lawrence Erlbaum.

Brummett, B. (1976). 'Some Implications of 'Process' or 'Intersubjectivity': Postmodern Rhetoric.' *Philosophy and Rhetoric*, 9, pp. 21-51.

Cherwitz, R.A. and J.A. Hikins (1983). 'Rhetorical Perspectivism.' *Quarterly Journal of Speech*, 69 (August), pp. 249-266.

Chesebro, J.W. (1968). 'The Comparative Advantages Case.' *Journal of the American Forensic Association*, 5 (Spring), pp. 57-63

Chesebro, J.W. (1971). 'Beyond the Orthodox: The Criteria Case.' *Journal of the American Forensic Association*, 7 (Winter), pp. 208-215.

Dauber, C.E. (1988). 'Through a Glass Darkly: Validity Standards and the Debate Over Nuclear Strategic Doctrine.' *Journal of the American Forensic Association*, 24 (Winter), pp. 168-180.

'Debate Paradigms.' *Journal of the American Forensic Association*, 18 (Winter), pp. 133-160.

Eemeren, F.H. van, R. Grootendorst, S. Jackson and S. (1993). *Reconstructing Argumentative Discourse*. Tuscaloosa: University of Alabama Press.

Ehninger, D. (1970). 'Argument as Method: Its Nature, Its Limitations, and Its Uses.' *Communication Monographs*, 37 (June), pp. 101-110.

Ehninger, D. and W. Brockriede (1963). *Decision by Debate*. New York: Dodd, Mead.

Fadely, L.D. (1967). 'The Validity of the Comparative Advantages Case.' *Journal of the American Forensic Association*, 4 (Winter), pp. 28-35.

Farrell, T.B. (1976). 'Knowledge, Consensus, and Rhetorical Theory.' *Quarterly Journal of Speech*, 62 (February), pp. 1-14.

Fisher, W.R. (1987). *Human Communication as Narration*. Columbia: University of South Carolina Press.

Goodnight, G.T. (1991). 'Controversy.' In: *Argument in Controversy: Proceedings of the Seventh SCA/AFA Conference on Argumentation*, ed. Donn W. Parson. Annandale, Va.: Speech Communication Association, pp. 1-13.

Goodnight, G.T. (1980). 'The Liberal and the Conservative Presumptions: On Political Philosophy and the Foundation of Public Argument.' In: *Proceedings of the [First] Summer Conference on Argumentation*, ed. Jack Rhodes and Sara Newell. Annandale, Va.: Speech Communication Association, pp. 304-337.

Goodnight, G.T. (1982). 'The Personal, Technical, and Public Spheres of Argument: A Speculative Inquiry into the Art of Public Deliberation.' *Journal of the American Forensic Association*, 28 (Spring), 214-227.

Gottlieb, G. (1968). *The Logic of Choice: An Investigation of the Concepts of Rule and Rationality*. New York: Macmillan.

Hamblin, C.L. (1970). *Fallacies*. London: Methuen.

Hovland, C.I. (ed.) (1957). *The Order of Presentation in Persuasion*. New Haven: Yale University Press.

Hovland, C.I. and I.L. Janis (eds.) (1959). *Personality and Persuasibility*. New Haven: Yale University Press.

Hovland, C.I. and M.J. Rosenberg (eds.) (1960). *Attitude Organization and Change*. New Haven: Yale University Press.

Hovland, C.I., I.L. Janis, and H.H. Kelley (1953). *Communication and Persuasion: Psychological Studies of Opinion Change*. New Haven: Yale University Press.

Hunter, A. (ed.) (1990). *The Rhetoric of Social Research*. New Brunswick, N.J.: Rutgers University Press.

Ivie, R.L. (1987). 'The Ideology of Freedom's 'Fragility' in American Foreign Policy Argument,' *Journal of the American Forensic Association*, 24 (Summer, 1987), pp. 27-36.

Jackson, S. and S. Jacobs (1980). 'Structure of Conversational Argument: Pragmatic Bases for the Enthymeme.' *Quarterly Journal of Speech*, 66 (October), pp. 251-265.

Jackson, S. and S. Jacobs (1981). 'The Collaborative Production of Proposals in Conversational Argument and Persuasion: A Study of Disagreement Regulation.' *Journal of the American Forensic Association*, 18 (Fall), pp. 77-90.

Johnstone, Jr., H.W. (1959). *Philosophy and Argument*. University Park: Pennsylvania State University Press.

Johnstone, Jr., H.W. (1970). *The Problem of the Self*. University Park: Pennsylvania State University Press.

Kaplow, L. (1981). 'Rethinking Counterplans: A Reconciliation with Debate Theory.' *Journal of the American Forensic Association*, 17 (Spring), pp. 215-226.

Kellner, H. (1989). *Language and Historical Representation.* Madison: University of Wisconsin Press.

Lewinski, J.D., B.R. Metzler and P.L. Settle (1973). 'The Goal Case Affirmative: An Alternative Approach to Academic Debate,' *Journal of the American Forensic Association*, 9 (Spring), pp. 458-463.

Lichtman, A.J. and D.M. Rohrer (1980). 'The Logic of Policy Dispute.' *Journal of the American Forensic Association*, 16 (Spring), pp. 236-247.

Lichtman, A., C. Garvin, and J. Corsi (1973). 'The Alternative-Justification Affirmative: A New Case Form.' *Journal of the American Forensic Association*, 10 (Fall), pp. 59-69.

McCarthy, T. (1978). *The Critical Theory of Jurgen Habermas.* Cambridge, Mass.: MIT Press.

McCloskey, D.N. (1985). *The Rhetoric of Economics.* Madison: University of Wisconsin Press.

McKerrow, R.E. (1980). 'Argument Communities: A Quest for Distinctions.' *Proceedings of the [First] Summer Conference on Argumentation*, eds. Jack Rhodes and Sara Newell. Annandale, Va.: Speech Communication Association, pp. 214-227.

McKerrow, R.E. (ed.) (1993). *Argument and the Postmodern Challenge: Proceedings of the Eighth SCA/AFA Summer Conference on Argumentation.* Annandale, Va.: Speech Communication Association.

Natanson, M. and H.W. Johnstone, Jr. (eds.) (1965). *Philosophy, Rhetoric, and Argumentation.* University Park: Pennsylvania State University Press.

Nelson, J.S., A. Megill and D.N. McCloskey (eds.) (1987). *The Rhetoric of the Human Sciences: Language and Argument in Scholarship and Public Affairs.* Madison: University of Wisconsin Press.

O'Keefe, D.J. (1977). 'Two Concepts of Argument.' *Journal of the American Forensic Association*, 13 (Winter), pp. 121-128.

O'Keefe, B.J. and P.J. Benoit (1982). 'Children's Arguments.' *Advances in Argumentation Theory and Research*, eds. J. Robert Cox and Charles Arthur Willard. Carbondale: Southern Illinois University Press, pp. 154-183.

Infante, D. (1981). 'Trait Argumentativeness as a Predictor of Communicative Behavior in Situations Requiring Argument.' *Central States Speech Journal*, 32 (Winter), pp. 265-273.

Patterson, J.W. and D. Zarefsky (1983). *Contemporary Debate.* Boston: Houghton Mifflin.

Perelman, C. (1963). *The Idea of Justice and the Problem of Argument.* trans. John Petrie, London: Routledge and Kegan Paul.

Prelli, L.A. (11989). *A Rhetoric of Science: Inventing Scientific Discourse.* Columbia: University of South Carolina Press.

Putnam, L.L., S.R. Wilson, M.S. Waltman and D. Turner (1986). 'The Evolution of Case Arguments in Teachers' Bargaining.' *Journal of the American Forensic Association*, 23 (Fall), pp. 63-81.

Rawls, J. (1971). *A Theory of Justice.* Cambridge, Mass.: Harvard University Press.

Rieke, R.D. and M.O. Sillars (1975). *Argumentation and the Decision Making Process.* New York: Wiley. in 3rd ed. with title *Argumentation and Critical Decision Making.* New York: HarperCollins, 1993.

Sherif, M. and C.I. Hovland (1961). *Social Judgment.* New Haven: Yale University Press.

Simons, H.W. (ed.) (1990). *The Rhetorical Turn: Invention and Persuasion in the Conduct of Inquiry.* Chicago: University of Chicago Press.

Toulmin, S. (1958). *The Uses of Argument.* Cambridge: Cambridge University Press.

Toulmin, S. (1992). 'Logic, Rhetoric & Reason: Redressing the Balance,' *Argumentation Illuminated*, eds. Frans H. van Eemeren, Rob Grootendorst, J. Anthony Blair, and Charles A. Willard, Amsterdam: SICSAT, pp. 3-11.

Toulmin, S. (1972). *Human Understanding.* Princeton: Princeton University Press.

Warnick, B. and E.S. Inch (1989). *Critical Thinking and Communication: The Use of Reason in Argument.* New York: Macmillan.

Wenzel, J.W. (1993). 'Cultivating Practical Reason: Argumentation Theory in Postmodernity.' In: McKerrow, R.E. (ed.) (1993). *Argument and the Postmodern Challenge: Proceedings of the Eighth SCA/AFA Summer Conference on Argumentation.* Annandale, Va.: Speech Communication Association.

Willard, C.A. (1976). 'On the Utility of Descriptive Diagrams for the Analysis and Criticism of Arguments.' *Communication Monographs*, 43 (November, 1976), pp. 308-319.

Willard, C.A. (1978). 'A Reformulation of the Concept of Argument: The Constructivist/ Interactionist Foundations of a Sociology of Argument.' *Journal of the American Forensic Association*, 14 (Winter, 1978), pp. 121-140.

Willard, C.A. (1983). *Argumentation and the Social Grouns of Knowledge.* Tuscaloosa: University of Alabama Press.

Willard, C.A. (1989). *A Theory of Argumentation.* Tuscaloosa: University of Alabama Press.

Woods, J. (1992). 'Who Cares About the Fallacies?' In: *Argumentation Illuminated*, eds. Frans H. van Eemeren, Rob Grootendorst, J. Anthony Blair and Charles A. Willard, Amsterdam: SICSAT, pp. 23-48.

Woods, J. and D. Walton (1982). *Argument: The Logic of the Fallacies.* Toronto: McGraw-Hill.

Zarefsky, D. (1969). 'The 'Traditional Case'-'Comparative Advantage Case' Dichotomy: Another Look.' *Journal of the American Forensic Association*, 6 (Winter), pp. 12-20.

Zarefsky, D. (1980). 'Product, Process, or Point of View?' *Proceedings of the [First] Summer Conference on Argumentation*, ed. Jack Rhodes and Sara Newell, Annandale, Va.: Speech Communication Association, pp. 228-238.

Zarefsky, D. (1982). 'Persistent Questions in the Theory of Argument Fields.' *Journal of the American Forensic Association*, 28 (Spring), pp. 191-203.

Zarefsky, D. (1992). 'Argument as Hypothesis Testing.' In: *Advanced Debate: Readings in Theory, Practice, and Teaching*, ed. David A. Thomas and John P. Hart, 5th ed. (Chicago: National Textbook, pp. 252-262.

Argumentation studies and dual-process models of persuasion

Daniel J. O'Keefe,
University of Illinois at Urbana-Champaign

Abstract

This essay discusses some interconnections between argumentation studies and persuasion effects research. Persuasion effects research is social-scientific work concerned with how and why persuasive messages have the effects they do; expressed broadly, such studies are concerned with identifying the factors influencing the effectiveness of persuasive messages and with constructing explanations of such effects.

The focus of this essay is an attractive general picture of how persuasive messages work that has emerged from research on persuasion effects: "dual-process" models. I first describe this emerging general picture, and then display some interconnections between it and argumentation studies, discussing both what it has to offer to argumentation, and what argumentation has to offer to it.

Dual-process models of persuasion

This description of the "dual-process" image of how persuasion works begins with a general overview, which is followed by a somewhat more detailed account. The description actually represents an amalgam of two different theoretical viewpoints--the elaboration likelihood model (ELM) of Richard Petty and John Cacioppo (Petty & Cacioppo 1986a, 1986b) and the heuristic-systematic model (HSM) of Shelly Chaiken (Chaiken 1987). But for my purposes here we can happily run these together.[1]

Overview of dual-process models
Dual-process models of persuasion are based on the idea that, under different conditions, receivers will vary in the degree to which they are likely to engage in systematic issue-relevant thinking--that is, thinking about issues and arguments relevant to the persuasive issue at hand. (This issue-relevant thinking is termed "elaboration" in the ELM version of this general approach.)

Thus sometimes receivers will engage in extensive elaboration, extensive issue-relevant thinking: they will attend closely to a presented message, carefully scrutinize the arguments it contains, reflect on other issue-relevant considerations (e.g., other arguments recalled from memory, or arguments they devise), and so on. But sometimes

---

[1] This description of the dual-process approach draws from an earlier treatment (O'Keefe 1990).

receivers won't undertake so much issue-relevant thinking; no one can engage in such effort for every persuasive topic or message, and hence sometimes receivers will display relatively little elaboration.

Elaboration continuum. The degree to which receivers engage in issue-relevant thinking thus forms a continuum, from cases of extremely high elaboration to cases of little or no elaboration. And these dual-process models suggest that this continuum is crucial in determining how persuasion works in any given circumstance--that is, the nature of persuasion varies as the degree of issue-relevant thinking varies. To bring out the variation in the nature of persuasion, dual-process models offer a broad distinction between two different persuasion processes (hence "dual-process"), sometimes described as two different "routes to persuasion": a "central" and a "peripheral" route.

Two routes to persuasion. The "central route" to persuasion represents the persuasion processes involved when elaboration is relatively high. Central-route persuasion comes about through extensive issue-relevant thinking: careful examination of the message's information and arguments, consideration of other issue-relevant material (e.g., arguments recalled or devised by the receiver), and so on. In short, persuasion through the central route is achieved through the receiver's thoughtful examination of issue-relevant considerations. (Hence this central route is also sometimes referred to as a matter of "systematic" processing by the receiver.)

The "peripheral route" represents the persuasion processes involved when elaboration is relatively low. Peripheral-route persuasion comes about because the receiver employs some heuristic principle, some simple decision rule, to evaluate the advocated position. For example, receivers might be guided by whether they like the communicator, or by whether they find the communicator credible. Thus in such cases receivers are said to engage in heuristic (rather than systematic) processing; instead of engaging in extensive issue-relevant thinking, they employ decision-making short-cuts.

The two "routes to persuasion" are not conceived of as exhaustive and mutually exclusive categories or kinds of persuasion. The two processes simply represent convenient idealized extremes on the underlying continuum of issue-relevant thinking. So, for example, at intermediate levels of elaboration, one expects to find some combination of systematic and heuristic processes. But it's convenient, for expositional purposes, to talk in terms of "two processes" or "two routes."

The general idea thus is that with variations in the degree of issue-relevant thinking (the degree of elaboration), different kinds of persuasion processes are engaged--systematic central-route processes for high elaboration, heuristic peripheral-route processes for low elaboration. And (as will be seen shortly) because different kinds of persuasion processes are engaged, the factors that make for persuasive success vary (that is, what makes for successful central-route persuasion is different from what makes for successful peripheral-route persuasion).

Given that the degree of elaboration is so important (to determining how persuasion works in any given case), the question naturally arises: what influences the degree of elaboration (the degree of issue-relevant thinking engaged in by a receiver)?

Factors affecting the degree of issue-relevant thinking
There are two broad classes of factors influencing the degree of elaboration that a receiver will likely undertake in any given circumstance. One of these classes concerns the receiver's motivation for engaging in elaboration, the other the receiver's ability to engage in such elaboration.

Influences on motivation. Although a number of different factors can influence a person's motivation for engaging in issue-relevant thinking, here I want to mention two leading factors. One is the personal relevance of the topic to the receiver (this is often glossed as a matter of the receiver's degree of "involvement" with the issue). As a given issue becomes increasingly personally relevant to a receiver, the receiver's motivation for engaging in thoughtful consideration of that issue increases (see, e.g., Petty & Cacioppo 1984; Petty, Cacioppo, & Goldman 1981; Petty, Cacioppo, & Schumann 1983).

A second is the receiver's level of need for cognition. "Need for cognition" refers to a person's tendency to engage in and enjoy thinking. This tendency varies among persons: that is, some people are generally disposed to enjoy and engage in effortful cognitive undertakings, whereas others are not. As one might suppose, people higher in need for cognition have generally greater motivation for engaging in issue-relevant thinking than do persons lower in need for cognition (see, e.g., Axsom, Yates, & Chaiken 1987; Haugtvedt, Petty, Cacioppo, & Steidley 1988).

Influences on ability. There are two notable factors influencing a person's ability to engage in issue-relevant thinking. One is distraction in the persuasive setting, that is, the presence of some distracting stimulus or task accompanying a persuasive message. (Researchers have used distractions such as having an audio message be accompanied by static or beep sounds, or having receivers monitor a bank of flashing lights.) Obviously, under conditions that would otherwise produce relatively high elaboration, distraction will interfere with such issue-relevant thinking (for a general discussion, see Petty & Cacioppo 1986a: 61-68).

A second factor influencing elaboration ability is the receiver's prior knowledge about the persuasive topic. The more extensive such prior knowledge, the better able the receiver is to engage in issue-relevant thinking (see, e.g., Wood 1982; Wood & Kallgren 1988).

Summary. As a way of summarizing these influences on the degree to which receivers are likely to engage in systematic issue-relevant thinking, consider these two circumstances: (a) a low-need-for-cognition receiver, listening to a persuasive message on a topic that he doesn't know much about, and that isn't very relevant to him personally, while there's simultaneously some distraction going on, versus (b) a high-need-for-cognition receiver, listening to a persuasive message on a personally-relevant topic where she's very knowledgeable and undistracted. Plainly, the latter case is likely to produce much more systematic thinking about the persuasive message.

The reason why these variations in the degree of issue-relevant thinking are important, according to these dual-process models, is that depending upon the degree of

issue-relevant thinking, different kinds of persuasion processes are activated--central-route in the case of high elaboration, peripheral-route in the case of low elaboration. I now turn to a somewhat more extensive description of each of those different kinds of persuasion processes, focussing on what's key to persuasive success in each process.

Central-route persuasion

Key: elaboration direction. In central-route persuasion (when elaboration is high), what's key to persuasive success is the evaluative direction of the receiver's elaboration (the evaluative direction of the receiver's issue-relevant thinking). That is, persuasive effects will depend upon the predominant valence of the receiver's issue-relevant thoughts: to the extent that the receiver is led to have predominantly favorable thoughts about the advocated position, the message will presumably be relatively successful; but if the receiver has predominantly unfavorable thoughts, then the message will presumably be relatively unsuccessful. Thus the question becomes: given relatively high elaboration, what influences the direction (the valence) of elaboration?

Influences on elaboration direction. Two particular factors stand out as influences on the direction of receivers' issue-relevant thinking. The first is whether the message's advocated position is proattitudinal or counterattitudinal. When the advocated position is one toward which the receiver is already favorably inclined--that is, when the message advocates a "proattitudinal" position--the receiver will presumably ordinarily be inclined to have favorable thoughts about the position advocated. By contrast, when the message advocates a counterattitudinal position, receivers will ordinarily be inclined to have unfavorable thoughts about the point of view being advocated. That is to say, everything else being equal, one expects proattitudinal messages to evoke predominantly favorable thoughts, and counterattitudinal messages to evoke predominantly unfavorable thoughts.

But if this were the whole story, then nobody would ever be persuaded by a counterattitudinal message. And we know that at least sometimes, people are persuaded by the arguments contained in counterattitudinal communications, and hence dual-process models suggest that a second influence on elaboration direction is the quality (the strength) of the message's arguments. Under conditions of extensive issue-relevant thinking, receivers are able to carefully examine the message's arguments. Unsurprisingly, then, the direction of receivers' elaboration depends (at least in part) on the results of such scrutiny: the more favorable the reactions evoked by the close scrutiny of message material, the more effective the message is.

So if a receiver's examination of the message's arguments reveals shoddy arguments and bad evidence, there's likely to be little persuasion; but if the message contains powerful arguments, sound reasoning, good evidence, and the like, the message will be more successful. Hence under conditions of high elaboration the quality (the strength) of the message's arguments influences the direction of elaboration (and thus influences persuasive success). (For examples of relevant research results, see Heesacker, Petty, & Cacioppo 1983; Petty & Cacioppo 1984; Petty, Cacioppo, & Goldman

1981; Petty, Cacioppo, & Schumann 1983.) Under conditions of high elaboration, then, it turns out to be important for persuaders to have good arguments (high quality evidence from well-qualified sources, discussion of important issues, evidence that is relevant to the conclusions drawn, and so forth).

Summary: central routes to persuasion. Under conditions of systematic issue-relevant thinking, the outcome of persuasive efforts depends upon the direction of receivers' elaboration: where a persuasive message leads receivers to have predominantly favorable thoughts about the position being advocated, persuasive success is likely. And the direction of receivers' elaboration depends (at least in part) on the quality of the message's arguments.

Peripheral-route persuasion

Key: heuristic principles. Dual-process models of persuasion suggest that under conditions of relatively low elaboration, the outcomes of persuasive efforts will turn not on the receiver's careful consideration of the message's arguments, but will instead be much more influenced by the receiver's use of heuristics, simplifying decision rules. These heuristics require little information processing, and are activated by peripheral cues, that is, by extrinsic features of the communication situation such as communicator characteristics (e.g., credibility). These heuristic principles are ordinarily not consciously articulated, but there is indirect evidence (of various sorts) that people do rely on these heuristics.

Two heuristic principles. A number of different heuristic principles apparently operate in persuasion; here I want to discuss just two relatively more prominent ones: the credibility and consensus heuristics.

One heuristic principle, the credibility heuristic, is based on the apparent expertise of the communicator, and amounts to a belief that "statements by credible sources can be trusted" (for alternative ways of putting this idea, see Chaiken 1987: 4; Cialdini 1987: 175). Thus, higher-credibility sources generally have greater persuasive impact. But--consistent with the dual-process image--the communicator's credibility has been found to have greater impact on persuasive outcomes when the receiver's degree of issue-relevant thinking is relatively low (e.g., when the issue is not very relevant to the receiver; see, e.g., Johnson & Scileppi 1969; Kiesler & Mathog 1968; Petty, Cacioppo, & Goldman 1981; Ratneshwar & Chaiken 1986; Rhine & Severance 1970). That is, the peripheral cue of credibility has been found to have greater impact on persuasive outcomes when elaboration is relatively low.

A second heuristic principle is the consensus heuristic, which is based on the reactions of other people to the message. This heuristic might be expressed as a belief that "if other people believe it, then it's probably true" (for variant phrasings, see Chaiken 1987: 4; Cialdini 1987: 174). When this heuristic is employed, the approving reactions of others should enhance message effectiveness (and disapproving reactions should reduce effectiveness). A number of studies have revealed the operation of such a heuristic; for example, several investigations have found that receivers are less

persuaded when they overhear an audience expressing disapproval (as opposed to approval) of the communicator's message (for a review, see Axsom, Yates, & Chaiken 1987).

Summary: peripheral routes to persuasion. Under conditions of low elaboration, the outcome of persuasive efforts depends less upon the direction of receivers' issue-relevant thinking than upon the operation of heuristic principles, simple decision rules activated by peripheral cues in the persuasion setting. Where receivers are unable or unmotivated to engage in extensive issue-relevant thinking, their reactions to persuasive communications will be guided by simpler principles such as the credibility and consensus heuristics.

Summary of dual-process models

The dual-process model is a convenient way of displaying the variation in persuasion processes, but it's important to keep in mind the underlying continuum of issue-relevant thinking. One way of crystallizing this idea is to see that (in considering what influences persuasive outcomes) there is something of a *tradeoff* between the impact of peripheral cues and the impact of elaboration (issue-relevant thinking): as elaboration increases, the effect of peripheral cues declines, and the effect of the receiver's issue-relevant thinking increases. For example, as variations in argument quality make more and more difference in outcomes, variations in communicator expertise make less and less (e.g., Petty, Cacioppo, & Goldman 1981).

Hence these "dual-process" models do not claim that (for instance) variations in argument quality will make no difference when elaboration is low, or that variations in communicator credibility will make no difference when elaboration is high. Rather, the suggestion is that, broadly speaking, the relative impact of elaboration and peripheral cues will vary as elaboration varies. With greater elaboration, persuasive effects come to depend more and more on the direction of elaboration (and less and less on peripheral cues); as elaboration decreases, the impact of peripheral cues increases (and that of elaboration declines).[2]

These dual-process models do not offer *the* definitive picture (for all time) of how persuasion works; they are not without flaw or immune to criticism. This general approach is only one of a number of different theoretical avenues to understanding persuasion; it does not explain everything about persuasion, and it certainly has defects. But this is plainly a very useful general picture (arguably the best in hand), and it certainly is an important step forward in our understanding of persuasive effects. For instance, one attractive feature of dual-process models is their ability to account for

---

[1]   Actually, this description of dual-process models is not quite accurate, as one key issue dividing different dual-process models is precisely whether there is inevitably this sort of tradeoff between heuristic and systematic processing. The description given here represents the viewpoint of one dual-process model, the elaboration likelihood model (ELM): these are taken to be opponent processes. But the heuristic-systematic model (HSM) takes a different view, namely that where these processes co-occur they may produce additive or interactive effects (see Bohner, Chaiken, & Hunyadi 1994; and Chaiken & Maheswaran 1994).

apparently-conflicting findings in earlier research. Why is it that the communicator's credibility sometimes exerts a large influence on persuasive outcomes, and other times very little influence? Because (the dual-process models suggest) the degree of issue-relevant thinking varies, and (correspondingly) so does the degree of reliance on a simple decision rule such as the credibility heuristic.

With this general description of dual-process models in place, we can now turn to a consideration of some interconnections between these dual-process models and argumentation studies, beginning with what these models have to offer to argumentation.

What dual-process persuasion models offer argumentation

I think that dual-process models of persuasion have two offerings to bring to argumentation studies: some reassurance that normatively good argument matters, and an expanded conception of rationality.

Comfort and reassurance
The existence (and powerfulness) of central-route processes should give some solace to anybody, including argumentation scholars, concerned with normatively good argument. Anyone who has tried to teach argument-analysis skills (or critical-thinking skills, or the like) has had at least one moment of utter despair about the human condition and its perfectibility. In fact, as Willard (1989) has pointed out, there's a common theme in argumentation-related pedagogy to the effect that, left to their own devices, people will be "intuitive, lazy, and impulsive, swayed this way and that by their attitudes, prejudices, and pieties" (1989: 183). Hence the importance of teaching people appropriate skills and principles: "argument principles are seen as remedies to passivity, apathy, and ignorance" (1989: 198). Without such training, the argument-consumer-in-the-street "escapes from freedom, shuns the political arena, wallows in lethargy or cussedness, and succumbs to 'the forces of nonreason'" (1989: 199).

But the research associated with these dual-process models of persuasion has made it clear that people *do* operate in a familiarly "rational" fashion, at least sometimes. Argument quality *can* matter, *does* matter. Central-route persuasion *does* work. People really are (at least sometimes) more swayed by the force of the better argument.

I think that sometimes persuaders are inclined to think that they have to choose: either they can be successful in persuasion (by using various underhanded tricky manoeuvres), or they can make normatively good arguments (and be unsuccessful). But these dual-process models have made it plain that in fact making normatively good arguments doesn't necessarily mean having to sacrifice practical persuasive success. In various ways, then, those who are concerned with normatively good argument can find a good deal of comfort and reassurance in the findings associated with these dual-process models.

Expanded conception of rationality

This general dual-process picture (of how persuasion works) also suggests the usefulness of an expanded conception of rationality with respect to the processing of persuasive messages--expanded in two ways.

Heuristic use. First, this image of persuasion suggests a conception of rationality that is expanded to include the idea that the use of specific heuristics can be rational. Consider, for example, the credibility heuristic. In a world in which knowledge is so specialized, there will inevitably be experts and non-experts, and it's no good pretending otherwise. Particularly in circumstances in which a person is not inclined to give much systematic attention to the argumentative details, invoking the credibility heuristic is arguably a very rational thing to do.

Now of course this point should be heard as very much connected with recent discussions in the fallacy literature concerning argument-from-authority (*argumentum ad verecundiam*). Argument-from-authority is (now) not treated as inevitably a mistake in reasoning. In fact, sometimes the presumption almost seems to have been reversed, as when Douglas Walton (1989b: 21) writes: "appeal to expert opinion is, in itself, a legitimate form of argumentation, but one that can be employed wrongly." (For a similar view, see Willard 1990.)

So now, instead of treating reliance on authority as automatically illegitimate, the question has become one of specifying the conditions under which appeal to authority is or isn't fallacious. For example, Walton (1989c: 60) summarizes "six requirements to be met for an appeal to expertise to be reasonable. First, the judgment put forward by the expert must actually fall within his field of competence. Second, the cited expert must be a legitimate expert, and not merely a celebrity, or someone not an expert. A third factor is the question of how authoritative an expert is, even if he is a legitimate expert in a field. Questions of specialization within fields of expertise are relevant here. Fourth, if several qualified experts have been consulted, there should be some way of resolving inconsistencies and disagreements that may arise. Fifth, if objective evidence is also available, this should be taken into account. In particular, an expert should be able to back up his opinion, if queried, by citing evidence in his field. The sixth requirement is that the expert's sayso must be correctly interpreted." (For another effort at identifying conditions for the non-fallacious use of appeal to authority, see Govier 1992: 385.)

But I'm trying to come at this question--the question of when it's sensible to invoke authority--from a slightly different direction. These efforts (at specifying conditions under which appeal-to-authority is or isn't fallacious) indicate the sorts of considerations that ought to be taken up when one is engaged in intensive scrutiny of expert claims ("go see what other experts say," "check the objective evidence," and so on); that is, the concern is with the appropriate conditions for the use of authority-based reasoning in *systematic* processing. My point is a different, and perhaps more extreme one, namely, that the *non*-systematic reliance on expertise (as embodied in the use of the credibility heuristic) is arguably rational.

The very same points can be made with the consensus heuristic. Since it's at least possible that some other people have been able to give some more thought to the matter than one has been able to oneself, the presence of a consensus is at least a plausible guide to belief and action, and hence reliance on the consensus heuristic is (at least sometimes) arguably rational.

Again, one might point to parallels in the recent treatment of the *ad populum* fallacy. Commonly, appeals to the popularity of a belief or product are seen to be a fallacious basis for acceptability. Consider, for example, Govier's (1992: 170-171) characterization: "Many arguments are based on popularity. Someone tries to show that a product is good because many people select it or that a belief is correct because many people hold it. Such arguments are extremely flawed because the merits of something are one matter and its popularity another. The problem is that things can be popular for many reasons, and only one of these is their good quality."[3]

But--paralleling the treating of authority appeals--it is now being recognized that popularity-based reasoning is not inevitably illegitimate. For example, Walton (1989b: 106) writes that forms of reasoning such as "everybody accepts that A is true, therefore A is true" are "weak but sometimes reasonable forms of argument. For example, if a proposition is widely accepted and you have no evidence against it, then if you have to make a decision, it could be much more reasonable to presume that it is true than to presume that it is false."

And so, correspondingly, there has been some effort at distinguishing fallacious ad populum appeals from related but more defensible arguments. For instance, Walton (1989a: 172) notes that "appeals to popular views or presumptions taken to be widely plausible for a given audience or cultural group are a legitimate part of reasoned argument in a democratic political system." And Govier (1992: 181, n. 14) emphasizes that "appeals to the popularity of beliefs should not be confused with the notion of common knowledge [as a basis of argument] ... The difference is that the belief whose popularity is appealed to is not universal in a culture, nor is it basic and elementary. Typically, its content is somewhat controversial, speculative, or normative, but it is claimed to be popular."

Again, though, notice: I want to approach this matter from a slightly different angle. These discussions of *ad populum* are concerned with the appropriate conditions for the rational use of popularity-based reasoning in systematic processing. My point concerns

---

[3] Now Govier (1992: 181, n. 13) does acknowledge "It can happen that things are popular because they are, in some respect, good. But this is not always the case and, in any event, the point at issue here is whether things can be shown to be good because they are popular." There are two points to be made here. (1) In the present discussion, the point at issue is not whether things can be *shown* to be good because they are popular, but whether--given that one is not inclined or able to engage in systematic argument processing--popularity might not be a useful heuristic basis on which to make a decision. (2) Of *course* it's "not always the case" that things are popular because they're good--but this is only to acknowledge the fallibility of heuristic procedures. However, as discussed below, the fact that heuristic procedures are fallible is not necessarily a reason not to use them.

the rationality of the *non*-systematic reliance on popularity (as embodied in the use of the consensus heuristic).[4]

In a way, then, certain recent developments in the fallacy literature and the point I'm making here (about heuristic use) are coming at related phenomena from different directions.[5] In each case, the idea is that certain reasoning practices, practices that might be condemned as normatively indefensible, are being suggested to be rational practices, at least in some circumstances. But where recent discussions of argument-from-authority and *ad populum* have tried to emphasize their appropriate use under conditions of systematic scrutiny, these dual-process models can be taken to suggest that even the non-systematic use of authority-based and popularity-based reasoning is defensible.

I do want to emphasize that it's no strike against heuristics that sometimes they lead to bad decisions, or to less-than-ideal decisions. After all, even the most carefully constructed and applied systematic argument-evaluation procedure isn't guaranteed to produce good outcomes. All we have are fallible procedures. Some may be more fallible than others, but the fact that a procedure is fallible isn't necessarily a good reason not to use it. The fallibility of heuristics, that is to say, is no strike against their rationality.

And, as a related point, notice: the fact that heuristic-based decisions may be *more* fallible than ones based on systematic processing is also not necessarily a strike against the use of heuristics. True enough, everything else being equal, one will prefer whatever procedure is least fallible. The problem is, everything else *isn't* always equal--and that brings me to the second way in which these dual-process models point to an expanded conception of rationality. (The first way, it will be recalled, is that the use of specific heuristics can be rational.)

Having two processes. These dual-process models suggest a conception of rationality that is broad enough to encompass the general idea of *having both* heuristic and systematic modes of processing. That is, the existence of both central-route processes and peripheral-route processes is an arguably rational arrangement. People have limited capacities for issue-relevant thinking, and hence they need some way of

---

[4]   Walton (1989b) in fact discusses popularity-based reasoning in a way that is similar to the point I'm trying to make out of these dual-process models. Concerning arguments of the form "everybody accepts that A is true, therefore A is true" and "nobody accepts that A is true, therefore A is false," he writes that these "are weak arguments in some cases that nevertheless have some plausibility value in directing a person toward a particular line of action when objective knowledge of the facts is lacking, yet a practical decision must be made. For example, if I am late for my train and do not know where the train platform is located, I may be guided by seeing everybody else in the area heading toward a tunnel." (1989b: 89-90)

[5]   Fallacies and heuristics aren't precisely the same thing. Fallacies, as usually conceived, are particular types of arguments; heuristics are cognitive decision-making guides. But there is an underlying commonality here, as can be seen by considering argument-from-authority and the credibility heuristic. An argument-from-authority consists of a speaker S's asserting "expert E says X, therefore X." The credibility heuristic, as applied to a particular case, yields a receiver's reasoning that "this expert E (the speaker) says X, therefore X." That is, the underlying reasoning is the same. (A similar identity underlies *ad populum* and the consensus heuristic.)

allotting attention (allotting processing time)--and some way of handling issues that don't get so much concerted attention. What better (more rational) way than heuristics?

To step to the side for a moment, there's a general point to be made here concerning the relationship of normative and descriptive aspects of the study of communication, namely: a helpful/useful normative model of communication (of any given communication practice) is one that is responsive to the descriptive realities of communication. (The interplay of descriptive and normative aspects of argumentation has recently been explored by van Eemeren, Grootendorst, Jackson, & Jacobs 1993.) The connection to the current discussion is this: since people don't have unlimited capacities for issue-relevant thinking, our normative guides need to respect that limitation. It's easy to construct normative models that begin "assume you have all the resources you need, including unlimited time"--but it's harder to start from more realistic premises.

A similar point has been made by Schellens (1991: 389) in discussing argument-from-authority and *ad hominem* arguments as "acceptable fallacies." Schellens notes that "argument from authority is not acceptable" in an "ideal discussion" (in which "the partners are equal, have maximum opportunities to verify assertions," and so forth), but in circumstances involving "epistemic dependency of the participants amongst themselves or collectively from external sources," then "norms for a reasonable discussion ... cannot exist without the authority and *ad hominem* arguments."

In any case, the point I want to emphasize is that one should not think of peripheral-route persuasion as somehow intrinsically non-rational, or as less rational than central-route persuasion. The general idea of using heuristics (sometimes) is quite sensible (sensible, that is, as decision guides in circumstances in which concerted attention is not possible or desirable). So notice the larger rationality of persuasion here--in some circumstances receivers engage in close scrutiny, in others they (quite sensibly and rationally) don't, but overall they proceed in quite reasonable ways.

Approached in this way, an important question arises: what is a (normatively) *good* basis for distinguishing issues as meriting systematic or heuristic processing? Research on these dual-process models, of course, is simply aimed at describing what *is* the basis of such differentiation--that is, what influences whether one or another route is pursued in any given case. A separate question, of course, is what the basis of choosing *ought* to be.

To make this connection slightly differently: one of argumentation's traditional central concerns is enhancing people's capacities for systematic argument processing. What I'm suggesting here is that it may also be useful to enhance people's capacity to *choose* when to engage in such argument scrutiny. This point, too, can be expressed as a matter of an expanded conception of rationality. Instead of assuming that proceeding rationally inevitably involves extensive issue-relevant thinking, one might alternatively consider that proceeding rationally inevitably involves instead deciding *whether* extensive issue-relevant thinking is appropriate. After all, someone who

devotes just as much thought and attention to every single decision (which candy bar to buy, which career path to follow) is not acting sensibly.

In short, then, these dual-process models of persuasion seem to me to offer argumentation studies both some solace (that normatively good argumentation does matter) and some considerations for reflection (about an expanded conception of rationality).

But now I want to turn to traffic in the other direction, to a consideration of what argumentation studies might offer to these dual-process models.

What argumentation offers dual-process persuasion models

The problem of argument quality

In the earlier description of these dual-process models, I (purposefully) side-stepped an important problem--namely, the *definition* of "argument quality" (argument strength) in this research area. The problem is that in this research, "argument quality" has been defined *empirically*, in terms of observed persuasive effects.

Specifically: to obtain experimental messages containing "strong" or "weak" arguments, these researchers pre-test various messages; a "strong-argument" message is defined as one that elicits predominantly favorable thoughts when receivers think carefully about the message, whereas a "weak-argument" message is one that yields predominantly unfavorable thoughts under such conditions. Thus, as two of the most prominent dual-process researchers have explicitly acknowledged, these researchers "have ignored the specific qualities that render some arguments cogent and others specious" (Petty & Cacioppo 1986a: 32). Obviously, this is not a defensible treatment of argument quality; "argument quality" in this research is not defined by reference to some independent set of normative standards. (For a somewhat amplified discussion of this problem, see O'Keefe 1990: 110-111.)

In fact, however, if one examines the "strong-argument" and "weak-argument" messages, it's apparent that these *do* differ in normative quality--the "strong-argument" messages in fact do make normatively better arguments than do the "weak-argument" messages. These messages differ in (for example) the relevance of the evidence to the conclusions drawn, in the apparent self-interest of cited evidence sources, in the desirability of the benefits claimed to attach to the advocated position, and so on. (For sample messages, see Petty & Cacioppo 1986a: 54-59.)

So, on the one hand, in fact one cannot yet say that (under conditions of systematic processing) normatively-better arguments are more persuasive (than their poorer counterparts). One can't say this, because the research doesn't have some independently-justified normative standard for argument.

But on the other hand, that is certainly the most plausible hypothesis at present (for explaining the observed effects). That is, the most plausible current hypothesis is precisely that what makes those "strong-argument" messages more persuasive is that they have normatively better arguments. But if one is to sustain the belief that what

makes them more persuasive is that they are normatively better, then one will need some independently-motivated account of argument quality--some way of establishing the contrast between high- and low-quality argumentation that does not depend upon the observed effects of the messages under conditions of systematic processing.

The contribution of argumentation studies

And here, obviously, is where argumentation studies is in a position to be helpful. It has become clear that what's needed now, for further advance in this line of work, is analyses of message content that attend to normative considerations. Plainly, the developed message-analytic equipment of argumentation studies may prove very useful.

Indeed, it can be useful in a couple of ways. First, it can be useful in analyzing the messages used in previous research, with an eye to describing their features in ways that are sensitive to normative questions of argument quality. A normatively-guided analysis of these messages may offer some insights into just what aspects of the messages may be contributing to the observed effects.

Second, the conceptual apparatus of argumentation studies can be useful in offering general criteria for normatively good arguments, and correspondingly useful in suggesting message construction principles that might guide the creation of experimental materials for subsequent research. That is to say, once one has an independently-motivated account of argument quality, it is possible to undertake empirical work that directly explores the relationship of argument quality to persuasive effects. ("Directly," that is to say, without the conceptual problems of research to date.)[6]

My suggestions here do not require that there be some grand, far-reaching agreement in argumentation studies about what constitutes *the* correct formulation of normative standards for argument (which is just as well, since there's not such agreement). There is, of course, a rough-and-ready consensus achievable about certain low-level descriptions (agreement, say, that this argument is better than that one), even if there's substantial disagreement about just how to formulate the larger theoretical housing (the higher-level descriptions). But given the current state of dual-process-model research, *any* independently-motivated account of argument quality will represent an advance.

There is an additional complexity to be mentioned. In the dual-process-model research that's been conducted thus far, the strong-versus-weak-argument contrast has

---

6   These two different aspects of the usefulness of argumentation studies correspond to what are actually two distinct research questions. One question is: what is it about those "strong-argument" dual-process messages that makes them persuasive under conditions of systematic processing? A second question is: what really *does* make for high-quality arguments, and how do such arguments figure in persuasion? These two questions will be closely related only to the extent that the basis for the effectiveness of the "strong-argument" messages really *is* their normative superiority. Because the basis for that effectiveness might be something else (that is, because the dual-process hypothesis--the hypothesis that it's the normative superiority of the "strong-argument" messages that produces their greater effectiveness under conditions of scrutiny--could be mistaken), it's important to see that these are distinct questions.

been formed in a way that confounds a number of *different* message features (the relevance of evidence to claim, the apparent trustworthiness of cited sources, and so on). That is, the design of the research makes it impossible to disentangle the separate effects of these various elements. Sorting out the roles played by different aspects of normatively-good argument is obviously going to be a substantial puzzle. (A preliminary attempt has been made by Areni & Lutz 1988.)

Distinguishing the effects of different variations is important, because it's surely *not* the case that (under conditions of systematic processing) people *never* make mistakes in reasoning, *never* misapprehend argument quality. On the contrary, it seems plausible to suppose that people might ordinarily be sensitive to some aspects of normatively-good argument, but not to others. Consider, for instance, that it appears that even under conditions of systematic processing, people give the single example (as opposed to statistical summaries of multiple examples) more weight than it is due (Taylor & Thompson 1982 provide a general review). So the question arises: under conditions of systematic processing, just what sorts of argumentative flaws *are* people more or less sensitive to?

Having evidence that bears on this question can be helpful for two reasons. First, it may illuminate why persuasive messages have the effects they do (under conditions of systematic scrutiny). One way of expressing this idea is to say that this research may clarify the implicit normative argument standards that persons ordinarily use--by indicating that people are sensitive to this normatively-significant feature, but not to that other one. And this, in turn, provides a basis for explaining why people react favorable to one message, but unfavorable to another.

Second, it can be used to inform the design of pedagogical interventions, used to adapt instruction so as to maximize the improvement in argument-analytic skills. Instead of starting from the assumption that people have no ability to distinguish good and bad argumentation, one might instead start from the idea that people are (under the right conditions) commonly able to see certain sorts of flaws but are generally unskilled in seeing others. Once one has a better grasp of just which skills need bolstering, one's instruction can be appropriately adapted. (For an example of an effort at addressing such questions, see Ryan & Norris 1991.)

Plainly, then, students of argumentation are well-situated to make useful contributions to dual-process persuasion research. Most of the researchers currently engaged in this work are in no position, because of their professional training, to undertake the relevant work. But students of argumentation will be on familiar ground.

Conclusion

There is plainly much prospect for useful interchange between argumentation studies and persuasion effects research, with benefits both directions. And these inter-connections underscore the importance and value of international, interdisciplinary

conferences where persons with differing outlooks and background knowledge can come together to discuss matters of common concern.

After all, the increasing specialization of knowledge is not simply some theoretical problem of interest to analysts of discourse in the public sphere. It is also a real and practical problem we all face in our professional lives. One suspects or knows that there is relevant work out there somewhere, with possibilities for mutual enrichment, and yet our customary disciplinary pathways do not make it easy to exploit such possibilities. The signal value of these quadrennial conferences is precisely that they encourage interfield connections, at a time when such connections are increasingly important.

References

Areni, C.S., R.J. & Lutz (1988). 'The role of argument quality in the elaboration likelihood model'. In: M.J. Houston (ed.), *Advances in Consumer Research* (Vol. 15, pp. 197-203), Provo, UT: Association for Consumer Research.

Axsom, D., S. Yates, & S. Chaiken (1987). 'Audience response as a heuristic cue in persuasion'. *Journal of Personality and Social Psychology* 53, pp. 30-40.

Bohner, G., S. Chaiken, & P. Hunyadi (1994). 'The role of mood and message ambiguity in the interplay of heuristic and systematic processing'. *European Journal of Social Psychology* 24, pp. 207-221.

Chaiken, S. (1987). 'The heuristic model of persuasion'. In: M. Zanna, J.M. Olson, & C.P. Herman (eds.), *Social Influence: The Ontario Symposium* (pp. 3-39), Hillsdale, NJ: Lawrence Erlbaum.

Chaiken, S., & D. Maheswaran (1994). 'Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgment'. *Journal of Personality and Social Psychology* 66, pp. 460-473.

Cialdini, R.B. (1987). 'Compliance principles of compliance professionals: Psychologists of necessity'. In: M.P. Zanna, J.M. Olson, & C.P. Herman (eds.), *Social Influence: The Ontario Symposium* (pp. 165-184), Hillsdale, NJ: Lawrence Erlbaum.

Eemeren, F.H. van, R. Grootendorst, S. Jackson, & S. Jacobs (1993). *Reconstructing Argumentative Discourse*. Tuscaloosa, AL: University of Alabama Press.

Govier, T. (1992). *A Practical Study of Argument* (3rd ed.). Belmont, CA: Wadsworth.

Haugtvedt, C., R.E. Petty, J.T. Cacioppo, T. & Steidley (1988). 'Personality and effectiveness: Exploring the utility of need for cognition'. In: M.J. Houston (ed.), *Advances in Consumer Research* (Vol. 15, pp. 209-212), Provo, UT: Association for Consumer Research.

Heesacker, M., R.E. Petty, & J.T. Cacioppo (1983). 'Field dependence and attitude change: Source credibility can alter persuasion by affecting message-relevant thinking'. *Journal of Personality* 51, pp. 653-666.

Johnson, H.H. & J.A. Scileppi (1969). 'Effects of ego-involvement conditions on attitude change to high and low credibility communicators'. *Journal of Personality and Social Psychology* 13, pp. 31-36.

Kiesler, S.B. & R.B. Mathog (1968). 'Distraction hypothesis in attitude change: Effects of effectiveness'. *Psychological Reports* 23, pp. 1123-1133.

O'Keefe, D.J. (1990). *Persuasion: Theory and Research*. Newbury Park, CA: Sage Publications.

Petty, R.E. & J.T. Cacioppo (1984). 'The effects of involvement on responses to argument quantity and quality: Central and peripheral routes to persuasion'. *Journal of Personality and Social Psychology* 46, pp. 69-81.

Petty, R.E. & J.T. Cacioppo (1986a). *Communication and Persuasion: Central and Peripheral Routes to Attitude Change.* New York: Springer-Verlag.

Petty, R.E. & J.T. Cacioppo (1986b). 'The elaboration likelihood model of persuasion'. In: L. Berkowitz (ed.), *Advances in Experimental Social Psychology* (Vol. 19, pp. 123-205), New York: Academic Press.

Petty, R.E., J.T. Cacioppo, & R. Goldman (1981). 'Personal involvement as a determinant of argument-based persuasion'. *Journal of Personality and Social Psychology* 41, pp. 847-855.

Petty, R.E., J.T. Cacioppo, & D. Schumann (1983). 'Central and peripheral routes to advertising effectiveness: The moderating role of involvement'. *Journal of Consumer Research* 10, pp. 135-146.

Ratneshwar, S. & S. Chaiken (1986). 'When is the expert source more persuasive? A heuristic processing analysis'. In: T.A. Shimp, S. Sharma, G. John, J.A. Quelch, J.H. Lindgren, Jr., W. Dillon, M.P. Gardner, & R.F. Dyer (eds.), *1986 AMA Educators' Proceedings* (p. 86), Chicago: American Marketing Association.

Rhine, R.J. & L.J. Severance (1970). 'Ego-involvement, discrepancy, source credibility, and attitude change'. *Journal of Personality and Social Psychology* 16, pp. 175-190.

Ryan, J. & S.P. Norris (1991). 'High school students' understanding of argumentative text'. In: F.H. van Eemeren, R. Grootendorst, J.A. Blair, & C.A. Willard (eds.), *Proceedings of the Second International Conference on Argumentation* (Vol. 1B, pp. 1127-1134), Amsterdam: SICSAT.

Schellens, P.J. (1991). 'Ad verecundiam and ad hominem arguments as acceptable fallacies'. In: F. van Eemeren, R. Grootendorst, J.A. Blair, & C.A. Willard (eds.), *Proceedings of the Second International Conference on Argumentation* (Vol. 1A, pp. 384-390), Amsterdam: SICSAT.

Taylor, S.E. & S.C. Thompson (1982). 'Stalking the elusive "vividness" effect'. *Psychological Review* 89, pp. 155-181.

Walton, D.N. (1989a). 'Dialogue theory for critical thinking.' *Argumentation* 3, pp. 169-184.

Walton, D.N. (1989b). *Informal logic: A Handbook for Critical Argumentation.* Cambridge: Cambridge University Press.

Walton, D.N. (1989c). 'Reasoned use of expertise in argumentation.' *Argumentation* 3, pp. 59-73.

Willard, C.A. (1989). *A Theory of Argumentation.* Tuscaloosa, AL: University of Alabama Press.

Willard, C.A. (1990). 'Authority.' *Informal Logic* 12, pp. 11-22.

Wood, W. (1982). 'Retrieval of attitude-relevant information from memory: Effects on susceptibility to persuasion and on intrinsic motivation'. *Journal of Personality and Social Psychology* 42, pp. 798-810.

Wood, W. & C.A. Kallgren (1988). 'Communicator attributes and persuasion: Recipients' access to attitude-relevant information in memory'. *Personality and Social Psychology Bulletin* 14, pp. 172-182.

Indicators of independent and interdependent arguments: 'anyway' and 'even'

Francisca Snoeck Henkemans,
University of Amsterdam

Abstract

In this paper, a connection is established between the semantical descriptions of 'anyway' and 'even' given by linguists such as Ducrot & Anscombre, Bennett, Kay and others and the pragma-dialectical characterization of independent and interdependent arguments. By combining linguistic insights with insights from argumentation theory, a more systematic explanation of the indicative function of 'anyway' and 'even' can be given.

1. The analysis of argumentative discourse

When analysing argumentative discourse, the analyst attempts to get a clear overview of the relevant elements in the text and of the relations between these elements. Crucial steps in the analysis of an argument are, *first*, establishing what the communicative function (or illocutionary force) of the different elements is (for instance, whether an utterance should be considered as a standpoint or as an argument) and, *second*, giving a characterization of the relations between the various speech acts (for instance of the relations between the arguments that are advanced).

In the pragma-dialectical approach to argumentation, the ideal model of a critical discussion serves as a heuristic tool for the analysis, or reconstruction, of argumentative discourse. The model specifies which elements are relevant to the resolution of a dispute, in what stage they are situated and what their contribution is to resolving the disagreement. Thus, the idealmodel offers the analyst an analytical perspective. To put it more simply, it specifies what he should look for when analysing argumentative discourse. However, not just any analysis that conforms to the ideal model will do. It should also be justified. In this endeavour, the analyst can refer to textual features and contextual facts. In any case, there should be some empirical support for any particular analysis.[1]

---

[1] For a more detailed account of how normative presumptions and empirical considerations can be reconciled in the analytic reconstruction of argumentative discourse cf. van Eemeren, Grootendorst, Jackson and Jacobs 1993.

It is generally recognized that the verbal presentation of argumentation plays an important role in the identification of standpoints and arguments and the relations between arguments. Nevertheless, until recently, argumentation theorists have not payed much attention to a methodical exploration of the clues provided by the verbal presentation. Some authors of textbooks on argumentation do indeed provide a list of expressions that can be indicative of the communicative function of arguments and standpoints. Usually, they restrict themselves to the most obvious indicators, such as 'because', 'therefore' and 'since'. Authors who mention indicators of *relations between arguments* that can be used to establish how the argumentation is structured, are thin on the ground. Van Eemeren and Grootendorst are an exception: they cite indicators of independent arguments such as 'by the way', 'anyway', 'moreover' and 'needless to add that', and also indicators of interdependent arguments, such as 'when it is also remembered that' and 'in addition to the fact that' (1992: 75-76, 80-81).

It is clear that numerous other indicators could be added to this list. More importantly, a systematic description should be given of the way in which various words and expressions may function as indicators of standpoints, arguments and argumentative structure. By making use of the semantical and pragmatic descriptions of different types of argumentative connectives and operators given by linguists such as Ducrot and Anscombre, Bennett, Fillmore, Kay and others, a more complete and systematic account can be given of the way in which various words and expressions may function as indicators of argumentative relations.

The type of linguistic research I am referring to serves a further theoretical purpose. It can be used to show that the theoretical concepts and distinctions of argumentation theory have their counterpart in ordinary language use.

In this paper, I shall discuss two potential argumentative indicators: *anyway* and *even*. These two adverbs are treated as indicators of argumentative structure both by pragma-dialecticians and by linguists. Van Eemeren and Grootendorst consider *anyway* to be a clear indicator of independent arguments, or - in their terminology - of multiple argumentation. They regard *even* as a somewhat less clear indicator of, again, independent arguments.

In order to clarify the concepts of *independent* and *interdependent* arguments, I shall make use of the dialogical analysis of argumentation structures that I have given in *Analysing Complex Argumentation* (Snoeck Henkemans 1992). The reason for restricting myself to discussing these two types of argumentation structure is that in the literature, *anyway* and *even* are only mentioned in connection with independent and interdependent arguments. I shall attempt to show that characterizing these two argumentation structures dialogically, provides an adequate starting-point for accounting for the indicative function of *anyway* and *even*.

2. Multiple and coordinative argumentation in an argumentative discussion

In most textbooks on argumentation, a distinction is made between interdependent and independent arguments. However, the definitions of the terms 'independent' and 'interdependent' differ considerably from author to author. Often they are also unclear or ambiguous.

In *Analysing complex argumentation*, I have attempted to solve some of these problems of definition, by giving a dialogical characterization of the notions 'independent' and 'interdependent'. In my characterization, arguments are interdependent if the speaker attempts to overcome doubts or answer criticism raised by one or more of his other arguments. They are independent if the arguments are not designed to fulfil such a 'repairing' function with respect to each other.

Coordinatively compound argumentation results from an arguer's attempt to advance additional arguments in order to remove his opponent's doubts or criticism concerning the sufficiency of the argumentation. For this purpose, he can make use of two different kinds of defensive strategy: he can undertake a direct defence or he can give an indirect defence. A direct defence consists of adding further evidence and it results in *cumulative* argumentation. The result of applying this strategy is, that the argumentation consists of a number of arguments that are each individually too weak to support the standpoint, but might in combination provide adequate support:

(1)   The dinner was a paragon of organization, as the hall was the perfect size for the company, the table arrangement was perfectly planned, and the service was excellent.

An indirect defence consists of adding a counterargument to the opponent's objection that the arguer's argument has insufficient weight, and it results in *complementary* argumentation. In both cases, the arguments that are put forward must be thought of as combined, because the arguer can only convince his opponent of the acceptability of the standpoint if he succeeds in removing his doubts, or criticism, regarding the sufficiency of the argumentation. An example of complementary argumentation is:

(2)   We had to go out for dinner, as there was no food at home, and the stores were already closed.

A reaction to the first argument ('there was no food at home') might be that we could have bought some food. The second argument ('the stores were already closed') clearly invalidates such a response.

In multiple argumentation, the only connection between the arguments is that they are all advanced as a defence of the same standpoint. Each of them is a separ-

ate attempt to defend the standpoint, and they are all motivated by the failure, or potential failure, of a previous attempt. The arguments do not require each other to lend adequate support to the standpoint. On the contrary, the reason for undertaking a new attempt to defend the standpoint is that the previous attempt has failed, or might fail. An example of multiple argumentation is:

> (3)   Of course you should buy the laptop computer. It's not really expensive, and equipment for daily use can easily cost a bit more.

In (3), there are two alternative (and even somewhat contradictory) attempts to defend the standpoint. By giving his second argument, the arguer makes it clear that he anticipates that his first attempt at defending the standpoint might fail.

3. A semantical analysis of 'anyway'

In order to account for the fact that the adverb *anyway* may function as an indicator of multiple argumentation, I shall make use of the analysis of the French word *d'ailleurs* given by Ducrot in *Les mots du discours* (Ducrot et al. 1980). *D'ailleurs* is roughly equivalent to the English *anyway*. I shall give a brief and simplified summary of the main characteristics of *d'ailleurs* mentioned by Ducrot.

According to Ducrot, the complete semantic scheme of all uses of *anyway* is visible in example (1):

> (1)   I don't want to rent this room (r): it is too expensive (P) and anyway, I don't like it. (Q).

According to Ducrot, *anyway* always functions in an argumentative context. It is used to present an argument (Q) that is added to one or more other arguments (P). P and Q are advanced in support of the same conclusion (r), but are argumentatively independent[2].

The argument Q constitutes what Ducrot calls a second discursive movement. By using *anyway*, the arguer gives the impression that he first only wanted to give the argument P, but that he has had second thoughts about it and decided to give the argument Q as well. P and Q are directed at different people or, rather, at different conceptions of the same person. Ducrot explains this as follows: the argument P is directed at an opponent for whom the argument P should suffice as support for the conclusion r. By adding *anyway*, the arguer indicates that he doubts the legitimacy

---

[†]   Luscher (1989: 118-119) points out that *d'ailleurs* can also be used non-argumentatively by a speaker, to comment upon or correct a preceding speech act.

of his own assumption. He anticipates the possibility that the empirical opponent might not be willing to accept *r* on the basis of *P* (1980: 217).

From this brief description, it should be clear that by the use of *anyway* the arguer can indicate that the dialogical situation he is anticipating is precisely the situation which, according to the pragma-dialectical analysis that I have presented here, gives rise to multiple argumentation: the situation in which the arguer decides to undertake a new and separate attempt to defend the standpoint, because he expects that a previous attempt might fail.[3] By using *anyway*, the arguer makes it clear that his first argument should have sufficed to convince his opponent. All the same, he advances a new argument, thus making it clear that he foresees that his first argument might after all not be convincing.

4. A semantical analysis of 'even'

In the French linguistic literature, *even* is generally associated with interdependent arguments. Roulet, for one, whose analysis is based on the work of Anscombre and Ducrot, thinks that one of the crucial differences between *anyway* and *even* is that *anyway* introduces independent arguments, and *even* interdependent arguments. In order to take a more well-founded decision on whether *even* is to be regarded as an indicator of multiple or as an indicator of coordinative argumentation, I shall compare the analysis of *even* given by some American linguists with the analysis of the French linguists of *même*, which is virtually equivalent.

In the semantical and pragmatic analysis of *even* and of its French counterpart *même*, two elements recur. In the words of Kay:

(2) A (textual) sentence containing *even* always depends on a contextual sentence (expressed or implied) which is, intuitively speaking, less 'extreme' (1990: 92)

A correct use of *even* thus requires the implicit or explicit presence of one or more other sentences or propositions.[4] Furthermore, the sentence containing *even* indicates a more extreme case than the other sentences or propositions.

---

[3] Since multiple argumentation consists of separate, and often even alternative attempts at defending a standpoint, it may happen that the arguments are such that they cannot all be intended to convince the same opponent (As is the case in my example 3). As is explained in Snoeck Henkemans (1992: 141), one of the arguments may be irrelevant for an opponent who accepts the other argument. The same phenomenon is mentioned by Ducrot (Ducrot et al. 218) in his analysis of *d'ailleurs*.

[4] Adler (1992: 26) does not think this is a necessary requirement. He offers the following counterexample: 'If Arthur is one of the worst students in my class and his is the first paper I grade, I might say [...] Even Arthur got an A. But I surely would not implicate that others beside Arthur received A's. I have not yet looked at their tests'. However, I doubt whether 'even' would be correctly used in such a context.

It differs from author to author what exactly is to be understood by 'more extreme'. For Fillmore (1965), Bennett (1982) and Lycan (1991), more extreme means: less expected, or more surprising. After having given the sentence 'Even Max tried on the trousers', Bennett, for instance, remarks:

(3)   One thinks of this as felicitously said in a situation where Max did try on the trousers, and so did someone else, and it is more surprising that Max did than that the other person did (1982: 404-405)

Unlike Bennett and others, Kay, just as Anscombre and Ducrot, considers the more surprising character of the sentence containing *even* neither a necessary nor a sufficient condition for a felicitious use of *even*.[5] For Kay, 'more extreme' means: 'more informative'. This, in turn, means that the sentence with *even* unilaterally entails the other sentence or sentences:

(4)   In the analysis presented here, the intuition of being more extreme is ex-plicated as greater informativeness, in turn defined by unilateral entailment in a scalar model (1990: 92).

What Kay means by unilateral entailment in a scalar model, can be explained by looking at his analysis of sentence (5):

(5)   The whole family showed up for Christmas, even aunt Irma.

When interpreting this sentence, it should be assumed that aunt Irma is located at the lowest point of a scale (for instance a scale of regular visitors), and that therefore her showing up a fortiori entails the other members of the family showing up. Intuitively, this may seem a bit strange: one would think that the entailment should go in a different direction, namely that the showing up of the whole family entails aunt Irma's showing up, and not conversely. According to Kay, this only shows that it is not entailment per se that we are concerned with, but entailment in a scalar model. That the whole family's showing up entails aunt Irma's showing up, is an entailment which, in Kays words, 'hold(s) in a context structured by a scalar model but which owe(s) nothing to the scalar model' (80). Such fortuitous entail-ments should be distinguished from entailments that exploit the scalar property of the model.

5   Lycan, although he uses the notion of 'unexpectedness' in his own analysis of the scalar properties of *even*, also points out that *even* does not necessarily have to introduce a more surprising event: 'Though all this talk of 'expectedness', 'likelihood', 'surprisingness' etc. is standard in the literature (...), it is misleading. Whatever scalar notion really is in play here is not always so fortrightly epistemic' (1991: 122).

Anscombre and Ducrot (1983) also analyse *even* as a scalar operator. According to them, the sentence containing *even* is not more informative, but it has greater argumentative force. When evaluating their proposal, one should bear in mind that Anscombre and Ducrot use the term 'argumentative' in a broader sense than argumentation theorists generally do. They consider as argumentative all utterances that lead the listener or reader, often implicitly, to a certain conclusion. Any evaluative meaning that transcends the purely informative (quantitative) meaning of a sentence is regarded as 'argumentative' by these authors.

I shall not concern myself further with the question as to whether or not *even* always has an argumentative function, since it is indisputable that it *can* be used in an argumentative context to indicate that the argument that follows is the strongest argument for a conclusion. According to Kay (1990: 91), the fact that *even* marks an assertion as more informative makes it particularly suited to be used in service of argumentative goals, but this does not mean that it can be used exclusively for argumentative purposes.

In an argumentative context, the conjunct with *even* usually introduces the strongest argument. However, as both Anscombre and Ducrot and Kay point out, there is an exception to this rule. There are cases in which *even* does not refer to the last argument or proposition, but tells something about the group of arguments as a whole. An example is (6):

(6) George drank a little wine, a little brandy, a little rum, a little calvados, and even a little armagnac.

According to Kay, in a sentence like this, *even* does not indicate that the last proposition, in this case, drinking armagnac, is the more extreme:

(7) The final conjunct is not interpreted as more extreme than the preceding one(s), rather the whole sentence is seen as more extreme than the initial conjunct(s) (1990: 74)

The full conjunction, including armagnac, is more informative, or stronger, in a scalar model than the conjunction which lacks armagnac. If (6) were used as support for the claim that George drank a large quantity or a large diversity, mentioning five beverages would produce a stronger argument than mentioning just four. In cases like this, *even* seems to have a similar meaning as the word *plus*.

According to Anscombre and Ducrot (1983), *even* can only be used to refer to the group of arguments as a whole, if each of the arguments by itself can lend some support to the standpoint. Since this condition is not fulfilled, sentence (8) is not acceptable:

(8)   *John can speak exactly three languages: English, French and even German.

In a case like (8), the standpoint requires a purely quantitative interpretation. None of the arguments by itself can be seen as an argument for the standpoint that John can speak exactly three languages. If *exactly three* were replaced by a more vague and qualitative expression like *quite a few*, the use of *even* would be appropriate.

One would think that in cases in which *even* serves as an indication of the argumentative force of the combined arguments, as in example (6), the order in which the arguments are presented is of no importance. Since the final argument is then not interpreted as more extreme or more strong, any other order of the arguments should also be acceptable[6]. In example (6) this seems to be true. However, there are also cases which have the same property as example (6) that none of the arguments by itself can support the standpoint independently, but where the order of the arguments is nonetheless not arbitrary. This is illustrated by the examples (9a) and (9b):

(9a)   *Mary can speak quite a few languages: she speaks French, Chinese and even English.
(9b)   Mary can speak quite a few languages: she speaks English, French and even Chinese.

In a context where Mary is a native speaker of English, (9a) seems to be less acceptable than (9b). According to Anscombre and Ducrot, examples such as these make it clear that the primary function of an expression like *quite a few* is not to indicate a certain quantity, but rather to give an evaluation. Being able to speak Chinese is presented as a stronger argument for a positive evaluation of Mary's talent for learning languages than the other arguments that are advanced.

According to Anscombre and Ducrot (1983: 63), the relations between the arguments in contexts in which *even* is used may vary from complete interdependence (as in example (6)) to something between interdependence and independence (as in example (9b)). Even if the argument that is introduced by *even* is presented as the strongest argument for the standpoint, it can still be the case that the arguer considers the combination of the arguments as stronger than the strongest argument seen in isolation.

Apart from the feature that *even* either introduces the strongest argument, or refers to the arguments as a whole, Anscombre and Ducrot also mention the feature that

---

[6]   This is, in fact, Fauconnier's (1976: 262) argument for attributing a different effect to the presence of 'even' in the final conjunct in such cases: the effect is then not to indicate that the final conjunct is improbable to the highest degree, but 'to underline the increasing improbability of each proposition, given the presentation of the preceding one'.

the proposition in which *even* is introduced and the propositions preceding this proposition, should be located on one and the same scale. How one can establish what type of scale this ʾshould be they do not discuss in any detail. They only indicate that the standpoint supported by these arguments is an important factor. This can be shown by comparing the examples (10a) and (10b):

(10a) Christmas was awful this year: my parents came to visit and even aunt Irma came.

(10b) Christmas was wonderful this year: my parents came to visit and even aunt Irma came.

In (10a), the dimension involved is one of awfulness, and aunt Irma's visiting is considered to be more awful than the visiting of the parents. In (10b) the dimension involved is one of wonderfulness, and aunt Irma's visit is presented as a stronger argument for the success of the occasion than the visit of the parents.

In my opinion, there is still another factor that influences the determination of the scalar dimension at issue. This factor consists of the arguments themselves. My point can be illustrated with the help of the examples (11a) and (11b):

(11a) My cat has barely eaten anything for two days, and today not even anything, so it must be ill.

(11b) My cat hasn't eaten anything for two days, and today it hasn't even drunk anything, so it must be ill.

In example 11a, the amount of food is presented as an indication of illness, and eating nothing is considered to be a stronger argument for the cat's being ill than eating barely anything. In 11b, the general consumption of the cat is under consideration, and not drinking is taken to be a more reliable sign of illness than not eating.

5. Conclusion

I would now like to return to the issue of whether the operator *even* should be considered as an indicator of independent arguments or as an indicator of interdependent arguments. In my opinion, if the different semantic properties of *even* are taken into consideration, it becomes apparent that *even* can function as an indicator of interdependent or - in the pragma-dialectical terminology - of coordinatively compound arguments.

I have shown that there are two ways in which *even* can be used. First, it can be used to indicate that the sentence as a whole is a stronger argument for the conclusion than the sentence minus the conjunct containing *even*. In that case, it is clear that the arguments should be taken together, and are thus interdependent: the

combination of arguments is stronger than any of the arguments considered in isolation, or any other combination of the arguments.

Second, *even* can be used to indicate that a particular argument constitutes the strongest evidence for the conclusion. In theory, it could be the case that the argument concerned is so strong that it could, by itself, be a sufficient defence for the standpoint. By using *even*, however, the arguer indicates that he has more arguments that are situated on the same scale. The final argument may tip the scale in favour of the standpoint, but the other arguments still play a reenforcing role. Unlike in the case of multiple argumentation, the arguments are not of a different order or type. Therefore, I think that in that case, analysing the arguments as interdependent is also to be preferred.

Even can be an indicator of *cumulative* argumentation: then each of the arguments lends some support to the conclusion by itself, but only the combination of arguments can be intended by the arguer as a sufficient defence of the conclusion.

Even can also be used as an indicator of *complementary* argumentation, as in the following example given by Kay:

(12)   He worked hard, and the boss wasn't even there.

The implicit conclusion that he deserves praise, is supported by two arguments. The argument introduced by *even* can be seen as an attempt to counter the objection that he only worked hard because his boss was around.

I hope to have shown that by combining the semantical analyses of linguists such as Ducrot and Anscombre and Kay with the theoretically motivated distinctions of the pragma-dialectical theory of argumentation, a more systematic explanation of the function of specific indicators of argumentation and argumentation structure can be given. *Anyway* does indeed seem to be particularly suited to function as an indicator of multiple argumentation, whereas *even* should rather be seen as an indicator of coordinatively compound argumentation.

References

Adler, J.E. (1992). 'Even-arguments, explanatory gaps and pragmatic scales'. *Philosophy and Rhetoric* 25, 1, pp. 22-44.
Anscombre, J.C. & O. Ducrot (1983). *L'argumentation dans la langue*. Bruxelles: P. Madarga.
Bennett, J. (1982). 'Even if'. *Linguistics and Philosophy* 5, pp. 403-418.
Ducrot, O. et al. (1980). *Les mots du discours*. Paris: Les éditions de minuit.
Eemeren, F.H. van & R. Grootendorst (1992).*Argumentation, Communication and Fallacies. A pragma-dialectical perspective*. Hillsdale, NJ: Lawrence Erlbaum.
Eemeren, F.H. van, R. Grootendorst, S. Jackson & S. Jacobs (1993). *Reconstructing Argumentative Discourse*. Tuscaloosa, Alabama: The University of Alabama Press.

Fauconnier, G. (1976). *Etude de certains aspects logiques et grammaticaux de la quantification et de l'anaphore en français et en anglais.* Paris: Champion.

Fillmore, Ch.J. (1965). *Entailment rules in a semantic theory.* POLA Report 10. Columbus: Ohio State University.

Kay, P. (1990). 'Even'. *Linguistics and Philosophy* 13, pp. 59-111.

Luscher, J. (1989). 'Connecteurs et marques de pertinence. L'example de d'ailleurs'. *En deça et au-delà du discours. Cahiers de Linguistique Française* 10, pp. 101-145.

Lycan, W.G. (1991).'*Even* and *even if*'. *Linguistics and Philosophy* 14, pp. 115-150.

Roulet et al. (1985). *L'articulation du discours en français contemporain.* New York/Paris: Lang.

Snoeck Henkemans, A.F. (1992). *Analysing Complex Argumentation; the Reconstruction of Multiple and Coordinatively Compound Argumentation in a Critical Discussion.* Amsterdam: Sicsat.

Accounting for transformations in the dialectical reconstruction of argumentative discourse

M. Agnès van Rees,
University of Amsterdam

Abstract

In this paper, an argument is made for the importance of taking into account the social inter-actional aspect of argumentative discourse, when reconstructing such discourse as a critical discussion. The analysis of a particular problem-solving discussion shows how dialectical trans-formations can be accounted for by an appeal to social practices, especially with regard to the maintenance of a status balance between the participants.

In order to enable an adequate evaluation, in pragma-dialectics, argumentative discourse is subjected to a dialectical reconstruction, highlighting those elements which the evaluation will address. Dialectical reconstruction entails looking at argumentative discourse from a particular, theoretically motivated point of view: the discourse is viewed as an attempt to attain the rational resolution of a conflict of opinion. The reconstruction is guided by a conception of what is necessary for the rational resolution of a conflict of opinion, represented in an idealized model of critical discussion. It abstracts those (and only those) elements in the discourse which are relevant with regard to this particular goal. The reconstruction results in an analytic overview in which the differences of opinion, the distribution of dialectical roles, the expressed and unexpressed premises which make up the arguments, the argumentation structure and the argumentation schemes of the arguments are laid out (Van Eemeren, Grootendorst, Jackson, and Jacobs 1993).

In order to arrive at such an analytic overview, a number of dialectical trans-formations are carried out on the discourse, which bring into focus those elements in the discourse which potentially contribute to the resolution of a conflict of opinion. These transformations are: deletion, addition, permutation, and substitution. The transformation of deletion selects those elements that are immediately relevant to the resolution, omitting what is irrelevant to this goal. The transformation of addition makes explicit those elements that are immediately relevant to the resolution but which have been left implicit in the discourse. The transformation of permutation rearranges elements in the discourse in such a way as to mirror the order in which the resolution

ideally is attained. The transformation of substitution, finally, reformulates relevant elements, in such a way as to most clearly show up their function in the resolution.

It is important to realize that pragma-dialectical reconstruction, like, indeed, any reconstruction, necessarily is an abstraction, even if it is a legitimate one for an analyst who is interested in evaluating discourse with a view to its dialectical rationality. The discourse is regarded as directed at the attainment of one particular goal, the rational resolution of a conflict of opinion. In actual fact, discourse usually is aimed at realizing a multitude of goals.

In this paper, I will argue that in reconstructing argumentative discourse, it is important to be aware of the existence of these other goals. I will argue that such an awareness is necessary to account for the presence of other elements in the discourse than those that are dialectically relevant, and so to justify their deletion in dialectical reconstruction, as well as to account for the fact that elements which are dialectically relevant sometimes do not look as if they are, and so to justify their reconstruction as relevant elements through substitution or addition.

I will focus on one particular type of goal, namely, one that is a corollary from the fact that argumentative discourse, being discourse, is a form of social interaction.[1] For a perspective on what this implies, we may look to the literature on conversational interaction.

Cheepen (1988) is a particularly relevant source here, since she specifically focusses on the social aspect of conversational interaction. In her study of informal spontaneous conversation, she convincingly argues that the establishment and monitoring of an appropriate interpersonal framework account for much of the linguistic work done by speakers. In her view, 'the interpersonal component is the basis on which other strands of meaning are built' (1988:3).

Of central importance in this interpersonal component is the concept of status. Cheepen holds that status, or the power relationship obtaining between the participants in an interaction, is central to the way in which the discourse is developed. Controlling the direction of the talk, for example, is the prerogative of the superior speaker. This may be done by such means as changing the topic or performing a framing move.[2] Or, to mention another example, repair actions, often profoundly influencing the subsequent course of the conversation, have to be undertaken when the status balance is disturbed. This happens, for instance, when one participant bluntly tells the other one what to do, or openly mocks him or disagrees with him.

---

[1]   In the literature about problem-solving discussions we find a reflection of the insight that discourse serves social and interactional goals in addition to other goals, in the distinction which traditionally is made between task-related and socio-emotional goals (Bales 1958, Maier 1963, Fisher 1980). While some authors view the social-emotional preoccupations of participants as a potential danger for the achievement of the task-related goals (e.g. Maier 1963), others, such as Fisher (1980) take a more positive view.

[2]   Examples of framing moves are: 'Right!', 'Okay'.

In this paper, I will examine a fragment of a real-life problem-solving discussion in which the participants are trying to solve a conflict of opinion. Generally, the interactional work participants do is most conspicuously present in the opening and closing stages of the interaction, when they are exchanging greetings etc. But I am particularly interested in the more task-related parts of the discussion in which the problem-solving and conflict resolution activities proper are being conducted. It will soon become clear that, there too, status work is all-pervasive.

The discussion which I will examine is one of a series of conversations in which two members of the management of a hospital deliberate with an outside PR adviser on the best strategy to pursue in negotiations with a nearby hospital, Verana, regarding future collaboration. Pressure is being exerted by the Ministry of Health to arrive at some form of collaborative agreement in the short term.

The discussion bears many of the features of ordinary everyday conversation. There is no institutionally determined, predictable sequence of events as might be found in, say, a court hearing or local council meeting. There is no chairman allocating turns to speak, there is no fixed order in which speakers have the floor, there is no pre-determined agenda, and there are no particular rights or obligations regarding who is entitled to perform which speech acts. The conversationalists are all of equal status. They are on Christian-name terms.

There are, however, a number of functional differences between the participants, relating to their role in the negotiations and the nature of their work: participant A is conducting the negotiations on behalf of the hospital management and is writing a draft plan of collaboration which will be the hospital's basis for negotiation, participant C is the second member of the hospital management, and participant B is the external PR adviser who is not a part of the hospital's management structure, so a relative outsider. Most of the conversation takes place between A and B.

The discussion centers on a problem raised by B: there is a risk[3] that the negotiators on the other side will deliberately try to delay the collaboration, so that the Ministry will gain the impression that 'the whole thing is too complicated' and accordingly impose amalgamation.[4] B raises this problem at the beginning of the fragment (lines 18-41).

---

[3]   The word *risk* which B uses, is an indication of the fact that a problem is being raised (cf. Jordan 1984). Amalgamation is evidently regarded by the conversationalists as an undesirable option.

[4]   In problem-solving discussions, of which the present one is an instance, the participants try to reach a solution to a problem through discussion. During the various phases of the problem-solving process, participants have to resolve various differences of opinion. These differences of opinion can relate to all stages of the problem-solving process: the participants may disagree on whether a problem exists at all, what it is (if it exists), what the potential solutions might be, by what criteria these solutions ought to be judged, and what the judgement ought to be. The present discussion concerns the first of these questions.

18 B: but what what you were saying Frits, (.) I thought you well that we would get
19   to that, the risk (.) that they er that maybe they deliberately, by er now once
20   again well just as we were saying dragging their heels in the sand and slowing
21   things down and ⌈er not doing
22 A: ⌊yes
23 B: anything (.) that i-if you found yourself in the situation that even at the ministry
24   you had to say that you still hadn't really got anywhere at all with the
25   collaboration, huh, with filling in the details of the collaboration, and still had
26   no more than what er whatsisname er (.) Boom called that sherry er
27   ⌈agreement
28 A: ⌊yes yes, that's what I told them again this afternoon. yes
29 B: that at some stage a sort of er (.) unease or uneasiness or or er irritation will set
30   in for people who actually do have something to say about the collaboration,
31 A: yes
32 B: and they say okay, cut the cackle, this is getting no-one anywhere, you people
33   want (.) to set up a form of collaboration that no-one's ⌈got any experience of at
34 A: ⌊that's not going to earn
35   anything
36 B: all, er er that's much too complicated, and and you clearly haven't got anything
37   down on paper, er it's not going to work, one party is working in quite a different
38   direction from the other, (.)
39   ⌈cut the cackle,   ⌈amalgamation
40 A: ⌊(yes but)   ⌊yes yes but
41 B: hey, at least we know that model

During B's introduction of this problem, A several times throws in a 'yes', and twice provides a supportive elaboration. Now, how should we reconstruct A's contribution? Should we, for example, reconstruct it as an expression of agreement?

There are several reasons not to do so. For one thing, in what follows, we shall see that A does not at all agree with B's claim that there is a problem. Now of course, maybe in that case we should impute inconsistency to A. But I don't think so.

To begin with, A's utterances are in accord with what the general principles of turn-taking in conversation require. A's behavior is the conventional way of showing listenership, that is, displaying his understanding of B taking an extended turn and his willingness to let him do so (cf. Sacks, Schegloff, and Jefferson 1974, Schegloff 1982, Bublitz 1988).

In the second place, A's acquiescence is in accord with what is known about the various stages of the social aspects of the process of decision-making. As, for example, Fisher (1980) has shown, in the first stage of this process, the orientation stage, contributions of the participants are aimed at avoiding open conflict and keeping the social climate friendly. They are not meant to express standpoints which the speaker is prepared to defend.

Finally, there is a third consideration: in these contributions, in addition to the work just described, A is doing status work as well: he is demonstrating his being in the

know and in control. This becomes especially clear in lines 28 and 34-35. In line 26, A emphatically agrees with B, adding that he told 'them' so several times himself. In line 34-35, A himself supplies the information on what the Ministry might say, which B is in the course of providing.

So all in all there is good reason not to reconstruct A's contributions as an expression of agreement, but to consider them as relevant to other goals than that of resolving the conflict of opinion, and so to delete them in the reconstruction.

Once A takes over the floor (after a couple of interruptions by C which are not rendered here), we can clearly see that his position is one of putting forward an opposing standpoint: he argues (in lines 59-65) that the problem B raises does not exist.

59 A: but don't worry, if I can just if I can just talk about him
60 C: yes
61 A: Egberts has tackled that point very well. he was clearly looking strictly for
62   simplification, of the formula. for the collaboration. and to start with it wouldn't
63   work out the way we wanted but later it did, and you'll see it will end up a véry
64   simple, clear, lucid, binding formula. and the rest of it is all verse eighty-three,
65   and *that* formula will go there and *then* their lordships will be satisfied.

Contrary to B's allegation, A claims that the Ministry will not think things are too complicated. To support this contention he advances the argument that Egberts will work out a very simple formula for collaboration in agreement with the wishes of 'our' side, which will be presented to the Ministry.

Clearly, a lot of reconstruction work is required in order to represent A's contribution in this way. We have to reconstruct 'but don't worry' as the standpoint 'there is no such problem', and 'their lordships will be satisfied' as the argument 'the Ministry will not think things are too complicated'. How can we warrant such a drastic move?

One very obvious justification is that we may assume A is trying to make his contribution be one that is relevant to the ongoing course of the talk - in accordance with Grice's Cooperative Principle (Grice 1975). And since B has just put forward the standpoint that there is a risk that the Ministry will think things are too complicated, a relevant sequel would be either to agree or to disagree with this standpoint.

But we can add to this justification by pointing out that the fact that A formulates his contribution in this particular way, can be accounted for in terms of status work. 'Don't worry' is an instantiation of a particular status-gaining strategy which consists, as Cheepen suggests, in displaying other-attentiveness while the other person is not in a position to reciprocate. A, in other words, is patronizing. 'Their lordships will be satisfied' is another formulation in which A is taking a superior stance, in this case, towards the Ministry. This same superior stance is manifested in the way in which he phrases the argument as a whole, which is one big display of control and superior knowledge: evaluating Egberts's behavior, elaborating on the development of the

negotiations which he himself conducted, assuring B ('you'll see'), and, finally, predicting and evaluating how it will go in the end. All these are status-raising techniques, not immediately relevant to the resolution of the conflict of opinion as such.

In the reconstruction, then, A's contribution would be stripped of these authoritative overtones, and pared down to the core of his standpoint and the arguments that he puts forward in support for it.

Turning to B's reaction to A's opposition against his standpoint, we find that it also stands in need of reconstruction. B's contribution runs like this:

```
70     (---)
71  B: ye-es, that is of course ⌉ I think it is very important
72  A:                          │ yes
73  B: that that ⌈ Egberts should stay on our side
74  A:           │ and that part that's exactly his line and we agree to that and Bob's
75     your uncle. yeah you have to ram it down their throat that's all I can
76     say about it
```

At first sight, when looking at B's reaction, we might think we have to do with an expression of agreement. But there are several reasons for not reconstructing it so.

B's initial reaction to A's argument that the problem is non-existent is silence (line 70). Moreover, once he embarks on a reply, he does so by starting out with a concession, signalled by the hesitantly drawn out 'ye-es' and the expression 'of course'. In addition, this concession refers to only part of A's argument, namely that Egberts is taking the same line as 'us'. B says nothing about whether Egberts's support offers a solution for the problem he has presented.

Silence, concessive start and the absence of explicit, direct agreement all are in accordance with a general conversational strategy for expressing disagreement. The strategy is aimed at minimizing the threat to the social face of the interlocutor which is inherent in producing a dispreferred second pair-part such as disagreement (cf. Pomerantz 1984).

In other words, there is a clear justification for reconstructing B's utterance, not as agreement, but as disagreement with A's argument that the problem no longer exists.

Unfortunately for the analyst, B is interrupted by A and unable to finish his contribution. But then, the interruption itself lends support to the above reconstruction, because A most certainly appears to interpret B's contribution as an expression of disagreement: he advances support for his assertion regarding Egberts's position ('that's exactly his line').[5]

---

[5]   Note, by the way, how, here too, A takes a superior role, in unilaterally closing the subject through his concluding generalization in lines 75-6 (cf. Polyani (1985) for this technique of topic closing).

Moreover, B's subsequent actions are consistent with this reconstruction. The first time B gets the floor again is in line 103. At that point A had been responding to C, who had brought up a concern of his own, not related to B's warning. C had been saying that 'we' don't have to talk to 'them' at all. A was objecting that we can't avoid talking to them because there will be a tug-of-war about the outpatient departments.

96 A: we've alréady worked out what [they] can have. (.) that's easily worked out, it's
97   twenty outpatients departments. I'm not familiar with their hospital. that may be
98   crazy but I'm not familiar with it. not, that's something they've never revealed,
99   in figures. but if you ask me they're looking for well over twenty outpatients
100   departments. in all, and *that*'s when the tug-of-war starts. because then *we*'ll
101   have too many.
102   (-)
103 B: well ex*act*ly, yes but *that*'s the point at which it *is* relevant again what they're
104   going to do with all those outpatients departments
105 A: ye-es
106 B: and for *them*, the *way* the collaboration is worked *out*
107   in detail is *also* ⌈ going to be important again,
108 A: ⌊ *yes*
109 B: ⌈ if they haven't yet ⌈ realized it ⌈ then they soon will
110 A: ⌊ *yes*   ⌊ *yes*   ⌊ *yes*

Latching onto A's response to C, B, after a short pause, begins his turn with 'well ex*act*ly'. Initial *well* is usually a means of distancing oneself from a previous position, but it is followed here by a strong expression of agreement, which would lend support to reconstructing it as agreement to A's position in regard to C's claim. However, this initial agreement is immediately followed by an opposition-indicating 'yes but', and the assertion that at *that* point for *them* it *is* going to be important how the collaboration is worked *out* in detail. In other words, B is agreeing with A's statement, while at the same time distancing himself from the direction the discussion is taking. He uses A's assertion as a lever for taking the discussion back to his own earlier point about the other side's evil intentions regarding the collaboration. The two *again*s (lines 103 and 107) indicate that these are the same intentions which he alluded to before, when he raised the problem for the fist time.

Further support for this analysis can be found in B's reaction, in line 124, to A's subsequent response (which is not rendered here).

124 B: yes but let me put the question differently, okay,
125 A: yes
126 B: d'you think, ⌈ (-) maybe they've somewhere (.) got a
127 A: ⌊ yeah
128 B: secret agenda after all that they actually want a full amalgamation?

The issue B raises here ('d'you think maybe they've got a secret agenda that they actually want a full amalgamation?') ties up with his own initial introduction of the problem in which he indicated that the other side was deliberately obstructing in order to get the Ministry to impose an amalgamation. By means of this question B attempts to elicit as a concession from A the assertion that Verana is in effect trying to bring about a full amalgamation, thus getting A to formulate the main argument for B's own position. With 'let me put *the question differently*', he lets it be known that his previous contribution was intended to be interpreted in the same vein.

All this lends support for reconstructing B's contributions as providing support for his initial standpoint, that there is a problem, and as an effort at refutation of the argument which A brought forward in opposition to that standpoint.

Now, of course, in none of the three utterances of B which we just examined this position is explicitly or directly present. At no point does B provide an explicit link to his initial standpoint or to A's opposing standpoint.

Yet, the substitution and addition transformations which the reconstruction requires can be justified, if we take into account that this implicitness and indirection of B serve social, interactional goals. They are an instantiation of the general face-saving strategies which were mentioned above, in the discussion of lines 70-6. B is at pains to avoid a direct expression of his dissatisfaction with the way the discussion is going and of his disagreement with A's opposition to his standpoint, because direct and explicit expression would mean a strong threat to A's social face.

In what follows, a lengthy debate arises about whether the other side is aiming for amalgamation. B keeps asking questions trying to elicit a concession on the part of A that the other side is actually striving for amalgamation, and A consistently keeps answering them in the negative. During this debate, A once again broadly displays his superior knowledge and control by detailed elaborations on the course of the negotiations and his part therein, once again determining the direction of the conversation to a large degree by unilateral topic closures and interruptions.

But I will refrain from discussing these sections in detail and focus instead on the moment when A finally concurs with B (line 300, 'but actually you're right you know').

300 A: but actually you're right you know, if er friend Van Denen happens to say,
301   during any other business, let's just think about this regional OR for a moment,
302   then that means that we haven't *yet* given up that point about training you
303   know, not even for those three years (.) forget it (.). I sometimes think in that
304   subcommittee, with Van Denen, er that can easily get derailed. (--) what I
305   mean is, that's just when Van Denen will start saying that kind of thing about
306   amalgamating, and er
307 C: no you're absolutely right
308 A: hey, you know? all that kind of pushing
309 C: yes

A certainly does not formulate his agreement as one acceding to the argument of his opponent. Rather, he brings it forward as one who knew it all to begin with, providing an argument of his own, based on his experience in the negotiations. Does that mean we should reconstruct this contribution as the expression of an evaluative standpoint for which argumentation is advanced?

I don't think so. Again, there are various reasons for this.

For one thing, it is a well-documented conversational procedure, when expressing agreement, not to just claim agreement, but to demonstrate it as well, for example, by formulating considerations of one's own (cf. Sacks 1992, Houtkoop-Steenstra 1987). This is what A is doing here.

Second, according to what is known about the pattern in which the social process in decision-making discussions evolves, in the final stage of the discussion, the confirmation stage, participants aim to express and strengthen agreement. Argumentation in that stage does not serve to overcome disagreement, but to confirm agreement (Fisher 1980).

And, finally, A's particular way of phrasing his agreement, once again, can be accounted for as the result of status work. Status-gaining strategies can be seen to be at work in formulations like 'actually', claiming to possess superior information about how things really are, and 'friend Van Denen', showing ironical condescension, and in A's detailed elaboration of his experiences in the negotiations which he conducts.

All these considerations support reconstruction of A's contribution, not as a speech act belonging to the opening or argumentation stages of a discussion, but as a speech act belonging to the concluding stage of a discussion, *in casu* the withdrawal of doubt and of an opposing standpoint.

In the above, I have justified particular dialectical transformations in reconstructing contributions of the participants to a problem-solving discussion, by showing that the way in which these contributions take shape can be accounted for in terms of the interpersonal work which the participants are doing.

So far, this explanation has been quite general in nature, pointing to general face-saving and status-establishing strategies which are operative in all conversation. But a more specific explanation for the verbal behavior of the participants to the discussion can be given, as well. This verbal behavior can be tied to the differing interests which are implied by the different positions which the participants hold in the organization to which they belong. A, who as a member of the board of directors of the hospital carries authority in the negotiations which he conducts, requests the advice of B, who, as an external PR advisor, doesn't stand in any hierarchical relationship to A and carries his own weight in PR matters. For B, this means that he has to be careful not to infringe on the authority and substantial right of decision which A possesses. B thus has to manoeuvre carefully. Hence the indirectness and implicitness of his contributions. A, on his part, has to show that as a negotiator he is capable and informed.

Hence his taking ample time to expand on the negotiations and his own role therein. Hence, also, his presenting himself as someone who has everything under control.

What we can learn from all this, is that in all forms of discourse, language use is geared towards serving several purposes at once. Some of these are social in nature, such as establishing and maintaining a balance of status, some representational, such as trying to resolve a difference of opinion in a rational way. Only the latter kind provides a context for dialectical reconstruction and evaluation.[6] But, in carrying out these tasks, we cannot afford to ignore the former.

---

[6]   To be sure, status interactions as such need in no way form an impediment to dialectical rationality: B and A don't try to gloss over or conceal their conflict of opinion, but try to resolve it through a regular exchange of arguments and critique. B finds the 'solution' A thinks to have provided unsatisfactory and tries throughout the discussion to show that the problem still exists. His concern to keep the status balance undisturbed does not induce him to cover up the lack of agreement or to abandon the issue. A, for his part, does not let his concern to protect his status prohibit him from giving up his standpoint in the end.

References

Bales, R.F. (1958). 'Task Roles and Social Roles in problem-solving groups'. In: E.E. Maccoby, T.M. Newcomb & E.L. Hartley (eds.), *Readings in Social Psychology* (pp. 437-447), New York: Holt, Rinehart & Winston.

Bublitz, W. (1988). *Supportive Fellow-Speakers and Cooperative Conversations. Discourse Topics and Topical Actions, Participant Roles and Recipient Action in a Particular Type of Everyday conversation.* Amsterdam: Benjamins.

Cheepen, C. (1988). *The Predictability of Informal Conversation.* London: Pinter publishers.

Eemeren, F.H. van, R. Grootendorst, S. Jackson & S. Jacobs (1993). *Reconstructing Argumentative Discourse.* Tuscaloosa: The University of Alabama Press.

Fisher, B.A. (1980). *Small Group Decision Making: Communication and the Group Process.* New York: McGraw-Hill.

Grice, H.P. (1975). 'Logic and conversation'. In: P. Cole & J.L. Morgan (eds.), *Syntax and Semantics, III: Speech Acts* (pp. 41-58), New York: Academic Press.

Houtkoop-Steenstra, J.P. (1987). *Establishing Agreement. An Analysis of Proposal-Acceptance Sequences.* Dordrecht: Foris.

Jordan, M.P. (1984). *Rhetoric of Everyday English Texts.* London: Allen & Unwin.

Maier, N.R.F. (1963). *Problem-solving Discussions and Conferences: Leadership Methods and Skills.* New York: McGraw-Hill.

Polyani, L. (1985). 'Conversational Story-telling'. In: T.A. van Dijk (ed.), *Handbook of discours analysis, III: Discourse and Dialogue* (pp. 183-201)), London: Academic Press.

Pomerantz, A. (1984). 'Agreeing and disagreeing with assessments: some features of preferred/ dispreferred turn shapes'. In: J.M. Atkinson & J. Heritage (eds.), *Structures of Social Action. Studies in Conversation Analysis* (pp. 57-102), Cambridge: Cambridge University Press.

Sacks, H. (1992). *Lectures on Conversation* (G. Jefferson, ed.). Oxford: Blackwell.

Sacks, H., E.A. Schegloff & G. Jefferson (1974). 'A simplest systematics for the organization of turn-taking in conversation'. *Language*, 50, 4, pp. 696-735.

Schegloff, E.A. (1982). 'Discourse as an interactional achievement: some uses of "uh huh" and other things that come between sentences'. In: D. Tannen (ed.), *Georgetown University Roundtable on Languages and Linguistics* (pp. 71-93), Washington, DC: Georgetown University Press.

Fallacies and heuristics

Sally Jackson,
University of Arizona

Abstract

In this paper, an explanation is proposed for the persuasiveness of fallacies. Most discourse is dependent on a broad presumption of acceptability that limits evalutation of claims to noticed trouble-spots: evaluation is triggered by identifiable symptoms of something wrong. Furthermore, the evaluation of claims occurs at varying levels of depth, depending on the level of suspicion aroused and the amount of effort the evaluator is willing to spend on the evaluation. The main implication of this is that fallacies are not incorrect argument schemes, or correct argument schemes applied incorrectly, but products of evaluation heuristics that can be given good defense as diagnostic tools.

Fallacies have long been understood as forms of argument that gain assent for assertions without authentic justification: patterns of argument that are persuasive without being sound. From a certain point of view, a form of argument that gains assent without deserving assent is worse than even a transparently invalid argument: not only does it fail to justify its conclusion, but it also conceals its own failure to do so, leading hearers into error along with speakers.

Although many scholarly and pedagogical treatments of fallacy seek to explain what is wrong with certain patterns of argument, few give serious attention to what is persuasive about these patterns--why, given that they are defective, they often gain assent. Yet the question of why fallacies are persuasive is certainly as interesting as the question of why they are incorrect, and moreover, a practical approach to the improvement of argumentation depends not only on some idealization of how things ought to look but also on some insight into why real-life circumstances deviate from that idealization.

As an explanation for the persuasiveness of fallacies I propose the following sketch: first, that most discourse, argumentative and nonargumentative, is dependent on a broad presumption of acceptability that limits evaluation of claims to noticed trouble-spots; second, that evaluation of claims is triggered by identifiable symptoms of something wrong; third, that this evaluation occurs at varying levels of depth, depending on the level of suspicion aroused and the amount of effort the evaluator is willing to spend on the evaluation; and fourth, that many fallacies are incidental products of evaluation heuristics that can be given good defense as diagnostic tools.

Before elaborating this sketch for the case of some familiar fallacies, I will review some of what we know about the role and functioning of presumptions in ordinary discourse.

Minimalism in ordinary discourse

From work on the organization of argument in conversation (Jackson 1987, 1992; Jackson & Jacobs 1980; Jacobs 1987, 1989; Jacobs & Jackson 1983, 1989), the following more-or-less empirical observations can be advanced:

1. The performance of any speech act creates an open-ended and indeterminate disagreement space, consisting of anything reconstructible as a belief the speaker can be assumed to hold (van Eemeren et al. 1993: 95-102; Jackson 1992).
2. This disagreement space is an opportunity for argument, but most of the time argument does not occur; the normal case is for hearers to assume that whatever the speaker believes is in fact true and defensible. This assumption figures in Grice's analysis of conversational cooperativity (Grice 1989) as the Quality Maxim. Argument occurs only selectively, so it will be important to give attention to how participants decide when to make the cooperative assumption and when to challenge another speaker's apparent beliefs.
3. Argument is about repairing disagreement in a 'locally managed' way; it expands speech act sequences only as necessary to fix something noticed as amiss (Jackson & Jacobs 1980; Jacobs & Jackson 1989).
4. Enthymeme is the normal form of argument (Jackson & Jacobs 1980); even when reasoning for one particular conclusion is laid out explicitly in a 'logically complete' form, the premises will typically just be statements both parties are willing to stipulate as acceptable.

The main implication to be drawn from these four observations is that in conversational argument, minimalism is the rule and departures from minimalism are the exceptions. Anything a speaker says, implies, or implicates is potentially a standpoint, and every reconstructible standpoint is in principle arguable. This is not to say that in performing speech acts, the deep structure is a complex of full-blown arguments and the surface structure is some partial representation filled out through response to challenges or disagreement, but rather to say that full-blown arguments are interactionally emergent just in case a standpoint does provoke challenge or disagreement. The grounds for any standpoint are likely to require excavation; in putting forward a view a speaker may never have given a moment's thought to what grounds would be required should that view be challenged (Jackson 1992). Nor is this a special defect of undisciplined argument; indeed, this seems as true of, say, scholarly discourse as of ordinary conversational interaction.

Like dialectical positions generally, the view of argument advanced here depends heavily on the concept of presumption. Presumption may be understood as the "default" position on a question, when nothing in particular is known about the circumstances to which the question refers. The most important presumptions shaping the organization of conversational argument are those underwritten by the Cooperative Principle. These standing presumptions mean that an assertion advanced in conversation is assumed to be acceptable unless there is something weighing against it, such as independent reason to doubt the assertion itself, independent reason to doubt the cooperativity of the speaker, or contextual information suggesting that the assertion is regarded as arguable by the speaker.

Much of what we accept, positively or provisionally, is accepted on no grounds other than that someone else has been presumed to have adequate grounds for having accepted it. Scott Jacobs and I argued that to account for certain facts of conversational organization, one must posit a "Reason Rule": an obligation to align one's utterances with the beliefs and wants of others. According to this Reason Rule, "One party's expressed beliefs and wants are a prima facie reason for another party to come to have those beliefs and wants and, thereby, for those beliefs and wants to structure the range of appropriate utterances that party can contribute to the conversation. If a speaker expresses belief X, and the hearer neither believes nor disbelieves X, then the speaker's expressed belief in X is reason for the hearer to believe X and to make his or her contributions conform to that belief" (Jacobs & Jackson 1983: 57).

An explicit outline of this background of presumption might look like the following:

i)   Speaker S asserts or implies proposition P to be true.
ii) [By the Cooperative Principle] S may be assumed to believe that P is true and to believe that there is adequate basis for that belief.
iii) [By the Reason Rule] S's belief that P is true is reason for Hearer, H, to believe P, unless contradicted by other evidence or presumptions.
iv) P is not contradicted by other evidence or presumptions.
v)   P should be accepted presumptively by H.

In the ordinary run of things, that is, when nothing triggers an examination of P, P would be accepted presumptively, as a matter of course. Note that the absence of other contradictory evidence or presumptions is part of the ground on which P is accepted. But absence of contradictory evidence is impossible to establish systematically, and as an empirical matter, people seem to depend not on any sort of in-depth search for reasons to disagree but on a few diagnostic tools organized around their standing concerns for communicative and interpersonal values (Jacobs et al. 1991).

Reasoning of the sort outlined above is of course quite suspect if evaluated against any sort of normative model. But we do not generally notice the pervasive occurrence of fallacious reasoning underwritten by the Reason Rule. Unless a presumptively

accepted proposition has been somehow formulated as a debatable claim, we would not want to say that any sort of fallacy has occurred, even though it is evident that S's belief in P is evidence in P's favor only under certain unexamined conditions having to do with S's abilities and motives, and even though it is evident that the absence of evidence against P is no assurance that P is true. This broad and unremarked reliance on presumption, which certainly leads us into error more often than does our reliance on authority, popular opinion, or other explicitly formulated "appeals," will figure heavily in our analysis of both formal and informal fallacies.

Informal fallacies: authority dependence

Recent work in informal logic suggests that many fallacious arguments gain their persuasiveness from resemblance to argument schemes with legitimate usefulness. Appeal to authority, for example, is a fallible but widely useful form of argument. Walton (1989a, 1989b) describes it as a type of "plausible argument," a type of argument sufficient to establish a presumption in favor of a conclusion, so long as there is no better evidence to suggest that the conclusion is false. Appeal to authority is considered fallacious only when it is used to close down discussion of a claim, to answer genuine controversy with an implicit claim that better minds have already settled the issue. The explanation for the persuasiveness of fallacious arguments *ad verecundiam* would be that these arguments gain plausibility from the hearer's failure to differentiate illegitimate from legitimate uses of the pattern.

Another contemporary account of the persuasiveness of these argument forms can be drawn from empirical research on cognitive processing of persuasive messages. Many attitude theorists now espouse one version or another of the theory that people process messages in more or less depth depending on contextual factors such as the importance of the issue or personal factors such as prior knowledge about the topic. The theory as articulated by proponents of the Elaboration Likelihood Model (Petty & Cacioppo 1986) or the Systematic/Heuristic Processing Model (Eagly & Chaiken 1993) is that people generally rely on superficial cues to guide their response to messages unless specially motivated to examine and evaluate the quality of the message content. Among "persuasion cues" identified in the social psychological research literature are such things as source credibility (Petty, Cacioppo & Goldman 1981) and response of other audience members (Axsom, Yates & Chaiken 1987). Hearers *not* motivated to engage with an argument are much more swayed by source credibility and by the reactions of their fellows than are hearers motivated to engage. In other words, fallacies are persuasive because audiences use them as shortcuts to avoid careful thinking about issues, whenever the cost of careful thinking exceeds what the hearer thinks the issue is worth.

Integrating these two contemporary accounts, we may interpret research on cognitive response to persuasive discourse as showing that fallacies are not mere logical errors but interpretive strategies with defensible design features and important

communication functions. Specifically, many textbook cases of informal fallacy (notably *ad populum* and *ad verecundiam*) have been shown to function as simplifying strategies (termed heuristics) used by audiences under certain conditions to substitute for "reasoning from scratch." These heuristics are not arbitrary rules of thumb, nor are they mere habits of thought. On the contrary, each can be defended as a plausible way of approximating decisions that would be made under ideal conditions of rational discourse. For example, *ad verecundiam* (a virtual prototype of plausible argument) can be given a very good defense based only on the assumption that well-qualified sources are less likely to make mistakes in their conclusions than poorly-qualified sources.

Measured against the empirical properties of argumentation, these accounts have much to recommend them. The view of informal fallacies as overextensions or misapplications of plausible reasoning strategies contains the important insight that an argument may aim only to assign presumption to one side or the other in a potential controversy. The idea that recipients of persuasive messages often evaluate the conclusion using simplifying heuristics in place of careful analysis contains the important insight that informal fallacies may describe not the materials presented by a speaker but the interpretive and reconstructive choices of the hearer. But taken together, these accounts have a soft spot: They assume that evaluation of a conclusion waits on evaluation of whatever material is offered as support.

A different account follows from the premise that argumentation functions as repair of disagreement within a system that presumes agreement. On this premise, argument is a collaborative production in which the recognition of some sort of disagreement stimulates the search for a resolution. This association of argumentation with interactional repair is a fundamental departure not only from traditional logical approaches but also from the contemporary approaches of informal logic and attitude theory. Instead of assuming that evaluation of conclusions ordinarily waits on evaluation of materials, this view assumes that in the ordinary run of things evaluation of materials waits on (preliminary or provisional) evaluation of the conclusion, and that further, the evaluation of materials is not always aimed at arriving at a judgment about the conclusion but often premised on a fixed judgment that the conclusion is wrong. In ordinary conversational circumstances, people search out and examine the grounds for conclusions only when there is some reason for disagreeing or some reason for thinking that disagreement might be in the offing.

Consider argument *ad verecundiam*. Analyzing *ad verecundiam* as a pattern of plausible argument, Walton outlines its form as follows (Walton 1989a: 193):

> E is an expert in domain D.
> E asserts that A is known to be true.
> *A is within D.*
> ∴ A may (plausibly) be taken to be true.

In Walton's discussion, this pattern is not in itself fallacious, but is prone to intrinsic weaknesses associated with the three premises: E may not be a real expert; A may not

be an accurate rendition of what E really said or may not in fact be known to be true; or A may not belong to the domain in which E is expert. The fallacy of *ad verecundiam*, according to Walton, occurs when such an appeal is "pressed too hard in a persuasion dialogue" (1989a: 197), specifically when the appeal to authority is used to close off debate over the impersonal grounds for belief in A.

It will be helpful to note that while Walton's pattern might describe the structure of a persuasive message, this pattern of reasoning is not restricted to cases in which a speaker argues from authority but occurs as well any time an audience evaluates a speaker's conclusion taking the speaker's expertise into account: when, for example, the speaker's carefully reasoned position is accepted or rejected not on the merits of the argument but on the speaker's own credibility, or when the speaker presents a summary of expertise-based arguments all of which are ignored in favor of information on the source of the arguments. In other words, the occurrence of an *ad verecundiam* fallacy does not necessarily involve a speaker making an *appeal* to authority; appeal to authority is a speaker's formulation of a much more general class of acts involving authority dependence in one form or another.

Some argumentation theorists (e.g., Willard 1990) suggest that reliance on authority is both widespread and reasonable, especially so when directly relevant evidence is unavailable or inaccessible. But in the social psychological study of attitude change, there is substantial evidence to suggest that people apply inferential patterns of this kind even when the materials available to them include *both* evidence supporting a conclusion and information concerning the source. In Walton's outline of appeal to authority, all that is included is information on the source and association of the assertion with that source. But in many actual instances in which the *ad verecundiam* fallacy may be said to occur, the material available to the audience includes impersonal grounds for belief in the assertion.

Audiences are said to employ the "credibility heuristic" when they substitute assessment of the source of a conclusion for assessment of the grounds the source might have for that conclusion (O'Keefe 1990: 182). Notice that the credibility heuristic is a method audiences use to evaluate a conclusion, not a pattern of argument speakers use to justify a conclusion. Walton's outline of appeal to authority is quite a good description of this method, at least of its careful employment.

Hence, there is something lacking in Walton's outline, whether intended as a description of materials advanced as an argument or as a description of the underlying logical structure of the credibility heuristic. What is lacking, empirically, is the other material a speaker may have presented or the other material an audience might have taken into account. These other materials need to be represented in any discussion of appeal to authority, because these other materials have something to do with why people rely on authority--and also something to do with the difference between legitimate and illegitimate appeal to authority.

In Willard's analysis of authority-dependence, the defense of argument from authority is built from the impenetrability of expert fields and the incompetence of the

audience to evaluate the evidence available to experts; argument from authority is a purposeful delegation of responsibility for conclusions in expert domains. In the entirely independent experimental work on the credibility heuristic within social psychology, it has been found that people are most likely to rely on source judgments instead of direct evidence when the importance of the conclusion is low relative to the effort required to evaluate it.

To adequately represent authority-dependence, to explain its occurrence, and to differentiate its legitimate and illegitimate forms, we need to add to Walton's outline some representation of the role of the invisible other materials that might have been taken into account. Consider the following revised outline, in which the speaker, E, may be said to have advanced grounds G in support of assertion A:

> E asserts A based on grounds G.
> *[The adequacy of G is unknown.]*
> ∴ A should be accepted or rejected depending on E's expertise in the relevant domain.
>
> The domain relevant to A and G is D.
> *E is an expert in domain D.*
> ∴ A should be accepted.

In this revised outline, the construction of a text is not what is at issue, but the reconstruction of the text by an audience. Of course some texts do contain good examples of appeal to authority, examples that look very like Walton's outline. The point is that explicit appeals to authority are one manifestation of authority-dependence in reasoning, and if we are to understand their role in disagreement management, we need to understand them in relation to a more general willingness on the part of audiences to select information on source and even prefer that information to other, objectively better, evidence.

Why do people accept information on source in lieu of directly relevant information, and more importantly, why do they seem to rely on evaluation of the source when directly relevant information is available to them? In understanding the role of authority in argument, it is very important to see that the point of relying on authoritative reasoning is to get to a conclusion with the information at hand, and sometimes to avoid having to conduct any deeper examination of the conclusion. As an empirical matter, people rely more heavily on authority when unable or unmotivated to evaluate the grounds on which the authority's conclusions are based. And this is true whether the grounds are disclosed to them or not; even when the quality of argument offered is much better than mere appeal to authority, audiences unprepared to evaluate grounds for a claim often reduce the information available to something like the form Walton gives as an outline of argument from authority. Note that if the audience is in fact unable to make a competent evaluation of the grounds for an authority's conclusions, it is quite sensible to treat authoritative opinion as a basis for strong presumptions.

To understand both legitimate and illegitimate appeals to authority, we need to position appeal to authority between, on the one hand, assertions offered with no defense at all, and on the other hand, assertions offered with impersonal grounds for

belief. As compared with no defense of a claim at all, an appeal to authority has two interesting design features: first, it implies that the conclusion is such that the hearer is *not* expected to accept it presumptively, and second, it invokes expertise or some similar quality as backing for a limited presumption in favor of the claim. The first of these features provides for some sort of problematization of the claim. The second makes the quality of the source available as a diagnostic cue, along with whatever else might have been available in any case.

Many textbook cases of fallacious appeals to authority involve invocation of irrelevant authority or bogus authority. On a presumption based account, the persuasiveness of such appeals--the fact that they work, while appeals to some other irrelevant or bogus authorities would not--can be explained as a consequence of a use of source information as diagnostic. When a claim is argued on someone's say-so, that someone is checked, but the checking is subject to the same presumptions as anything else stated or implied in conversation. Unless something negative is known about the source, two presumptions operate: first, that the speaker appealing to the authority believes that the authority's views are relevant to the truth of P, and second, that the authority committed to P believes that there is adequate basis for P. When nothing is known against the authority, a superficial check should result in acceptance of the authority. Note a paradoxical implication, though: an appeal to a patently unreliable "authority" is likely in some circumstances to be less persuasive than a completely unsupported assumption.

Formal fallacies: the atmosphere effect

On a characterization of fallacy as "an argument that seems valid but is not," a large and homogeneous class of fallacies may be identified among categorical syllogisms. Consider the following form:

> No A are B.
> Some B are C.
> ∴ Some A are not C.

This form is invalid, of course, but it is also exceedingly likely to pass as valid, along with all of the following (likewise invalid) forms:

> Some A are B.
> Some B are C.
> ∴ Some A are C.

> No A are B.
> No B are C.
> ∴ No A are C.

Some A are not B.
All B are C.
∴ Some A are not C.

What do all of these forms have in common, and why are they more likely to mislead than the following transparently invalid forms?

Some A are B.
All B are C.
∴ All A are C.

No A are B.
No B are C.
∴ All A are C.

Over 60 years ago, experimental psychologists took up the problem of explaining why people persist in accepting certain invalid forms as valid but readily recognize invalidity in other similar forms. Woodworth and Sells (1935) hypothesized that people do not really reason carefully about such materials, but instead extract certain superficial features from the premises and use them to predict the sort of conclusion that can be drawn. The pattern of errors to be explained was termed "the atmosphere effect," and I will refer to the hypothesis offered to explain the errors as the atmosphere hypothesis.

According to the atmosphere hypothesis, people evaluating categorical syllogisms take note of the logical features of the premises, without regard for the actual relationships among the categories, then apply simple heuristic rules to the extracted features. The four types of statements involved in categorical syllogisms are completely described in terms of two features, quality (affirmative or negative) and quantity (universal or particular). Woodworth and Sells did not suppose that untrained people thought about statements in the special technical vocabulary of formal logic, but they did assume that these features were noticed spontaneously and used to compute conclusions. Woodworth and Sells suggested that people compute or evaluate conclusions on the basis of the following simple rules: (1) if the premises (as a set) are affirmative, the conclusion must also be affirmative; any negation in the premises requires negation in the conclusion; and (2) if the premises (as a set) are universal, the conclusion must also be universal; any particularity in the premises requires particularity in the conclusion. If a pair of premises has a valid conclusion about category A, that conclusion will be generated by the atmosphere rules.

The atmosphere *effect* is a tendency for people to accept conclusions that "match" the features of the premises, and the atmosphere *hypothesis* is the idea that the way this comes about is by computing an appropriate conclusion through application of these rules to the features of the premises. A variety of competing explanations for the pattern of errors have been proposed, but these are not my concern at present.

Two observations can be made at this point. First, the atmosphere hypothesis explains why people accept invalid arguments as valid, not why people reject valid arguments. We do not seem to need a theoretical explanation for the rejection of valid conclusions, for these do not seem to occur in a regular pattern; for materials like those we are considering, people accept many invalid syllogisms but rarely reject valid ones. As noted earlier, the study of fallacies is for all practical purposes the study of underjustified assent.

Second, and more importantly for our stance toward fallacies, the atmosphere effect itself can be washed out by creating concrete substitution instances of the invalid forms, *if* the substitution instance involves an easily evaluated falsehood. Compare the three syllogisms below. The first is extremely likely to be accepted as valid, by which I mean if you present it to a classroom full of people, a lot of them will judge it as valid; the second, though formally identical, is extremely unlikely to be accepted as valid; the third, despite the presence of familiar, concrete content, is as likely as the first to be accepted as valid.

> All A's are B's.
> Some B's are not C's.
> ∴ Some A's are not C's.
>
> All flowers are living things.
> Some living things are not plants.
> ∴ Some flowers are not plants.
>
> All vegetables are healthful.
> Some healthful things are not tasty.
> ∴ Some vegetables are not tasty.

The crucial difference between the two concrete examples of this form is that the conclusion "Some flowers are not plants" is obviously false, while the conclusion "Some vegetables are not tasty" is almost certain to be taken as true. It is important to know that people are not just assuming that arguments with false conclusions are invalid. People can recognize arguments with false conclusions as valid and correctly diagnose the trouble as reasoning from false belief, at least in cases like the following:

> All women are mothers.
> No mothers are athletes.
> ∴ No women are athletes.

Why does the atmosphere effect appear? The original atmosphere hypothesis suggests that it is because people process premises superficially and apply correspondingly superficial rules to the task of inference or evaluation. But this hypothesis is conceptually and empirically unsatisfactory. Conceptually, there is no independent evidence, other than the atmosphere effect itself, for thinking that people use premise

features in any active way to arrive at or evaluate a conclusion. Empirically, the atmosphere hypothesis offers no explanation for why the atmosphere effect *fails* to appear in certain circumstances: any sign of trouble, be it manifestly false content or criticism by other evaluators, will disarm this interesting class of fallacies (see, e.g., Jacobs, Allen, Jackson & Petrel 1985).

With a minor repair, however, this hypothesis dovetails nicely with the notion that people use a variety of heuristics to simplify the task of evaluating what they hear. It happens that every conclusion that fails to match the features of the premises will turn out, on inspection, to be invalid; in other words, a type mismatch between conclusion and premise set is sufficient, but not necessary, for invalidity. The same is true of falsity; when a conclusion is false, it is sure that either the premises are false or that the argument is invalid, though not every unsound argument has a false conclusion. The point of heuristics is to make evaluation easier; if an evaluator with no pronounced reason to suspect anything amiss checks for obvious symptoms and finds nothing, then the evaluator will be led into error for any argument whose particular defects do not appear as visible symptoms of the sort the evaluator habitually notices.

What I am suggesting is that the atmosphere effect is a byproduct of a sort of triage, in which any claim encountered in discourse is judged loosely against some set of indicators of trouble, but evaluated carefully only when one of these indicators comes up positive. The default decision is to accept a claim, and in fact, it would be better to say that the default decision is to presume the truth of the claim, explicitly deciding to accept or reject it only when something triggers an in-depth evaluation of the claim and its grounds.

Implications

The main implication of what I have said is that fallacies are not incorrect argument schemes or correct argument schemes applied incorrectly, but failed diagnostic strategies. The search for some rule that will differentiate reasonable appeals to authority from illegitimate appeals to authority will always come up empty, because the problem is that once one accepts the diagnostic strategy as a general method for screening claims, one must accept asymptomatic cases as the cost of doing business in this way.

Many standard fallacies can be re-thought in terms of presumptions applied without awareness that the case at hand is an exception. That social interaction depends on a broad presumption of acceptability does not mean that it is reasonable to cling to these presumptions regardless of circumstances. Presumptions are not, in themselves, pernicious. Presumptions without reliable methods for recognizing exceptions *are* pernicious. Many patterns of reasoning that appear to arise from inability to recognize flaws in argument may be better understood as side consequences of a broad and unremarked presumption of acceptability controlled by attentiveness to certain kinds of diagnostic cues.

Within the view of fallacy proposed here, the signs people use to recognize exceptions become very important; critical thinking on such a view requires not memorization of suspect patterns but development of a sense of when to be suspicious. Although it is almost certainly the case that people can be made more critical through instruction and practice, it will be useful to start by inventorying the sorts of things people use spontaneously to diagnose trouble.

Most obvious is the recognition that a proposition reconstructed from discourse is directly contradictory or otherwise inconsistent with a previously held belief. People are notoriously tolerant of fallacies in the case for their own side of a controversy; this is not so much a matter of applying looser standards to arguments they favor as a matter of applying no standards at all when not doing evaluation. If argument functions as repair of misalignment in belief, there is no purpose in evaluating argument once the conclusion has been accepted, except in those special discourses structured by a contrived skepticism (such as debate and academic argument).

But a search for problems can be triggered by many other circumstances. One such circumstance is overjustification. When a conversationalist gives explicit defense of what would otherwise appear uncontroversial, other conversationalists search for explanations, as for any other violation of Gricean maxims: one possible implicature drawn from such occurrences is that the speaker expects the position to be controversial, and when hearers draw this implicature, they have reason to search for what might be the problem. Conversation prefers under-elaboration of the grounds for belief rather than over-elaboration (Jackson & Jacobs 1980); the enthymeme is a rational strategy for controlling disagreement space.

Yet another such circumstance is awareness of controversy. A listener who knows an issue to be controversial, or who is warned that a message soon-to-be-presented concerns a controversial topic, will be more attentive to possible troubles than a listener not forewarned of controversy. Mere forewarning makes people more resistant to persuasion (O'Keefe 1990: 182); this interesting psychological fact is easily explained in terms of disturbance of the presumption the upcoming message would have enjoyed without the forewarning. Forewarning ought not lead to rejection of conclusive argument, and so far as I know, there is no empirical evidence to suggest that it does; a presumption-based account predicts that if there is any weakness in an argument, it will be more likely to be noticed when the audience is put on alert and more likely to be overlooked when the audience is allowed to respond on the basis of the general cooperative presumption.

A search for problems may lead immediately to evaluation in depth of the grounds for a conclusion, but it may lead instead to a quick survey of diagnostic cues. Alerted to the possibility of controversy, a quick check that all is in order might include assessment of all sorts of easily noted features such as the general trustworthiness of the source, the availability of evidence to support key contentions, the apparent orderliness of the argument, the apparent response of other audience members, and so on.

Even alerted to the chance of disagreement, a listener may orient to superficial diagnostic cues to aid in the decision about whether to go any further in evaluation of the argument. Moreover, these diagnostic cues can be used in diverse ways: to predict soundness, but also to project the difficulty of a direct, in-depth evaluation of argument quality. When an argument depends on technical material or very complicated reasoning, for example, an audience may take the difficulty of the material both as an indication that they will be unable to assess the evidence directly and as indirect evidence that the source is a knowledgeable person.

One general implication to be taken from the association of fallacies with heuristics is that people *choose* to reason fallaciously. This sounds ridiculous, but in effect, when a person chooses a general strategy with some acknowledged risk of error, the person is buying something at a cost: efficiency, for example, at the cost of occasional blunders. When an issue is important enough to justify effort, heuristics decline in importance; when the issue is important and the argument disagreeable, fallacies become extraordinarily easy to spot.

References

Axsom, D., S. Yates & S. Chaiken (1987). 'Audience response as a heuristic cue in persuasion'. *Journal of Personality and Social Psychology* 53, pp. 30-40.

Eagly, A.H. & S. Chaiken (1993). *The Psychology of Attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich.

Grice, P. (1989). *Studies in the Way of Words*. Cambridge: Harvard U. Press.

Eemeren, F.H. van, R. Grootendorst, S. Jackson & S. Jacobs (1993). *Reconstructing Argumentative Discourse*. Tuscaloosa, AL: U. of Alabama Press.

Jackson, S. (1987). 'Rational and pragmatic aspects of argumentation'. In: F.H. van Eemeren, R. Grootendorst, J.A. Blair & C.A. Willard (eds.), *Argumentation: Across the Lines of Discipline* (pp. 217-227), Dordrecht-Holland: Foris.

Jackson, S. (1992). '"Virtual standpoints" and the pragmatics of conversational argument'. In: F.H. van Eemeren, R. Grootendorst, J.A. Blair & C.A. Willard (eds.), *Argumentation Illuminated* (pp. 260-269), Amsterdam: SICSAT.

Jackson, S. & S. Jacobs (1980). 'Structure of conversational argument: Pragmatic bases for the enthymeme'. *Quarterly Journal of Speech* 66, pp. 251-265.

Jacobs, S. (1987). 'The management of disagreement in conversation'. In: F.H. van Eemeren, R. Grootendorst, J.A. Blair & C.A. Willard (eds.), *Argumentation: Across the Lines of Discipline* (pp. 229-239), Dordrecht-Holland: Foris.

Jacobs, S. (1989). 'Speech acts and arguments'. *Argumentation* 3, pp. 345-365.

Jacobs, S. & S. Jackson (1983). 'Speech act structure in conversation: Rational aspects of pragmatic coherence'. In: R.T. Craig & K. Tracy (eds.), *Conversational Coherence: Form, Structure, and Strategy* (pp. 47-66), Newbury Park, CA: Sage.

Jacobs, S. & S. Jackson (1989). 'Building a model of conversational argument'. In: B. Dervin, L. Grossberg, B.J. O'Keefe & E. Wartella (eds.), *Rethinking Communication: Paradigm Exemplars* (pp. 153-171), Beverly Hills/Newbury Park, CA: Sage.

Jacobs, S., M. Allen, S. Jackson & D. Petrel (1985). 'Can ordinary arguers recognize a valid conclusion if it walks up and bites them on the butt?' In: J.R. Cox, M.O. Sillars & G.B. Walker (eds.),

Argument and Social Practice: Proceedings of the Fourth SCA/AFA Conference on Argumentation (pp. 665-674), Annandale, VA: Speech Communication Association.

O'Keefe, D.J. (1990). *Persuasion: Theory and Research.* Thousand Oaks, CA: Sage.

Petty, R.E. & J.T. Cacioppo (1986a). *Communication and Persuasion: Central and Peripheral Routes to Attitude Change.* New York: Springer-Verlag.

Petty, R.E., J.T. Cacioppo & R. Goldman (1981). 'Personal involvement as a determinant of argument-based persuasion'. *Journal of Personality and Social Psychology* 41, pp. 847-855.

Walton, D.N. (1989a). *Informal Logic: A Handbook for Critical Argumentation.* Cambridge: Cambridge U. Press.

Walton, D.N. (1989b). 'Reasoned use of expertise in argumentation'. *Argumentation* 3, pp. 59-73.

Willard, C.A. (1990). 'Authority'. *Informal Logic* 12, pp. 11-22.

Woodworth, R.S. & S.B. Sells (1935). 'An atmosphere effect in formal syllogistic reasoning'. *Journal of Experimental Psychology* 18, pp. 451-460.

The straw man fallacy

Douglas Walton,
University of Winnipeg

Abstract

In this paper, an analysis is given of the straw man fallacy as a misrepresentation of someone's commitments in order to refute that person's argument. With this analysis a distinction can be made between straw man and other closely related fallacies such as *ad hominem, secundum quid* and *ad verecundiam*. When alleged cases of the straw man fallacy are evaluated, the speaker's commitment should be conceived normatively in relation to the type of conversation the speaker was supposed to be engaged in.

The straw man fallacy appears to be a modern addition to the list of traditional informal fallacies covered in the logic textbooks. No mention of this fallacy as a distinct type of fallacy in the standard treatment, or as a historical item, is made by Hamblin (1970). The first inclusion of it we can find in a textbook as an informal fallacy is in Chase (1956: 40).

Aristotle did not include the straw man fallacy in his list of sophistical refutations, although he does indicate, in several passages, an awareness of something very close to it. Evans (1977: 81) mentions that in Aristotelian dialectical refutation, where the dialectician refutes another party's views by deducing *adoxa* (implausible propositions, generally held to be false) from them: "Aristotle requires of the serious dialectician ... fidelity [according with the real or expressed views of the other party] in representing the views of others ..." Aristotle indicates in several places (*Topics* 105 b 6; *On Sophistical Refutations* 174 b 21) how this principle of fidelity for genuine refutation could be exploited in sophistical refutation, by only giving the appearance of the real view of the other party as the basis for your refutation.[1] This comes fairly close to a recognition of what would nowadays be called the straw man fallacy.

---

[1] In *Topics* (105 b 6), Aristotle writes of a useful method for forming propositions to refute an opponent: "choosing not only opinions actually received but also opinions which resemble these ..." In *On Sophistical Refutations* (174 b 21), Aristotle writes of the tactic of looking for contradictions between "the answerer's views and either his own statements or the views of those whose words and actions he admits to be right ..." This tactic sounds more like what we would call a form of the circumstantial *ad hominem* attack (see section 5, below). But it also has elements of awareness of the straw man tactic, as well. Further (174 b 34), Aristotle suggests, "One should also sometimes attack points other than the one mentioned, excluding it if one can make no attack on the position laid down ..." This tactic might nowadays be classified under *ignoratio elenchi* (wrong conclusion), or it could also be a reference to the straw man fallacy.

DeMorgan (1847: 281) also indicated an awareness of the kinds of faulty inferences associated with misrepresenting another party's views in argumentation. But he, like Aristotle, did not use the term 'straw man fallacy,' or some comparable expression, to classify a single category of error of this type.

Hence the historical question of how straw man first entered the logic curriculum, as a distinctive fallacy, remains open. But as shown in this paper below, it is now in (at least a few) leading textbooks, and is definitely a very important fallacy in its own right, in the logic curriculum.

In this paper, the goal is to give a practically useful analysis of the straw man fallacy that can be applied to real cases in everyday argumentation, and a theoretically clear and exact enough analysis that is adequate to distinguishing between straw man and several closely related neighboring fallacies.

1. Initial account of the fallacy

Johnson and Blair (1983: 71) define the straw man fallacy as committed "... when you misrepresent your opponent's position, attribute to that person a point of view with a set-up implausibility that you can easily demolish, then proceed to argue against the set-up version as though it were your opponent's." They cite the following three conditions, for a pair of arguers M and N, and a pair of positions, Q and R: (1) M attributes to N the view or position, Q; (2) N's position is not Q, but a different one, R; and (3) M criticizes Q as though it were the view or position actually held by N. According to their analysis, the straw man fallacy can be defined, in general, by the meeting of these three characteristic conditions (1983: 74). The framework here, as Johnson and Blair put it (1983: 70), is one of an *adversary context* where two participants in dialogue, M and N, are arguing with each other. That is, one is attacking the other (has the aim of refuting or criticizing the other), and each is trying to defend his or her own position from the attacks of the other.

This is a very clear account of the logical structure of the straw man fallacy. But how does one define the variable Q, representing the arguer's position? The way advocated in this paper is to define it as the total commitment set of a participant in a dialogue. This way of defining an arguer's position utilizes the device of a *commitment set* (Hamblin 1970: 264), a set of propositions listed, e.g. on a sheet of paper, or in a computer data base, representing what an arguer in a dialogue has committed herself to, as a result of moves (like asking questions, or making assertions) she has made during the course of that dialogue.

But even if we can define 'position' normatively and abstractly, in general, it is another question to determine what it amounts to in a specific case. According to Govier (1992: 157), the straw man fallacy is committed "when a person misrepresents an argument, theory, or claim, and then, on the basis of that misrepresentation, claims to have refuted the position that he has misinterpreted." Govier brings out some of the main practical difficulties in dealing with the problem posed by the straw man fallacy

in real-life argumentation. The problem is to know, or be able to prove, that an arguer's position has been misrepresented in a given case. How should this be done? It depends on interpreting what someone means to say, on determining what their real position is on an issue. But this can be a hard determination to make, in some cases. So the second thing we need is a way of determining what an arguer's position is, or at least can fairly be interpreted to be, in a particular case where the straw man fallacy has been alleged, or is a danger.

Let us begin with a fairly standard type of case of the kind commonly found in textbook examples, as a first step towards grasping the nature of the problem involved in analyzing the straw man fallacy.

The following case is a brief instance that can be used to illustrate the gist of what is involved in this fallacy.

Case 1: Bob and Arlene are arguing about environmental laws that regulate industrial pollution, and Bob has taken a moderate "green" position. Arlene argues, "People like you want to make the planet into the pristine place it was hundreds of years ago. You preservationists don't want to let anybody do anything to the land that could possibly have ecological consequences. Therefore, what you are committed to is the elimination of all private property and all industrial manufacturing. Imagine the unemployment and social destruction of private homes implied by this."

Arlene attributes to Bob the so-called preservationist position, which is generally taken to represent an extreme version of the green position, allowing for very little to be done on preserved lands. But did Bob in fact advocate any of the viewpoints characteristic of this extreme position? There is no evidence given in the case that this is so, and in fact we are told that Bob's ecological position is a moderate green stance.

In evaluating this case, everything depends on what Bob said before in the argument, and what this discourse may rightly be taken to imply about his commitments on the subject. Let us say that in fact Bob's position was nowhere near the extreme recreation of it portrayed by Arlene's rebuttal. Here Arlene can be said to have committed the straw man fallacy by exaggerating Bob's position to make it appear much more radical than (let's presume) it really was, as Bob presented it.

Of course, to provide a more realistic case study of this fallacy, we would have to provide details of the example that recounted enough of Bob's actual wording of his earlier argumentation to provide enough evidence for us to reconstruct his stated and implied commitments. Then we would have to compare this reconstructed position with Arlene's simulated version of it. And then we would have to arrive at an evaluation of how far the one position is from the other. Textual evidence would have to support all the claims.

But this is enough for our initial account of the straw man fallacy. It has now been defined clearly enough, as an abstract logical structure, and illustrated in a graphic enough way by presenting a typical example, so that we can identify it as a specific fallacy. We now go on to study a number of borderline and more problematic cases that require further clarifications of the fallacy as a distinctive type of argumentation.

2. General positions

The first thing to be emphasized is that the given text of discourse, the exact words of a speaker (quoted in proper context), should be the ultimate evidence and guideline used to determine the arguer's position.

But what happens if we do not have a record of what the arguer actually said in the past discourse, e.g. a transcript, tape recording, etc.? Here the problem of determining a position is more acute, as Govier (1992: 157) notes:

> The straw man fallacy is more difficult to detect when the views being criticized are not quoted explicitly. This happens when the positions discussed are general ones, not identified with the stated ideas of any single specific person, such as the environmentalist position on DNA research, feminism, evolutionary theory, the capitalist position on free markets, the belief in free will, and so on. In these contexts, you have to depend on your own background knowledge to determine the real context of the position.

These cases are more difficult, because a representation of an arguer's position may have to be extrapolated by presumption, on a basis of what is generally known or expected about how this position is standardly advocated by others who share roughly the same viewpoint (Walton 1992). Such interpretations, however, if not based on the arguer's exact words, as recorded, may be highly presumptive and conditional in nature, e.g. "Since she said A in context C, we may presume (by assumption), that she is also committed to B." By their nature, however, such inferences are tentative and subject to default (should the speaker be around to rebut them).

Chase (1956) defines the "straw man" tactic as the following kind of argument: "You take a few stray characteristics, build a dummy around them, and then briskly demolish it." (1956: 40). Chase classifies the straw man fallacy as a species of over-generalization. And we can easily see why there is a justification for seeing it this way. The straw man tactic is essentially to take some small part of an arguer's position, and then treat it as if that represented his larger position, even though it is not really representative of that larger position. It is a form of generalizing from one aspect to a larger, broader position, but not in a representative way.

In some cases, we are talking about the position a person has presumably taken on in virtue of belonging to some group. This is more complex, because although you belong to a group, like the Conservative Party, it does not follow that your views will be conservative in every respect. One problem is that you might have different subgroups, more radical and more moderate positions in the same general group identified as a position.

DeMorgan (1847) recognized this complication, and drew attention to the more subtle type of straw man fallacy where you have two different subgroups who take different subpositions within the same general position. For example, in a political debate, among those who take a broadly liberal position, you may have a group that represents the unions (a workers' group), and those who are more middle-of-the-road, and see their interests as more allied to business. The fallacy DeMorgan points out is

a kind of straw man fallacy that draws a conclusion from one premise from each group (1847: 281).

> Again, as to subjects in which men go in parties, it is not very uncommon to take one premise from some individuals of a party, another from others, and to fix the logical conclusion of the two upon the whole party: when perhaps the conclusion is denied by all, some of whom deny the first premise by affirming the second, while the rest deny the second by affirming the first.

This is a subtle form of straw man fallacy that involves the notion of a subposition within a broader, or more inclusive position on an issue.

Another problem is that key words used to characterize a position are often used in such a way that they can only be defined in relation to an arguer's point of view who has already adopted a positive or negative group position. For example, if a church group describes an opponent's position as "heretical," all this really means is that the opponent's view is against the position of the church who characterized it as "heretical" in the first place.

3. Ad hominem

This brings us to a consideration of the relationship between the straw man fallacy and the *ad hominem* fallacy. Terms broadly used to define group positions that contain political and ethical implications, like 'communist' and so forth, are commonly used in *ad hominem* attacks. These terms are used (rightly or wrongly) to sum up an arguer's position, and the *ad hominem* argument then draws negative implications out of the attributed position.

The account of the *ad hominem* argument given in Locke's *Essay Concerning Human Understanding* (1690), quoted in Hamblin (1970: 160), makes it clear how this type of argument against *ex concessis* an opponent is essentially based on the opponent's position, and what is inferred from it. Locke describes the *argumentum ad hominem* as a move "to press a man with consequences drawn from his own principles or concessions." This broad view of the *ad hominem* as an argument from an opponent's position has been extensively analyzed by Johnstone (1959).

Straw man is particularly closely related to the circumstantial type of *ad hominem* argument in many cases. Consider the following example from Walton (1989: 154).

Case 2: George: The notorious problems we have been having with postal strikes means that there is no longer reliable mail service provided by the government. I think we ought to allow private, for-profit mail-delivery companies to compete on an equal footing with the Post Office.

Bob: But George, you are a communist.

This case was used in Walton (1989) as an inconclusive but basically reasonable circumstantial *ad hominem*, on the assumption that George is an avowed communist who, in the past, has been known to base his argumentation on many standard com-

munist principles and positions. If so, Bob has a good point - how can George consistently argue for a for-profit mail system run by private enterprise, if, in the past, it is just this sort of arrangement he has consistently and vehemently argued against?

On the other hand, if one were to adopt different assumptions in describing this case, or filling it out further, it is not hard to see how it could be an instance of the straw man fallacy. Suppose, for example, that George was not really a communist at all. Or suppose that George had advocated some pro-communist views in the past, but had also taken a very mild form of communist position that left a good deal of room for private enterprise in some sectors of the economy.

This close affinity between the straw man and circumstantial *ad hominem* fallacies may be one reason why van Eemeren and Grootendorst (1987: 286) categorize as a species of straw man fallacy a type of case they describe as "referring to views of the group to which the opponent belongs," as illustrated by their example (1987: 286).

Case 3:   That may be what he says now, but as a communist he naturally does not mean a word
           of it.

In the standard textbook treatment of fallacies, this case would normally be treated as a circumstantial type of *ad hominem* argument, perhaps even of the "poisoning the well" variety. The proponent is engaging in a personal attack on the other party's sincerity in engaging in collaborative dialogue by arguing that, since he is a communist, you can't really trust him to speak the truth, for he will always just revert to the communist ideology and propaganda as his method of argument.

Since this is so clearly a classical case of the *ad hominem*, why would van Eemeren and Grootendorst classify it as a case of the straw man fallacy? What is revealed here is the connection between the two fallacies. The case cited, depending on further details of the dialogue, could quite naturally involve elements of both.

The circumstantial *ad hominem* argument basically works by one party utilizing the other party's position in a dialogue to cite some conflict between that position and what the other party presently advocates, by what she says or how she acts now, for example (Walton 1985). Thus this type of *ad hominem* attack is essentially based on some representation of the other party's position. And hence, you can easily see that it could, in many cases, also involve the straw man argument. The straw man would be part of the means of carrying out the *ad hominem* attack.

4. Dialogue at cross purposes

Vernon and Nissen (1968) define the straw man fallacy as being committed "when a position being attacked is first stated in a distorted and hence more vulnerable form." (1968: 160). Their analysis of the fallacy is particularly interesting, because they give a good explanation of what is basically wrong with straw man arguments, in the sense of their being obstructive or counterproductive in argumentation (1968: 160):

This kind of reasoning is not only fallacious and unfair, but may also be very unwise from a purely pragmatic point of view. The latter can be the case where political ideologies, for example, are concerned. If you base your opinion of an opposing ideology on an oversimplified and distorted version of that ideology which can easily be made to look ridiculous, then you are making the serious mistake of underestimating your opponent. Any ideology or program with a large following over a period of years must have *some* merit in order to attract and hold such a following. One cannot hope to argue effectively against such a doctrine unless he understands it well enough to be able to state it in its strongest form, for the real issues will be found only at this level.

The failure to engage with the real position of your opponent in a type of dialogue like a political debate, in a way, defeats the whole purpose of your argument. It is what Aristotle would classify as a failure of real refutation. From this perspective, the outcome is that your opponent's (real) position has not been challenged at all by your argument. It is a kind of failure of an argument to succeed in its real purpose of refuting or critically questioning the opposed point of view.

To see the importance of the straw man fallacy, it is necessary to appreciate that in many instances of argumentation used in everyday conversation, a proponent's premises used in her argument are based on the commitments of the respondent. To be successful and useful in the conversation, these propositions must really represent the position of the respondent.

For example, in a persuasion dialogue (Walton 1989: 5), one of the two main kinds of argument used is the *internal proof*, meaning proof by a proponent of a claim, constructed by inferring that claim from the other participant's concessions in a dialogue. Internal proofs can take a positive form, where the proponent has the aim of proving a claim to the respondent, based on premises that are commitments of the respondent. Or they can take a negative form, where the proponent has the aim of refuting or criticizing the respondent's position by drawing a conclusion from it that is unacceptable or questionable.

For these reasons then, one can see how a straw man argument is obstructive to, and tends to defeat the whole purpose of a persuasion dialogue. The critical discussion is a type of persuasion dialogue where the purpose is to resolve a conflict of opinions by means of reasonable argumentation. However, if arguments used by the one party do not represent the real position of the other, this will interfere with the resolution of the conflict of opinions, or even make it appear that it has been achieved when really it has not.

5. Secundum quid

Traditionally, according to (Hamblin 1970: 28), *secundum quid* (meaning "in a certain respect; *para to pe*, in Greek) is the fallacy of neglecting qualifications that should properly be attached to a generalization. It is the fallacy of taking a proposition that has a qualified meaning, and using it as though it were an absolute principle or generalization. In the analysis given in Walton (1992: 75-80) the *secundum quid* fallacy is shown to be a confusion between, or a trading on the confusion between, two

different types of generalizations-the absolute (universal, exceptionless) generalization, and the qualified (defeasible) generalization of a kind that is inherently open to exceptions.

It is easy to see how this fallacy relates to the straw man fallacy. The latter often works by exaggerating and absolutizing an opponent's position in argumentation, making the opponent appear to be a kind of perfectionist who takes an absolutistic view. Such a portrayal makes the opponent's position much easier to criticize or refute. Indeed, van Eemeren and Grootendorst (1987: 286) explicitly define one type of straw man fallacy as a species of distorting someone's point of view by absolutizing it, or omitting qualifications from it.

Distorting someone's standpoint
- oversimplification by omission of his nuances or qualifications
- exaggeration   by absolutizations or generations
                     of his statements

Certainly what this indicates is that there is a very close connection between the straw man and *secundum quid* fallacies. It indicates that, in many cases, evaluating an instance of a straw man argument depends very much on a judgment of just how absolute or qualified an arguer was, when laying out his position in the prior sequence of argumentation.

As our analysis above has already indicated, the straw man fallacy brings to the fore the applied nature of informal logic. Whether the fallacy has been committed in a given case depends on how a text of discourse in that case is interpreted, in the conversational context it was supposed to be a contribution to. This is very much a contextual question of how an argument was used in a given case.

The straw man fallacy is made even more tricky to pin down in many cases by another factor. In these cases, an arguer's unstated presumptions or nonexplicit premises or conclusions may be the only indications we have of one or more of his commitments. This brings us to the question of enthymemes, or unstated premises. When attributing enthymemes, especially to an opponent, it can be very tempting to exaggerate the opponent's position by filling in a missing premise of the form 'Generally things that have property F also have property G, subject to exceptions' with an absolute, or strict generalization, of the form 'All things that have property F also have property G, without exception.' This kind of move is a form of the *secundum quid* fallacy, meaning that qualifications have been ignored. But the same move may also be a case of the straw man fallacy, the tactic of misrepresenting an opponent's position by making it seem stronger, or stricter than it really is, in order to more easily refute it.

The same kind of tactic is involved when an argument is wrongly taken to be a different type of argument than it was meant to be, the way the speaker put it forward.

Again, the tactic is to take the argument as being of a stronger kind than the speaker really meant.

For example, suppose a proponent puts forward an argument based on an appeal to an analogy, and is correctly interpreted as claiming that two situations tend to be similar in certain respects. Suppose the analogy is imperfect, and subject to default, but nevertheless qua argument from analogy, it is a fairly reasonable argument, and not without merit. Seen as an argument based on an analogy then, this argument is rightly interpreted as inherently presumptive and defeasible, open to exceptions and qualifications. But what if a critic portrays the speaker's way of putting forward the argument, unjustifiably, as one that was meant to be deductively valid. If we accept the assumption that the missing premise in question has to make her argument deductively valid, then we will not find that missing premise in the given text of discourse, and that could seem like a decisive criticism. In the case of an argument from analogy, we would take the argument as claiming that the two situations in question must be exactly equal, in every respect, for the argument to be any good. But this attribution is based on a misinterpretation, and commits a variant of the straw man fallacy by taking the argument in a much stricter way than a charitable interpretation of how it was used in the discourse would support.

6. Limits of the straw man fallacy

We have seen that the straw man fallacy is closely related to several other important fallacies. But it is a distinctive type of fallacy in its own right, and can be distinguished from these other fallacies.

It is different from *ignoratio elenchi* because in this fallacy, it is specifically the thesis of the other (and not her whole position, or set of commitments as a whole) that is misrepresented or gotten wrong. It is different from the circumstantial *ad hominem* because this type of argument cites a conflict between an arguer's position and his specific argument of the moment, and uses this supposed conflict to attack the arguer. This tactic can often involve a straw man fallacy, but that is only part of the argument, and is not essential to the circumstantial *ad hominem* as a fallacy. See Walton (1985) for an extensive analysis of the circumstantial *ad hominem* as a distinctive type of fallacy in its own right.

Straw man is also closely related to the *secundum quid* fallacy, because in the straw man argument, as van Eemeren and Grootendorst pointed out, the attacked party's point of view is often absolutized, making it appear more extreme and simplistic than it really is. But clearly this *secundum quid* element is only one aspect of the straw man fallacy. An arguer's position in the straw man fallacy can also be misrepresented and distorted in other ways.

Wrenching from context is another one of those means used to distort an arguer's position in the straw man fallacy. In this kind of case, the superfallacy is straw man

and the subfallacy (the means of carrying out the other tactic) is the fallacy of wrenching from context.

In other cases, however, we can have instances of the fallacy of wrenching from context that do not involve the straw man fallacy. For example, if I cite the quoted opinion of some third party whose opinion is used to support some part of my point of view, but wrench it out of context in a misleading way, then I have committed the fallacy of wrenching from context. But it is not a case of the straw man fallacy, unless I have used that quoted opinion to attack or criticize your (my opponent's) position in the argument.

Of course, generally, any argument I put forward will be opposed to one of yours (if we have a conflict of opinions as the basis of the dispute). But unless the wrenching or misquotation is directly used as a misrepresentation of *your* position, the fallacy should not be classified as an instance of straw man.

Another qualification should be noted, as well. The straw man fallacy is not simply the misrepresentation of someone's position, but the use of that misrepresentation to refute or criticize that person's argument in a context of disputation. The same qualification should be made for the fallacy of wrenching from context, which should only be judged a fallacy when done to misrepresent their view as part of an argument.

Care is needed here, because there is a tendency on the part of students to identify any cases of misquotation, misrepresentation of a position, or wrenching from context, as instances of the straw man fallacy (or some related fallacy), without carefully examining the case to see how the misrepresentation has been used. To correct this tendency, it is worthwhile to remember the three-part analysis of the straw man fallacy given by Johnson and Blair in section 1, above.

Straw man is also related to another fallacy, the *argumentum ad verecundiam*. When an appeal is made to the claimed opinion of an expert as an authority to back up an argument, there is a danger that the authority may be misquoted or misinterpreted. An example of a failure to meet this requirement for a reasonable appeal to authority is given by Salmon (1963: 64).

Case 4: The authority of Einstein is sometimes summoned to support the theory that there is no such thing as right or wrong except insofar as it is relative to a particular culture. It is claimed that Einstein proved everything is relative. As a matter of fact, Einstein expounded an important physical theory of relativity, but his theory says nothing whatever about cultures or moral standards. This use of Einstein as an authority is a clear case of misinterpretation of the statements of an authority.

This type of failure is pretty close to the straw man fallacy, and could perhaps, even be thought to be a species of it. The difference is that in the straw man argument, a proponent distorts or misrepresents the position of the respondent (the opponent in the dispute). But in the variant that relates to the *ad verecundiam*, as exemplified in case 4 above, the proponent misrepresents the position of the authority whose alleged opinion is being used to back up the proponent's own argument.

In both cases, the basic underlying fault is the misrepresentation of somebody's position in relation to an argument between opposed points of view. Both are cases of misquotation or misrepresentation of the position of another participant in argumentation. But there the similarity ends. In the straw man argument, by definition, the misrepresented position of an arguer is used to attack, to criticize or refute the point of view of that arguer. In these other cases, the misquotation or misrepresentation is used for different purposes in argumentation.

Hence it is important to recognize that misquotation and other forms of misrepresenting someone's position in an argument are not always fallacies of the straw man type. Although there are similarities in the method of argument used, these failures of argumentation should be classified under the headings of fallacies or errors other than that of straw man.

7. Analysis of the fallacy

The fallaciousness of the straw man argument needs to be seen as a pragmatic failure-the problem is that such an argument goes at cross purposes to the goal of a conversational exchange. Because the deception or error may not be seen, the destruction of the argumentation in a dialogue can deeply effect a conversational exchange. But the fallacy also has a logical structure as a characteristic sequence of reasoning from premises to a conclusion.

There are three parts to the straw man fallacy. First, the structure of reasoning in the fallacy is displayed in the three-part account of the straw man fallacy given by Johnson and Blair (1983: Section 1). Second, the explanation of why the straw man type of argument interferes with the basic goal of a critical discussion, and is therefore normatively at cross purposes with this type of dialogue, is that the resolution of a critical discussion requires the use of argumentation by one party that is based on premises that represent the real position of the other party.

The precise reason why the straw man is normatively counterproductive in a critical discussion is that for the critical discussion to succeed in resolving a conflict of opinions by reasonable argumentation, it is necessary that each party argues against the other party's side by using premises that represent the commitments (position) of that other party. Otherwise the dialogue is at "cross purposes." This requirement applies to other types of dialogue, like negotiations, as well as to the critical discussion.

But there is also a third aspect needed to explain why the straw man argument is a distinctive species of fallacy in its own right. Because of the various kinds of problems and trickiness in determining what an arguer's position really is in a given case, it can be easy to get this wrong, and to mistake an arguer's real position for something else that is not her real position, but only appears to be. This is the essence of the deception or error inherent in the straw man fallacy as a distinctive type of sophistical tactic.

The key here is the realization that attacking an opponent in argumentation, by drawing negative implications from her position on the issue of the dispute, is inherently reasonable as a type of argument.[2] But it can go wrong in a number of ways, resulting in a sophistical refutation or fallacy. One of these ways is to get the premise wrong, by distorting or misrepresenting that arguer's position, even though the negative conclusion drawn may be by a valid inference. It is this deceptive shift that is the essence of the straw man fallacy.

The straw man fallacy is committed where the proponent in a critical discussion misrepresents the position of the respondent with a simulated position, in order to appear to refute the respondent by carrying out a refutation of the simulated position. This tactic typically works by attributing to the respondent a simulated position that is implausible and easy to refute, and then, the simulated position is shown to have some absurd or unacceptable consequence that is a sufficient basis for repudiating it. The pretense or deception is to argue against the simulated position as though it really were the respondent's position that he has maintained or supported judging by his discourse in the previous sequence of dialogue. What is suggested then is that the arguer's real position implies the absurd consequence. Thus it appears that the real position has been refuted by *modus tollens*, the consequence being false. Thus the fallacy involves a misrepresentation of an arguer's real position or point of view, and the use of that misrepresentation to give the false appearance that the arguer has been refuted by valid reasoning.

If the respondent is actually present when the charge of fallacy is to be evaluated, then the case is quite different from the situation where he is not available for comment. If he is present, then he is in a privileged position to pronounce on what his present position is on the issue. However, even if he is present, he is still bound by what he said before, when we determine what his commitments were, as expressed at that point in the dialogue.

In a case where the respondent is present, it may not be too difficult for him to reply to the charge of fallacy by insisting that his position is not what the proponent has pictured it as. If the proponent continues to press the charge, the two can resolve the problem by going back over the record what the respondent actually said in the previous dialogue (to the extent that this was recorded, or can be recalled), and discuss exactly what his commitments on the issue should be taken to be, given what is now known of his remarks at the time.

It is important to realize that the job of determining what an arguer's commitments really are, or may fairly be taken to be, in a real case, is by no means trivial. Indeed, in some cases, this judgment itself can be a subject of intense argument between two parties. We are all familiar with cases of familiar disputes where one party claims, "You remember when you said that!" and the other party replies, "No, I never said

---

[2]   Probably the most familiar kind of case in modern logic would be the kind of argument called *reductio ad absurdum*.

that!" If the conversation was never witnessed or recorded, or if there is no other record of it, independent of the sayso of the two primary disputants, the issue may not even be possible to resolve.

In a normative (ideal) model of dialogue, commitments are recorded or retained in a commitment store (Hamblin 1970: 264). In the real world of everyday argumentation however, disputes can arise because this is in fact not the case, or because memory, or even a written transcript, is subject to dispute.

If the respondent is not present, as is typically the case with the kinds of cases cited as examples of the fallacy in the logic textbooks, and evaluated in a logic class, or case study, then the evaluators should be required to go very strictly by the existing discourse, using the principle of charity in fairly interpreting that text of discourse. Here, the respondent must be given the benefit of the doubt, where competing interpretations may be more or less plausible.

Analyzing and evaluating an allegation of straw man fallacy in a particular instance comes down to a question of determining fairly, by the evidence, what the commitments of a respondent can fairly be taken to amount to as explicit propositions. This judgment is arrived at by examining what was said, and how it was said, in the given context of dialogue. It depends on what type of conversation the speaker was supposed to be engaged in, when he put his original argument forward. If it was a critical discussion, then the first thing that needs to be determined is what thesis the speaker was supposed to be arguing for. Another thing of importance may be how the speaker has qualified his support for that thesis. A third factor in determining a speaker's commitments is the detailed, more localized record of what the speaker actually said as he developed his point of view, and argued against the other party's opposed point of view in the dialogue.

Commitment, as a critical and normative concept appropriate for use in evaluating cases of alleged fallacies, is not a psychological notion. It should be conceived normatively in relation to the requirements of the type of dialogue a speaker is supposed to be engaging in. There are various types of dialogue with distinctive goals and other features that define them as familiar contexts of conversation. See Hamblin (1970), van Eemeren and Grootendorst (1984), and Walton (1992). The concept of commitment is precisely defined for several different types of dialogue in which argumentation takes place by Walton and Krabbe (1994).

The key to evaluating particular cases where the straw man fallacy is alleged to have been committed is to be sought in the evidence furnished by the text of discourse and the context of dialogue, as known in that case. Of course, in some cases, there is not enough evidence to determine what an arguer's position really is, or may fairly be taken to be. In these cases, the best evaluation should be a conditional one, and the charge of fallacy judged relative to the given evidence.

However, as we have shown, the normative tools for aiding us to evaluate evidence of this kind in judging cases of the straw man fallacy have now been well enough developed to yield a clear and useful analysis of this fallacy.

References

Aristotle (1939). *Topica*. Transl. by E.S. Forster, Loeb Classical Library, Cambridge, Mass.: Harvard University Press.

Aristotle (1938). *On Sophistical Refutations*. Transl. by E.S. Forster, Loeb Classical Library, Cambridge, Mass.: Harvard University Press.

Chase, S. (1956). *Guides to Straight Thinking*. New York: Harper and Row.

DeMorgan, A. (1847). *Formal Logic*. London: Taylor and Walton.

Eemeren, F.H. van and R. Grootendorst (1984). *Speech Acts in Argumentative Discussions*. Dordrecht: Foris.

Eemeren, F.H. van and R. Grootendorst (1987). 'Fallacies in a pragma-dialectical perspective'. *Argumentation*, 1, pp. 283-301.

Eemeren, F.H. van and R. Grootendorst (1992). *Argumentation, Communication, and Fallacies*. Hillsdale, N.J.: Lawrence Erlbaum.

Evans, J.D.G. (1977). *Aristotle's Concept of Dialectic*. Cambridge: Cambridge University Press.

Govier, T. (1992). *A Practical Study of Argument*. 3rd ed., Belmont, California: Wadsworth.

Hamblin, C. (1970). *Fallacies*. London: Methuen.

Johnson, R. and A. Blair (1983). *Logical Self-Defence*. 2nd ed., Toronto: McGraw-Hill Ryerson.

Johnstone, H.W., Jr. (1959). *Philosophy and Argument*. University Park, Pa.: The Pennsylvania State University Press.

Little, J.F., L.A. Groarke and C.W. Tindale (1989). *Good Reasoning Matters*. Toronto: McClelland and Stewart.

Salmon, W.C. (1963). *Logic*. Englewood Cliffs, N.J.: Prentice-Hall.

Vernon, T.S. and L.A. Nissen (1968). *Reflective Thinking: The Fundamentals of Logic*. Belmont, California: Wadsworth.

Walton, D.N. (1985). *Arguer's Position*. Westport, Connecticut: Greenwood Press.

Walton, D.N. (1989). *Informal Logic*. Cambridge: Cambridge University Press.

Walton, D.N. (1990). 'Ignoring qualifications (*secundum quid*) as a sub-fallacy of hasty generalization'. *Logique et Analyse*, 129, 30, pp. 113-154.

Walton, D.N. (1992). *Plausible Argument in Everyday Conversation*. Albany, New York: State University of New York Press.

Walton, D.N. and E.C.W. Krabbe (1995). *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*. Albany, New York: State University of New York Press.

Can we ever pin one down to a formal fallacy?

Erik C.W. Krabbe,
University of Groningen

Abstract

In this paper, some circumstances are discussed in which it is possible to track down a formal fallacy. Charges of formal fallaciousness often seem impotent as instruments of argument evaluation and criticism. In a special dialogical setting, however, it does seem possible to pin down a formal fallacy. In order to show that, the Oliver-Massey asymmetry needs to be neutralized.

1. Introduction

Among charges of fallacy, that of having committed a formal fallacy seems particularly intractable. Whenever one tries to lay one's hands on what at first seems a flagrant case of objectionable formal invalidity, one is confronted with a plethora of devices that allow the alleged perpetrator to escape from logical criticism. In this paper I hope to show that in some special dialectical circumstances the charge may nevertheless hold water. First I shall briefly summarize the multifarious defenses that one may put up to parry the charge (Section 2). Then I shall indicate what notions of 'fallacy' and 'formal' are presupposed in this paper (Section 3). In Section 4, the dialogical setting that may give rise to the charge of having committed a formal fallacy will be described. However, upon second thoughts, we have to admit that all formal fallacies, as far as they are cases of invalidity, are cases of *non sequitur*, and that *non sequitur* is the better label for vicious invalidity in general, whereas formal fallacies in the strict sense constitute a special type of case. In order to deal with charges of invalidity, we need to refute or at least to neutralize the Oliver-Massey asymmetry thesis according to which, though we do have bona fide methods of establishing the validity of arguments, we do not have any satisfactory method of establishing their invalidity (Section 5). The question in the title is answered in Section 6.

2. Nothing but excuses

The argument that since it rains, we won't go out, is invalid. The reason being that it is quite possible to go out in the rain. The form of this argument is '*P*, therefore *Q*', so how does the arguer escape a charge of having committed a formal fallacy?

In fact the arguer is confronted with an embarrassment of riches. One way is to plead that the argument is to be given a charitable reading: quite obviously, there is, by conversational implicature, an implicit premise to the effect that whenever it rains, we won't go out. In general, the application of even a moderate type of charity in argument interpretation provides a strong pull towards a reconstruction that makes the argument a valid one. Of course this may mean that the trouble now resides in the premise, but at least the arguer is off the hook as far as formal fallaciousness is concerned. This strategy, which Massey called the enthymematic ploy (1975a), is reasonable up to a point. As the pragma-dialecticians have shown us, Grice's Cooperative Principle provides a starting point for theories that allow us to reconstruct the so-called missing elements of given arguments (van Eemeren and Grootendorst 1984: Ch. 6, 1992: Ch. 6). Govier (1987: Ch. 7: 'A new approach to charity') discusses the extent to which charity can be used in argument interpretation, without making it a ploy that does away with each and every flaw in argument.

Sometimes a more straightforward defense is available. For instance, if one argues that since everyone does what Simon does, and since everyone is doing a handspring, Simon is doing a handspring, one does not need to have recourse to the enthymematic ploy in order to rebut a charge of asserting the consequent. That is, if the critic adduces the invalid form '*for all x: if Dsx then Ex, Ef,* therefore *Dsf*' (with *Dxy*: *x* does *y*; *Ex*: everyone does *x*; *s*: Simon; *f*: the act of doing a handspring), and claims this to be the best paraphrase, the arguer may retort that the critic's analysis of the argument is too shallow to bear out its validity. An alternative analysis shows the conclusion to follow from the second premise alone: '*for all x: Fx,* therefore *Fs*' (with *Fx*: *x* does a handspring; *s*: Simon). So, by monotonicity (addition of premises cannot undo validity), the argument is valid. Generally, it is possible to counter a charge of having committed a formal fallacy by giving a refined analysis, either within the same, or within some other acceptable system of logic. That a paraphrase showing the argument to be an instance of an invalid form does not suffice to establish the invalidity of an argument is of course well-known. It is the correct part of the Oliver-Massey asymmetry thesis (Oliver 1967; Massey 1975a, 1975b, 1981).

A further line of defense consists of the destruction of the counterexamples yielded by the critic's formal analysis. For instance, let the argument be: 'if it doesn't rain, my parents are coming, it doesn't rain, therefore my father is coming'. The critic's claim is that the form of this argument is '*if not R then P, not-R,* therefore *F*' (using obvious abbreviations), that this paraphrase is sufficiently detailed and, moreover, that it is formally invalid, because of the counterexample assignment of truth to *P* and falsity to both *R* and *F*. This counterexample assignment, however, shows the invalidity of the form, but not the invalidity of the original argument, since if we revert to the assigned meanings for the variables, it turns out that the counterexample situation would have to be such that it does not rain, that my parents are coming, but that my father is not coming, and this is quite impossible. Thus invalidity (of the argument itself) has not been shown. In fact, since there are no other counterexample assign-

ments to the form, the destruction of the one possible counterexample shows that the original argument was valid. Thus the critic is hoisted with his own petard: his formal analysis is used against him to prove validity, instead of invalidity.

Another way out, for the arguer, would be to short-circuit the discussion. If it is admitted that the argument is deductively incorrect (invalid), but, at the same time claimed that the argument has some other virtue, such as inductive strength, or presumptive force, the show is over as far as formal fallaciousness is concerned.

Again, if all defenses fail (the argument really is invalid, there is no conversational implicature of a helpful missing premise, and the deductive intent is announced clearly by words like 'necessarily') a last resort might be to claim that no fallacy was committed, because the argument does not even *seem* valid. Or because this type of reasoning does not occur frequently, or does not constitute a serious problem for argumentation or cognition. And so on. Anyhow, it was an error, not a fallacy.

3. Fallaciousness and formality

The considerations given above make charges of formal fallaciousness seem almost impotent as instruments of argument evaluation and criticism. Yet, I would argue, to give them up altogether would be rash. Before we can make any advance, however, it must be made clear what the present notion of formal fallaciousness amounts to. There are so many notions of 'fallacy' and 'formal' that we cannot hope to deal with them all in this essay. No precise definitions are to be given here and now, and so an element of vagueness will remain, but a choice has to be made with respect to the types of concepts that we want to use.

The notion of fallacy used here is the pragma-dialectical one of a violation of a code for what is called 'critical discussion' or 'persuasion dialogue', i.e., a code of conduct for rational discussants whose goal is the resolution of a conflict of expressed opinions (cf. van Eemeren and Grootendorst 1984: 182, 1992: 104). The element of vagueness is that this code will not here be specified.[1] A fallacy, then, is a transgression of a rule of persuasion dialogue, whereas acts that conform to the rules, but are strategically inferior, are to be characterized as errors or blunders. The distinction can be drawn sharply only with reference to a specific set of rules. As long as our model of dialogue is incomplete, all we mean by calling an act fallacious is *that we expect it to contravene the rules, once the latter are fully specified*. Another noteworthy aspect of this notion of fallacy is that it primarily pertains to (argumentative) acts in dialogue, and only derivatively to the arguments as products (texts or recordings).

---

[1]   Ultimately, a code is to give us 'a stylized picture of how people reason or should reason'. It is to display empirical realism (to be plausible), yet it must also show 'normative bite' (Walton and Krabbe (forthcoming): Section 5.1). Thus codes of conduct for dialogue are to be based both upon the empirical study of dialogue (empirical logic) and upon normative considerations.

The term 'formal' refers to the schematic aspect of language, more specifically to the schemata that show how complex expressions are grammatically constructed from simpler expressions. (This is the second sense of 'formal' in Barth and Krabbe 1982: 15.) In a broad sense, any fallacy that somehow hinges upon the schematic aspect of language could be called formal. But this is too indefinite and would give us too much (*petitio principii, ad ignorantiam, ad hominem,* and many other types of fallacy have formal aspects). Our concept of formal fallaciousness is to be tied to the concept of formal validity. Formal validity refers (primarily) to a specific premise-conclusion relationship in an elementary or basic argumentative step (with premises $P1,..., Pn$, and conclusion C). Hence it does not refer to argumentation structures in general (trees built up from basic arguments), or to other global aspects of argumentation.

A counterexample to a basic argument ($P1,..., Pn/C$) is a situation, actual or fictitious, (a possible world, if one wishes) such that in that situation all the premises are true and the conclusion is false.[2] The argument is valid if there is no counterexample to it. The concept of what constitutes a counterexample, that is, of what is deemed to constitute a possibility and what not, depends on context and is, moreover, subject to the vicissitudes of intellectual history. Hence, the same holds for the notion of validity. But this does not mean that anything goes: we have to reckon with the conceptual possibilities and concepts of validity of our own age and with the context in which we appeal to them.

An argument is formally valid in logic L, if it can be correctly paraphrased in L such that its schema (or form) is valid in L. It can easily be seen that whenever an argument is shown to be formally valid in classical propositional or predicate logic, it has to be valid in the more general sense of not allowing a counterexample. If it were not, the counterexample would give us all the structural features necessary to define a countermodel to the argument form (in propositional logic it would give us in-structions of how to assign truth and falsity). Hence, this form would have to be formally invalid, and could not have been used to show the formal validity of the argument. In any case, a system of formal logic that did not have this property (the property that formal validity as defined by the system implies unqualified validity), would not be acceptable. Hence we may conclude that formal validity, i.e., validity on the strength of some extant and acceptable system of formal logic, implies validity. (We shall not speculate about future possible logics, and gladly accept the historical relativism in the notion of formal validity.) That the converse implication does not hold

---

[2] It is important to distinguish between this notion of a counterexample (situation) and the notion of a countermodel or counterexample assignment (also often called 'counterexample') as it occurs in formal logic. The first pertains to basic arguments, the second (set-theoretical) notion pertains to basic argument *forms.* Other notions of counterexample, not discussed in this paper, pertain to universal statements, or, more generally, to theories.

is shown by the argument about the rain, the parents, and the father in Section 2: it is valid, but not formally valid.[3]

A formal fallacy, in dialogue, is committed as soon as a party presents a formally invalid (i.e., not formally valid) argument that violates the code of conduct for the dialogue. This stipulation both opens the possibility of a (formally or simply) invalid argument that is presented without there being a violation of the code (and hence without fallaciousness), as well as that of a valid argument whose presentation nevertheless violates the code, so that it must be condemned as fallacious. The former situation occurs if the code proscribes only a subset of the (formally or simply) invalid arguments, for instance, only those that violate certain rules of thumb. The latter situation occurs if the code banishes a subset of the valid arguments, for instance those that are formally invalid. But, in order to keep matters relatively simple, let us from now on assume that only invalid arguments are proscribed by the code.

4. A profile of dialogue

When and where can a charge of having committed a formal fallacy function in dialogue? With a specific code or dialectic system in hand, answers would be easily forthcoming. But one needs to gain some insight into the various possibilities and their consequences, before a code is fixed. This can be achieved without going too far into technical details, and may, moreover, help us make up our minds when we try to get to a plausible system of normative rules of dialogue. Also, there is a simple method available that helps one to structure a discussion of various dialectical possibilities. I refer to the method of profiles used by Douglas Walton in his discussions of the spouse-beating question (Walton 1989a: 68, 69, 1989b: 37, 38) and elsewhere applied to a discussion of fallacies of relevance (Krabbe 1992). This method will now be applied to the present discussion.

Let the context of dialogue be one of persuasion dialogue. There are two parties: Wilma and Bruce. Wilma has advanced a thesis *T* (move 1). Bruce has challenged this thesis (move 2). Thus, after two moves the dialogue has yielded a conflict of expressed opinions. This conflict is *pure* or simple or nonmixed, since only Wilma has a thesis to defend, and it is *single*, since there is only one thesis (Barth and Krabbe 1982: 56; van Eemeren and Grootendorst 1984: 80; 1992: 17). Now the goal of the discussion is to reach conflict resolution, and this goal is shared by the two participants. In this sense the dialogue is cooperative. However, in order to reach that goal Bruce has to put his efforts into criticism of the thesis, whereas Wilma is to defend the thesis to the

---

[3]   Notice that a formally invalid argument is not an invalid argument of a special kind: it may very well be valid. For arguments, formal invalidity does not imply invalidity. The class of invalid arguments is just a subset of the class of formally invalid arguments, not the other way around.

best of her ability. Thus the aims of the two parties within the discussion are opposed, and in this sense the dialogue is competitive.

In the next move (move 3) Wilma is to give some reason R (possibly a conjunction of a number of statements). The dialogue has now yielded one argument: 'R, therefore T'. This is Wilma's first defense. It is now up to Bruce to make a move. There are various reasonable possibilities for the profile to display (see Figure 1).[4]

(4a)   Bruce agrees and accepts the thesis. End of discussion, conflict resolved.

(4b)   Bruce challenges R (or a conjunctive part of it). This is *tenability criticism*. It brings about a subdiscussion focused on R.

(4c)   Bruce challenges the connection between R and T. This is *connection criticism*. Wilma is to give more reasons in her next move in order to strengthen the argument.


```
                        (1)  W:  T(hesis)
                                  |
                        (2)  B:  Why T?
                                  |
                        (3)  W:  R(eason)
        _____/  /  |  \  _____
(4a) B: OK          /        |     \              (4e) (fallacy
              (4c) B: Why (R→T)?     \              criticism)
  (4b) B: Why R?                      (4d) (active
                                       criticism)
```
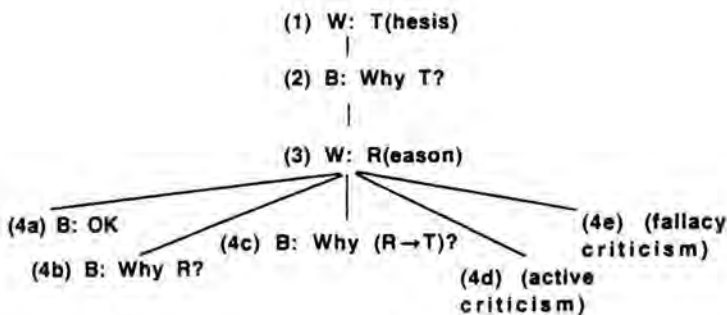
Figure 1.

(4d)   Bruce admits that Wilma was entitled to bring up this argument, but nevertheless claims that the argument is wrong or mistaken or insufficient. This is *active criticism*. Bruce has taken it upon himself as a burden of proof

---

*   Other authors have shown how various ones of these moves generate specific argument structures (Freeman 1991; Snoeck Henkemans 1992).

to show exactly how the argument went wrong or why it is insufficient. From now on Bruce is not only putting critical questions, but also arguing for certain points of view. One possibility is that Bruce offers a counter-argument, that is, an argument for *not-T*. If *T* happens to refer to a matter that can only be decided upon a balance of considerations, Bruce may do so without denying the worth of Wilma's initial argument, as far as it goes. But there are various other ways to actively criticize an argument (Finocchiaro 1980: Ch. 15, 17).

(4e)   Bruce denies that Wilma was entitled to bring up this argument. The argument is claimed to be inadmissible. This is *fallacy criticism*. Again Bruce has taken it upon himself as a burden of proof to show what is wrong with the argument. But now it will not suffice to show that the argument went wrong or is insufficient to make the conclusion acceptable. A point of order has been made (Hamblin 1970: Ch. 9, esp. 283, 284) and the discussion moves up one level. Bruce has to show that Wilma's argument does in fact violate the rules of dialogue. A lawsuit is on in which Wilma stands accused.

By way of illustration, suppose that Wilma and Bruce are discussing whom to appoint as a staff member, Jack or Jill. Suppose that Wilma argues that Jill is to be preferred, since the number of women on the staff is deplorably low. Bruce could agree (4a), or question the data about the number of women (4b), or ask for additional reasons to select Jill (4c), or bring up the consideration that Jack is better qualified for the job (4d), or decry Wilma's argument as totally irrelevant (4e). Which of these reactions would be most suitable depends, of course, upon the circumstances of the case.

The only part of the profile where a charge of having committed a formal fallacy could have a place is that engendered by move 4e. So let us consider that part in more detail. Bruce has claimed that the argument presented by Wilma is inadmissible (move 4e). This is Bruce's thesis. Wilma challenges this thesis (move 5). This gives us another nonmixed and single conflict of opinion, but this time it is located on the metalevel. The first defense move on Bruce's part consists of a specification of the type of fallacy. If it is just one rule that has been violated, Bruce could state that rule. Alternatively, Bruce could use a fallacy label, provided the meaning of this label is ultimately analyzed in terms of rule violations. So Bruce could claim that *R* has nothing to do with *T* (*ignoratio elenchi*) or that *R* is a red herring or *ad hominem*, and so on. In the present case let the claim be that there is an inadmissible lack of validity (formal fallacies must be found in that area), and let the label be *non sequitur* (move 6). In Section 2 we had a preview of various ways to wriggle out from under accusations of invalidity. Which ways of defense would it be reasonable to allow Wilma to use?

There is one fast way that should be allowed: Wilma may grant that her argument is not deductive (move 7a). In that case Wilma is to be acquitted on the spot of the

charge of having committed a *non sequitur*. However, this speedy acquittal involves some cost: Wilma may be forced to restate her thesis with a qualifier like 'probably', or 'presumably'. The rules of dialogue are to contain adequate provisions for this. After such an amendment has been made the discussion may shift to some other branch of the profile.

If Wilma does not move out, in the way just described, we shall assume that she holds to the deductive validity of her argument. But, in this branch of the profile, Wilma, as the accused, does not have to prove this deductive validity. She does not have to prove her innocence. Hence, as an alternative to move 7a, it suffices to ask a critical question, for instance: why would my argument be deductively invalid? (move 7b). See Figure 2.

(4e) B: This argument is inadmissible.

(5) W: Why inadmissible?

(6) B: It is a *non sequitur.*

(7a) W: Not deductive!   (7b) W: Why invalid?   (7c) W: Why
viciously invalid?
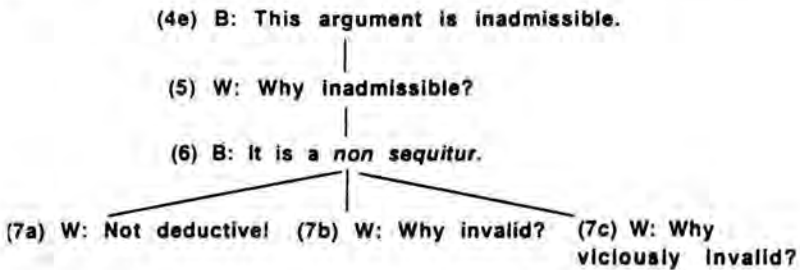
Figure 2.

It is now upon Bruce to fulfil a burden of proof. How can he show Wilma that her argument is invalid? Notice that all Bruce has to do is to convince Wilma of this, and that 'proof' in an absolute sense has no role to play in the present discussion. Three techniques Bruce could use are the following (see Figure 3):

(7b) W: Why invalid?

(8a) B: (counterexample)   (8b) B: (logical   (8c) B: (formal
analogy)   paraphrase)

(9a) W: Why this   (9b) W: Why is
paraphrase?   this form
invalid?
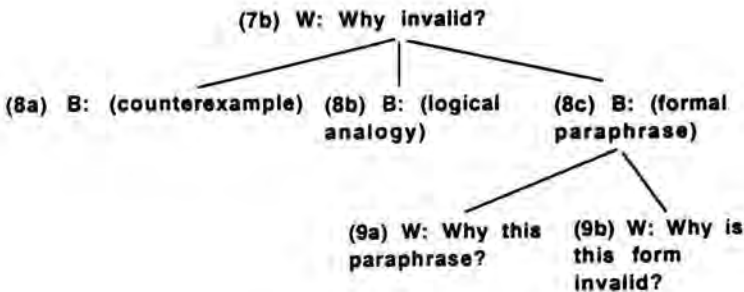
Figure 3.

(8a)   The method of *counterexample*. This is the royal road of showing invalidity. A definitive proof of invalidity to which Wilma cannot refuse her assent is actually feasible, even if it is not the case that the premises are all true whereas the conclusion is false. According to Massey this is not possible, but Massey, as we shall see, is wrong. I shall return to this move in the next section.

(8b)   The method of *logical analogy*. This technique consists of drawing up another, formally analogous, argument such that it can be shown (to Wilma) that its premises are true, whereas its conclusion is false. This may induce Wilma to admit that her own argument was invalid 'by parity of reasoning' (Woods and Hudak 1989). However, for the purpose of undercutting a pretence of deductive validity, the method is less straightforward than that of giving a counterexample.

(8c)   The method of *formal paraphrase*. This presupposes the presence of some logical skills among the participants. Bruce is to paraphrase Wilma's argument in some formal logical system. The reason that the argument is invalid is expressed as follows: 'this paraphrase captures the gist of your argument (meaning: the ground for its presumed validity), and this paraphrase constitutes an invalid logical form'.

Wilma may, in branch 8c, go on and question why this would be a paraphrase capturing the gist of the argument (move 9a). And Bruce may have a hard job, if he wants to press this on. But remember that in this branch the participants are supposed to avail themselves of some logical skills. Presumably, there are cases in which Wilma would be willing to grant that a certain paraphrase is adequate. After all, those who refuse to grant concessions with unreasonable stubbornness are, in the long run, checked by the company of discussants! Once this point has been granted, there is still the question of whether the invalidity of the logical form in question can be sustained (move 9b). Bruce has to show this by logical techniques, for instance that of defining a counter-model to the form in question.[5] If Bruce succeeds, Wilma must admit that her original argument was invalid.

If Wilma cannot be brought to concede the invalidity, she stands acquitted of having committed a *non sequitur*. But even if she does concede this, the verdict is not complete. She may ask Bruce to show her that this particular type of invalidity is vicious, i.e., that it is indeed proscribed by the code (move 7c). What Bruce's task now amounts to depends heavily on the details of the code. For instance, the code may stipulate that a case of invalidity, in order to count as a fallacy, must be discoverable

---

[5]   A countermodel is often called a counterexample, but then it is a counterexample in a sense different from that presupposed in the description of move 8a. Cf. Note 2.

by a number of standard checks, so that Wilma may be held responsible for not having gone through these checks. Let us now suppose that Bruce manages to show that these conditions are indeed fulfilled. Then the verdict *non sequitur* can be pronounced, and Wilma may be sentenced to the retraction of her argument. She may also be held accountable for the costs of the procedure. But if she is acquitted Bruce will have to pay these costs.

To end this section, let us return to the concept of a formal fallacy. Since all *non sequitur* arguments are invalid, and since all invalid arguments are formally invalid (i.e., not formally valid), all cases of *non sequitur* are, trivially, cases of formal fallacy. But the formal aspects are most prominently found in branch 8c. Perhaps it is better to reserve the verdict *formal fallacy* (in a strict sense) to those cases of *non sequitur* that are shown to be invalid by a discussion along that branch. But that will make the notion of a formal fallacy dependent upon the way the discussion is carried on. By these lights, there aren't any formal fallacies per se, yet there are techniques that may lead one to concede having committed such a fallacy, that is, to have misused logical forms in one's thought process.

5. Counterexample and invalidity

According to Massey there is only one legitimate method of establishing the invalidity of an argument: show it to have true premises but a false conclusion. This he calls 'the trivial logic-indifferent method' (Massey 1975a: 64, 1981: 494). Massey writes:

> Apart from the trivial logic-indifferent method, I claim, there is *no method whatsoever of establishing invalidity* that has theoretical legitimacy. To falsify this claim a single counter-instance would suffice. To date my critics have failed to provide any. (Massey 1981: 494)

The reason he gives to support his claim is, roughly, that arguments can be paraphrased in many different ways and that, therefore, an argument that fails to yield a valid form in one system of logic, may still yield a valid form in some other system, the class of systems being open-ended.

If Massey's claim were true, it might still be possible to gain someone's concession that his argument was invalid, but it would be *very* hard to pin the recalcitrant down to their fallacies of *non sequitur*. Except, of course, in those cases where the trivial logic-indifferent method works. But matters are different. There is another method, and this method, the method of counterexample, provides a strong and solid instrument for showing invalidity. Consider the following argument (Massey 1975a: 65, 1981: 495):

If Harrisburg is the capital of Pennsylvania, then Pittsburgh is not.
Pittsburgh is not the capital of Pennsylvania.

Harrisburg is the capital of Pennsylvania.

Even though the propositional form of this argument can be shown to be invalid by truth tables, the invalidity of the argument itself cannot be thus established. After all, there might be some other system of logic in which the argument could be paraphrased so as to exhibit a valid form. Since the conclusion is true, Massey's trivial logic-indifferent method cannot be used either. The method of counterexample, however, works. Suppose that Philadelphia were the capital of Pennsylvania. This fictitious situation is what I call a counterexample. To be more precise: there is in this situation a unique capital of Pennsylvania, and this capital is Philadelphia. So Harrisburg would not be the capital of Pennsylvania in this situation, hence the conclusion would be false, and the first premise would be true, since it now would have a false antecedent. Also, Pittsburgh would not be the capital of Pennsylvania, hence the second premise would be true. Thus the premises would be true, and the conclusion would be false. The imagined situation constitutes a counterexample, and, by definition, the argument is invalid. (Therefore, it must also be *formally* invalid, according to any respectable system of logic.)

In general, the method of counterexample works as follows: find an obviously consistent set of logically simple and perspicuous sentences that together demonstrably entail the truth of the premises and the falsity of the conclusion, i.e., find the description of a counterexample. A counterexample may be fictitious, therefore it is not required that these sentences be true. To find the required set, logical analysis may be very helpful. E.W. Beth's method of semantic tableaux, in particular, is an effective instrument for the discovery of counterexamples. But in order to convince one's opponent, one need not expound the techniques used in the discovery of the counterexample. It suffices to convince her that these sentences describe a possible situation, and then derive the required truth values for the premises and conclusion of the original argument.

The method of counterexample does not merely reduce the problem of invalidity to the problem of consistency. These problems are, of course, equivalent, and a reduction of one to the other would not, by itself, constitute any advance. The method, however, does more: it reduces complicated cases of invalidity (or consistency) to relatively simple cases of consistency and entailment. Massey's trivial logic-indifferent method constitutes just a special case: the case where we present (part of) the actual world, instead of some fictitious world. But the actual world is just one of the possible worlds. Why would it be theoretically illegitimate to use our imagination?

6. Conclusions

We saw that charges of having committed a formal fallacy can be made in persuasion dialogue, but that a more prominent place may be given to the charge of having committed a *non sequitur*. The metadialogue in which the critic tries to substantiate a case of *non sequitur* is concerned with two claims: on the one hand, there is the claim that the original argument was invalid, on the other hand there is the claim that this

is a vicious type of invalidity, that it constitutes a violation of dialogue rules. The critic has to establish both claims in order to make the charge stick.

To show invalidity the critic may avail himself of several techniques: counter-example, logical analogy, and formal paraphrase. The use of these techniques is, by the way, not restricted to a metadialogue that follows upon a charge of fallacy. The very same techniques can also be used in active criticism (branch 4d of the profile), where there is no charge of fallacy.

The most vigorous technique is that of presenting a counterexample. The Oliver-Massey thesis propounding the asymmetry between the methods of proving validity and the methods of proving invalidity can not be taken to imply that there is no robust method of showing invalidity, and hence does not stand in the way of establishing cases of *non sequitur*.[6] The method of counterexample seems even strong enough to warrant charges of *non sequitur* fired at 'monolectic' arguments, such as argumentative texts (which according to a pragma-dialectical tenet are to be viewed upon as implicit discussions). That is, the method could be used to show that any reasonable recon-struction of a certain (monolectic) argument is invalid, and that this invalidity is of a type that a responsible author should have avoided.

Whether a fallacy of *non sequitur* is to be called a formal fallacy in the strict sense depends upon the way the invalidity claim was defended in dialogue. It is proposed that this label be restricted to cases where the, relatively sophisticated, technique of formal paraphrase has been used. This technique, in order to be successful, presupposes a fair degree of cooperativeness on the side of the accused: at a certain point she is to stop quibbling about the paraphrase.

It seems that the answer to the question whether we can ever pin the opponent down to a formal fallacy must be: yes we may be able to pin her down, but only if she lets us.

---

[6]   Cf. Govier (1987): Ch. 9: 'Four reasons there are no fallacies?', Section 2: 'Formal invalidity as no story'.

References

Barth, E.M. & E.C.W. Krabbe (1982). *From Axiom to Dialogue. A Philosophical Study of Logics and Argumentation*. Berlin/New York: Walter de Gruyter.

Eemeren, F.H. van & R. Grootendorst (1984). *Speech Acts in Argumentative Discussions. A Theoretical Model for the Analysis of Discussions Directed towards Solving Conflicts of Opinion*. Dordrecht/Cinnaminson: Foris.

Eemeren, F.H. van & R. Grootendorst (1992). *Argumentation, Communication, and Fallacies. A Pragma-Dialectical Perspective*. Hillsdale, New Jersey/Hove/London: Lawrence Erlbaum Associates.

Finocchiaro, M.A. (1980). *Galileo and the Art of Reasoning. Rhetorical Foundations of Logic and Scientific Method*. Dordrecht/Boston/London: Reidel.

Freeman, J.B. (1991). *Dialectics and the Macrostructure of Arguments*. Berlin: Foris.

Govier, T. (1987). *Problems in Argument Analysis and Evaluation*. Dordrecht/Providence, RI: Foris.

Hamblin, C.L. (1970). *Fallacies*. London: Methuen.

Krabbe, E.C.W. (1992). 'So what? Profiles for relevance criticism in persuasion dialogue'. *Argumentation* 6, 2, pp. 271-283.

Massey, G.J. (1975a). 'Are there any good arguments that bad arguments are bad?'. *Philosophy in Context* 4, pp. 61-77.

Massey, G.J. (1975b). 'In defense of the asymmetry. Questions for Gerald J. Massey'. *Philosophy in Context* 4, Supplement, pp. 44-56.

Massey, G.J. (1981). 'The fallacy behind Fallacies'. In: P.A. French, Th.E. Uehling, Jr. & H.K. Wettstein (eds.), *The Foundations of Analytic Philosophy* (pp. 489-500), Minneapolis: University of Minneapolis Press, Midwest Studies in Philosophy 6.

Oliver, J.W. (1967). 'Formal fallacies and other invalid arguments'. *Mind. A Quarterly Review of Psychology and Philosophy* 76, pp. 463-478.

Snoeck Henkemans, A.F. (1992). *Analysing Complex Argumentation. The Reconstruction of Multiple and Coordinatively Compound Argumentation in Critical Discussion*. (PhD thesis). Amsterdam: SICSAT.

Walton, D.N. (1989a). *Question-Reply Argumentation*. New York: Greenwood Press.

Walton, D.N. (1989b). *Informal Logic. A Handbook for Critical Argumentation*. Cambridge: Cambridge UP.

Walton, D.N. & E.C.W. Krabbe. *Commitment in Dialogue. Basic Concepts of Interpersonal Reasoning*. State University of New York Press. Forthcoming.

Woods, J. & B. Hudak (1989). 'By parity of reasoning'. *Informal Logic. Reasoning and Argumentation in Theory and Practice* 11, 3, pp. 125-139.

Informal factors in the formal evaluation of arguments

Maurice A. Finocchiaro,
University of Nevada-Las Vegas

Abstract

In this paper, it is argued that the Oliver-Massey asymmetry - the asymmetry between showing that a given argument is formally valid, and showing that it is formally invalid - does not hold. Both formal validation and invalidation can be justified to a greater or lesser degree. However, both processes are based on formally invalid arguments.

1. Aim

The question I whould like to explore in this paper is whether there are significant differences between the process of showing that a given argument is formally valid and the process of showing that a given argument is formally invalid. These two processes may be labeled, for short, the formal validation and the formal invalidation of arguments; furthermore, a significant difference between the two may be labeled an asymmetry. Thus, the problem may be formulated as asking whether there is an asymmetry between the formal validation and the formal invalidation of arguments, and what are the implications of such an asymmetry or lack thereof. In other words, the question is whether there is a logical and epistemological asymmetry in the concepts of formal validity and invalidity.

The motivation underlying this undertaking is two-fold. On the one hand, some philosophers have advanced the thesis that there is an asymmetry between formal validation and formal invalidation, and this strikes me as an intriguing claim whose correctness deserves further scrutiny. On the other hand, I believe that this problem is a fruitful and instructive one for anyone concerned with understanding the relationship between logic and argumentation.

2. Theoretical Context and Conceptual Framework

Before plunging into the details of this problem of formal validation versus formal invalidation, it will be useful to sketch the theoretical, conceptual, and methodological context underlying my investigation. In this inquiry, both formal validation and formal invalidation are regarded as special cases of formal evaluation; moreover, formal validation is also taken to be a special case of positive evaluation, and formal in-

validation is also considered to be a special case of negative evaluation. Similarly, these three kinds of evaluations (formal, negative, and positive) may be regarded as special cases of the evaluation of arguments in general.

Evaluation is here something I would distinguish from the construction and the interpretation of arguments, without however separating these three activities. In other words, the construction, the interpretation, and the evaluation of arguments are interrelated, but that is not to say that they are the same. In short, this distinction is meant to avoid confusion and conflation, not to establish a separation or bifurcation. At the terminological level, I think it would be proper to refer to what I have in mind by means of other labels such as appraisal, assessment, and judgment. I am reluctant to add the term "criticism" to this list of near-synonyms because I think criticism tends to have a negative connotation, in the sense of negative evaluation; nevertheless, to the extent that this connotation can be avoided, I would have no objection to speaking also of criticism.[1]

My distinction between negative and positive evaluation is meant to refer to the difference between good and bad arguments, or right and wrong, correct and incorrect, sound and unsound, valid and invalid, logical and illogical, convincing and unconvincing, plausible and implausible, erroneous and not, fallacious and not, and the like. One could also speak of the difference between favorable and unfavorable evaluation, approving and disapproving, appreciative and depreciative, or constructive and destructive. Of course, at a subsequent and more refined level of analysis one may want to explore the differences among all the pairs just mentioned. I do not mean to pre-empt the possibility of such refinements; all I am saying is that they do not significantly affect the issue I am concerned with here.

Similarly, I do not mean to deny the propriety of discussing whether the evaluation of arguments is a viable and feasible activity.[2] To be sure, I am inclined to think that the anti-evaluation stance is ultimately untenable, partly because it is itself the evaluative expression of a thesis about evaluative arguments, and partly because I am too realistically and empirically minded to be blind to the fact that arguments are getting evaluated all the time. However, for the purpose of the present discussion I am simply taking for granted that there is such a thing as the evaluation of arguments and that many examples exist. On this assumption, I then want to explore the nature of the difference between two types of evaluative verdicts.

One other preliminary clarification is in order. Much of what claims to be, or is often taken to be, evaluation of arguments is really evaluation of individual claims, theses, beliefs, theories, hypotheses, and the like. Here, my point is that I want to be strict and literal, so that by the evaluation of an argument I do not mean simply the evaluation of a claim, but rather the evaluation of a claim together with the supporting

---

[1]   For a discussion of some non-negative aspects of "criticism," see Bailin (1988); cf. Finocchiaro (1989b; 1990).

[2]   For some insightful discussions of this cluster of issues, see Willard (1983).

reasons or justifying premises. This point is worth stressing despite the fact that one type of evaluation of a claim would consist of examining whether it is justified, and to do the latter one would have to evaluate actual or potential justifications of the claim, namely arguments; the reason is that I want to deal with the evaluation of arguments directly and explicitly, not indirectly and implicitly. In other words, in evaluating an argument favorably, scholars and laymen often advance some other argument in support of the conclusion of the original argument; and analogously, in evaluating an argument unfavorably, they often give a refutation of the original conclusion, namely a counterargument, namely another argument whose conclusion is a denial of the original conclusion. The study of the evaluation of individual claims is certainly instructive in itself and would have to be part of a general theory of evaluation; and it is relevant to the present problem because there are important relationships between the evaluation of arguments and the evaluation of the conclusions of arguments. But, again, relationships ought not to be turned into confusing conflations, any more than distinctions should become bifurcating separations.

Finally, one last distinction will allow us to go to the heart of the problem, and not merely enable us to clarify a different aspect of the matter. The evaluation of arguments may be conceived as involving two main things: an evaluative claim about some argument, and the articulation of the rationale for the evaluative claim, namely a metalevel argument about the original argument. Now, we could not go very far in understanding the nature of the evaluation of arguments without understanding how one justifies such evaluations. In other words, we need to study the logic (and epistemology) of a special class of arguments, namely evaluative arguments, and in particular evaluative arguments about arguments; and the question we want to examine is whether these meta-arguments exhibit a significant difference when their conclusions express favorable verdicts from when they express unfavorable verdicts.

3. Formal Validation versus Formal Invalidation

In what follows I focus on a special case of the general problem which has been discussed by Gerald Massey and his critics.[3] Their concern has been with evaluation from the point of view of formal logic, namely with assessing the formal validity or the formal invalidity of an argument; thus, as suggested above, one could speak here of the formal validation and the formal invalidation of arguments. Massey has elaborated the thesis that, whereas formal logic does provide techniques for proving arguments formally valid, it does not provide any techniques for proving arguments formally invalid; and he has argued that, therefore, there is a fundamental asymmetry between

---

[3]   Massey (1970, pp. 93-94; 1975a; 1975b; 1976; 1980; 1981a; 1981b; 1987); Bencivenga (1979); George (1983); and McKay (1984, 1987). Other noteworthy critiques of Massey are Govier (1987, chapter 9), Johnson (1989), Krabbe (1995), and Woods (1989; 1995); these critiques raise important issues and are incisive, but their focus is different from the present one, or else they appeared while this paper was in press and could not be taken into account in a substantive manner.

formal validation and formal invalidation, the asymmetry being that the former is feasible, but the latter is not. Using argument-theoretical language, and echoing the title of one of Massey's papers, one might say that he has tried to show that, while it is possible to give good arguments that good arguments are good, it is not possible to give good arguments that bad arguments are bad.

Massey himself has occasionally presented this thesis as a direct attack on informal logic. However, the conception of informal logic presupposed in that polemic is the one which equates the field with the study of fallacy in general, and of so-called informal fallacies in particular. Now, whatever justification there may have been in the past for such a conception, the field has lately advanced much beyond that,[4] and so we need not say anything more about that aspect of the polemic.

Massey's primary objective has really been to exploit the asymmetry in order to motivate and justify his own program for a philosophical linguistics which would simultaneously unify logic and grammar and provide what could plausibly be called a "natural logic";[5] this program is meant to provide principles for the analysis of the formal structure of argumentation in natural language, and such that formal validity and formal invalidity would be consequences of grammatical-logical well-formedness and the failure of well-formedness. Despite the attractiveness of this empirical bent and of this emphasis on natural language argumentation, the formalism of the program is so ambitious that I feel no inclination to discuss further this alleged consequence of the alleged asymmetry.

At any rate, the asymmetry has been criticized from both directions. That is, Ermanno Bencivenga (1979) has objected that the formal validation of arguments is as questionable as their formal invalidation, and Thomas McKay (1984, 1987) has objected that the formal invalidation of arguments is not as hopeless and arbitrary as Massey makes it sound. In the one case, formal validation and invalidation are equally unfeasible, strictly speaking; in the other case, they are equally feasible, loosely speaking. But in both cases they stand on an equal plane, and the asymmetry disappears.

Massey himself has recently credited to someone else the discovery of the asymmetry thesis, that is, to the lead paper in the October 1967 issue of *Mind*, by American philosopher James W. Oliver (1967). Therefore, Oliver's contribution should also be taken into account.

As suggested above, the problem of formal validation versus formal invalidation is an important one, and now I can also suggest some reasons for this importance. First, my resolution of this problem will lead to an appreciation of some of the limitations of formal logic, and these point in the direction of informal logic. Moreover, this type of asymmetry seems to be a special case of a general phenomenon or cluster of prob-

---

[4]   See, for example, Doss (1985), Finocchiaro (1980; 1984), Freeman (1994), and Johnson and Blair (1985). See also the criticism in Johnson (1989).

[5]   Massey (1975a, pp. 74-76) regards this as being partly in the tradition of Lakoff (1970).

lems which have been studied from many quarters and which cannot be easily dis-missed.[6] Finally, it should also be kept in mind that the literature on formal validation versus invalidation has prima facie an added, self-referential relevance insofar as it constitutes for us something of a case study of the problem: we have Oliver's original argument; then Massey's argument could be regarded as a positive evaluation of that original argument; on the other hand, the objections by Massey's critics (Bencivenga, McKay, and George) could be regarded as negative evaluations of Oliver's argument.

4. Oliver on Showing Invalidity

Let us now focus on Oliver's argument since, as Massey (1987) himself has pointed out, it antedates Massey's account by about a decade. Oliver's main conclusion is that there is one and only one rigorously correct method of showing that an argument is formally invalid: it is to show that its premises are true and that its conclusion is false. This claim has two parts: one negative, the other positive. The positive part asserts that

(1)   showing that the premises are true and the conclusion false is a rigorously correct method of showing that the argument is formally invalid.

The negative part asserts that

(2)   there is no other rigorously correct method of showing that an argument is formally invalid.

---

[6]   For more details on this, see Finocchiaro (1980, pp. 332-41; 1981, pp. 17-18; 1988, pp. 28-29, 121-22, 141, and 245-48; and 1992). Here one should add two other apparent asymmetries which emerged in discussions at the Third International Symposium on Informal Logic, University of Windsor, Canada, 15-18 June 1989, where parts of this paper were first presented. One was hinted at in the panel discussion on legal reasoning, and it involves situations where considerations of the burden of proof are important; it seems that in Anglo-Saxon jurisprudence guilty verdicts must be proved more stringently than innocent verdicts, that is, legal proofs of guilt are more easily criticizable than legal proofs of innocence. Another type of asymmetry was mentioned by Michael Scriven (1989) in his plenary lecture, and it involves a significant difference between the negative evaluation and the positive evaluation of at least products. He pointed out that there is an important class of evaluation situations where the identification of a particular flaw is sufficient to justify the rejection of the product (for example, demonstrably unsafe brakes in an automobile), whereas the product can be deemed accept-able only when it possesses a whole list of necessary qualities; this suggests that negative evaluative conclusions are easier to justify than positive ones; or, in argument-theoretical terminology, we might say that good arguments with negative evaluations as conclusions are easier to construct than arguments with positive evaluative conclusions. This type of asymmetry certainly deserves further study, partly to determine its exact nature, and partly to determine its exact relationship to the others. For example, its direction seems to run counter to that of some of the other asymmetries because here the justification of negative evaluations seems easier than the justification of positive ones, whereas the asymmetry suggested above seemed to involve the reverse. Moreover, in his lecture Scriven himself discussed the case of what he called evaluative arguments whose inferential soundness is not affected by the falsity of some premises; in these cases no one quality can outweigh the cumulative effect of all the others, and therefore even the present type of asymmetry does not seem to apply.

Oliver's argument in support of the positive part of his thesis is brief and un-controversial. It is based primarily on the following definition of formal validity:

(1.1)   a formally valid argument is one which instantiates a valid argument form; and
(1.2)   a valid argument form is an argument form which has no counterexamples; where
(1.3)   a counterexample to a given argument form is an argument which instantiates that form and has true premises and false conclusion.

The rest of the argument would be the following: whatever forms the given argument instantiates, none of those forms are valid because they all have a counterexample, namely the given argument in question.

These definitions and this argument are only a slightly more pedantic manner of saying the following: a formally valid argument is one such that it is impossible for its premises to be true and its conclusion false; now if the given argument has true premises and false conclusion, then it is obviously possible for its premises to be true and its conclusion false. QED.

Oliver's argument in support of the negative part of his thesis (2) is more round-about, and ultimately can at best be only inductively correct since it depends on an empirical or imaginary examination of various methods together with a disqualification of each as not being rigorously correct. His examination consists of three types of con-siderations:

(2.1)   no method is rigorously correct if it depends on the false principle that (2.11) a formally invalid argument is one which instantiates an invalid argument form;
(2.2)   all apparently rigorous methods, such as truth tables, syllogistic rules, and Venn diagrams, are methods for showing the invalidity (or validity) of various argument forms rather than of arguments;
(2.3)   if one examines the textbooks, none of the methods one finds are in fact rigor-ously correct: they either use the false principle (2.11) just mentioned, or they equivocate between arguments and argument forms.

Let us focus on this principle (2.11) that a formally invalid argument is one which instantiates an invalid argument form, for which we may adopt the elegant label of "pseudo-principle of illogical form" given to it by Massey (1987). Now, as Oliver points out, the first thing to understand about this principle is that, although it looks equivalent to the definition of formal validity, it really is not. That is, it looks like one can obtain this principle by starting with the definition and replacing the term "valid" by the term "invalid." The two principles are indeed isomorphic transformations of each other, with the terms "valid" and "invalid" interchanged. However, such a trans-formation is itself invalid. In particular, the invalidity pseudo-principle does not follow validly from the definitional principle of validity. In other words, the following meta-argument is not formally valid:

(3)   a formally valid argument is one which instantiates a valid argument form,

(4)   so, a formally invalid argument is one which instantiates an invalid argument form.

Whatever plausibility this argument has, it stems from the fact that it appears to instantiate an argument form such that the premise is a biconditional and the conclusion is another biconditional whose two components are denials of the components of the former. That is, the premise may be interpreted as the biconditional that

(3.1)   an argument is valid iff it instantiates a valid argument form;
         and the conclusion may be interpreted as the biconditional that

(4.1)   an argument is invalid iff it instantiates an invalid argument form.

And then it looks as if the transition from the premise to the conclusion instantiates the form:

(3.2)   p iff q, so (4.2) not-p iff not-q,

or at least the form:

(3.3)   (x)(Fx iff Gx), so (4.3) (x)(-Fx iff -Gx).

However, such instantiations do not work out. For example, if "Gx" symbolizes the expression "x instantiates a valid form" then "-Gx" would symbolize the expression "x does not instantiate a valid form," and the latter is by no means synonymous with the expression "x instantiates an invalid form." The difference between these two expressions is that the first means that there is no valid form instantiated by x, whereas the latter means that there is an invalid form instantiated by x. Now, going back to the argument in question, what this shows is that it does not instantiate either one of the two valid forms mentioned. But this does not prove its invalidity, since to do this one would have to show that the argument does not instantiate any valid form. What this analysis does do is to explain the semblance or appearance of validity.

At this point it would be natural to try to find a form which the argument does instantiate. The premise is an instance of the form:

(3.4)   (x){if Ax then [Vx iff (Ey)(Fy & Ixy & Vy)]},[7]

which may me read as:

---

[7]   Here and elsewhere in this paper, the existential quantifier is symbolized by the letter 'E' rather than the more usual "backward-E".

(3.5)   an argument is formally valid iff there is some argument form such that the argument instantiates it and this form is valid.

And the conclusion instantiates the form:
(4.4)   (x){if Ax then [-Vx iff (Ey)(Fy & Ixy & -Vx)]},

which may be read as:

(4.5)   an argument is not formally valid iff there is some argument form such that the argument instantiates it and this form is not valid.

Now, this meta-argument form is itself invalid, but to conclude from the invalidity of this form that the original argument is invalid would be to commit the error which Oliver is trying to expose.

Therefore, he resorts to what he regards as the only correct method of proving in-validity, mentioned above. He first points out that the premise is true by definition, and then he refutes the conclusion by giving some counterinstances, that is, instances of arguments which instantiate invalid forms but are valid. One of these is the following instance of affirming the consequent:

(5)   "If something is red, everything is red.
        Everything is red.
        Therefore, something is red" (Oliver, 1967, p. 463).

In regard to the apparently rigorous methods of proving invalidity, for the cases of truth tables and Venn Diagrams, it is obvious that they refer to argument forms and not to arguments, and that therefore the gap remains between the invalidity of the forms and the invalidity of the original natural language arguments. The case of the syllogistic rules is not so obvious. Oliver's argument here is that there are many syllogisms which instantiate invalid forms but are valid. Consider for example:

(6)   "Some men are non-self-identical.
        Some Parisians are non-self-identical.
        Therefore, no Parisians are men" (Oliver, 1967, p. 471).

This instantiates the syllogistic form:

(7)   some A are B; some C are B; so, no C are A,

which violates all the main rules of the syllogistic theory. Yet the original argument is formally valid because the premises are both logically false. Or consider the argument:

(8)   "Nothing that is blue or not blue is square.
       Nothing that is red or not red is square.
       Therefore, something that is red or not red is blue or not blue"
       (Oliver, 1967, p. 472).

This syllogism instantiates the form:

(9)   no A is B; no C is B; so, some C is A,

which is obviously invalid and violates several rules. Yet the original argument is form-ally valid because the conclusion is logically true.

5. Massey's Asymmetry Thesis

Although in his latest paper on the topic Massey (1987) has credited Oliver with having anticipated the asymmetry thesis, Oliver does not in fact speak of asymmetry. Indeed, he says almost nothing about methods for proving validity, and so he is making no comparison and contrast on the basis of which he might have inferred an asymmetry. This contrast and this inference were Massey's own contributions in his *earlier* papers on the topic.[8] Working at that time independently of Oliver, Massey did three main things. He explained how the definition of formal validity yields a viable method for proving that arguments are valid: this is the familiar technique of finding a form which the argument may be said to instantiate and which is valid. He strengthened Oliver's conclusion that there is one and only method of proving formal invalidity, namely the combined verification of premises and falsification of conclusion, which Massey (1975a, p. 64) labels "the trivial logic-indifferent method."[9] And then on the basis of these two theses he explicitly drew the conclusion that there is a fundamental asymmetry between formal validation and invalidation.

Massey's strengthening of the point about the difficulties with formal invalidation is as follows. As we have already seen, the basic definition of formal validity is essentially a biconditional one side of which states that

(10)   if an argument instantiates a valid argument form, then it is formally valid;

---

[8]   I emphasize *earlier* because in his latest paper on the topic, Massey (1987) seems to abandon the asymmetry thesis and to focus on the indeterminacy of translation, which he claims applies to both formal validity and formal invalidity; however, he claims (1987, p. 6) that the indeterminacy of translation applies asymmetrically to the cases of validity and invalidity, and so a new version of his asymmetry thesis emerges. The examination of this new alleged asymmetry is beyond the scope of this paper.

[9]   Massey also provided the following more elegant counterexample to the pseudo-principle of formal invalidity: "If something has been created by God, then everything has been created by God. Everything has been created by God. [Therefore,] Something has been created by God" (1981a, p. 492).

the other side may be stated as saying that

(11)   if an argument does not instantiate a valid argument form, then it is formally invalid.

The latter principle could be regarded as providing a method of showing invalidity by showing that the argument fails to instantiate a valid form.

The essential difficulty with this latter principle is that the class of valid argument forms is not closed; that is, we cannot provide a complete list of all valid argument forms. Formal logic is an open-ended science, and formal logicians are constantly adding to the known list of valid argument forms. This is illustrated even in the pedagogy of formal logic, which usually proceeds from truth functions, to monadic predicates, to relational predicates, to identity theory, and so on. Thus, if a valid syllogism is symbolized merely with the resources of the truth-functional calculus, it will fail to instantiate any valid form, but that will not render it invalid, but merely *truth-functionally* invalid. Similarly, if a valid argument whose validity depends on identity is analyzed with the resources of predicate calculus without identity, it will fail to instantiate a valid form, but that will not render it invalid.

Further, in the context of logical theorizing, Massey gives two interesting examples. Consider the argument:

(12)   John took a walk by the river; so, John took a walk.[10]

This argument is indeed intuitively valid, and could even be shown to be valid with the help of some meaning postulates. However, its *formal* validity was in question until Davidson (1968) devised an argument form which this argument instantiates. Now, if the above principle had been taken to refer to known logical forms, before Davidson's analysis it would have declared this argument invalid. Similar remarks apply to an argument first shown valid by Leonard and Goodman (1940) by devising a calculus of individuals to enrich ordinary predicate calculus. The argument is:

(13)   Tom, Dick, and Harry are partners; so, Tom and Harry are partners.[11]

Here the essential point is that, just because one has not found a valid argument form, one cannot be certain that someone else will not find it, or that someone else will not invent some new logical system which will allow us to devise an appropriate form.

---

[10]   Massey (1981a, p. 495); cf. Massey (1987, pp. 7-8) and Davidson (1968).

[11]   Massey (1981a, p. 495); cf. Massey (1976), and Leonard and Goodman (1940).

6. The Formal Validation of Arguments

Criticism of the Oliver-Massey thesis has tried to show that on the one hand formal validation is not as rigorous as they make it sound, and that on the other hand formal invalidation is more rigorous than they make it sound.[12] Let us begin with formal validation.

One criticism is that the formal validation of an argument depends on the existence of valid argument forms, but the validation of the latter is a less rigorous affair than it may seem.[13] Recall that to say that an argument form is valid is to say that it has no counterexamples. Now, although finding even one counterexample will invalidate the form, not finding it may be due to lack of ingenuity. Of course, one will have a validation if one finds a proof in a consistent and sound logical system, but many logical systems (including relational predicate calculus) are not decidable, and so finding a proof is not a mechanical task. In other words, in special cases of decidable logical systems like the truth-functional calculus, the validation of an argument form is a mechanical procedure; but in general, since many logical systems are undecidable, the validation will depend on the construction of a proof in a sound and consistent system. Such a construction will provide a rigorous validation, but finding it is not a rigorous process.

The plausibility of Massey's thesis about formal validation derives partly from the simplicity and triviality of his examples. One of these involves an argument of the form "P & Q; so, Q,"[14] another the form "all A are B; x is A; so x is B."[15] The issue could have been much more difficult if we were dealing with a form like the following:

(14)   -(Ex)(Ey)(Fx & Gy & Hxy);
         (x)([if Fx then (Ey)(Fy & Hxy)];
         (x)(y)(if Hxy then Hyx);
         so, -(Ex)(Fx & Gx).[16]

---

[12] This useful classification is due to Massey himself (1987). There is also the criticism of George (1983), objecting that Massey presupposes an untenable concept of argument, as well as the critiques in Johnson (1989), Govier (1987, chapter 9), Krabbe (1995), and Woods (1989; 1995), which raise other issues or are too recent to take into account in this paper.

[13] This point is similar to one made by Bencivenga (1979, pp. 249-50).

[14] Cf. Massey (1975a, p. 63).

[15] Cf. Massey (1975a, pp. 64-65).

[16] See Kalish, Montague, & Mar (1980, pp. 249, 261), where this form is instantiated by the argument: no teacher is married to a student; every teacher is married to a teacher; marriage is a symmetrical relationship; so, no teacher is a student.

Another criticism of the above-mentioned thesis about the formal validation of arguments involves a fact admitted by Massey himself (1980, p. 321), and exploited by him for other purposes. That is, the validation of *arguments* involves linguistic intuitions to the effect that certain statements are synonymous with certain others. This applies even to the trivial example referred to in the previous paragraph. The original argument was:

(15)   Sam and Sue are doctors; therefore, Sue is a doctor.[17]

Here the premise must be intuited to be synonymous with the statement that "Sam is a doctor and Sue is a doctor"; only then can we regard it as an instance of the conjunction "P & Q." The crucial importance of linguistic intuition is clearly and vividly shown by another example given by Massey where the same translation would be wrong: "Tom and Dick weigh 200 kilograms" (1980, p. 320). This obviously does not mean that Tom weighs 200 kilograms and Dick weighs 200 kilograms.

7. The Formal Invalidation of Arguments

Let us now consider some criticism of the Oliver-Massey thesis about formal invalidity. As mentioned before, this criticism tries to show that formal invalidity can be justified without reliance on the nonsequiturs and equivocations they are bent on exposing.

It will be useful to focus on the argument we examined earlier, which goes from the definition of formal validity to the pseudo-principle of illogical form. The meta-argument was as follows:

(16)   an argument is formally valid iff it instantiates a valid argument form;
        so, an argument is formally invalid iff it instantiates an invalid argument form.

Earlier, following Oliver, we invalidated this argument by pointing out that the premise is true and the conclusion is false. Moreover, it certainly would be formally invalid to argue that this argument is invalid *simply because* it instantiates the following invalid form:

(17)   (x){if Ax then [Vx iff (Ey)(Fy & Ixy & Vy)]}
        so, (x){if Ax then [-Vx iff (Ey)(Fy & Ixy & -Vy)]}.

However, as Thomas McKay has argued, "instancehood" is not their only relationship, for we can also show that the form "represents all details likely to be relevant to the validity of the argument" (McKay, 1984, p. 99). Notice that this is something that would not be true if we were to interpret a valid syllogism by the mere resources of the sentential calculus, which would yield the argument form:

---

[17]   Cf. Massey (1975a, p. 63).

(18)   P; Q; so, R.

Notice also that this is a form also instantiated by every argument with two premises and one conclusion. Similarly, recall that earlier we mentioned an argument which is valid even though it instantiates the form of affirming the consequent; that form did not represent all relevant logical details.

Therefore, the pseudo-principle of illogical form should be modified to read:

(19)   an argument is formally invalid iff it instantiates an invalid argument form *and* this form "represents all details likely to be relevant to the validity of the argument."[18]

Now, this modified rule of formal invalidity is not as neat as the pseudo-principle, and its satisfaction is both a matter of degree and subject to revision. However, all that this means is that the formal invalidation of an argument is an empirical, pragmatic, or informal business.[19] It remains true that by using such a rule arguments cannot be formally invalidated by means of formally valid meta-arguments, but it is equally true that their formal invalidation *can* be justified. We might add that this difficulty with this part of the Oliver-Massey thesis is that they were restricting themselves to what might be called deductive arguments, in a situation where all we can ever hope for is inductive arguments.

A second criticism would involve a reappraisal of what is called the "method of counterexample" by some (Salmon 1984, p. 21) and the technique of "refutation by logical analogy" by others.[20] This is the technique of invalidating an argument by formulating another argument with the same logical form as the first and with obviously true premises and obviously false conclusion. Oliver (1967, pp. 469-70) explicitly criticizes this technique by interpreting it as being identical to the technique of using the pseudo-principle of illogical form. I suppose he is thinking that to say that two arguments have the same form is to say that the two arguments instantiate the same form. On this interpretation, his criticism would indeed follow.[21]

---

[18]   Again, the quoted clause is from McKay (1984, p. 99).

[19]   Bencivenga (1979) also reaches conclusions that point in a pragmatic and empirical direction.

[20]   Copi (1986a, pp. 187-90; 1986b, pp. 289-91); Govier (1985); and Oliver (1967, p. 469). McKay (1987) also discusses this same technique, although without using the label. Since writing these words, I have been convinced by Krabbe (1995) that the method of logical analogy should be distinguished from the method of counterexample, although, as Krabbe also clarifies, the latter should also be distinguished from the formal method of countermodel; his important paper deserves careful study.

[21]   Oliver's interpretation was perhaps partially justified since the target of his criticism seemed to be Copi, who does indeed have that in mind. In fact, although Oliver (1967, p. 469) deliberately omits a specific bibliographical reference for the quotation he is criticizing, and although he must have been using an earlier edition of Copi's book, even the seventh edition of the latter reads almost exactly like Oliver's quotation; cf. Copi (1986b, p. 289).

However, this technique ought to be interpreted as a way of bypassing the problem of having to deal with a logical form to attribute to the arguments in question and to be instantiated by them. The really important thing would be the counterexample, which is another actual argument. Talking of the same form of two arguments thus would be a way of saying that they are formally isomorphic or analogous, that is, that there is a one-to-one correspondence between various elements of the two arguments; it would not be a way of saying that each of the two arguments has some unique logical form and that the logical form of each is identical.[22]

Let us apply the technique to the same meta-argument discussed above. When presented with this argument, one could respond that it is no more valid than the following argument, which is obviously invalid:

(20)   a person is a doctor iff he has received a doctoral degree;
        so, a person is a nondoctor iff he has received a nondoctoral degree.[23]

This argument is obviously invalid because its premise is obviously true, whereas its conclusion is obviously false: a nondoctor is a person who has not received *any* doctoral degrees, whereas doctors usually receive other nondoctoral degrees before their doctoral one.

Now, let us ask why they could be said to have the same logical form. Clearly it would be incorrect to say that the original argument (16) and the counterexample (20) have the same logical form because they instantiate the same argument form; we have already seen that, for example, all syllogisms may be said to instantiate the form "P, Q, so R," but this does not even begin to give them the same form. Equally obviously, it would be correct to say that these two arguments have the same form because they instantiate the same argument form (17) *and* this form represents all the details likely to be relevant to their validity; however, this type of consideration would merely repeat

---

[22]   The move I am suggesting is analogous to Quine's move about meaning and sameness of meaning: the latter need not presuppose the existence of mysterious entities called meanings which words have, but may be conceived as a relationship of pairs of linguistic expressions, which may be labeled synonymy; see Quine (1961, pp. 11-12, 22, and 48). I also think that my move is practically identical to the one suggested by George (1983), although I am not sure because his argument is unnecessarily complicated. Working in the context of Bolzano's logical theory, George argues that for Bolzano the form of an argument is a set of arguments generable from it, and that this avoids Massey's asymmetry because "the fact that invalid forms can have valid arguments as elements is a matter of no significance, since the only form we look to in assessing validity or invalidity is that generated from the argument itself" (George, 1983, p. 321). Moreover George criticizes Massey by arguing that his examples are not well-formed arguments, since there is more to defining an argument than specifying a < {premise}, conclusion > pair, namely a third element amounting to "understanding what the argument is"; in short, Massey's examples are examples of ambiguous arguments from the point of view of Bolzano's theory. It should be mentioned that Massey (1987) takes George's criticism into account and ends up partially agreeing with him.

[23]   Another example would be: a homeowner is a person who owns a home; so a nonhomeowner is a person who owns something which is not a home.

the point made earlier, when we modified the pseudo-principle of illogical form into a workable rule for formal invalidity.

If we want to make a different point, perhaps we could say something like the following. In each of the two arguments one is moving from a biconditional premise to a biconditional conclusion, and in the process a particular term which appears in both clauses of the premise is replaced by a contradictory term in both clauses.

Let us see why this sort of consideration would not apply for the case of the valid instance of affirming the consequent discussed earlier. Oliver's example was:

(21)   "If something is red, everything is red.
        Everything is red.
        Therefore, something is red" (Oliver, 1967, p. 463).

Suppose we tried to invalidate this argument by advancing the following alleged counterexample:

(22)   if Ronald Reagan lives in San Diego, then he lives in California;
        Ronald Reagan lives in California;
        so, Ronald Reagan lives in San Diego.

These two arguments have many similarities, but also one crucial difference: in the original argument the conclusion is a special case of the second premise, but this is not so in the alleged counterexample. We may conclude that in order to have the same form, two arguments must share *all* relevant structural details, and not just some.[24]

More generally, we may say that two arguments have the same form iff whenever the first instantiates a given form so does the second and viceversa; that is, two arguments have the same form iff there is no argument form which one of them instantiates but the other does not. For example, in the two arguments just discussed, (21) and (22), the following form is instantiated by the first but not by the second:

(23)   if $(Ex)Rx$ then $(x)Rx$; $(x)Rx$; so, $(Ex)Rx$.

---

[24] Incidentally, one consequence of this would be that some instances of the same conditional argument forms would not really have the same form. For example, the following are traditionally regarded as having the same form because they both instantiate modus tollens: (A) if Richard Nixon lives in Florida then he lives on the East Coast; Richard Nixon does not live on the East Coast; so, Richard Nixon does not live in Florida; (B) if Richard Nixon lives in Florida then he does not live in New York; Richard Nixon lives in New York; so, Richard Nixon does not live in Florida. In general, each of the four types of conditional propositional argument would have four subtypes, depending on the occurrence or nonoccurrence of negations in the antecedent or the consequent. This corresponds to the way such arguments are experienced at the psychological level, as experimental psychologists have demonstrated (Evans, 1972a, 1972b, 1982, 1983a, 1983b, 1983c).

Similar remarks would apply to the above mentioned valid syllogisms which instantiate invalid forms. For example, argument (8) could not be invalidated by advancing the following alleged counterexample:

(24)   no man is a woman;
        no inanimate object is a woman;
        so, some inanimate object is a man.

Both this argument and the earlier one (8) instantiate many forms, but the following is instantiated only by the earlier one (8) and not by this one:

(25)   no A or not-A is a B;
        no C or not-C is B;
        so, some C or not-C is A or not-A.

However, this is still too formalist.[25] I believe that ultimately we should take more seriously the suggestion implicit in the label which refers to analogy. That is, ultimately this method of invalidation should be conceived as analogical reasoning about arguments,[26] that is, as a meta-argument which concludes that the given argument is invalid because the counterexample argument is invalid and the two arguments are analogous. Then the alleged analogy could be discussed in the usual ways, by examining the extent and nature of the similarities and the dissimilarities between the two arguments. Here too we would have a type of inductive reasoning about deductive or formal arguments, or to be more exact, inductive reasoning about the formal and the deductive evaluation of arguments.

8. Conclusion

The analysis in the previous section shows that, although the formal invalidation of arguments is not an easy matter, it is a task which can be justified to a greater or lesser degree. This was also found to be the case for formal validation, though we arrived at such a conclusion by approaching from the other side, that is by criticizing the thesis that formal validations were generally susceptible of rigorous demonstration. Therefore,

---

[25]   Further formalist developments could perhaps be articulated by adapting some of the ideas contained in Hitchcock (1994).

[26]   I owe this idea in part to Govier (1985), who comes close to saying just this, and to Copi (1986a) who very suggestively includes the technique of refutation by logical analogy in the chapter on analogical reasoning.

the two processes seem to be more similar than dissimilar, and the asymmetry alleged by Oliver and Massey evaporates.[27]

More importantly and more positively, we have seen that even the formal validation and invalidation of arguments involve what may be called informal argumentation. Formal validation depends on the determination of the validity of argument forms, which in turn depends on the discovery and invention of proofs, which is ultimately a process of informal argumentation; and formal validation also depends on the translation or interpretation of the original natural language sentences involved, which depends in turn on linguistic intuitions about the synonymy of various natural language sentences; and the latter is an inherently informal process of argumentation (*pace* Massey). Similarly, formal invalidation depends partly on the claim that a given argument form reflects all relevant logical details, which is best regarded as an inductive generalization; and formal invalidation also depends on the comparison between the argument in question and a proposed counterexample argument, and such comparison is best regarded as an instance of analogical reasoning, with all the advantages and the pitfalls pertaining thereto.

Further informal-logic and argumentation-theory implications stem from the case-study aspect of our investigation. That is, once one reconstructs Oliver's account as an argument trying to show the formal invalidity of justifications of typical invalidity-verdicts, Massey's own account becomes primarily a second argument using Oliver's main conclusion as a premise to arrive at the further conclusion of asymmetry, rather than a positive evaluation of Oliver's original argument accompanied by a supporting justification. Moreover, I presented two criticisms of the Oliver-Massey argument by adapting and adding to the objections raised by Bencivenga, McKay, and George; and these criticisms were negative evaluations of Oliver's main conclusion about formal invalidation and of Massey's conclusion about formal validation. Now, it is perhaps debatable whether or not these criticisms were directed at the arguments rather than at the conclusions, but it is clear that the negative evaluations were not simple invalidity-verdicts, if at all; for the criticisms were that the formal validation of arguments is a less formal affair than the Massey asymmetry thesis claims, and that the formal in-validation of arguments is a less hopeless affair than Oliver's thesis claims. All of this is, I believe, typical of philosophical argumentation, which goes to show that the latter is not significantly governed by formal-logical consideration even when the topic involves concepts of formal logic.

Of course, this general positive conclusion ought to come as no surprise to the argumentation theorist and informal logician; the most it may do is to give explicit articulation to their basic intuitions. However, from a pedagogical point of view, a perplexing problem emerges from our exercise of evaluating the Oliver-Massey argument. Given, as we have seen, that the actual arguments on which one bases

---

[27] Woods (1995) also argues for symmetry, but does so in a context somewhat different from the present one; his paper will deserve further serious reflection in the future.

verdicts of formal validity or invalidity are typically formally invalid, is it proper to pretend to teach students how to argue formally about the subject matter that makes up the usual examples of formal-logic textbooks? Would it not be better to begin arguing formally about such concepts as validity, invalidity, argument form, and the like? I believe this can be done only by completely mathematizing the subject and teaching formal logic purely as a branch of mathematics. But there is another way, which turns out to be another way to teach logic and argumentation. It is to teach about such concepts as formal validity and invalidity, argument form, and so on, by emphasizing the nonformal, informal, and inductive considerations discussed above. To do this would be to teach the informal logic of formal logic,[28] as it were; to teach informal logic by using formal logic as substantive content. But then the question arises whether or in what contexts such a substantive content is appropriate. How does it compare with current events, advertising copy, newspaper editorials and reports, and magazine articles?[29] How does it compare with the arguments contained in classic texts in the history of thought?[30] These are not meant to be merely "rhetorical" questions, but their resolution is beyond the scope of this paper.[31]

[28] This idea may be taken as an inversion of, or complement to, the idea of "the necessity of formalization in informal logic" (Woods 1989); on this issue, cf. also Freeman (1994) and MacPherson (1995). However, my idea is in accordance with a project in which I have also explored the informal logic of science (Finocchiaro 1988a), and it stems from my belief that is is generally instructive to explore those aspects of a given phenomenon which seem to be antithetical to the way it appears on the surface (see Finocchiaro 1988b).

[29] In regard to this material (favored in recent elementary informal-logic textbooks), I am inclined to share the reservations expressed by Woods (1989).

[30] This is the material I would tend to favor, in regard to which see Finocchiaro (1987; 1989a; 1991; 1993; and 1994); it also corresponds to the trend exemplified in Fisher (1988), Fogelin (1987), and Garver (1985).

[31] A shorter version of this paper was presented at the Third International Symposium on Informal Logic, University of Windsor, Windsor, Ontario, Canada, June 1989. Another version, largely equivalent to the present one, was published in a volume stemming from that symposium (Johnson and Blair 1994).

References

Bailin, S. (1988). *Achieving Extraordinary Ends*. Dordrecht: Kluwer.

Bencivenga, E. (1979). 'Good and Bad Arguments.' *Journal of Philosophical Logic* 8:247-59.

Copi, I. M. (1968). *Introduction to Logic*. 3rd ed. New York: Macmillan.

Copi, I. M. (1986a). *Informal Logic*. New York: Macmillan.

Copi, I. M. (1986b). *Introduction to Logic*. 7th ed. New York: Macmillan.

Davidson, D. (1968). 'The Logical Form of Action Sentences.' In: N. Rescher (ed.), *The Logic of Decision and Action*. Pittsburgh: University of Pittsburgh Press.

Doss, S. (1985). 'Three Steps Toward a Theory of Informal Logic.' *Informal Logic* 7:127-35.

Evans, J. St. B. T. (1972a). 'Interpretation and 'Matching Bias' in a Reasoning Task.' *Quarterly Journal of Experimental Psychology* 24:193-99.

Evans, J. B. T. (1972b). 'Reasoning with Negatives.' *British Journal of Psychology* 63:213-19.

Evans, J. St. B. T. (1982). *The Psychology of Deductive Reasoning*. London: Routledge.

Evans, J. St. B. T. (1983a). 'Introduction' to *Thinking and Reasoning*, ed. J. St. B. T. Evans, London: Routledge, pp. 1-15.

Evans, J. St. B. T. (1983b). 'Selective Processes in Reasoning.' In: J. St. B. T. Evans (ed.), *Thinking and Reasoning*. London: Routledge, pp. 135-63.

Evans, J. St. B. T. (ed.) (1983c). *Thinking and Reasoning: Psychological Approaches*. London: Routledge.

Finocchiaro, M. A. (1980). *Galileo and the Art of Reasoning: Rhetorical Foundations of Logic and Scientific Method*. Dordrecht: Kluwer.

Finocchiaro, M. A. (1981). 'Fallacies and the Evaluation of Reasoning.' *American Philosophical Quarterly* 18:13-22.

Finocchiaro, M. A. (1984). 'Informal Logic and the Theory of Reasoning.' *Informal Logic* 6(2):3-8.

Finocchiaro, M. A. (1987). 'An Historical Approach to the Study of Argumentation.' In: F.H. van Eemeren e.a. (eds.), *Argumentation: Across the Lines of Discipline: Proceedings of the Conference on Argumentation 1986*. Dordrecht: Foris Publications, pp. 81-91.

Finocchiaro, M. A. (1988a). 'Empiricism, Judgment, and Argument: Toward an Informal Logic of Science.' *Argumentation* 2:313-35.

Finocchiaro, M. A. (1988b). *Gramsci and the History of Dialectical Thought*. Cambridge: Cambridge University Press.

Finocchiaro, M. A. (1989a). 'Methodological Problems in Empirical Logic.' *Communication and Cognition* 22:313-35. [Reprinted in: E.M. Barth, J. Vandormael, & F. Vandamme (eds.), *From an Empirical Point of View: The Empirical Turn in Logic*. Ghent, Belgium: Communication & Cognition, 1992, pp. 21-45.]

Finocchiaro, M. A. (1989b). 'Siegel on Critical Thinking.' *Philosophy of the Social Sciences* 19:483-92.

Finocchiaro, M. A. (1990). 'Critical Thinking and Thinking Critically: Response to Siegel.' *Philosophy of the Social Sciences* 20:462-66.

Finocchiaro, M. A. (1991). 'Induction and Intuition in the Normative Study of Reasoning.' In: E. Eells and T. Maruszzewski (eds.), *Probability and Rationality: Studies on L. Jonathan Cohen's Philosophy of Science*. pp. 81-95. Amsterdam (Holland) and Atlanta (GA): Rodopi. (*Poznan Studies in the Philosophy of the Sciences and the Humanities*, vol. 21).

Finocchiaro, M. A. (1992). 'Asymmetries in Argumentation and Evaluation.' In: F. H. van Eemeren e.a. (eds.), *Argumentation Illuminated*. Amsterdam: SICSAT, pp. 62-72.

Finocchiaro, M. A. (1994). 'Two Empirical Approaches to the Study of Reasoning.' *Informal Logic* 16:1-21.

Fisher, A. (1988. *The Logic of Real Arguments*. Cambridge: Cambridge University Press.

Fogelin, R. J. (1987). *Understanding Arguments: An Introduction to Informal Logic*. 3rd ed. San Diego, CA: Harcourt, Brace, Jovanovich.

Freeman, J. B. (1994). 'The Place of Informal Logic in Logic.' In: R. H. Johnson and J. A. Blair (eds.), *New essays in Informal Logic.* Windsor, Ontario: Informal Logic, pp. 36-49.

Garver, E. (1985). 'Teaching Critical Thinking as a Discipline.' *Informal Logic* 7:166-68.

George, R. (1983). 'A Postscript on Fallacies.' *Journal of Philosophical Logic* 12:319-25.

Govier, T. (1985). 'Logical Analogies.' *Informal Logic* 7:27-33.

Govier, T. (1987). *Problems in Argument Analysis and Evaluation.* Dordrecht: Foris.

Hitchcock, D. (1994). 'Validity in Conductive Arguments.' In: R. H. Johnson and J. A. Blair (eds.), *New essays in Informal Logic.* Windsor, Ontario: Informal Logic, pp. 58-66.

Johnson, R. H. (1989). 'Massey on Fallacy and Informal Logic: A Reply.' *Synthese* 80:407-26.

Johnson, R. H., and Blair, J. A. (1985). 'Informal Logic: The Past Five Years 1978-1983.' *American Philosophical Quarterly* 22:181-96.

Kalish, D., Montague, R., and Mar, G. (1980). *Logic: Techniques of Formal Reasoning.* 2nd ed. New York: Harcourt Brace Jovanovich.

Krabbe, E. C. W. (1995). 'Can We Ever Pin One Down to a Formal Fallacy?' In: F. H. van Eemeren e.a. (eds.), *Proceedings of the Third ISSA Conference on Argumentation,* vol. II, *Analysis and Evaluation.* Amsterdam: SICSAT, pp. 333-44.

Lakoff, G. (1970). 'Linguistics and Natural Logic.' *Synthese* 22:151-271.

Leonard, H., and Goodman, N. (1940). 'The Calculus of Individuals and Its Uses.' *Journal of Symbolic Logic* 5:45-56.

Massey, G. J. (1970). *Understanding Symbolic Logic.* New York: Harper.

Massey, G. J. (1975a). 'Are There Good Arguments That Bad Arguments Are Bad?' *Philosophy in Context* 4:61-77.

Massey, G. J. (1975b). 'In Defense of the Asymmetry.' *Philosophy in Context* 4(Supplement):44-56.

Massey, G. J. (1976). 'Tom, Dick, and Harry, and All the King's Men.' *American Philosophical Quarterly* 13:89-107.

Massey, G. J. (1980). 'Logic and Linguistics.' In: E. Agazzi (ed.), *Modern logic -- A survey.* Dordrecht: Reidel, pp. 311-29.

Massey, G. J. (1981a). 'The Fallacy Behind Fallacies.' *Midwest Studies in Philosophy* 6:489-500.

Massey, G. J. (1981b). 'The Pedagogy of Logic.' *Teaching Philosophy* 4:303-36.

Massey, G. J. (1987). 'Asymmetry, Fallacy, and Indeterminacy.' Paper presented at the Symposium on 'Informal Logic: Asymmetry and Fallacy,' American Philosophical Association, Pacific Division, San Francisco.

McKay, T. J. (1984). 'On Showing Invalidity.' *Canadian Journal of Philosophy* 14:97-101.

McKay, T. J. (1987). 'Shifting the Burden of Proof.' Paper presented at the Symposium on 'Informal Logic: Asymmetry and Fallacy,' American Philosophical Association, Pacific Division, San Francisco.

MacPherson, B. (1995). 'Formalizing Informal Logic.' In: F. H. van Eemeren e.a. (eds.), *Proceedings of the Third ISSA Conference on Argumentation,* vol. II, *Analysis and Evaluation.* Amsterdam: SICSAT, pp. 366-73.

Oliver, J. W. (1967). 'Formal Fallacies and Other Invalid Arguments.' *Mind* 76:463-78.

Quine, W. V. O. (1961). *From a Logical Point of View.* 2nd ed. New York: Harper.

Salmon, W. C. (1984). *Logic.* 3rd ed. Englewood Cliffs, NJ: Prentice-Hall.

Scriven, M. (1989). 'The Philosophy of Ordinary Logic.' Paper presented at the Third International Symposium on Informal Logic, University of Windsor, Canada.

Willard, C. A. (1983). *Argumentation and the Social Grounds of Knowledge.* University, Alabama: University of Alabama Press.

Woods, J. (1989). 'The Necessity of Formalization in Informal Logic.' *Argumentation* 3:149-67.

Woods, J. (1995). 'Fearful Symmetry.' In: H. H. Hansen and R. C. Pinto (eds.), *Fallacies: Classical and Contemporary Readings.* University Park, PA: Pennsylvania State University Press, pp. 181-93.

The relation of argument to inference [1]

Robert C. Pinto,
University of Windsor

Abstract

This paper attempts to clarify the relationships that hold between 1) arguments and inferences, 2) the normative study of arguments and the normative study of inference, 3) logic as the normative study of inference and the study of argumentation. It aims to provide an alternative to (a) standard formal logic textbook accounts of reasoning or inference, (b) the pragma-dialectical theories of van Eemeren and Grootendorst and (c) the views presented by Doug Walton in his paper "What is logic? What is reasoning?". More particularly, it argues that there is a normative study of inference that does not coincide either with formal logic or with the study of argumentation, but which must be presupposed by the study of argumentation.

I'm going to use the word *inference* for the mental act or event in which a person draws a conclusion from premises, or arrives at a conclusion on the basis of the consideration of a body of evidence. I'm going to use the word *argument* for a set of statements or propositions that one person offers to another in the attempt to induce that other person to accept some conclusion. And I'm going to use the word *argumentation* for an interactive social process involving two or more people, in which the principal goal is to induce belief or agreement through the presentation of arguments. [2]

My concern in this paper is to clarify the relationships that hold between

1) arguments and inferences
2) the normative study of arguments and the normative study of inference
3) logic as the normative study of inference and the study of argumentation.

My aim is to develop an account of these matters that will provide an alternative to (a) standard formal logic textbook accounts of reasoning or inference, (b) the pragma-dialectical theories of van Eemeren and Grootendorst and (c) the views presented by Doug Walton in his paper "What is logic? What is reasoning?" (Walton 1990). More

---

particularly, my thesis is that there is a normative study of inference that does not coincide either with formal logic or with the study of argumentation, but which must be presupposed by the study of argumentation.

1. Arguments and inferences

Books on logic--both formal and informal logic--typically claim to deal with *thinking* or *reasoning*. Yet even the most cursory examination reveals that they talk mostly about what they call *arguments*. There is a more or less standard rationale for this practice. Reasoning is identified or equated with *inference* (with basing a conclusion on premisses) and it is claimed that "[c]orresponding to every possible inference is an *argument*, and it is with arguments that logic is chiefly concerned."[3] The identification and, what is more important, the appraisal of reasoning then becomes a matter of formulating and appraising the arguments that correspond to inferences.

Reasoning and inference
One ground for dissenting from this picture is to object to the identification of reasoning with inference. Ralph Johnson, for example, has balked at that identification:

> If reasoning is anything at all, it would seem to be the seeking, the having and the giving of reasons. (Johnson 1991: 4)

In line with this idea, Johnson classifies explaining, predicting, asserting, arguing, defining, and clarifying--and not just inferring--as instances or species of reasoning. Though I sympathise with Johnson's reservations on this point,[4] I am not going to follow up on them here. There is a long tradition, going back at least to the middle ages, in which words that could be translated into English as 'reasoning' are used to mean the act of drawing a conclusion. In classical phases of that tradition, reasoning was one of three "operations" of the intellect (Aquinas[5]) or four "actions of thinking"

---

[1]   The words are from Copi and Cohen 1990: 6.

[4]   I don't quite share Johnson's view of the relation of inference to reasoning. Johnson thinks of inference as one of many *species* of reasoning. I find it more perspicuous to see reasoning as a process in the course of which a variety of different sorts of acts or events occur—posing questions, propounding hypotheses, analysing concepts and meanings, etc. Inference is one of the kinds of act or event that occur in the course of the process of reasoning, but not the only one. On my view, the relation of inference to reasoning is a relation of part to whole, not a relation of species to genus.

[5]   See Aquinas, *Summa Theologica*, I, 85,5, for example. Aquinas uses the word *ratiocinare*.

(Arnauld and Nicole in the *Port-Royal Logic*[6]). The three operations Aquinas recognised were:

1. Simple apprehension (or conceiving of the essence of something),
2. Judging (affirming or denying one thing of another)
3. Reasoning (drawing a conclusion syllogistically).

To these Arnauld and Nicole added a fourth: ordering (ordonner), by which they meant arranging the ideas, judgements and reasonings one has about a certain subject in "the manner most proper for making that subject known" (Arnauld and Nicole 1965: 38). Something like Aquinas' theory survives even in recent lexicography: recent editions of the Merriam-Webster collegiate dictionary still explain the shared meaning element of 'think,' 'cogitate,' 'reflect,' 'reason,' 'speculate,' 'deliberate' in the following way: "to use one's powers of conception, judgement or inference" (*Webster's Ninth New Collegiate Dictionary* 1991: 1226).

Inference and argument
Since Frege's attack on what he called "psychologism," it has become standard to distinguish clearly between the psychological study of thought processes and the logical study of reasoning, and to portray the subject matter of logic as something other than mental events and processes. In its historical context, Frege's attack on psychologism was a response to late 19th century neo-Kantian attempts to reduce logic to empirical psychology, attempts that were seen as threatening to deprive logic of its status as a *normative* theory of reasoning.[7] A crucial element in Frege's strategy was to insist that the subject matter of logic is not acts of thinking, ideas, or anything subjective.[8] Logic is to be a "science of truth" (Frege 1977: 1-2), the "thoughts" (*Gedanken*) it deals with are things "for which the question of truth can arise" and are "senses of sentences" (p. 4). They are "neither things in the external world nor ideas," but belong to a "third realm," are independent of what anyone thinks and, if true, are timelessly true (p .17). When a thinker grasps a thought, it does not belong to the contents of the thinker's consciousness, but there must be something in consciousness that "aims at" the thought (p.25-26). In retrospect, it makes sense to construe Fregean "thoughts" as *propositions*, where propositions are expressed by declarative sentences and constitute the *content*

---

[6]   See Arnauld and Nicole 1965: 37-38. Arnauld and Nicole use the word *raisonner*: "On appelle *raisonner* l'action de notre esprit, par laquelle il forme un jugement de plusieurs autres; comme lorsqu'ayant jugé que la véritable vertu doit être rapportée à Dieu, & que la vertu des payens ne lui étoit pas rapportée, il en conclut que la vertu des payens n'étoit pas une véritable vertu."

[7]   In the context, at least, of the assumption that empirical sciences can discover only what is, and not what ought to be.

[8]   Though it is not *about* mental acts, it has a bearing on them: "From the laws of truth there follow prescriptions about asserting, thinking, judging, inferring" (p. 1).

of both assertion and belief.[9] Logic then becomes the study of certain relations that hold between propositional contents, and in particular, the relations that justify moving from one propositional content to another.

A variant on the Fregean approach is found in Copi and Cohen. Early in their exposition of what logic is they introduce the notion of inference:

> *Inference* is a process by which one proposition is arrived at and affirmed on the basis of one or more other propositions accepted as the starting points of the process. (Copi and Cohen 1990: 5)

But, as we saw above, they are quick to point out that to every inference there corresponds an argument. And they tell us,

> An argument, in the logician's sense, is any group of propositions of which one is claimed to follow from the others, which are regarded as providing support or grounds for the truth of that one. (Copy and Cohen 1990: 6)[10]

Again, logic becomes the study of the relations that hold among the *propositions* (designated as premisses and conclusion) that make up arguments.

When Doug Walton develops his conception of reasoning, a conception which has its roots in the tradition described in the preceding section, he also focuses our attention not on psychological processes, but on the relations among abstract propositional entities. His "first pass" at a definition of reasoning sounds like Copi and Cohen's account of inference:

> *Reasoning* is the making or granting of assumptions called *premisses* (starting points) and the process of moving towards conclusions (end points) from these assumptions by means of warrants. (Walton 1990: 403)

But Walton is quick to make it clear that he doesn't want to identify reasoning with the *psychological* process or act of drawing an inference. He says rather that he defines reasoning "as a kind of abstract structure" (Walton 1990: 401). Pointing out that reasoning can be studied from a psychological or a logical point of view, Walton offers the following definition, which seems to be the one he takes most seriously:

---

[9] This departs slightly from Frege's usage, in that he was anxious to deny that *Gedanken* comprised the content of any consciousness--in his view that threatened to render them objectionably subjective. Nevertheless, even for Frege, Gedanken were *what* is asserted and assented to. And in current philosophical parlance, it has become standard to call *what is asserted* or believed the content (or propositional content) of the belief or assertion.

[10] As a matter of fact, Copi and Cohen are quite clear on the point that an argument in their sense is not simply a set (or "mere collection") of propositions; it has a "structure" by virtue of which one of the members is a *conclusion*, and other are *premisses*. In this respect it very much resembles what Walton will define as "reasoning from the logical point of view."

From the logical, as opposed to the psychological point of view, *reasoning* can be defined generally as a sequence of steps from some points (premisses) to other points (conclusions). (Walton 1990: 404.)

Walton wants the word 'reasoning' to stand for an abstract structure of propositions (a sequence in the mathematical sense of an ordered n-tuple), a structure that can be used or instantiated in various contexts. As we shall see, Walton's conception of logic is broader than the standard post-Fregean conception, since he wants to include the *pragmatics* of reasoning as a part of logic. For him the pragmatics of reasoning is the study of the use of abstract propositional structures in contexts of discourse, and most especially in contexts of persuasion and of dialogue.

Walton, in a deliberate departure from Copi, defines argument as follows:

Argument is a social and verbal means of trying to resolve, or at least to contend with, a conflict or difference that has arisen or exists between two or more parties.... (Walton 1990: 411)

Walton treats *argument* as a "framework of use" in which reasoning can occur (Walton 1990: 411).[11] The important point here is that for Walton, in contradistinction to Copi and Cohen, the consideration of *argument* belongs to the *pragmatics* of reasoning.

Walton's conception of argument overlaps with the pragma-dialectical account (PDA) offered by van Eemeren and Grootendorst. At the heart of PDA is an analysis of the presentation of arguments in terms of speech act theory. Presenting an argument is an illocutionary act complex in which propositions are asserted in order to convince someone to adopt a given standpoint toward an expressed opinion. Van Eemeren and Grootendorst say, for example,

The essential condition for the illocutionary act complex argumentation is different for pro-argumentation and contra-argumentation [i.e. presenting an argument for the adoption of an opinion and presenting an argument for the rejection of an opinion]. For pro-argumentation it may be formulated thus:

Advancing the constellation of statements S_1, S_2 (......S_n) counts as an attempt by S to justify O to L's satisfaction, i.e., to convince L of the acceptability of O. (van Eemeren and Grootendorst 1984: 43)

There is an echo here of Copi-type arguments and Walton-type reasonings, in that the propositional contents of the statements advanced and of the opinion argued for correspond to the elements of the abstract structures Copi and Walton speak of. But

---

11   And the framework of argument may or may not occur in broader framework of dialogue.

van Eemeren and Grootendorst make clear, and rightly so I believe, that *advancing* statements for the purpose of *convincing or persuading* is what makes for argument. What is it to convince or persuade?

> What we understand by *convince* is: to use pro-argumentation to induce a listener to accept an expressed opinion, or to use contra-argumentation to induce a listener to reject an expressed opinion. (van Eemeren and Grootendorst 1984: 48)

Even van Eemeren and Grootendorst, in their own way, want to avoid falling into the pit of the psychological. The acceptance or rejection of an expressed opinion at which argumentation aims is not intended to be a psychological phenomenon:

> It [acceptance] amounts to no more nor less than *agreeing* to the point of view defended in the argumentation. Thus our term *accept* has a lesser extension than the expression "be convinced" may have in colloquial idiom, and it is free of any psychological (and philosophical) connotations." (van Eemeren and Grootendorst 1984: 69)

Arguments as invitations to inference
I want to side with Walton and with van Eemeren and Grootendorst on the question of whether arguments ought to be conceived simply as abstract structures of propositions or whether the use of certain linguistic materials in contexts of persuasion ought to be considered a necessary condition of the existence or occurrence of an argument. The word "argument," I want to hold, is appropriately applied to sequences of propositions only when they serve as instruments of persuasion.

But I want to differ with Walton and van Eemeren and Grootendorst, in a slight but I think significant way, on the question of what it is that arguments, as instruments of persuasion, aim at. I want to say that the typical goal of an argument is to effect an *inference* in the person to whom it's addressed (and not simply to effect acceptance of its conclusion). Notice that both arguments and inferences have premises and have conclusions. That is surely not coincidence. This commonality becomes intelligible if we view the premises that are put forward by the arguer as intended to elicit assent to the argument's conclusion by forming the basis of an inference drawn by the person to whom the argument is addressed. Indeed, imagine a situation in which the presentation of an argument *caused* assent to its conclusion but in which the addressee did not make an inference from the argument's premises to its conclusion. For example, the argument is actually too complicated for the addressee to follow, but worn down by its length and caught up by the arguer's charm, the addressee's resistance to the conclusion disappears. Would we count this as a case in which the addressee was *persuaded by the argument* to accept its conclusion? Caused, yes. But not, I maintain, persuaded.

If I am right, then, arguments are invitations to inference.[12] Arguments succeed when the persons to whom they are addressed accept their conclusions *on the basis of* their premisses. Arguments fail when the addressee either refuses to accept their premisses, or accepting their premisses does not draw the intended conclusion from those premisses.

Logic, formal logic and argumentation theory

When we appraise arguments, we can do so from several points of view. If I am right that arguments are invitations to inference, an important evaluative question will always be: *ought the addressee to make the inference* which the argument invites? And that will be quite a different question from: *ought the arguer to have offered this particular argument* to this particular audience. Moreover, this latter question can be considered from variety of different points of view. Was the argument likely to be effective? Was it morally right to offer such an argument? Was it in the arguer's interest to offer that argument at that time? Was it a good argument to offer from the point of view of advancing the goals of negotiation, or of critical dialogue, or of pedagogical dialogue? And so on.

Logical appraisal of an argument deals, I would suggest, with the issues raised by the question of whether the *inference* invited by an argument is an inference that ought to be made--and, more particularly, ought to be made by the person to who it is addressed.[13] If logic be the study which elaborates the concepts, categories, and principles requisite for the appraisal of inference, then the theory of argumentation will be different from logic and will in some sense presuppose it.

At the same time that logic--as the elaboration of the concepts, categories, and principles requisite for the appraisal of inference--is distinguishable from argumentation theory, logic does not coincide with formal logic either. By formal logic I mean both (a) classical modern logic (as instanced in the propositional calculus and quantification theory) and (b) alternate logics (e.g., nonmonotonic logics) which pattern themselves on classical modern logic.

---

[12] Scriven (1976: 55-56) reviews the "relationship between argument and inference." He says, "The *function* of the argument is to persuade you that since the premise is true, you must also accept the conclusion. The persuasion will be powerful if it is clear that the inference from the premise to the conclusion is sound, that the premiss does in fact imply the conclusion. (Other ways to put this are to say that you can legitimately infer the conclusion from the premise, or that the conclusion is in fact a consequence of the premise.)" Scriven's point isn't exactly the same as mine. But it's close.

[13] For purposes of this paper, I leave open the question of whether the *logical* appraisal of arguments ought to concern itself with the acceptability of premisses, as well as with the question of whether the premisses provide a suitable basis for drawing the conclusion that the argument invites. Those who take classical formal logic as a paradigm for logical appraisal typically view the evaluation of premisses as lying outside the scope of logic (see, for example, Copi and Cohen 1990: 53.) Informal logicians, on the other hand, frequently view the determination of the acceptability of premisses as an important part of the logical appraisal of arguments (see for example Johnson 1987). Though I happen to share the latter view, it is not essential to any of the points I try to make in this paper.

Classical modern logic investigates sets and sequences of propositions and, for an important subclass of propositions, has developed powerful techniques for ascertaining the presence of consistency, equivalence, entailment, etc.[14] The pertinent claim it can make about the relation of a set of premisses to a conclusion is that it entails the conclusion. But entailment is neither a necessary nor a sufficient condition for the premisses and conclusion of an argument or inference being suitably linked. Not sufficient, because an argument of the form "P, therefore P" meets the criterion of entailment but is hopeless as an argument.[15] Not necessary, because there are innumerable inductively strong arguments in which premisses do not entail conclusions. The abstract structures that classical logic studies just don't coincide with the factors that make *arguments* logically good.

I want to suggest, moreover, that what prevents classical logic from being a general theory of inference or reasoning may not lie simply in the fact that the only premiss-conclusion link that it considers is entailment. Basic to the classical conception is the assumption that the suitability of the relationship between premisses and conclusion can be appraised simply by examining the propositional content of the premisses and of the conclusion (and, in most versions of the classical conception, examining only the "logical form" of that propositional content). On such an assumption, consideration of relations that obtain between propositions *in abstraction from their occurrence in actual thinking or contexts of discourse* can yield insight into whether premisses and conclusion are suitably related. This approach to appraising the relationship between premisses and conclusion has, on the whole, worked splendidly for appraising mathematical and/or "deductive" inferences. But attempts to construct similar logics for other kinds of inference--inductive inference, conductive inference (see Wellman 1971), abduction or inference to the best explanation, as well as large stretches of what AI calls practical or everyday reasoning[16]--have yet to bear anything near final fruit. We do not know how successful we will be in developing models of such reasoning that resemble to a greater or lesser degree the structures developed by classical formal logic. But consider the possibility that in these types of reasoning *the content of the premisses is not by itself sufficient to warrant acceptance of the conclusion.*

---

[14] This logic succeeds only for a subclass of propositions because it lacks the resources to deal with semantic entailments.

[15] It can also happen that the premisses of an argument entail its conclusion, though no one is in a position to know this. E.g., if Golbach's conjecture is indeed derivable from the axioms of arithmetic, then those axioms entail Golbach's conjecture. But even if that is true, in the present state of mathematical knowledge an argument which consisted only of the axioms of arithmetic as premisses and Golbach's conjecture as conclusion would not be a good argument. And the problem with it would lie in the nature of the link between premisses and conclusion. See Pinto 1994, where these points are developed at greater length.

[16] And which is being explored in some of the newly developed nonmonotonic logics--especially default logics.

Writing about analogical reasoning, Evelyn Barker has said

> Like inductive reasoning generally, an inductive analogy is not based merely on its stated premises but on all our knowledge about the world. (Barker 1989: 187).

Suppose that something like this is true--that background information which motivates the move from premisses (or data) to conclusion, and which is essential to its warrant, cannot be rendered explicit because of its complexity and/or its character. Then it is hard to see how evaluative strategies patterned on classical modern logic could provide a pattern or paradigm on which to model the understanding of reasoning generally-- since those strategies locate the justification of a conclusion precisely in the propositional content of its premisses and the rules of inference which license the move from one propositional content to another.

2. Inference

If arguments are invitations to inference, what then is inference? And if an important dimension of the appraisal of arguments is appraisal of the inference that argument invites, on what is such appraisal to be based, if not on formal logic?

What is inference? A first try

In an inference, one belief or set of beliefs (called the premiss or premisses) "leads to" another belief (called the conclusion), which in some way or other is "based on" the premiss. What can we make of this? A tempting answer is that inference occurs when one belief *causes* another--that when I infer q from p, I believe q *because* I believe p. D. M Armstrong, for example, has attempted to elaborate a causal theory of inference (Armstrong 1968: 194-200.), and so in a way did C. S. Peirce before him.

But the mere fact that one of your beliefs causes or leads to some other belief does not mean that the second belief is inferred from or grounded in the first. One way of seeing this is to recall a puzzle from G. E. Moore (which Armstrong cites in developing his point). I want to go out, and my belief that it's raining causes me to search for my umbrella, as a result of which I find my umbrella in the hall closet and come to believe it's there. My belief that it's raining is one of the causes of my belief that the umbrella is in the closet. Yet we certainly wouldn't want to say that the first belief provides a premiss from which the second is inferred. Hence not every belief acquisition that's the result of a belief already held is an inference. How then to pick out the inferences from this broader group of belief acquisitions?

We might try to say that a person S infers q from p just in case S comes to believe q because S believes that p and also believes that the truth of p justifies the belief that q. Such an answer would require us to unpack the requisite sense of justification, and

might well turn out to be circular--if we could not explain what justification is except by reference to what correct or proper inference is.

Another, potentially more promising, way of explaining why Moore's example isn't a case of inference is to appeal to something like Peirce's notion of a "habit of mind" that leads us to "draw one inference rather than another."[17] In "The Fixation of Belief" Peirce writes:

> That which determines us, from given premises, to draw one inference rather than another is some habit of mind, whether it is constitutional or acquired....
> The particular habit of mind which governs this or that inference may be formulated in a proposition whose truth depends on the validity of the inferences which the habit determines; and such a formula is called a *guiding principle* of inference. Suppose, for example, that we observe that a rotating disk of copper quickly comes to rest when placed between the poles of a magnet, and we infer that this will happen with every disk of copper. The guiding principle is that what is true of one piece of copper is true of another. Such a guiding principle with respect to copper would be much safer than with regard to many other substances--brass, for example. (Peirce 1960: 5.227-228.)

Notice a couple of things about what Peirce says in this passage.

1)   Peirce is (consciously and deliberately, I think) refusing to take "What is true of one piece of copper is true of all" as an additional *premiss* in this inference. Rather, the guiding principle "formulates" the habit--lets us see *why* the conclusion is connected to the premiss.

2)   The guiding principle in this example is not a "law of logic." Qua "inference rule" it doesn't depend on the "logical form" of the propositions involved. It is what some authors call a "material inference rule"--an inference rule that is subject-matter dependent. As Peirce himself notes (Peirce 1960: 5.228), "almost any fact may serve as a guiding principle."

3)   Insofar as they are principles of inference that are subject-matter dependent, Peirce's "guiding principles" resemble the elements of arguments called *warrants* in Toulmin's theory of argument (see Toulmin 1958 and Toulmin et al. 1979).

---

[17]   Armstrong appeals to something resembling Peircean habits in Armstrong 1968: 198, but without mention of Peirce. See also Armstrong 1973.

One can see, I think, how the notion of a guiding principle or a habit of mind helps with Moore's puzzle. And one can also see how such an account of inference could provide a framework for the appraisal of inferences. Thus in "The Fixation of Belief" Peirce wrote:

> The object of reasoning is to find out, from the consideration of what we already know, something else which we do not know. Consequently, reasoning is good if it be such as to give a true conclusion from true premisses (Peirce 1960: 5.226).

And a few paragraphs after having introduced the notion of habits of mind and guiding principles, he says:

> The habit is good or otherwise, according as it produces true conclusions from true premisses or not; and an inference is regarded as valid or not, without reference to the truth or falsity of its conclusion specially, but according as the habit which determines it is such as to produce true conclusions in general or not. (Peirce 1960: 5.227-228)

Criticisms of the causal theory of inference
Now I think that there is something importantly right-headed in this theory of inference, but I don't think it will do in anything like the form in which it has been presented here.

First, four objections that I'll treat as relatively minor for purposes of this paper. (1) The upshot of inference can be, not the acquisition of new belief, but the anchoring of pre-existing belief (as when I find additional evidence or additional reasons for what I already believe). (2) On the causal account, the upshot of inference or reasoning is always *belief.* This is too narrow: reasoning or inference can lead, not to a firm belief that something is so, but to the suspicion that it is so, or to the retraction of the belief that it is so. Moreover, the upshot of inference can be, not the modification of a doxastic attitude toward a proposition, but a decision about what to do or the acquisition of a resolve to act in a certain way.[18] (3) A related but slightly different point: it seems natural and correct to speak of drawing inferences in the course of suppositional reasoning, but in such cases neither premisses nor conclusion are believed by the one who makes the inference or draws the suppositional conclusion. (4) Peirce says that the habit "determines us, from given premisses, to draw one inference rather than another." This can't be right as it stands. The explanation of why I draw the conclusion I do will surely have to include more than the fact that I believe the premisses together with the fact that I have such a "habit of mind." Many things which I currently believe *could* lead me to various conclusions in accord with my current

---

[18] See Pinto 1991 for analogous points concerning arguments.

habits of mind, but don't do so. Which conclusions I actually draw depends not only on the "guiding principles" that govern my thought, but also on my current interests, concerns, etc. The full story of the generation of belief by inference will have to be considerably more complex than the story that Peirce (or for that matter, Armstrong) has told so far.

These first four objections point out several respects in which the initial causal story is incomplete. The next two objections, I believe, pose even more serious problems for the initial story.

(5) Many of the most important inferences we make do not exhibit readily discernible conformity to the requisite patterns, rules or generalisations. The best work in the philosophy of science leaves little doubt that what are called simplicity and considerations of overall explanatory coherence play a crucial role in the inferences we make from a given body of actual or putative evidence. But no one has succeeded in reducing these factors to articulable pattern or to rule. At the point in our cognitive lives at which inference becomes most interesting and most fateful, doxastic transitions don't lend themselves to being understood along the lines set out in the initial causal story.

(6) All but one of the preceding objections try to show that the initial causal story does not capture a necessary condition of inference. One can also question whether that story captures anything like a sufficient condition of inference. Imagine that whenever Smith comes to believe another person to be Irish he will believe that person to be dull-witted, and that many of Smith's beliefs can be explained by reference to this tendency of his (e.g., his belief that O'Brien is dull-witted). But suppose that (a) when we ask Smith *why* he thinks O'Brien is dull-witted, he can't tell us and (b) getting Smith to acknowledge large numbers outstanding Irish literary figures, scientists, etc., has no effect at all on his tendency to assume, of individuals he knows to be Irish, that they are dull. Are we prepared to count these products of Smith's doxastic tendencies instances of inference? Are they instances of *reasoning*? Peirce himself has pointed out in other places that for inference (or reasoning) in the full sense to occur, it is *not* sufficient that the premisses cause the conclusion to be believed in accordance with some rule or principle. He insists that it is also required that the person making the inference *know* or *see* (or at least think) that an appropriate connection holds between premisses and conclusion. Thus Peirce (1960: 5.295) writes

> For this theory [pragmaticism] 'requires' that in reasoning we should be conscious, not only of the conclusion, and of our deliberate approval of it, but also of its being the result of the premiss from which it does result, and furthermore that the inference is one of a possible class of inferences which conform to one guiding principle.

Perhaps we should go a step further, and require that if a doxastic transition is to count as inference (or a part of reasoning), it must occur in the context of a cognitive economy that meets certain minimum standards of self-consciousness and of rationality.

What is inference? A second try

Let us use the phrase *proto-inference* to refer to the phenomena described by our first causal theory of inference. And let's suppose the first four objections can be met by an account of the causes of belief and of other mental states that is considerably richer than the one initially envisaged. I want to suggest that we could meet the potentially weightier objections--(5) and (6)--if we develop and enhance the account of proto-inference in certain definite and intelligible ways. The result would be to make *liability to criticism* an essential component of the very concept of inference. Imagine then a development that proceeded through the following six steps.

1) We start out with the concept of proto-inference--a postulated, ill-understood causal transition from belief in premises to belief in a conclusion, dependent "somehow" on the presence of a recognised pattern that embraces the premises and the conclusion.

2) Then we introduce the idea that such inferences are good or bad depending on whether the patterns are truth-preserving; or even better (and taking our cue from Peirce), that they are strong or weak to the extent that the patterns are truth-preserving. By patterns here we mean both logical form and material principles of inference that could function as Peircean guiding principles, Toulmin warrants, etc.

3) Next we teach those who make proto-inferences to identify (actual or potential) proto-inferences as good or bad by reference to such patterns, and to resist the proto-inferences that are bad. As a consequence of doing this, we have introduced a practice of criticism.

4) We then re-conceive inference as something which is, as such, subject to the critical practice introduced in step 3. This reconceptualisation involves more than the bare notion that inferences can be graded as good or bad (strong or weak). It involves, in addition, the idea that belief transitions not open to the influence of critical reflection are not inferences in the full sense of the word.[19]

5) Critical practice, as introduced in step 3, was a matter of discerning the pattern which the inference exemplifies and judging the extent to which that pattern is truth preserving. But 20th century epistemology--and in particular, 20th century philosophy of science--has made us aware that the goodness of many

---

[19] This idea may already present in Peirce's idea, noted above, that for reasoning proper to occur, the one who draws a conclusion must be aware of the guiding principle in terms of which he draws it. It is certainly present in Toulmin's idea that warrants require backing.

of our most fateful and highly prized inferences do not yield to any simple analysis in terms of pattern or guiding principle. And yet the value of those inferences is not something that is just arbitrarily accepted; rather it is something open to discussion and rational evaluation. We move, therefore, to a broadened conception of criticism, one not tied quite so closely to logical rules or material principles of inference, but modelled in part on the discussions of the probative value of evidence that occur in contexts where articulable rules are not available.

6)   Finally, we reconceptualise inference again as belief transition open to the influence of critical reflection in this broadened sense.

The upshot would be a conception of inference not subject to Moore's puzzle and not vulnerable to objections (5) and (6) in the preceding section. Moreover, this conception of inference already contains the idea of a critical evaluation of belief-transition that, I maintain, does not coincide with formal logic.

3. Logical pragmatics, argumentation theory, and the evaluation of inference

The study of argumentation that has developed over the last 15 years has made a major contribution to resuscitating and refurbishing aspects of the study of inference and argument that had either been forgotten or fallen into questionable repute. There can be no question but that the elaboration of empirical and normative models of critical dialogue by the pragma-dialectical school, and the reconsideration of the informal fallacies in the context of dialectic and argumentative dialogue by Walton, Woods, Krabbe and others, have deepened our understanding of argument and argumentation and have often resulted in analyses of the fallacies that are more rigorous and more subtle than anything we've seen before.

Despite my respect and indeed my enthusiasm for these developments, some of their proponents seem to me to divide things up in a way that threatens to lose sight of an essential ingredient of the phenomena they want to deal with. That ingredient is precisely the critical evaluation of inference that I alluded to in the preceding section.

The achievement of the pragma-dialectical school is the elaboration of a normative model of argumentative discourse, whose main components are an account of the phases of such discourse, of the speech acts that are appropriate to each of the phases, and most importantly a set of rules--essentially procedural--designed to maximise the possibility that such discourse achieves its goals. But the procedural rules which comprise the model presuppose and make explicit reference to *logical rules* that are presumed to be already at hand. Van Eemeren and Grootendorst state that the parties to the discourse "must have logical rules which they can apply in order to evaluate the validity of the protagonist's argument" (1984: 169). The 1984 presentation is of a decidedly deductivist cast. The presentation in van Eemeren and Grootendorst 1992 is more supple and not so unremittingly deductivist. Nevertheless, even there the sup-

position seems to be that a set of straightforward procedures for determining validity are already available and can be presupposed by the procedural rules which comprise the model.

My countersuggestion is that there needs to be a critical practice and a set of techniques for evaluating the inferences that don't fall under any articulable inference rule, and that this practice and these techniques cannot be defined or captured by any set of procedural rules either. For example, where simplicity emerges as a salient criterion of theory choice, one can rationally consider (monolectically as well as dialectically) which of two theories is simpler, but no rule or algorithmic procedure will settle the question.

Analogously to van Eemeren and Grootendorst, Doug Walton has written:

> Formal logic has to do with forms of argument (syntax) and truth values (semantics). At any rate, that is the traditional conception. Informal logic (or more broadly, argumentation, as a field) has to do with the uses of argumentation schemes in a context of dialogue, an essentially pragmatic undertaking. (Walton 1990: 417-418)

It is not as clear in Walton's case that he views all pragmatic considerations as essentially procedural in nature. However, to illustrate what he says in the paragraph just quoted, Walton writes:

> Only recently has it become more apparent that a pragmatic approach is absolutely necessary in order to make sense of informal fallacies. What are fallacies? They are violations of the rules of reasonable dialogue. But over and above this, they are also deceptive tactics used unfairly in argument to defeat an adversary in dialogue. If the study of fallacies is to be part of logic, clearly logic can make no headway in working toward its primary goal unless the pragmatic study of the uses of reasoning in argument (informal logic) is included as a legitimate part of the subject. (Walton 1990: 419.)

In this passage there is no suggestion that the *probative* force of the reasoning which occurs in argument is to be assessed by informal logic. It is easy to leave with the impression that for Walton, as for van Eemeren and Grootendorst, the strength with which premisses support a conclusion is to be judged by formal logic.[20]

If my characterisation of these authors' positions is correct, they are suggesting that to understand and appraise arguments and argumentation we need only (a) formal logic and (b) procedural rules for the use of arguments in the context of dialogue and/or

---

[20]   In much of his most recent writing, Walton allies himself even more closely with the pragma-dialectical approach to argument. See for example Walton 1992.

critical discussion. My countersuggestion is that, in light of the considerations adduced earlier, there is need for a logical study of the principles of inference that does not coincide with formal logic. Since there does not today exist a formal logic capable of functioning as a normative theory of inference in general, I believe that the burden of proof is on these authors to show that my countersuggestion is false.

References

Armstrong, D. M. (1968). *A Materialist View of the Mind*. London: Routledge and Kegan Paul.

Armstrong, D. M. (1973). *Belief, Truth and Knowledge*. London: Cambridge University Press.

Arnauld, A. and P. Nicole (1965). *La Logique, ou L'Art de Penser*. Édition critique par Pierre Clair et François Girbal. Paris: Presses Universitaires de France.

Barker, E.M. (1989). 'Beardsely's Theory of Analogy.' *Informal Logic* 9: 185-194.

Copi, I.M. and C. Cohen (1990). *Introduction to Logic*, 8th ed. New York: Macmillan.

Eemeren, F.H. van and R. Grootendorst(1984). *Speech Acts in Argumentative Discussions*. Dordecht: Foris.

Eemeren, F.H. van and R. Grootendorst (1992). *Argumentation, Communication and Fallacies*. Hillsdale NJ: Lawrence Erlbaum.

Frege, G. (1977). *Logical Investigations*, ed. by P. T. Geach. New Haven: Yale University Press.

Johnson, R. H. (1991). 'Walton's Inferentialist Concept of Reasoning.' A paper delivered at the meetings of the Canadian Philosophical Association, May 1991.

Johnson, R. H. and J. A. Blair (1983). *Logical Self-Defense*, Second Edition. Toronto: McGraw-Hill Ryerson.

Moore, G. E. (1962). *Commonplace Book*. London: Allen & Unwin.

Peirce, C.S. (1960). *Collected Papers of Charles Saunders Peirce*, ed. C. Hartshorne and P. Weiss. Cambridge: Harvard University Press.

Pinto, R.C. (1995). 'Logic, Epistemology and Argument Appraisal.' In: R. H. Johnson and J. A. Blair, *New Essays in Informal Logic*. Windsor, Ontario.: Informal Logic.

Pinto, R.C. (1991). 'Generalizing the Notion of Argument.' In: F. van Eemeren, R. Grootendorst, J.A. Blair and C.A. Willard (eds.) *Proceedings of the Second International Conference on Argumentation*, Amsterdam: SICSAT.

Scriven, M. (1976). *Reasoning*. New York: McGraw-Hill.

Toulmin, S. (1958). *The Uses of Argument*. Cambridge: Cambridge University Press.

Toulmin, S., R. Rieke and A. Janik (1979). *An Introduction to Reasoning*. New York: Macmillan.

Walton, D.N. (1990). 'What is Reasoning? What is an Argument?' *The Journal of Philosophy*, 1990, pp. 399-419.

Walton, D.N. (1992). *Slippery Slope Arguments*. Oxford: Oxford University Press.

Webster's Ninth New Collegiate Dictionary (1991). Markham, Ontario: Thomas Allen & Son.

Wellman, C. (1971). *Challenge and response: justification in ethics*. Carbondale : Southern Illinois University Press.

Semantic intuitions:
Conflict resolution in the formal sciences

John Woods,
University of Lethbridge

Abstract

Thanks to Russell's Paradox, there is no intuitive concept of set. Thanks to the Liar Paradox, there is no intuitive concept of truth. Since truth enters into the definition of semantic consequence, there is no intuitive concept of it either. The loss of intuitive concepts is ruinous for a certain conception of how philosophical analysis is to be done. In Moore's hands, philosophical analysis is the decomposition of complex concepts into simple, intuitive concepts which reveal the basic structure of reality as it really is. Judged from this Moorean perspective, it ought to be philosophically impossible to produce theories of sets or of truths or of consequence. In the case of sets, this was precisely Frege's position. Russell soldiered on in *Principles of Mathematics*, he pleaded a distinction between analysis by way of philosophical definitions and analysis by way of mathematical definitions. Russell's mathematical analysis of sets was an undeclared return to idealism (in what was supposed to be his first analytic book).

There is a moral to all this: The logical and semantic paradoxes push theorists, unannounced and often unaware, into idealism.

One of the attractions of the theory of argument in the present day is the care it takes with practical affairs. It is nothing but good that more or less sophisticated dialectical techniques now exist for the display and appraisal of arguments about things as important as income tax returns, auto repairs and the upkeep of public parks. If there is a dark side to such developments, it is the tendency of the practically-minded to overlook contentions that rage unresolved over strictly theoretical matters. In this chapter, I attempt modest redress by examining strategies of conflict resolution in the formal sciences. I shall attend to the disposition of disputants to invoke "intuitions" and "counterintuitions" in their efforts to resolve disagreements for which empirical considerations are unavailing if definable at all.

It is commonly supposed that Russell's Paradox destroyed the intuitive concept of set, and that the Liar Paradox destroyed the intuitive concept of truth. Since semantic consequence itself imbibes a truth property, there is no intuitive concept of

it either.[1] This makes problematic any attempt to adjudicate rival theories of sets, or truth, or consequence by appealing to intuitions and counterintuitions, never mind that this is precisely what the literature on such matters gives plenty of evidence of. This leaves us with two good questions. (1) What is the dialectical structure of conflict resolution in the formal sciences? (2) Leaving aside for now their apparent unavailability to disciplines that have lost their intuitive target concepts, what might we suppose theorists were about when they invoked, as they still do, intuitions and counterintuitions on either side of a non-empirical disagreement? For concreteness, and not withstanding the oddity just noted, we might keep in mind the noisy contentions that entailment-theorists throw themselves into over the status of *ex falso quodlibet*, which proclaims the coincidence of negation inconsistency and absolute inconsistency.[2] For those whose interests don't run to the consequence relation, there is no shortage of standing disagreements to keep in mind, as witness the abiding tension among argumentation theorists about whether fallacies are dominantly dialectical rather than dominantly logical in nature.

1. Formal Theories

We should remark the frequency with which it is claimed by people who don't like it that *ex falso* is counterintuitive. With some theorists, that is the whole case against *ex falso*. Perhaps it is always too slender a case, but the reverse mistake is equally bad. It is the mistake of ignoring the near indispensability of intuitions and counterintuitions in the non-empirical sciences, or, as I shall say hereafter, the *formal sciences.*[3]

In so saying, one of our questions recurs. For let T be any formal theory for which it is presumed provable that there are no intuitive concepts of its target properties. Then how, with respect to the adjudication of rival accounts of such properties, can appeals to intuitions and counterintuitions be in *any* sense "indispensable"?

It is obvious enough that we will go only a little way towards appreciating what counts as conflict resolution in the formal sciences until we have a serviceable

---

[1]   Russell said this about sets in the *Principles of Mathematics* (1903), xv-xvi, xviii and 2; Frege said it (or something like it) in his letter to Russell (reprinted in Jean van Heijenoort (ed.), *From Frege to Gödel* (1967), 127-128) and more forcefully in writings published after his death (*Posthumous Writings*) Hans Hermes, Friedrich Kambartel and Friedrich Kaulbach (eds.) (1979), 269- 270. The allied claim about truth Tarski forwards in *Logic, Semantics and Metamathematics*, 2nd edition, John Corcoran (ed.); In: Hackett (1983), 158, and, about consequence, in *op. cit.*, 429.

[2]   The coincidence derives from the fact that, if *ex falso* is true, a contradiction entails every statement whatever, which expressly provides for the entailment of apparently irrelevant statements.

[3]   In one way I greatly dislike this use of "formal". It suggests the deployment of logical forms. In my use of it here, the suggestion is cancelled. "Formal" is just lexical variation on "non-empirical".

understanding of "formal theories". Also required is a foray into the *dialectical epistemics* of conflict resolution. It is "dialectical" in as much as conflict resolution, minimally a two-party affair, is inherently a consensual arrangement whatever else it is. It is "epistemic" in as much as the conflicts whose resolution-strategies the theory specifies and pronounces upon are disagreements about what people say they know what they claim to be true, and whose resolution therefore may be expected to inherit something of this epistemic cachet. I shall mean by a formal theory T a set of claims K and a projective mechanism M which derives new sentences. Normally M's projective propensities will make it appropriate to speak of T as the closure of K under M. Where M is not a purely deductive mechanism, it is understood that some sort of deductive device is a proper submechanism of M. Thus M will force out of K the deductions appropriate to the kind of theory that T is. What makes T non-empirical is that the selection of K, as of M, is independent of both T's particular empirical design (if any) and of T's empirical consequences (if any). Independence, here, may be taken discretely or holistically. Discretely, K and M alike are proof against recalcitrant observations. Holistically, they are the last, or close to last, to submit to revision in the web of belief or in the laddered fabric of theory.

This is very rough, of course, and useful only as a loose expository convenience. Worse than that, it is circular, pending independent specification of "empirical designs" and "empirical consequences". But we can see what is meant, and for those who can't, here is an example from quantum physics. There are impeccable experimental results which suggest that the classical Distributivity Principle fails in the microdomain. In the physical theory, a vector space H (a Hilbert space) is correlated with every physical system X, and each basic or elementary physical statement is mapped to a subspace of H.[4] The mapping obeys the lattice-theoretic conditions

1. $X (p \vee q)$ = the *join* of spaces $X(p)$ and $X(q)$
2. $X (p \wedge q)$ = the *meet* of spaces $X(p)$ and $X(q)$
3. $X (\neg p)$ = the *orthocomplement* of $X(p)$.[5]

When H is two-dimensional, $\vee$ and $\neg$ obey Distributivity, which in skeletal form is:

$$\Phi \vee (\Psi \wedge \chi) \text{ iff } (\Phi \vee \Psi) \wedge (\Phi \vee \chi).$$

If H has a finite dimensionality greater than two, Distributivity fails.

---

[4] H is characterized as the set of total functions from the real numbers to the complex plane for which the appropriate relations--join, meet, etc.,--are specified.

[5] See Bas C. van Fraassen (1975), 577-607; especially 580-581.

As they are used in the physical theory, $\vee$, $\wedge$ and $\neg$ don't always obey Distributivity. This is (classically) fatal if join is classical disjunction, meet is classical conjunction, and orthocomplementation is classical negation. There is reason to resist their equations. The connections in question are carefully specified to reflect physical operations, and this allows us to think that meet is stronger than conjunction and that join is weaker than disjunction. The meet and join of quantum observations are, in part, a matter of what possible operations satisfy the von Neumann dynamics.[6]

Whatever is to be made of our present case, it is clear enough that if the inequivalence of meet and join with conjunction and disjunction is an arguable position, then it is also arguable that the Distribution Principle holds in M and is undisturbed by the experimental results that find their way into T. Distributivity could then be said to be part of the non-physical subtheory of quantum physics or, as it sometimes said, its pure logic, and this certainly is something like the distinction we are after.

Formal theories can be approached *abstractly* or *analytically*. In the first case, the specifications of K and M are independent of the theorist's belief that the K-elements are true. Sometimes they are chosen because they have properties which make them attractive things for the theorist to "play with".[7] Motivations for a theory can be

> numerous and complex, but one is surely logical *curiosity*. Suppose one takes claims about the underlying [Hegelian] dialectic (as opposed to classical logic) at face value[;] can a rigorous logical system be formed incorporating just these claims? (Compare Heyting's axiomatization of such a seemingly *esoteric* subject [as] intuitionistic logic).[8]

On the other hand, when a theory is approached analytically there is antecedent presumption of the theory-independent truth of K and the projective soundness of M, and for the view that jointly they constitute or portend an analysis of some target concept, such as *set* or *consequence*. There is good reason for thinking that Frege, and Russell too in the early days, took the axioms on sets to be a conceptual analysis in this sense, and that much of the contemporary disagreement about

---

6   See Peter Forrest (1988), 40-41.

7   For example: "Now that we all understand the virtues of a model-theoretic semantics satisfying general Montagovian standards of rigour and clarity, there is joy in *playing around* with virtually every specific detail of Montague's original paradigm. The following...illustrate various aspects of this new wave of *free speculation*". (Johan van Benthem (1983), 1-17; 1. Emphases added.)

8   J.F.A.K. van Benthem (1979), 333-347; 337. Emphasis added.

consequence or entailment is a disagreement about how the concept of entailment is to be analyzed.[9]

It should not be supposed that non-empirical ventures abstractly entered into are mere play. A good deal of pure research is highly and unapologetically abstract. Riemann's geometry was shown consistent if Euclid's is long before relativity theory furnished an empirically commanding model of it, and that alone effectively ended the *à priori* imperial pretensions of the old geometry.

Various disciplines are distinguished by their brisk and pacific pluralism, in which a hefty multiplicity of generally inequivalent and often incompatible theories and would-be theories co-exist at levels of discord and cross-examination that are noticeably not high. Pluralism of this pacific and tolerant kind is a valuable type-indicator of its constituent incompatible theories. But pluralism is not always so gentle a thing; it is sometimes a matter of high vexation and spiky noise. This, too, is an indictor of theory-type--analytic theories for noisy pluralism and abstract theories for gentle. Paragraphs ago, we wondered how intuitions could in any way be indispensable indicators of properties of which demonstrably there is no intuitive conception. There is a related question about what might be called *analytica pluralism*. It is, "How is analytical pluralism possible?" Ambiguation is very much the received answer. If T and T' are incompatible analytical theories of a target concept C, then T and T' marshall incompatible C-intuitions, each "indispensable" to the respective analytical account. If we ask how T and T' could both be true, it is said that they capture different senses of C. If we take pluralism to be the view that T and T' *can* both be true, our question now is whether an analytic theorist can coherently espouse pluralism about what he himself is analytically minded about. Suppose he can. Then he must be antecedently disposed to regard his own intuitions about C and showing C to be ambiguous. But in general this is not true of analytic rivals. This being so, the ambiguation strategy for conflicted analytical theories T and T' is a strategy not for the T-theorist and not for the T'-theorist, but for some third party who, let us note, need not himself be analytically minded about C. So I conclude that pluralism, tailor-made for abstract theories, is problematic for analytic ones.

Our distinction between analytic and abstract is rather slack and certainly neither exclusive nor exhaustive. There are those who judge the naive theory of sets to

---

[9]   For historical accuracy, it should be remarked impassing that Frege could not have thought that the Russell Paradox destroyed Frege's analysis of sets. Frege distrusted the concept of sets independently of the Paradox, and certainly had no theory of them. Frege acknowledged value-ranges, and he thought the Paradox as damaging to his value-ranges as to Russell's propositional functions. That Frege subscribed to the naive theory of sets is a confusion of commentators who see in Frege's value-ranges what Russell saw in intuitive sets. Even so, I shall conform my remarks here to the traditional view, in which Frege himself would later in his life seem to have acquiesced, that there is an early Fregean set theory and that the Paradox wrecked it. What this means is that *were* there an early Fregean theory of sets it would have been wrecked thus.

have been an analytic account (and a failed one). Of the pathological biconditional on which the Russell Paradox rests, Quine notes "the thundering heptameter that shattered naive set theory: the class of all those classes not belonging to themselves".[10] If the naive theory were not an analytic theory, Quine's artful description would be overblown; nor would there be occasion to say that "desperate accommodations were called for."[11] It is widely believed, but also contested, that successor theories (e.g. ZF) are not and do not purport to be analytic theories, but hardly anyone seriously conceives of ZF on the historical model of Riemann's own highly abstract entertainments. Where naive set theory can be seen as an attempt at an analysis of sets, ZF can be seen as a *strategic stipulation* on sets, judged at bottom by its mathematical fruitfulness and, as Russell himself said, by its propensity *to be believed.*[12] Seen this way, it becomes explicable how doing set theory was possible in the wake of Russell's paradox for anyone who thought of the naive theory as an analytic theory. At least it becomes partly explicable. If ZF is not an analytic theory, it doesn't matter that there is no intuitive concept of set for it to be a theory of.[13] Why doesn't this place ZF in the camp of purely abstract theories in which there is no presumption that anything there is true?

When you stipulate that $\alpha\Phi$'s you make $\ulcorner\Phi(\alpha)\urcorner$ true by stipulation. Many people suppose that stipulated truths stand to truth as fool's gold stands to gold. I propose a contrary assumption, on sufferance: that one way for a sentence to *be* true is to be made true by stipulation. If this is so, a distinction falls out between the pure forms of abstract and stipulative theories. Purely abstract theories leave nothing true that wasn't true before. Stipulative theories leave new truths, true by stipulation. If the distinction amounts to anything, it should be possible to specify differences between how an abstract theory's entertainments and a stipulative theory's stipulations are grounded and justified. I owe the reader a large debt. The pledge of truth by stipulation needs to be redeemed. This is essayed elsewhere in some details.[14] For now it is proposed that stipulative theories be understood casually enough to make of 'stipulative' itself a stipulation. Let us simply say that a theory is stipulative to the extent that it makes posits there was no inclination to make prior to theory, and

---

[10]   W.V. Quine (1987), 146.

[11]   Ibid, 148.

[12]   Although I shall say below why I think we must take Russell as meaning not *believed* but *accepted.*

[13]   "'Intuition' is in general not to be trusted... [It] is bankrupt, for it wound up in contradiction... The logician has had to resort to mythmaking". (W.V. Quine (1966), 27.) So "when we pursue general set theory, we must grapple with the paradoxes, whether by von Neumann's method of non-elements, or by Russell's hierarchy of types, or by some other probably equally artificial device". (W.V. Quine (1951), 138.)

[14]   John Woods, *Paradox and Paraconsistency: Conflict Resolution in the Formal Sciences,* in progress.

that it holds that making them is making them true. (I expand upon this idea in a later section).

In the real life and natural history of theories, the distinction between analysis and stipulation, often in any event a matter of degree, divides its terrain dischronically. Sometimes we see that what starts out as stipulation takes on an analytical gravamen in the fullness of time. For some mathematicians precisely this has been the fate of ZF. To some extent, it is a matter of learning conditions. Someone immersed in the culture of naive set theory is likely to have at most a stipulative appreciation of ZF. But if from the beginning one's exposure to sets has been ZFish, the greater the likelihood of thinking of ZF-structures analytically. We ought not be surprised by the generalizational shift from Black to Shoenfield and Martin. Here is Black:

> ...once we abandon our recourse to our intuitions and 'resort to myth making'... the need to bolster our myths by some justification other than pragmatic considerations of convenience for the technical needs of mathematics and the sciences becomes imperative.[15]

On the other hand, the Russell Paradox

> does not really contradict the *intuitive* notion of a set [...since for any set A,] A is not one of the possible elements of A; so the Russell paradox disappears.[16]

Indeed, sets from the ZF perspective just are

> the standard conception of set.[17]

There is a moral to this. It is insufficient to disqualify a theory for its failure to be or to attempt to be an analysis of its target concepts. I would say that any modal theory in which the Barcan formula[18]

$$\Diamond \exists \alpha \Phi \rightarrow \exists \alpha \Diamond \Phi$$

[13]   Max Black (1943), 264-275.

[14]   Joseph R. Shoenfield (1967), 238. Emphasis added.

[17]   Donald Martin, (1970), 113.

[18]   Which sanctions, among other things, this inference: Since it is possible that there are intelligent Martians, there exist things of which it is true that they might be Martians.

is proclaimed is a defective analysis of the concept of possibility, never mind that the formula is embedded in a number of successful modal theories, notably in the quantificational extension of S4. This is telling. It suggests a slope slippery enough to slide down. Short of giving up on a substantial part of modern modal logic, we must say that the goodness of no good formal theory depends wholly or dominantly on its being an analytically good theory of its target concept.

Readers will have noticed that analytical theories need not be seen as theories that give the *meanings* of terms for its target concepts.[19] For a theory to give an analysis of a concept, it suffices to specify truth conditions for arbitrary sentences which invoke or apply the concept in appropriate ways. What is required for analytic status is that the truths conditions that the theory specifies are conditions that obtain independently of their recognition by theory. Meanings aren't required for analysis unless truth conditions are meanings. What is essential is not that truth conditions constitute meanings but rather, even if they do, that they be theory-independent, and in that sense *objective*. Objectivity enters the picture in two different ways (and they may be related ways). If T is an analytic theory then T's truth Φ are true independently of the fact that $T \vdash \Phi$. And if T is an analytic theory whose domain of interpretation is D, then T's objects exist independently of the fact that they show up as members of D. Theories objective in the first way are "semantically objective". Theories objective in the second way are "ontically objective". In standard model theoretic semantics, realistically construed, the two kinds of objectivity coincide.

It is a matter of controversy whether analytic theories could be theories of meaning. What is not in doubt is that if there are any analytic theories worthy of the name, they are theories rooted in realist presumptions. They are theories in which it is intended to tell the objective truth about matters which objectively instantiate target concepts. It appears to be distinctive of formal theories that their Ks are heavily populated by intuitions. When a theory sanctions a consequence judged counterintuitive it is striking how readily one inclines to the view that the theory is wrecked or, more soberly, that it has something to try urgently to explain away. If we take a dominantly analytic view of theories, it is easy enough to specify a concept of *counterexample;* for what would a counterexample to a theory be if not a counterintuitive consequence of it? But if our approach is dominantly abstract, counterintuitiveness can be a more or less routine concomitant and the idea of counterexample becomes all the harder to get a grip on. One is left to wonder whether there is any serious methodological role to be played by considerations of intuitiveness in the dynamics of any theory that pretends to be dominantly non-analytic.

---

[19]   This is precisely what they do in fact do according to the ambiguation strategist. (Which is another reason to distrust ambiguation as a general pacifier of philosophic discord.)

2. Semantic Intuitions

We see in these considerations something of the methodological designs of philosophical skepticism about meanings. To admit meanings, so the skeptic says, is to assign them load-bearing roles in theory. If we suppose that a sentence is counterintuitive if it violates the meaning of embedded or implicationally adjacent terms, then meaning violation can be expected to play an explanatory role in the dialectics of counterintuitiveness appraisal. Consider an actual case. If a relevant logician attributes counterintuitiveness to *ex falso quodlibet*, a classical logician may accept the attribution or reject it. If he accepts it, he is unlikely to give to the counterintuitiveness any significance other than that of something surprising; but whatever the details of his view of what the counterintuitiveness positively consists in, his negative view will be dialectically recalcitrant: it is *not* a counterexample. On the other hand, if the classical logician refuses the attribution, his refusal is correct of incorrect. If correct, nothing more here needs saying. If it is incorrect then, by the assumption of the present case, the classicist does not know what "entails" means. The relevantist's invocation of counterintuitiveness is dialectically impotent. For if the classicist doesn't know what "entails" means how can pointing this out get him to see that *ex falso* is not true? This is not to deny that when confronted with *ex falso* a theorist may reject the theory of entailment which sanctions it. Nor is it to deny that if asked, the repenting logician may explain himself by saying that his now disabled theory let him down by implying something counterintuitive. This is not, however, the case under review, for which the would-be defacter is not in fact

(1) Your theory implies *ex falso*

but rather

(2) Your theory implies something *counterintuitive*.

This leaves the urger of (2) with no dialectical slack to cut. Of course, a complaint along the lines of (1) *may* get a theorist to lighten up, but if he doesn't there is no prospect of gain to be got by pressing (2) even *if it is true*. For, again, if it is true, anyone holding *ex falso* doesn't know what *ex falso* means, that is, doesn't know what

$$\Phi \wedge \neg \Phi \mid\vdash \Psi$$

means. And if that is the case,

(3) $\lceil \beta \Phi \wedge \neg \Phi \mid\vdash \Psi \rceil$ is counterintuitive

cannot serve as dialectically telling against anyone for whom (1) is not already a counterexample. We can generalize on this: Let T be a theory and Φ a counterexample to it. The T-theorist either recognizes that Φ is a counterexample or not. If he does, that is the end of the story. If he does not, then the perfectly true assertion, ⌜Φ is a counterexample to T⌝, cannot be availing. I conclude, then, that a charge of counterintuitiveness as presently conceived is dialectically underdetermining.

It is possible that I have played the theme of meaning-violation rather too lightly (some would say too tightly). Certainly, there is a substantial tradition in analytic philosophy which offers sharper purport for our metaphor. On this view, it remains true that if Φ is counterintuitive, Φ involves a violation of meaning; but in its turn, if Φ involves a violation of meaning then Φ is *false-by-meanings-alone* (or *m*-false), hence inconsistent with a statement true-by-meaning-alone (or *m*-true).

What now are the dialectics of a charge of counterintuitiveness directed to someone who resists the attribution? For the charge to be dialectically serious (something other than the exasperation of, "It's obviously false; why can't you see it this way, too?") the maker of the charge must specify the truth, true in virtue of meaning, with which he takes the counterintuitive sentence to be inconsistent. If it is Φ that is counterintuitive, then a sentence with which it is inconsistent is ⌜¬Φ⌝. But it is dialectically otiose to cite it, for it makes the attributor's case circular: "Φ is *m*-false because ⌜¬Φ⌝ is *m*-true". On the other hand if Ψ is some sentence distinct from ⌜¬Φ⌝ with which Φ is inconsistent, the attributor might cite Ψ. In this he might succeed in winning his opponent's concession or he might not. If he does succeed he has a "starter" for his complaint, and it suffices for this that his opponent grant Ψ's truth (and its inconsistency with Φ). For this to happen, it is unnecessary that the opponent be got to see that Ψ is true *by meanings*. If, on the other hand, the opponent resists Ψ, cannot see that Ψ is true, or cannot see that it contradicts Φ, this resistance will not be overcome by citing Ψ's truth by virtue of meaning. If the opponent fails to see that Ψ is true, he will fail to see that Ψ is true by meanings even if it *is* true by meanings. And if it also happened that Ψ could not be seen as true unless it were seen as true by meanings, it does not come to pass in the general case that attributing counterintuitiveness to Φ will get an opponent to see that Ψ, with which Φ is inconsistent, is true by meanings.[20] The same is true for the inconsistency of Φ with Ψ. If Φ is inconsistent with Ψ by virtue of meanings, and the opponent fails to see the inconsistency, the attribution of counterintuitiveness to Φ will not get him to overcome this failure. Perhaps his dialectical *vis-à-vis* will

---

[20]   There are apparent exceptions. If I doubt that Harry is pusillanimous while granting that he is fainthearted, you could direct me to a good dictionary and get me to see that "Harry is fainthearted" is true only if "Harry is pusillanimous" is true. But this is a case of your getting me to see the truth of "Harry is pusillanimous" by getting me to see what 'pusillanimous' means. It is not a case of getting me to see that "Harry is pusillanimous" is a truth true by virtue of meaning.

offer him a proof of Φ's inconsistency with Ψ. If so, he may have established to his opponent's satisfaction Φ's falsehood on Ψ's truth. But it is not necessary to the success of the proof of inconsistency that the opponent be made to see that the inconsistency obtains in virtue of meanings. I conclude that where it is not circular, a case against a sentence based on its counterintuitiveness stands or falls independently of an opponent's recognition of the *m*-truth (or *m*-falsehood) of any sentence germane to the construction of the case. Once again, a charge of counterintuitiveness is dialectically *underdetermining*. In this the meaning-skeptic is vindicated. Meanings fail the methodological conditions imposed on them by theory. They don't bear any weight.

There can be a similar problem with intuitiveness. It is widely conceded, and quite true, that we try to get our theories to conform to our intuitions, that we want their provisions to be intuitive. Suppose that we take intuitiveness on the model of counterintuitiveness: hence Φ is intuitive when if it is true by meanings. Consider a case. Theorist S registers his claim that Φ. S believes Φ and he may believe that Φ is true by meanings (and he might be right). Suppose that S is queried by S' who doubts that Φ. Suppose that S has no reason to give for believing Φ, that is, that Φ is *probatively bereft* for S, except for his belief that Φ is true by meanings. S says so to doubting S'. It is no use. If the doubter really does doubt that Φ, S's reply must be unsuccessful. Since S' doesn't believe Φ to be true he can hardly believe it to be true by meanings, even if it is. Invocation of truth by meaning is dialectically obtiose. If Φ's truth by meaning has any real work to do, it is work done only within a community of Φ-*believers*. Truth by meanings can there be invoked to explain why Φ has been proclaimed without proof or evidence. Truth-by-meaning triggers the transformation of no evidence into self-evidence, and that is something, but not much. It makes of self-evidence the conspicuous runt in the litter of evidence. Evidence is probative even across the grain of intersubjective disagreement. Self-evidence can survive no such journey. Consider a proposition Φ which I believe and which you disbelieve. If you are reasonable and open-minded you might ask for evidence and I might furnish some. It is not that we are antecedently guaranteed to agree that "my" evidence *is* evidence or that, if it is, it is adequate to sustain Φ. The point rather is that where the only evidence for Φ is self-evidence, then no one who doubts that Φ would or should ever consider the question of Φ's self-evidence, except negatively. In doubting that Φ, the doubter is committed to the view that self-evidence is precisely what Φ doesn't have.

As we have been conceiving of them so far, intuitions and counterintuitions are so much methodological dead weight. They get us in the dynamics of conflict resolution. They also bankrupt at least one historically prominent conception of analytic theories. In particular, they make it impossible for analytical theories to recover from the devastation of paradox. This suggests what an admirer of "Two

Dogmas of Empiricism" may think has long been obvious, that talk of intuitions and counterintuitions embeds a wholly useless distinction between the empirical and non-empirical or, more soberly and realistically, a wholly useless way of marking a far from useless distinction. But sound as far as it goes, it does not suffice to be told merely to quit the uncritical positivism that lurks in the preceding discussion. No serious formal theorist, save those devoted to pure play, doubts that he can proceed without intuitions or that he can succeed in utter indifference to counterintuitiveness. So, an important question for such a theorist is, what does he think he is talking about in saying such things?

3. Epistemic Ecologies

If what has been said just now is so, it becomes apparent that the significance, and value, of invoking intuitions is deeply a function of dialectical circumstance. This is as it should be, for reasons as obvious as the fact that we have been examining the rôle of intuitions in the context of conflict resolution in non-empirical theories. We would do well to flesh out this dialectical character of conflict resolution, although, partly for want of space and partly for want of insight, everything offered here will be approximations of modest ordinality. The notion of an epistemic community will prove useful. We begin by re-tailoring the idea of counterintuitiveness.

A statement is counterintuitive for a cognitive agent in an epistemic community.[21] I will say that Φ is a counterintuitive provision of a theory T in an epistemic community **ec** just in case it is believed false in **ec** *and* is believed false in the face T's theory's derivation of it. A proposition counterintuitive in **ec** is one judged to violate the community's "pre-theoretical intuitions", as is sometimes said. "Intuition" is an unfortunate turn of phrase for our present purposes. It suggests knowledge by special license, or knowledge with exotic origins, e.g., in meanings. The suggestions should be disarmed. An analytic theory's pre-theoretic intuitions in a community are the propositions in the theory's domain which members of the community are prepared to be realists about; they are propositions judged objectively true independently of their place in theory.

Counterintuitive disclosures are the stock in trade of interesting theories. When the theory is empirical, that it produces a counterintuitive result is, in general, nothing to complain of. It is widely recognized that nature is full of surprises. It is with theories of the formal sort, that counterintuitiveness seems to matter distinctively. For it seems entirely natural to take a person's attribution of counterintuitiveness as

---

a *complaint*. What is it to complain of counterintuitiveness? What is it to bring a charge against a theory on such grounds?

> **DefChgeCount**: An allegation in **ec** of counterintuitiveness against a theory T with respect to a proposition Φ is a valid *charge against* T iff that $T \vdash \Phi$ is a *reductio ad falsum* in **ec**.

I will say that a charge of counterintuitiveness is *self-sustaining* in **ec** if it satisfies DefChgeCount. Where a valid charge is not self-sustaining it is customary for it to be sustained by provision of a counterexample.

> **DefCountEx**: Ψ is a *counterexample* in **ec** to a theory T in respect of a proposition Φ, iff Ψ is held true in **ec**, $T \vdash \Phi$, Ψ is inconsistent with Φ and recognized to be so in **ec**.[22]

When it has been determined in **ec** that a charge of counterintuitiveness in self-sustaining or has been sustained by a counterexample, it must be judged that either T embodies a falsehood other than Φ itself,[23] or that T's projective machinery is defective. It is noteworthy that a successfully sustained charge of counterintuitiveness does not always have a unique solution in $T-\{\Phi\}$ or in M. I will say that successful charges usually do not uniquely "specify their pedigrees", that is, they do not, just so, *particularize* the error whose existence they allege. Successful counterexamples with respect to T are refutations of something in $T-\{\Phi\}$ or in M. To *finish* a sustained change of counterintuitiveness against T with respect to Φ, it is necessary to amend $T-\{\Phi\}$ or to amend M, or both. Knowing how to do this involves knowing the pedigree of the counterintuitiveness claim in question. It is an evident and recurrent feature of our epistemic lives that sustained charges are hard to finish. We will say that a refutation is *blinkered* in an **ec** if it cannot be finished there.[24]

---

[22] What is it for something to be recognized or believed in an epistemic community? I take it as given that universal consent is not a prerequisite, that Italians can love garlic without Luigi having to as well. There are two likely candidates. (1) It is believed in **ec** that Φ if most members of **ec** believe that Φ. (2) It is believed in **ec** that Φ if the members of **ec** believe that Φ. Case (1) offers us *plural quantification*, concerning which see e.g., Gila Sher (1991), and Johan van Benthem and Alice ter Meulen (eds.) (1985). Case (2) reflects the use of *generic* statements, concerning which see, e.g., E.M. Barth (1974), and Gregory N. Carlson and Francis Jeffry Pelletier, (eds.) (1995). Neither plural quantification nor genericity knows yet the comforts of settled theory; so here too I shall take the idea of a community's belief that Φ as primitive.

[23] Except where Φ is derived from itself, as an axiom, say.

[24] The only good account by a logician of finishing refutations is that adumbrated by Aristotle in the *Sophistical Refutations*. See John Woods, *Aristotle's Earlier Logic*, to be published.

Epistemic conflict at its most interesting is conflict in communities in which there is disagreement about whether a charge of counterintuitiveness is valid and about whether, if an attempt is made to sustain it, the purported counterexample is genuine. Epistemic communities, like all social organizations, are ecologies. Essential to the structural integrity of communities is affordable access to strategies of conflict management. One of the most economical of these strategies is to agree to disagree. In terms of our present definitions, this is tantamount to fairly controlled reciprocal optings out. The conflict in **ec** about whether the derivation of Φ in T constitutes a *reductio ad falsum* might prompt division of **ec** into **ec'** and **ec"**, in one of which Φ is judged to falsify T and the other of which not. If these are highly abstract communities, their disagreements will be far removed from the practical necessities of life. They will be sub-communities which can afford to pay one another no mind. For disagreements which model in approximately the ways presently under review, it can be predicted that on contested questions such as the counterintuitiveness of Φ in T, **ec'** and **ec"** will take different positions with respect to T, with T proclaimed in **ec'** and some variant T* proclaimed in **ec"**, and that by and large **ec'** and **ec"** *will* ignore one another with respect to the issues that divide them. Alternatively, they will routinely beg one another's questions, and no one will much care.

A familiar notion of paradox is extendable to epistemic communities.

> **DefPar^{ec}**: That T⊢Φ is judged true in **ec** is a *paradox* in **ec** iff Φ is counterintuitive in **ec** and either the charge of counterintuitiveness is unsustained or the counterexample that sustains it is unfinished (hence the refutation it provides is blinkered).

When a community satisfies the conditions of **DefPar^{ec}**, it is in a quandary with respect to T. For example, it doesn't know how to make "its" conception of *set* consistent.

Quandaries beget stalemates between communities. Thus

> **DefStlMt**: Let T⊢Φ be judged true in **ec'** and **ec"**. Let Φ be counterintuitive in **ec'** and intuitive in **ec"**. Let there be at least one non-blinkered refutation of Φ in **ec'**. Then **ec'** is in a stalemate with **ec"** with respect to T iff every non-blinkered refutation of Φ in **ec'** is rejected in **ec"**.

If we imagine **ec'** to be a community of relevant logicians and **ec"** a community of classical entailment theorists, the two communities would be in a stalemate in the theory, T, of entailment if they disagreed about (say) *ex falso* and disagreed about the success or failure of every measure deployed to elevate the charge of counterin-

tuitiveness into a mutually accepted finished refutation of something in T. Thus if the relevantist pleads the counterintuitiveness of *ex falso*, the classicist will agree or disagree; but if he agrees he will say that it is only a *surprise*. If the relevantist points out that the standard definition lacks a relevance-condition, the classicist will agree or disagree; but if he agrees he will not agree that relevance is a condition on entailment. If the relevantist attacks Disjunctive Syllogism (DS), the classicist will reply that DS fails on an interpretation of the connective 'v' which is not preserved in the classical account of entailment (and so the relevantist has "changed the subject"); and on and on.

It is one of the more interesting features of the natural history of formal theories in this century that the development of the theory of sets from Russell's failed attempt to present day versions of ZF was by and large a development unafflicted either by paradox in the sense of **DefPar**^{ec} or by stalemate, whereas in the history of entailment theory from (say) 1912 to the present, we see very little *but* stalemate.
Stalemates are nothing to despair over. They can be evaded or relaxed as when, for example, an unsustained charge counterintuitiveness is answered.

> **DefAcqChge**: A charge of counterintuitiveness in **ec** with respect to the fact that T ⊢ Φ is *acquitted* in **ec** iff for some Ψ believed true in **ec** it is established that Ψ is not true unless Φ is and **ec** persists in its belief in Ψ notwithstanding.

It is easy to see how **DefAcqChge** extends across the lines of disagreement between parties in **ec** or between subcommunities of it. When a charge of counterintuitiveness if acquitted, minds are changed. One's belief in the falsehood of Φ is outweighed by one's belief in the truth of some Ψ which acquits it. Acquittal is the natural antidote for quandary. When a charge of counterintuitiveness with respect to Φ is acquitted, it is the *charge* that is acquitted, not the counterintuitiveness. Φ may persist in seeming false from the beginning to the end of the acquittal process.

More strictly speaking, acquittal on a charge of counterintuitiveness against Φ is compatible with one's not believing Φ. Acquittal guarantees not a change of belief but a change of acceptance. To this day there are people, some of them quite sensible, for whom Cantor's Paradise cannot be a matter of belief. The notion of the transfinite, of actual but uncompletable infinities, is "too bad to be true", or counterintuitive enough to preclude positive belief. But if the transfinite goes, so do significant chunks of modern mathematics to which the transfinite-doubter gives his untroubled assent. This is an arrangement in which the charge of counterintuitiveness against the transfinite is acquitted, and it leaves in its wake not belief but acceptance.

From this we get a notion of surprise.

SurpΦ: Φ is a *surprise* in a theory T such that T⊢Φ in **ec** iff Φ is counterintuitively chargeable and acquittable in **ec** under conditions in which Φ is accepted in **ec** but not believed.[25]

Surprise, as we have it here, shares structural features with Moore's Paradox, the notorious "Φ, but I don't believe it". This alone tends to make us think that a surprising consequence counts against itself probatively. When the conditions of **Surp**Φ are met, the "surprisee" is committed to an utterance in the form

m: Φ, but I don't believe it.

This is Moore's Paradox in benign form. Uttering **m** does not constitute a blindspot for anyone to whom it is directed.[26] For in the context in which it is uttered, it is not impossible for the addressee to know either what the speaker is asserting or what he believes. Context makes it clear that what the surprisee does not believe is something he nevertheless accepts.

The factor of surprise calls seriously into question the utility of purely analytic theories. We said that T is an analytic theory when, for its K and M, T is the closure of K under M and K is a set of sentences believed objectively true. If there is to be *any* methodological rôle for counterintuitiveness with respect to formal theories, acquittal must be possible in principle. But with acquittal comes surprise, and with surprise comes sentences in the form **m**. If, as we said, a theory counts as an analysis of its target concepts, the analysis is given by the sentences of T believed objectively true. By these lights, it will be entirely commonplace that successful non-empirical theories will not qualify as analyses of their target concepts. This will be so to the extent that T carries consequences that are accepted but not believed.

Once the distinction between belief and acceptance is admitted to the consideration of such things as counterintuitiveness and surprise, it is easy and natural to grant to

---

[25]   In the community of mathematicians who accept the axiom of choice, the Tarski-Banach theorem is highly counterintuitive but is acquitted of the charge. It is a surprise. To see what the acquittal comes to, let us say that objects Σ and Γ are congruent by finite decomposition if they partition finitely into sets $\Sigma_1, ..., \Sigma_n, \Gamma_1, ..., \Gamma_n$, and $\Sigma = \cup\Sigma_1, = \cup\Gamma_1$, and Σ_i is congruent with Γ_i (for i a non-zero whole number). The Tarski-Banach theorem establishes that all *spheres* are congruent by finite decomposition. It is not surprising (no pun) that in some quarters the theorem is known as the Tarski-Banach Paradox. See, e.g. Thomas J. Jech (1977), 345-370; 351 ff.

[26]   See Roy A. Sorensen (1988): "Thus a proposition *p* is a blindspot relative to a given propositional attitude A and a given individual *a* (at time *t*) if and only if *p* is consistent but *a* cannot have attitude A towards *p*... Given the constraints imposed by certain desiderata of belief, I cannot believe that 'It is raining but I do not believe it' even though it is a consistent proposition" (52-53).

it a broader provenance. In particular, it is quite straightforward that the K of a theory need not be restricted to sentences believed objectively true, and that the M of a theory need not be confined to projection mechanisms that are believed objectively sound (think of M as an *abductive* device, for example). With this said, it becomes apparent that formal theories seriously worthy of the name will be hybrids of what is believed and what is accepted. Such theories cannot be analytic theories, though it may be thought that they contain analytic fragments. In this connection, we might mention our former distinction between analysis and stipulation, which is what hybrid theories appears to be hybrids of.

Abstract theories also become easier to characterize. In pure form, they are theories whose K-members are neither believed objectively true nor accepted as true, and are so in sufficient numbers as to make it powerfully odd to think of them as giving accounts of how target concepts are. When Riemann constructed his brilliantly peculiar geometry he constructed something within which and about which all sorts of things are objectively true or rationally acceptable. What Riemann did not think he was doing was giving a rationally believable or acceptable account of real space. In this he was wrong, of course, but that doesn't change the fact that the theory he thought he was constructing was an abstract theory in the sense at hand.

Stipulative, hybrid, and abstract theories are themselves natural antidotes to stalemate. They afford disputants the luxury of not minding over-much their deadlocked beliefs, and they make possible some fruitful poaching in the opponent's preserve. It is a good thing, too. Often a theorist's creative playing around with a theory he detests will cause him to lighten up. He may come to respect his opponent's insights. He may even make theoretical strides on his opponents' behalf, proving new theorems or streamlining old axioms, for example. The playfully motivated proofs of new theorems in the enemy camp may eventuate in the poacher's enlarged acceptance of the rival theory. They might even precipitate a full scale conversion.

If acceptance is what stipulation seriously contends for, it becomes important to specify conditions on a successful outcome. Such conditions are notoriously difficult to pin down with any exactitude or promise of exhaustiveness. I will mention only the standard ones: internal coherence, propensity to solve otherwise unsolvable problems, predictive facilitation, simplicity, enlargement of the research programme, and so on. Overall these are economic rather than objectively probative considerations. Their importance is such that some philosophers have been prepared to rewrite the definition of "probative", and to pragmatize the notion of truth and depsychologize the notion of belief. I will not comment on these deviations except to say that the revisions they encompass attest to the importance of economic considerations in the construction and appraisal of theories. These are considerations having a loosely cost-benefit character, "loosely" because there is little serious

prospect of the quantification of costs and benefits in such contexts. Here, too, I lack the space for detailed characterization. Even so, we may propose as a condition on the dialectics of theory-defence the following:

> **EcDefen**: For cognitive agents S and S' and some set of non-empirical sentences Σ proclaimed by S' and attacked by S, S' makes a satisfactory *economic defence* of Σ against S's attack by making the success of S's attack too costly for S.

> **CostAtt**: S's attack on Σ is *too costly* for S iff its success commits S to propositions he is not prepared to accept or to the rejection of propositions he is not prepared to reject, or to methodological adjustments he is not prepared to make.[27]

Thought of this way, defences are pure cases of Lockean *ad hominem* manoeuvres. This is nothing to complain of. Locke was never so silly as to call fallacious the "pressing of a man with consequences of his own principles or concessions". All the same, I will say that an S'-defence is *secure* rather than merely economic iff what S' considers costs would also be considered costs by S were he relevantly situated. One calls to mind Quine's defence of classical logic in the face of intuitionistic defections. We lose the law of Double Negation, said Quine, and therewith "classical negation".[28] But "[i]t is hard to face up to the rejection of anything so basic."[29] We may agree that Quine is citing costs that he himself is not prepared to bear, and that this may be said to be a sort of *autodefence*. But in as much as the rejection of classical negation is precisely what the intuitionist earnestly wishes for, Quine's remark is no defence of classical logic against the intuitionist, and is still less a secure defence.[30]

I am not able to be very specific about what epistemic communities are. They are, even so, handy things to make theoretical use of. Apart from the fact that they are, so to speak, the natural habitats of dialectical transactions, they also appear to be the natural medium for the changing of minds about conceptual matters. It is exceedingly difficult to see how it came to pass that from the collapse of set theory in 1901, ZF came to be, as Norman Martin says, the standard conception of set,

---

[27]   The definitions also extend to epistemic communities, in the obvious way.

[28]   *Philosophy of Logic*, 74.

[29]   Ibid, 85.

[30]   In fact, we may say that autodefence stands to defence is something like the relation in which self-evidence stands to evidence.

except by reference to mechanisms of adjustment and accommodation that are essentially social.[31]

4. Idealism

Frege was devastated by Russell's news of the Paradox. It led him to proclaim what elsewhere[32] I have called

Frege's Sorrow: There is no intuitive concept of set.

We have been asking, in effect, how it is possible for a theoretical setback to qualify for Frege's Sorrow.

The history of mathematics since the Russell paradox (and well before) displays a remarkable poise in the face of such difficulties and a readiness to soldier on. This suggests that Frege's reaction was a psychological peculiarity and a methodological anomaly. In fact it was neither.

To ask how Frege's reaction was possible is just to ask how it could be that a mathematician takes an analytic approach to his target concepts. It is to ask, "How is an analytic theory of sets explicable". As it happens there is a fateful concurrence between, in Jena, the death of psychologism in logical theory, and in Cambridge, the death of idealism in philosophy. As for the latter, it is convenient to mark the years between 1897 and 1903 as the period when Moore was converting Russell away from idealism and to something that was to be called "philosophical analysis". The year 1897 saw the publication of *An Essay on the Foundations of Geometry*, Russell's first and only purely idealist book. In 1903 there appeared *The Principles in Mathematics*, the first and in some ways the most aggressive work of his conversion.[33] Here are six years of such mathematical fruitfulness, especially in Germany, that it is easy to overlook philosophical developments at home, in Cambridge. Among the mathematical and dominantly German results of the period are: Frege's own masterwork, *Grundgesetze der Arithmetik*;[34] Cantor's "Beiträge

---

[31] Epistemic communities also are wanted for my theory of stipulative truth, which is, in turn, a kind of constructivism for formal theories. But I am getting ahead of myself. See *Paradox and Paraconsistency*, *op. cit.*

[32] Woods, *Paradox and Paraconsistency*, in progress.

[33] Bertrand Russell (1987) and (1903).

[34] Frege (1893-1903), partial translation as *The Basic Laws of Arithmetic: An Exposition of the System* (1964).

zur Begründung der transfiniten Mengenlehre" (1897);[35] Burali' Forti's "Una ques-
tione sui numeri transfiniti" (1899);[36] Hilbert's *Grundlagen der Geometrie*; [37] and
just a year away was Zermelo's "Beweis, dass jede Menge wohlgeordnet werden
kann".[38] Meanwhile, in Cambridge, Moore was revolutionizing the course of
Englishspeaking philosophy.

If we are to answer the question of how it is that the collapse of set theory could
have given rise to Frege's Sorrow, it will be necessary to solve the following
puzzle, which I shall call the

Riddle of Analysis:[39] Russell discovered the paradox that bears his name in
June 1901.[40] Doing so should have *either* (1) pre-empted entirely his conver-
sion to Moore's conception of philosophy *or* (2) been the occasion of his
abandonment of the philosophy of mathematics.

Our puzzle is set by (i) the philosophical changes that Russell underwent between
The Foundations of Geometry and the *The Principles of Mathematics*,[41] and (ii)
the discovery of 1901.

In the *Foundations of Geometry*, Russell attempts to reconcile two fundamental
insights. The first is that there is but one true geometry, which, in turn, is the one

---

[35] Parts one and two (1895-1897); translated as *Contributions to the Foundations of the Theory of Transfinite Numbers* (1915).

[36] Translated as "A Question on Transfinite Numbers", in van Heijenoort, op. cit.

[37] Translated as *The Foundations of Geometry* (E.J. Townsend, trans), 1902, 10th edition, with a Supplement by Paul Bernays translated under the previous English title (1971).

[38] Zermelo (1904); translated as "Proof That Every Set Can Be Well-Ordered", in van Heijenoort, op. cit.

[39] Not to be confused with the "Paradox of Analysis" which, later, would so greatly exercise Moore, and which was anticipated by Husserl.

[40] In this I concur with Ivor Grattan-Guinness (1978), 127-137: "[J]une [1901] appears to be the most likely date". (135). Judging from Hilbert's letter to Frege, it is apparent that Zermelo derived the paradox that bears Russell's name in 1899. See David Hilbert (1976). It would also appear the Cantor was aware of "Russellian" difficulties: "Cantor's 'definitions' [from 1899] only allow as sets those collections which are *wholes* and this does not at all imply that any collection can be a set. Nothing like the comprehension principle of so-called 'naive set theory' follows from Cantor's statements. If 'naive set theory' is charac-terized as set theory based on the comprehension principle, then this goes back, not to Cantor, but to Russell [1903]." Michael Hallett (1984); 38.

[41] A change that was mediated by the publication in mid-course of the book on Leibniz [Bertrand Russell (1900); republished 1937], concerning which see Peter Hylton (1990); 152-156.

true theory of space, some of whose axioms are known *à priori* and some by experience. The second is the paradigmatically idealist claim that the concept of space is inherently and incorrigibly contradictory, and hence, so too is geometry. Russell writes:

> After hypostatisizing space, as Geometry is compelled to do, the mind imperatively demands elements... But what sort of elements do we thus obtain? Analysis, being unable to find any earlier halting place, finds its elements in points, that is, in zero quanta of space. Such a conception is a palpable contradiction... A point must be spatial, otherwise it would not fulfill the function of a spatial element; but again it must contain no space, for any finite extension is capable of further analysis. Points can never be given in intuition, which has no concern with the infinitesimal: they are a purely conceptual construction, arising out of the need of terms between which spatial relations can hold. If space be more than relativity, spatial relations must involve spatial relata; but no relata appear, until we have analyzed our spatial data down to nothing.[42]

Noteworthy, is that the irremediable inconsistency of geometry prompted in Russell no Frege's Sorrow. Contradictions can be dealt with dialectically. If geometry is afflicted with contradiction then, like any theory so fated, geometry demands supplementation (with what, right at the beginning so to speak, Aristotle called "further qualifications")[43] under which the contradiction disappears. The supplementation of an old theory involves changing its subject matter to a different, but linked, subject matter. If the old theory is inconsistent, then a satisfactory supplementation of it must involve a change in subject matter that removes the inconsistency. In the case of geometry, the contradiction is subdued by abandoning the abstraction that underlies the old theory. The abstraction is that of conceiving of space as the mere possibility of diversity, that is, as the abstraction to empty space from the actual diversity of things in real space.[44] "[I]t is empty space...which gives rise to the antinomy in question; for empty space is a bare possibility of relations, undifferentiated and homogeneous, and thus wholly destitute of parts or thinghood".[45] In fact, "the relativity of space...renders impossible the expression

---

[42]   Russell (1987); 189-190.

[43]   Aristotle, *Metaphysics* Γ3, 1005ᵇ 19-13: "The firmest of all first first principles is that it is impossible for the same thing to belong and not to belong to the same thing at the same time in the same respect." However, *"we must presuppose, in face of dialectical objections, any further qualification which might be added."* Emphasis added.

[44]   Russell (1987); 128.

[45]   Ibid, 191.

of...[any theorem] of pure Geometry, in a manner which is free from contradictions."[46]

Given that the inconsistency arises from our conceiving of space as an abstraction from real things in space we must, says Russell, resolve "to give every geometrical proposition a certain reference to matter in general",[47] itself "a peculiar and abstract kind of matter, which is not regarded as possessing any causal qualities, as exerting or subject to the action of forces".[48] Here is the first glimmer of logical atomism, not fully born until 1918. In its adumbration here there is nothing to recommend the postulation of acausal atoms but their contribution to a rehabilitated geometry. True, "the mathematical antinomies...arise only in connection with empty space, not with spatial order as an aggregate [of acausal atoms]"[49] but, even so, we are driven to our inconsistent concept of space by "an *unavoidable psychological illusion*."[50] In as much as we are driven to our inconsistent notion of space by an unavoidable psychological illusion, it may be said that it is our *intuitive* conception of space that suffers the inconsistency .

Because a dialectical resolution of the inconsistency requires that we change the concept of space, the new concept will not be an intuitive concept and will involve the invocation of factors which we have no antecedent reason to accept and which we accept now because doing so (a) removes the contradiction and (b) preserves a link with the old concept of space. It will be, so to speak, as intuitive as consistency allows, namely, not very. Readers familiar with the reconstitution of set theory in the aftermath of Russell's paradox will notice the similarity to what Russell is proposing for the reconstitution of geometry, but I daresay that it may safely be supposed that they would be astonished to learn that Russell conceived of such reconstitution as a working out of the Hegelian dialectic (in McTaggart's slightly peculiar understanding of it).[51]

The author of *Foundations of Geometry* was an idealist. He engineered the rescue of geometry from its intuitive inconsistency in ways entirely faithful to his idealism.

---

[46]   Ibid, 128.

[47]   Ibid, 190.

[48]   Ibid, 191.

[49]   Ibid, 196.

[50]   Idem. Emphasis added.

[51]   cf. Ernst Zermelo (1930); 29-47; 47: "And so the set theoretic 'antinomies' properly understood instead of leading to a contradiction and a mutilation of mathematical science, lead rather to an *unsurveyable* unfolding and enrichment of that science". Emphasis added.

Russell thus subscribed to its four defining conditions. *First*, human knowledge is at least partly *creative*, the constitution of the human knower is at least partially constitutive of what the human knower knows. *Second*, truth is not absolute. It is *comparative* (and so some claims are more true than others), it is *partial* (and so nothing a human knower can know is completely true); and it is *mutable* (what is true today may be false tomorrow). *Third*, with the exception of a special kind of thinking, which idealists called "metaphysical", all (human) thinking is defective: incoherent , contradictory and incomplete. *Fourth*, something is real only to the extent that it is self-sufficient, i.e., not dependent on other things, and, failing this test, the objects of human knowledge are less than fully real.

It is perhaps worth noting in passing that the ideas for which idealists seem to have received persistent (and alarmed) attention are corollaries or appendages of those listed here. "Psychologism" is a word for the first claim, that human knowledge is at least partially constituted by the human knower, but it should be emphasized that Bradley was at least hostile as Frege to psychologism in logic.[52] Moreover, the infamous attack on internal relations is misnamed. Bradley's view was that all relations were unreal, external and internal, never mind that he made special arguments against internal relations. For consider the general case: if R is a relation such that for all objects x and y, which satisfy the condition that x bears R to y, the existence of R is bound up with the existence of its relata. Relations thus fail the self-sufficiency test. It is rather delightful that, if relations are taken extensionally, that is, as sets of n-tuples of relata, the dependency of relations on their relata is entirely obvious.

Idealists' exceptions to the law of Excluded Middle can be seen as a straightforward consequence of the relativity of truth. Here too, much nonsense has been written about the role of dialectic in idealistic thought. But in its barest essentials it is little more than the rescue of an inconsistent theory by the contrivance of a consistent near-thing. Then there is the Absolute, also the subject of more florid description than it ever deserved; it is best seen as that which is described by the limit on which dialectic converges.[53] Suppose then that all descriptions of the world, after much successive refinement, were free of inconsistency, provably coherent, and fully compliant with the law of Excluded Middle; then the totality of such descriptions could be said to be an utterly faithful record of Reality and thus *absolute*,

---

[52]   Bradley: "Both logic and psychology, if they are to exist at all, must remain each in principle independent. The undistinguished use of both at once must, even where instructive, remain in principle confusion. And the subordination of one to the other, whenever seriously attempted, will never, I think, fail to make manifest in its result the absurdity of its leading idea." (F.H. Bradley (1883); corrected impression 1922, 613).

[53]   Thus Peirce, for example, gave to his theory of truth a markedly idealistic cast in a series of six papers published in 1897 and 1888 in *The Popular Science Monthly*.

(from which the use of that very word as the name of Reality). There is ample opportunity to agitate over the conditions whose fulfilment constitutes this "end of dialectic", but the main idea is not all that exotic.

Idealism, therefore, is one of those philosophies in which most of the fun is in the details. The basic ideas are not quite as common as dirt, as the saying has it, but they are entirely without shock value: the human condition guarantees only limited and problematic access to how things really are and, in their congress, such as it is, with the world as it *is*, there is the constitutive impress of the human cognitive apparatus and repertoire.

Moore's conception of analysis, though not itself expressly defined, involves fundamentally and essentially the unqualified repudiation of our four basic conditions on idealism and of their corollaries and supplements. This is the negative characterization of analysis: (1) How the world owes nothing whatever to any presuppositions of our knowledge of it. Hence the objects of knowledge are wholly non-mental.[54] (2) Truth is not relative, and Excluded Middle holds without exception. (3) The concepts of ordinary thinking are not as such defective; in particular, they are not inconsistent. (4) That of which humans have knowledge, when they do, is fully real.

Analysis also has a positive dimension. The world is a totality of propositions, of which the constituents are concepts. Concepts are either simple or complex. If complex they decompose into simple concepts. Complex concepts are thus *analyzable* into simple concepts. Simple concepts have no analyses and are directly and accurately intuited by the knower. There is no difference between a true proposition and a fact; so Moore espouses not a correspondence "theory" of truth, but an identity "theory".[55]

5. Russellian Analysis

By 1903, Russell appears to have been well and goodly disabused of his former idealism. It was Moore's doing:

> On fundamental questions of philosophy, my position, in all its chief features, is derived from Mr G.E. Moore. I have accepted from him the non-existent nature of propositions (except such as happen to assert existence) and their independence of any knowing mind; also the pluralism which regards the world, both

---

[54]   Except, of course, for those parts of reality which are in their own right mental: minds and their contents.

[55]   See G.E. Moore (1899), 397-405; (1898), 176-193; and (1903), 433-453.

that of existents and that of entities, as composed of an infinite number of mutually independent entities, with relations which are ultimate, and not reducible to adjectives of their terms or of the whole which these compose. Before learning these views from him, I found myself completely unable to construct any philosophy of arithmetic, whereas their acceptance brought about an immediate liberation from a large number of difficulties which I believe to be otherwise irreparable. The doctrines just mentioned are, in my opinion, quite indispensable to any even tolerably satisfactory philosophy of mathematics...[56]

Russell credits Moore with the idea that a *philosophical definition* of a mathematical concept "professes to be, not an arbitrary decision to use a common word in an uncommon signification, but rather a precise analysis of the ideas which...are implied in the ordinary use of the term. Our method will therefore be one of *analysis*, and our problem may be called *philosophical*--in the sense, that is to say, that we seek to pass from the complex to the simple, from the demonstrable to its indemonstrable premisses."[57] Thus the method of philosophy as regards mathematics is analysis, and analysis decomposes target concepts into simple concepts which the knower directly intuits or of which he has a direct "non-sensuous" perception. On the face of it, the paradox of set theory is a genuine catastrophe. It means that we can't treat classes philosophically. This appears to bring Russell alarmingly close to Frege's Sorrow. There is no concept of class, or anyhow "I have failed to perceive any concept fulfilling the conditions adequate for the notion of *class*".[58] With that said, we have an explanation of the Riddle of Analysis. The Riddle, we said, is that upon discovery of the paradox Russell was in consistency bound either to revoke his commitment to philosophical analysis or to abandon all hope for a philosophy of mathematics. If reality is consistent, and if knowledge is knowledge of reality, and if knowledge of reality involves the direct apprehension of simple consistent concepts, then there can be no knowledge of classes.

Of course, Russell dissembled. He invoked a distinction between *philosophical analysis* and *mathematical analysis*. Having conceded the philosophical intractability of the concept of class, Russell makes room for a procedure which gives, for something that we might loosely and inaccurately call the concept of class,

---

[56]   Russell, *The Philosophy of Mathematics*, xviii.

[57]   Ibid, 2; emphasis added.

[58]   Ibid, xv-xvi.

"merely a set of conditions insuring its presence".[59] This "mathematical sense of *definition* is widely different from that current among philophers..."[60] In fact,

> it is necessary to realize that definition, in mathematics, does not mean, as in philosophy, an analysis of the idea of be defined into constituent ideas. This notion, in any case, is only applicable to concepts, whereas in mathematics it is possible to define terms which are not concepts.[61]

Moreover

> of the three kinds of definition admitted by Peano--the nominal definition, the definition by postulates, and the definition by abstraction--I recognize only the nominal....[62]

Russell opts for nominal definitions in mathematics in part because "definition by abstraction, and generally the process employed in such definitions, suffers from an absolutely fatal formal defect: it does not show that *only one object* satisfies the definition."[63]

A mathematical definition of something specifies a set of conditions which if satisfied at all is uniquely satisfied. The definition of Fregean sets cannot have been a philosophical analysis, because there is no concept of set. It cannot have been a mathematical definition either, because the paradox guarantees that the conditions on classes are not satisfied, hence not uniquely satisfied. For the Fregean set there is no "set of conditions" insuring its presence. Small wonder, then, that Russell's remedial treatment of sets via the theory of types should strike him as "harsh and highly artificial".[64] Even so, the method of mathematical philosophy can be seen

---

[59]   Ibid. Note that Russell says that the presence of a given class is "insured", not "ensured" or "assured". Russell was no illiterate; his choice of words is significant. In a *mathematical* definition of a class, the defining condition on the price for its "presence", they are the premiums which the theorist must pay.

[60]   Ibid, 15.

[61]   Ibid, 27.

[62]   Ibid, 112. It is interesting that Frege's Rule V for "sets" is sometimes called the Abstraction Axiom.

[63]   Ibid, 114. Emphasis added.

[64]   Ibid, 500.

as bearing correctively on the perception of indefinables. It can get us to recognize that "we are trying to perceive the *wrong thing*, and so redirect our attention."[65]

As we now see, Russell has a sort of solution to the Riddle of Analysis. Sets cannot be analyzed philosophically; since there are no sets, there is nothing for philosophers to be realists about in that connection. Even so, the *term* "set" (or "class") is amenable to mathematical "analysis", even when it produces results that are harsh and arbitrary. The arbitrariness in explicable by the fact that mathematical analyses are nominal definitions, and the harshness in explained by the unexpected difficulty of hitting upon a nominal definition that is consistently satisfiable. Thus Russell's implied answer to the Riddle of Analysis is this:

> **Russell's Answer:** The paradox of set theory doomed all prospects of an analytic theory of sets. However, it did not preclude, but rather simulated, a *stipulative* theory of sets, harshness, arbitrariness and all.

We now have the means to take the measure of Frege's Sorrow. In Frege's reply to Russell he shows concurrent disposition to what, in effect, are Russell's notions of philosophical analysis and mathematical definition. When Frege registered his concern that "not only the foundations of my arithmetic, but also the *sole possible* foundations of arithmetic seem to vanish,"[66] he shows himself possessed of the belief that the foundations of arithmetic are securable only by means of an analytic theory, which the paradox demonstrates is an entirely forlorn hope. But when in the very next line of his letter to Russell, Frege thinks that it may be possible to retain "the essentials of my proofs", perhaps, by reformulating his Rule V, he shows himself ready to turn a stipulative approach to sets.

Indeed this was precisely the approach that Frege did take. In the Appendix to volume two of the *Grundgesetze*, Frege replaced Rule V with the more complicated Rule V'. It didn't work. The inconsistency was still derivable, although it is evident that Frege was not aware of this at the time. Forty years later in the unpublished paper of 1924-25, Frege gives up on the set theory altogether: So it is evident that, as befits his own robust realism, in the end Frege's fidelity to analytic theories won the day, and lost it too.

Russell's technique of mathematical (as opposed to conceptual or philosophical) analysis is hardly reconcilable to his new-found affection for Moorean realism. In all essentials the theory of types is a reversion to idealism. Initially conceived of as fit subject for analysis, sets were stricken by paradox, and Russell's subsequent treatment of them qualifies as dialectical. One changes the subject matter of set

---

[65] Hylton (1990), 234. Emphasis added.

[66] Emphasis added.

theory by adding qualifications. One adds these qualifications in ways that preserve as much as is possible a linkage with the odd notion. In this nominal definitions are resorted to and, twice over, idealist themes are sounded. First, set theory if true at all is less true than a *consistent* analysis would have been. Second, the theorist's knowledge of sets is in part a matter of his own creation. Russell's recovery from the inconsistency of sets in *The Principle of Mathematics* is in all essentials indistinguishable from his recovery from the inconsistency space in the *Foundations of Geometry*.

6. Conclusion

Frege's Sorrow proclaims the death of analytic theories for sets, and with it the methodological disablement of intuitions and counterintuitions conceived of analytically. Historically, sets went stipulative among those who, like Russell, had surrendered to the exuberant blanishments of Moorean analysis. The way of stipulation is a way straight back to idealism, which is no bad thing perhaps. Whatever our general philosophical tastes, idealism not only makes sense of stipulation, it also offers hospitality to the dialecticized conceptions of intuitiveness and counterinuitiveness sketched in section 3. It gives a home to all that remains of these notions when *Frege's Sorrow* holds sway. One of the enchantments of idealism is that sets of beliefs satisfying its provisions do not routinely announce that they do. It is possible for a theorist whose theory is subject to idealist constraints to take his theory for real. This was the moral of the generation-shift in the history of sets, from those who saw sets as stipulations of those who now see them as winking away in Plato's (and Cantor's) Paradise. There is a a twofold moral in this. One person's stipulations may well be his granddaughter's analyses. There is little more to such transitions then the overtaking of acceptance by belief, something that befalls us like measles or a sneeze. The other part of the moral is that the granddaughter is always mistaken if idealism is true.

References

Aristotle (1964). *The Complete Works of Aristotle*, 2 volumes, ed. J. Barnes. Princeton: Princeton University press.
Barth, E.M. (1974). *The Logic of Articles in Traditional Philosophy*. Dordrecht and Boston: Reidel.
Barwise, J. (ed.) (1977). *Handbook of Mathematical Logic*. Amsterdam: North-Holland.
Benthem, J. van (1979). 'What is Dialogical Logic.' *Erkenntnis* 14 (1979).
Benthem, J. van (1983). 'Five Easy Pieces.' In: A. ter Meulen (ed.), *Studies in Model Theoretic Semantics*. Dordrecht: Foris.
Benthem, J. van and A. ter Meulen (eds.) (1985). *Generalized Quantifiers in Natural Language*. Dordrecht: Foris Publications.
Black, M. (1943). 'A Critical Review of [Quine's] Mathematical Logic.' *Mind*, 52 (1943).

Bradley, F.H. (1883). *The Principles of Logic*, two volumes (corrected impression: 1922). Oxford: Oxford University Press. Carlson, G.N. and F.J. Pelletier (eds.) (1995). *The Generic Book*. Chicago: University of Chicago Press.

Cantor, G. (1895-7). 'Beiträge zur Begründung der transfiniten Mengenlehre', parts 1 and 2. *Mathematischen Annalen* 46 and 49 (1895-7); translated as *Contributions to the Foundations of the Theory of Transfinite Numbers*. Chicago and London: Open Court (1915); reprinted New York: Dover (1952).

Forrest, P. (1988). *Quantum Metaphysics*. Oxford: Blackwell.

Forti's, B. (1897). 'Una questione sui numeri transfiniti.' *Rendiconti del Circolo mathematico di Palermo*, 11 (1897); translated as 'A Question on Transfinite Numbers'. In: J. van Heijenoort (ed.) (1967). *From Frege to Gödel*. Cambridge, MA: Harvard University Press.

Fraassen, B. van (1975). 'The Labyrinth of Quantum Logics.' In: C.A. Hooker (ed.). *The Logico-Algebraic Approcah to quantum Mechanics*, vol. 1. Dordrcht: Reidel.

Frege, G. (1893-1903). *Grundgesetze der Arithmetik*. Jena: Pohle; partial translation as *The Basic Laws of Arithmetic: An Exposition of the System*. Berkeley and Los Angeles: University of California Press.

Gilbert, M. (1993). *On Social Facts*. Princeton: Princeton University Press.

Grattan-Guinness, I. (1978). 'How Bertrand Russel Discovered His Paradox.' *Historia Mathematica*, 5 (1978).

Hallett, M. (1984). *Cantorian Set Theory and Limitation of Size*. Oxford: Clarendon Press.

Heijenoort, J. van (ed.) (1967). *From Frege to Gödel*. Cambridge, MA: Harvard University Press.

Hermes, H., et. al. (eds.) (1976). Gottlob Frege, *Wissenschaftliche Briefwechsel*. Hamburg:

Hermes, H., F. Kambartel and F. Kaulbach (eds.) (1979). *Posthumous Writings*. Chicago: University of Chicago Press.

Hilbert, D. (1976). 'Letter to Frege.' In: H. Hermes et. al. (ed.), *Wissenschaftlicher Briefwechsel*. Hamburg:

Hilbert, D. (1899). *Grundlagen der Geometrie*. Leipzig: Teubner; translated as *The Foundations of Geometry*. Chicago: Open Court (1902).

Hylton, P. (1990). *Russell, Idealism and the Emergence of Analytic Philosophy*. Oxford: Clarendon Press.

Jourdain, P.E.B. (trans.) (1915). *Contributions to the Foundations of the Theory of Transfinite Numbers*. Chicago and London: Open Court; reprinted New York: Dover (1952).

Martin, D. (1970). 'Review of Jean van Heijenoort (ed.), From Frege to Gödel.' *Journal of Philosophy*, 67 (1970).

Meulen, A ter (ed.) (1983). *Studies in Model Theoretic Semantics*. Dordrecht: Foris.

Moore, G.E. (1898). 'The Nature of Judgement.' *Mind* (1898).

Moore, G.E. (1899). 'Critical Notice of Russell's *An Essay on the Foundations of Geometry*.' *Mind* (1899).

Moore, G.E. (1903). 'The Refutation of Idealism.' *Mind* (1903); reprinted in G.E. Moore, *Philosophical Studies*. London: George Allen and Unwin (1922).

Peirce, C.S. (1888 and 1897). A series of papers published in *The Popular Monthly*, 1897 and 1888.

Quine, W.V. (1951). 'Review of Geach, Subject and Predicate.' *Journal of Symbolic Logic*, 16 (1951).

Quine, W.V. (1966). *Selected Logical Papers*. New York: Random House.

Quine, W.V. (1987). *Quiddities*. Cambridge, MA: Harvard University Press.

Russell, B. (1900). *A Critical Exposition of the Philosophy of Leibniz*. Cambridge: Cambridge University Press.

Russel, B. (1903). *Principles of Mathematics*. Cambridge: Cambridge University Press.

Russel, B. (1987). *An Essay on the Foundations of Geometry*. Cambridge: Cambridge University Press.

Sher, G. (1991). *The Bounds of Logic*. Cambridge, MA: MIT Press.

Shoenfield, J.R. (1967). *Mathematical Logic*. Reading, MA: Allison-Wesley.

Sorensen, R.A. (1988). *Blindspots*. Oxford: Clarendon Press.

Tarski, A. (1983). *Logic, Semantics and Metamathematics*, 2nd edition, J. Corcoran (ed.). Indianapolis, IN: Hackett.

Woods, J. (to appear). *Paradox and Paraconsistency: Conflict Resolution in the Formal Sciences*.

Woods, J. (to appear). *Aristotle's Earlier Logic*.

Zermelo, E. (1904). 'Beweis dass jede Menge wohlgeordnet werden kann.' *Mathematischen Annalen*, 59 (1904).

Zermelo, E. (1930). 'Über Grenzzahlen und Megenbereiche: neue Untersuchungen über die Grundlagen der Mengenlehre.' *Fundamentae mathematicae*, 16 (1930).

Argument Based Reasoning: some remarks on the relation between Argumentation Theory and Artificial Intelligence

Richard J.C.M. Starmans,
Tilburg University

Abstract

In this paper, the relation is discussed between modern argumentation theory and formal theories of commonsense reasoning as they have been developed in Artificial Intelligence. On the one hand, argumentation theory can benefit from the concepts of inference developed in non-monotonic logics, since logic plays a major role in argumentation and new concepts of logic and inference have been developed. Conversely, Artificial Intelligence has much to learn from modern argumentation theory. E.g., ideas from dialectics, debate, and the legal field can be welcome contributions to the work in defeasible logic.

1. Introduction

In this paper, I will discuss the relation between modern argumentation theory (AT) and formal theories of commonsense reasoning as they have been developed in Artificial Intelligence (AI). In the literature both disciplines are hardly related; they are studied rather isolated in separated communities with their own objectives, methods and intellectual background, without much cross-fertilization. Despite substantial differences, there are some interesting relations as well, some of which will be discussed here, by focussing on two consecutive developments in AI. The aim of this article is threefold.

First, I will argue that both AI and AT, despite their intrinsic differences, emerge from the same tradition in reasoning. This will be elucidated by referring to their attitude towards classical formal logic. However, a straightforward comparison between AI and AT is rather intricate, due to the fact that this concept of "commonsense" seems to be fairly ill-defined. Therefore, a brief conceptual analysis is required.

Secondly, I will make the relation between both disciplines more precise by focussing on two successive trends in AI: the rise of so called non-monotonic logics (NML) and subsequently, work from the more philosophically inspired "defeasible

reasoning" community (DR). The former gave rise to very radical new concepts of formal non-deductive inference, the latter invoked an approach of "dialectical" or "distributed" inference in AI, which we call Argument-Based Reasoning (ABR), and which gained a considerable impact in AI recently.

Thirdly and finally, I will briefly evaluate ABR by comparing it with some fundamental insights and results of modern argumentation theory. As will be shown, unfortunately many of these ideas seem to be rather neglected in ABR. This, despite the fact that researchers in the field of defeasible reasoning explicitly maintain to adhere to principles of argument, dialectics and debate. We will propose some minimal requirements, which AI-systems should meet, to deserve the predicate "argument based". More generally, it is argued that such a worked out account of argumentation is an excellent candidate to give shape to a conceptual model which should underlie each formalism.

2. "Resource-bounded" inference

Since the study of reasoning has exceeded the realms of philosophy, a great variety of disciplines takes interest in the subject as their primary or one of their primary objectives. But most noticeably, since World War II, several disciplines arose, predominantly concerned with the way reasoning is actually performed in everyday life. Performed by agents which are "resource-bounded" in several ways. Because they have to deal "rationally" with far from perfect knowledge; inconsistent, uncertain, incomplete or weakly defined data. But also because they have limited time, storage of information and retrieval techniques at their disposal.

Clearly, this approach opposes to more canonical and idealized concepts of knowledge and inference, as they have been adopted traditionally in logic and epistemology; concepts, usually presupposing consistent, well-defined knowledge and perfectly rational and introspective agents, neither constrained by limited resources, nor by biased perception. And, not to forget, agents with "belief-states" which are closed under logical consequence. As a result of this traditional approach, reasoning abilities can be judged fully according to *normative* models which are based on these idealized assumptions. Models like, for example, classical logic (Frege), classical probability-theory (Kolmogorov) and the axioms of utility (Von Morgenstern).

Unsurprisingly, this concern with "real-life" inference, invoked new approaches, challenging both these idealized models as well as their normative standards. Several interesting developments depict this.

A well-known example is the rapid emergence of decision-theory in psychology in the sixties. Ideas, taken from and inspired by utilitarism and mathematical economists, started obtruding psychology, a discipline until then mainly dominated by behavioristic approaches. The rise of cognitivism favored approaches determined by the point of departure that persons and groups essentially should be considered

and studied as " rational actors", making judgements and decisions, deliberating on possible choices and actions. However, numerous experiments in cognitive psychology showed that people usually do not perform reasoning tasks in accordance with Morgenstern's axioms of utility, Kolmogorov's axioms of probability and Frege's logic. This controversy between the proposed normative models and actual practice in reasoning and decision-making gave rise to all kinds of endeavors to explain or deny the phenomenon or to correct the "fallacious" reasoner. Many alternative approaches were launched, both formal and informal ones. Varying from general ideas of bounded rationality (replacing the idealistic notions) to influential theories like the "bias and heuristic" approach of (Kahneman, Tversky 1982). Their work stimulated further empirical research in reasoning, deepened our insight in the nature of inference. Among other things cognitive research showed us to what extent cognitive limitations restrict the choiceproblems, how cognitive and social stress determine the outcome of a process of reasoning, how preferences are manipulable by tricky representations ("framing"), and that intuitions which are quite plausible in isolation, appear to be inconsistent together, as well as other "Arrow"-like results. This enumeration can easily be extended, but in my opinion a more important feature of this research field is that it encouraged the aspiration to provide all kinds of practical and realistic models of inference, including the "Prospect Theory" of Kahneman and Tversky (1979, 1984), the "Conflict Model of Decision Making" by Janis and Mann (1977), Montgomery's "Dominance Search Theory" (1989), and Beach and Michell's "Image Theory" (1987).

But also AT and AI indubitably emerged from this tradition of "resource-bounded" inference. The former, first and foremost because of its point of departure that we should study real-life argument, as it has been produced in ordinary language in ordinary discourse. Thus, clearly neglecting positivistic ideas about inappropriateness of natural language. But, also because of the conviction that argumentation cannot be understood, modelled, or judged properly without specifying "social" parameters. In fact, the discipline arose from the criticism against the concept of argument as adopted and taught in traditional philosophical textbooks on logic and argument, in which these social aspects were commonly neglected and deductive validity was the only standard in the evaluation of argument.

Pioneering work of Toulmin, Perelman, Naess and Hamblin encouraged a (new) interest in concepts of law, rhetoric, dialectics. Unfortunately, their concern was predominantly determined negatively and the criticism of especially Toulmin and Perelman proposed a total and unnecessary rejection of classical logic. However, these insights about language use and social determination were continued when the discipline became institutionalized and mature in the seventies and eighties and developed as a highly interdisciplinary research area, studying argumentation from a general linguistic and communicative point of view. Regarding the role of logic, however, the situation isn't too obvious yet, despite a widespread agreement that

classical logic is not suitable as a or rather the model for argument. In America there is this informal logic movement. Some informal logicians advocate a total rejection of formal methods in the study and analysis of argument. Their attacks, however, are not totally free from the "strawman fallacy", crippling the scope of logic and neglecting almost entirely more current approaches and results in the field, some of which will be discussed here.

In general, a more mitigated point of view seems to be commonly accepted in argumentation theory, nonetheless. Formal techniques as well as classical logic do play an important role in the field. For example, take the plea of Woods (1989) for formal methods in the analysis of fallacies, or the theory of formal dialectics developed by Barth and Krabbe (1982). Or, consider the role of "reconstructive" deductivism in pragma-dialectics (van Eemeren and Grootendorst, 1984) and in Groarke (1992).

Thirdly and finally, AI can be put into this tradition; not just because of its practical objective to formalize and, more importantly, to simulate reasoning processes by implementing algorithms in computers. But mainly, because their main point of departure was the fairly obscure conjecture that computers should "follow" inferences made by human beings; and because human beings typically are supposed to possess and exhibit "commonsense", a new term was established, usually indicated as commonsense-reasoning. Now, this notion may generally be considered as highly ill-defined and ambiguous (we will discuss it in the next section), nevertheless, formalization of commonsense reasoning undoubtedly became the main topic in AI (and sometimes even is identified with it). And, more importantly, it initiated the development of numerous formal non-standard logics, due to the generally conceded point of view that classical (logical and probabilistic) formal models are inappropriate to formalize this "commonsense" reasoning.

So at a very general level decision theory in cognitive psychology, AT and AI, emerge from the same tradition; they all adhere to the same principles and refrain from an unconditional application of and addiction to classical normative standards. Despite their differences, neither AT nor AI does reject classical logic fully, nor do they accept it as the underlying model of inference. This role of logic will be discussed more in detail in the next sections. First, some remarks about the notion of commonsense must be made.

3. What is commonsense reasoning?

Clearly, this intention of AI to simulate reasoning as such, is far from surprising, because AI wants to simulate and mechanize intelligence and most definitions of this concept adopt the ability to reason as a crucial, necessary aspect of it. However, the addition of the attribute "commonsense" raises some problems. Due to this

problematic nature of the concept, a fully proper use of the term certainly would deserve an accurate conceptual analysis. However, here we have to restrict ourselves to some short remarks.

First and foremost, it must be observed that the notion particularly arises in AI and is used in a context where formalization and mechanization of reasoning are intended. So, commonsense usually arises when formalization of inference is meant! But of course, the question remains what kind of inferences, underlying the concept, we really try to capture. Is there a type of inference which can be characterized as not-commonsense? Is commonsense reasoning not just reasoning? Because of its importance in AI, it seems reasonable to demand that in order to formalize commonsense, first an intuitively satisfiable account of this type, or rather, these types of reasoning, must be available. An account, which describes these aspects and types, which can be distinguished cognitively and epistemically from other types of inference.

Unfortunately, this is not the case. The use of the term is mainly motivated by practical aims. It does not refer to well-distinguished type of inference, nor is there a general account or typology of reasoning, decision-making or rationality underlying the concept.

Sometimes, the term serves as a garb for all kinds of more or less lucid concepts like plausible reasoning, possibilistic reasoning, default-reasoning, non-monotonic reasoning, and so on. Because these concepts often overlap, are considered to be synonymous or complementary and often are even less well-defined, neither such an enumerative definition seems satisfiable. Usually, one of the aspects of "resource-bounded" inference, we gave in section 1, is intended. So, what all these notions have in common, is that they do not indicate well-defined and distinguished types of inference, but rather primitive, self-evident and directly observable aspects of reasoning, which cannot be modelled properly with the traditional formalisms. The term denotes all kinds of knowledge as well as reasoning tasks. Sometimes, it refers to the state of the data (incomplete, inconsistent, vague), sometimes it refers to the type of data (pre-scientific, things everybody knows), sometimes it directs to the techniques (heuristics) people use. Due to the fact that the notion of commonsense as such doesn't provide us with a clear account of reasoning, we will have to borrow it elsewhere. As we will argue in the next sections ideas from both AT and decision-theory in psychology will play a role here. Now, I will make the relation between both AT and AI somewhat more precise by focussing on two succeeding developments in AI.

4. Non-monotonic logic

4.1 What is non-monotonic logic?
Non-monotonic logics were developed to model patterns of commonsense reasoning. Systems like default-logic, developed by Reiter (1980), circumscription,

launched by McCarthy (1980) or auto-epistemic logic (Moore, 1985) are especially intended to capture types of inference which are considered to be "in accordance with commonsense". And, more in general, all non-monotonic logics are usually characterized in the well-known textbooks like Lukaszewicz (1990) and Brewka (1991) as "formalization of commonsense". Due to the characteristics of knowledge, we described above, agents sometime have to jump to conclusions and, consequently, have to withdraw these conclusions, once faced with additional data. The basic assumption of non-monotonic logicians is that these patterns cannot be modelled properly by classical logic, because of its monotonicity.

If a conclusion φ is derivable from a theory Σ, where Σ represents a set of premises, then φ is also derivable from every superset of Σ. So new information cannot invalidate old conclusions. Of course, many intelligent tasks do require this possibility. People use representation conventions for efficient storage of information, they must be able to handle rules with exceptions and generic sentences, they must be able to deal with inconsistencies in a reasonable way. It can easily be verified that a straightforward application of classical logic as the underlying model of reasoning, immediately runs into trouble. Therefore an inference-procedure is required which is to a certain extent inconsistency-tolerant, context-sensitive, allows sophisticated representation conventions, and admits generic information. So new, often ingenious, non-monotonic (and therefore non-deductive!) inference relations had to be developed and many, quite dissimilar formalisms have been proposed. For a concise but lucid overview, see Brewka (1991). It goes without saying that we cannot discuss them here and indeed a more global and indefinite characterization of NML is hardly imaginable, but here we will concentrate on two aspects of NML which are relevant for the field of AT.

The first facet is predominantly methodological in nature. All kinds of objections against classical logic, which have been put forward by Toulmin and Perelman in the past and more recently by informal logicians, show a plain resemblance with the points of departure in NML; the rigidity of the entailment relation, the small applicability of deductive validity, the impossibility to deal with exceptions and less defined information. A closer inspection shows that for example Scriven's plea for a "probative" logic (Scriven 1987) as well as Johnson's objective to "naturalize" logic (Johnson 1989) perfectly match the objections of NML against monotonic logic as they have been put forward in the early work of Reiter (1980) and McCarthy (1980). Next, the study of fallacies has to be mentioned. In AT patterns of reasoning like the "ad verecundiam", the "ad ignorantiam", the "ad consequentiam" or the hasty generalization are no longer necessarily fallacious; their acceptability depends on the purpose of a certain argumentative situation or process. But these are exactly the types of inference one is trying to formalize in AI. For example, from this point of view Moore's autoepistemic logic is nothing but a formal account of the ad ignorantiam "fallacy", whereas default logic, which underlies jumping to the conclusion, matches the hasty generalization. And according to Walton (1989)

the underlying principles of traditional models for expertsystems are those which have been recognized in AT as the "ad verecundiam". But, in our opinion there is a second reason why NML cannot be ignored in AT. NML induced new concepts of formal logic, which go beyond deductivism. So in a discussion about the role of logic in a theory of argument, just the conviction that classical monotonic logic is less important, cannot be a reason to reject formal logic in general, since NML gave shape to many alternative formal systems, built on analogous underlying principles. Rather than specifying these relations extensively, we restrict ourselves to one example, which deals with the beforementioned representation conventions in NML and the problem of unexpressed premises in AT.

4.2. Unexpressed premises and the qualification-problem.

4.2.1 Unexpressed premises
The analysis of unexpressed, implicit or "hidden" premises is one of the most complex issues in argumentation theory. First and foremost, there seems a widespread agreement that their occurrence is a quite natural and normal phenomenon, that matches perfectly with for example Gricean and Searlean principles of co-operativity and indirect speech acts. But on the other side, in the analysis of argumentation, the phenomenon gives rise to a variety of questions, thus rather establishing a complex of problems, than a distinguished one. Nevertheless, the following questions seem to appear in many discussions.

1) Is there truly something missing, or hidden, something which must be clarified or made explicit by some new "premises", behaving like "gap-fillers" ?
2) Given the fact that there is some information which has to be made explicit, how and to what extent must this be done and what linguistic or meta-linguistic expression or construction is suitable for these ends?
3) Can this process of making information explicit be performed in a neutral, un-ambiguous way, following a generally acknowledged procedure?

It goes without saying that we cannot attach all implications of these problems here. Regarding the first question, despite all controversies, there seems an agreement that there is information which has been used by the sender, though it has not been articulated or represented. Information, that definitely must be acknowledged by the receiver, to fully comprehend or evaluate the argument. Obviously, in the analysis of argumentation, the structure of this "link" must be made explicit, whether one is interested in arguer's commitments (like in pragma-dialectics), in the arguer's intentions (as in epistemics) or maybe in the relation between unexpressed premises and presuppositions.

With respect to the second question we restrict ourselves to the statement that often a conditional is required, an if...then-construction, which makes explicit the

"hidden" information. The feasible successive question is of course, what logical properties this conditional is supposed to exhibit. The dominant, though not universally accepted deductivistic approach will require a material implication, thus settling validity in a classical sense.

Also question 3 is far from trivial. It seems rather difficult to develop some procedure to find the implicit premise. Usually, several candidates are available and there is no reason to assume in advance that every arbitrary person will "derive" the same rule.

Now, in order to demonstrate the significance of these three questions in the field of AI, we briefly discuss one of the main problems in knowledge representation, the so-called qualification problem.

4.2.2. The Qualification-problem

This phenomenon that implicit information is used as well as the conviction that it must at least be possible to supply additional information to conceive a fully correct and deductively valid inference, is nothing new. Even in more traditional applications of logic in computer science it can be found. For example, in database theory. A database is built for the effective storage and quick retrieval of huge numbers of data. Therefore, the knowledge must be represented efficiently. As an example consider Δ to be a database, containing the following facts:

$$\Delta := \{\, likes(john, money), likes(mary, art)\,\}$$

Given the query "likes(peter, football)?" the intended and expected answer is of course "no". However, this inference goes far beyond classical derivability; it is not a logical consequence of theory Δ. If we consider the database as a logical theory and we demand a proper inference, additional axioms are required. We need the Unique Name Assumption (UNA), which states that different names (constants in the language) denote different objects or entities in the domain. Furthermore, we require the Domain Closure Assumption (DCA), which states that all individuals are assumed named. Finally, we need the Closed World Assumption (CWA), by which all instances of the relation "likes" are assumed to be derivable from the theory.

(UNA) $(john \neq money) \wedge (john \neq mary) \wedge (mary \neq art) \wedge ...)$

(DCA) $\forall x [\, x = john \vee x = mary \vee x = money \vee x = art\,]$

(CWA) $\forall x \forall y [\, likes(x,y) \rightarrow (x = john \wedge y = money) \vee (x = mary \wedge y = art)]$

Now, from $\Delta \cup \{\, UNA, DCA, CWA\,\}$ we can derive correctly that Peter doesn't like football. Clearly, these premises play a role, though they have not been modelled explicitly by the knowledge engineer. In fact, many non-monotonic formalisms (for example circumscription) are based on this idea of "repair"; a non-monotonic inference based on Σ can be reduced to a monotonic inference based on a superset of Σ.

Thus, assuming that it is possible and worthwhile to represent this information in the same language as the other premises.

However, usually it is not so easy to "repair" this inference; it often seems impossible or at least not useful to specify all information explicitly in the same language. Suppose we have knowledge about starting the engine of a car. We know that performing this action can only be successful if many preconditions are met. There must be an engine in the car, the car must not be stolen, the engine must be installed properly, there must be enough fuel, and so on. Firstly, it must be noticed that though we are quite aware of the relevance of all these preconditions, we do not check them all. For this seems fairly impossible. The list of conditions can be infinitely long. We assume, by default, that all the requirements have been fulfilled, as long as we have no information to the contrary. Perhaps we only check some of the most vital preconditions. Secondly and more importantly, if we want to represent our knowledge about starting an engine, we are unable to list all these conditions as well! Obviously, this is a very fundamental issue in knowledge representation, which is known as the qualification problem: one cannot specify explicitly all qualifications, which are required for successfully performing an action.

Suppose, for the sake of argument, that we want to represent the information in a rule. The most suitable way, of course is to specify all the preconditions in the antecedens and the conclusion ("the engine will start") in the consequens. However, as soon as we learn an exception to the rule, the negation of this formula has to be added to the antecedens, as well as the exceptions to the exception and so on. The problem arises if one tries to express "generic" knowledge by means of a conditional. Even if it would be possible to list all relevant qualifications, this would seriously damage the modularity of the system (which is of vital importance in every knowledge representation language). Furthermore, this is not what we want, if we develop a rule; we want to use it without the obligation to derive the negation of all those exceptions first.

The relation between the problem of unexpressed premises and this qualification-problem will be apparent, if we take the questions of section 4.2.1 into account.

Regarding the first question, again we use tacit information in our inference. Information about how we might jump to conclusions, information (assumptions) about completeness of data or information about significance of certain preconditions and contextual information. The analogy is clear. The arguer is fully licensed to use this unexpressed premises, but he is committed to them as well. The knowledge engineer doesn't need to represent explicitly some information, though it must be encoded someway in the system and it certainly has to be made explicit if a certain inference has to be motivated or explained!

As far as the second question is concerned, the qualification problem impedes a straightforward application of the material implication, because this conditional leaves no room for generic knowledge or exceptions. Traditional use of this

conditional demands a full enumeration of all these preconditions (qualifications) in the antecedens. Once these preconditions have been fulfilled, the conclusion (a statement, an action, a decision) is a logical consequence. Such a conclusion cannot be withdrawn.

So, logicians in NML developed several logics with alternative if...then-constructions. For example, default logic appeared a very powerful tool in modelling these items and certainly is successful in representing several types of commonsense knowledge.

These so-called defaults are a kind of inference-rules which enable us to infer a conclusion without a full specification and check of all preconditions. So the question is not just if there is something which has to be made explicit, but rather to what extent it must be made explicit. With respect to the second question the relation with the problem of unexpressed premises is straightforward as well. The from a logical point of view conceivable question is of course: Should one actually try to make a formally invalid argument, deductively valid by adding premisses? Those who are not satisfied with the deductive approach should be aware of the fact that other logics, with different conditionals, are available and perhaps useful, for example default logic. In our opinion, this example shows how AT can take advantage of the new concepts of inference, as they have been developed in NML. Conversely, we believe that, for entirely different reasons, AI has to pay attention to developments in the field of AT.

5. Defeasible Reasoning

Roughly spoken, work in the defeasible reasoning community is based on principles which are almost canonical in traditional philosophy (dialectics, epistemology), rhetoric, legal theory, dialogue-logic, and theory of argumentation, but only recently gained attention in AI. The idea is that (many types of) reasoning can be considered as a process of constructing, comparing and weighing arguments for and against a certain conclusion. Human beings typically construct and evaluate arguments when they explore available knowledge, make a decision, try to persuade an opponent, jump to conclusions, or have to deal with inconsistent information. Arguments are meta-linguistic constructions, a kind of "defeasible" and "non-demonstrative" proofs, that give a certain support for a conclusion, but not a definite warrant. This defeasibility is due to the fact that arguments (unlike proofs in mathematics!) typically interfere. They can be questioned, attacked by counterarguments, "overruled" by "better" arguments with more conclusive force, defeated or reinstated.

A feasible presupposition for this 'deliberate' reasoning is a distributed environment, i.e. the introduction of two or more agents, performing the process of reasoning. Agents with possibly dissimilar dialectical roles, a non-symmetrical distribution of the burden of proof and maybe even with different, conflicting norms and standards about rationality and appropriate argument. From this point of view, more than the proposition, the notion of argument is the key-notion in the

study of reasoning. Consequently, every system that wants to formalize reasoning has to deal with this concept. It should indicate what arguments are, how they interfere and how they should be weighted or evaluated. Obviously also related notions like rationality, burden of proof, (violation of) norms, behavior of agents, and (types of) debate must be incorporated in any theory or framework for reasoning.

Despite both its intuitive nature and its longstanding tradition in philosophy, the notions of argument and Argument-Based-Reasoning (ABR) do not play a significant role in classical logic, nor in most of current non-monotonic formalisms.

Indeed, to a certain extent, ABR can be considered as a reaction of philosophically oriented researchers in AI on NML. And, it seems more related to ideas of inference and rationality, as put forward in decision-theoretic models of cognitive psychologists. Many objections against NML can be put forward, both epistemical and computational ones. Here we restrict ourselves to those which were launched by John L. Pollock and Ron Loui. According to the former "current theories of non-monotonic reasoning coming out of AI are simplistic and overlook much of the fine structure of defeasible reasoning" (Pollock 1987: 482). He argues that such a theory must be philosophically adequate, but not just that. In his opinion a "satisfactory theory of defeasible reasoning ought to be sufficiently precise that it can be implemented in a computer program. Constructing such a computer program and seeing that it does the right thing, will be a useful test of the theory, and simultaneously a contribution to AI". Ron P. Loui is even more explicit in his habitually highly polemical essays (Loui 1990, 1991). In his opinion NML and defeasible reasoning have few things in common and emerge from two logical traditions. He is convinced that the approach of NML will appear unfertile in the end, for the reasons we listed in the beginning of this section.

Besides the two features we already mentioned, the adherence to arguments and counterarguments (deliberation), as well as the distributed environment, ABR can be characterized by its focussing on procedural aspects of reasoning and its objections against the domination of declarativism in logic.

Other researchers seem to be more inspired by the work on epistemology, conditionals and dialectics, than by NML as well. Among others, Nute (1988), Pollock (1987), Konoligue (1987), Vreeswijk (1993) and Simari (1991) adopted the notion of argument as the cornerstone of reasoning and they gave shape to this ABR-approach. But also in the legal field important contributions have been made, for example Gordon (1993) or Hage (1992).

In a certain sense this tendency of the defeasible reasoning community towards epistemic and dialectical principles, has caused a further deviation from the more classical logical approach of reasoning. Unlike the standard non-monotonical formalisms one does not try to capture intuitive correct plausible inferences, by slight modifications of inference-relations, nor by defining preference-relations on

the models of first order theories. Also the adherence to the process of reasoning, and consequently the limited importance of model-theoretic semantics, is remarkable. This deviation from classical logic is not unproblematic of course and it has forced many of the beforementioned authors 'to begin from scratch'. Moreover, many questions can be raised against their ambitions. Do these formalisms deserve the attention Loui claims and are they free from the drawbacks, that characterize NML. Can these ABR-formalisms, which definitely are studied less intensively than those in NML, overcome the computational problems? Are the underlying principles and presuppositions about logic and reasoning in the defeasible reasoning community indeed incommensurable with for example Reiter's default-logic, McCarthy's circumscription, or Moore's auto-epistemic logic? Or can these formalisms be modified, extended or generalized and finally adopted in an ABR-framework? Apart from these aspects, here we restrict ourselves to the concept of argumentation that has been used in ABR. Due to this adoption of the concept of argument, the aspect of deliberation, the distributed environment and the adherence to procedural aspects, ABR seems, even stronger than NML, meet objections of some researchers in the field of AT against logic.

6. What's wrong with ABR?

Unification-oriented researchers in the field of reasoning may be pleased with this development, since the "new" paradigm seems to bridge many gaps. Indeed, we believe that Loui, Pollock and others did succeed in establishing highly important items; mainly the points of deliberation, the distributed environment as well as their adherence to procedural aspects. Nevertheless, we believe some critical remarks are required here as well. In our opinion, the high expectations of ABR as a promising paradigm in commonsense reasoning cannot be fulfilled by a naive eclecticism of unconnected "philosophical concepts" about argument and debate, neither by a revival of ancient rhetoric, a reintroduction of Heracleitean, Hegelian or Popperian views on dialectics, nor by a new application of the Toulmin-model. Also the in itself quite important contention that the types of inference AI-researchers are trying to capture, can best be modeled by a process of constructing and weighing arguments and counterarguments, is not satisfiable. A more solid foundation is required.

A formalism that is really supposed to model some types of reasoning demands or rather presupposes a general theory of reasoning as an underlying model.

So analogously, any formal system that claims to be argument based, should be built on a general theory of argument, that consequently serves as a conceptual model. With "general" we intend to express that we need a proper account of the phenomenon of argumentation, the concept of argument, its role in a theory of reasoning, its structure, its purpose, its scope and a description of the social context in which it arises, including all *communicative, social* and *decision-theoretical*

parameters. And, obviously a proper account of related concepts that arise in this environment as well; relevance, rationality, burden of proof. By explicitating all kinds of methodological, theoretical or other assumptions, this conceptual model can determine all kinds of *design decisions* that must be made in developing the formalism.

With that end in view, it must be noticed that none of the existing studies in ABR is based on ideas from modern AT! The most recent bibliographic reference is usually Toulmins famous study "The Uses of Argument", which is clearly unsatisfiable due to the rather primitive concept of argument Toulmin adopts. Even features and assumptions which underlie fairly all research in AT, are commonly more or less disregarded in ABR. In general one could state that the majority of current ABR-systems lacks:

A) a Functional Theory of Language-use. Due to the generally accepted point of view in AT that argument is a form of language-use and consequently should be studied at the language-level, a sufficiently rich theory of language and language-use is needed.
B) a concept or Rationality or Reasonableness. Despite the fact that it has been generally acknowledged that the concept of truth, as developed by Tarski, has to be replaced by a notion of rationality, few ABR-researchers elaborate this concept. It is a fairly primitive concept, without any critical-rationalistic, or decision-theoretic justification.
C) a notion of Relevance: Notwithstanding the fact that this concept is commonly considered to be one of the essential problems in the process of argumentation, as far as we know, no analyses of this topic are available in the ABR-community.
D) extra-logical criteria for soundness of argumentation. In ABR, there is a strong tendency, to use syntactic criteria in determining whether an argument has been defeated by a counterargument. Often these considerations are based on specificity. However, as has been put forward by others, for example Vreeswijk (1993), the scope of this device is limited. In our opinion, a full specification of social parameters is required. We need information about the (type of) debate, the initial knowledge of the agents, as well as a specification of their individual aims and the collective goal, the side-effects of the argument-moves, in order to determine whether an argument is permissible, successful or warranted.
E) Procedural Rules. Clearly this point is closely related to the previous one. If one agrees that the underlying concept of ABR is a debate, then a formal account of this is required. In order to guarantee a successful debate, several rules have to be specified. Some of these are general rules for a discussion, whereas others are highly domain specific and determined by the social para-

meters we gave in D. Unsurprisingly, in the formal work in the legal field interesting results on this subject matter have been established, due to the highly institutionalized character of legal reasoning. Again, with respect to this aspect much work has to be done.

Now both the absence of references to recent literature in AT as well as the small concern with beforementioned items, indicate a more general problem. Usually, ABR-formalisms lack a general theory/framework of argument, underlying the formalism. Partly due to this attitude, some notions have not been developed properly yet and sometimes lack conceptual clarity. And due to this lack of a general theory, much research still uses individual and isolated intuitions about some benchmark-problems as a startingpoint, developed to describe intuitive "would be" plausible patterns of reasoning (whether they are about flying penguins, pacifistic republicans or employed students).

Building a formalism on a general theory of argument, we are not primarily interested in those specific isolated "benchmark-problems". A main motivation for this is that in our opinion the most important application of ABR can be found in situations which are far more complicated than those benchmark-problems; situations with several opposite interests and goals, inconsistent information and preferences, as well as different procedures for obtaining specific data. Situations in which, as the decision-theoretical literature teaches us, a full reliance on what is supposed to be evident, reasonable, commonsense or "in accordance with intuition", is very hazardous and unsatisfiable.

Now, of course there is no reason to exclude in advance any worked out theory of reasoning to serve as such a conceptual model, including the theories in the field of psychological decisiontheory, which we mentioned briefly in section one. Here we only maintain that adhering to argument or debate demands a theory of argument that is sufficiently rich and general in the sense that we described above.

In our doctoral dissertation (to appear) we make a careful attempt to give shape to such a conceptual model by adopting the theory of pragma-dialectics based on van Eemeren and Grootendorst (1984, 1992) and we investigate to what extent these ideas are both general and specific enough as a basis for an ABR-formalism.

The main motivation for this choice is the fact that pragma-dialectics does provide a general and detailed account of argument. It has been applied in several distinct domains and by adopting it many of the beforementioned requirements can be fulfilled rather easily.

It gives four parameters of the concept of argument (functionalizing, externalizing, dialectifying, socializing), it provides standards of rationality, relevance, problemsolving validity. Furthermore, the way pragma-dialectics functionalizes, externalizes, socializes and dialectifies argumentation, has some practical ad-

vantages. The idea of performing a debate to resolve a conflict of opinions in a rational way captures several distinct reasoning tasks.

The goal-oriented and procedural character of argumentation, seems a natural "instance" of the procedures of search which are desired in AI. Also the adherence to commitment (based on performed speech-acts) rather than to intensionality (a much less applicable concept) must be mentioned. Finally, the theory indicates how several types of rules (a code of conduct) and higher order conditions can be applied, used or added to guarantee a successful debate.

However, it is quite obvious that we only raised some of the most elementary items here and most work in the field of ABR has to be done yet. The overall dichotomy seems clear. Can we find a concept which matches both insights and requirements from the field of AT, as well as those of AI? The success of this approach will depend on the extent to which logicians and AI-researchers will succeed in developing useful ABR-applications in a well-defined domain, rather than constructing new and dull toy-problems, which still seem so dominant in AI.

7. Conclusion

Though we are quite aware of the general and global character of this paper, we believe it suffices to show that despite or maybe even due to the vague character of the concept of commonsense, there are interesting relations with research performed in AT. Summarizing, we can state that these relations can be made specific at two levels. In my opinion, AT can benefit from the concepts of inference developed in NML. Since logic still plays a major role in argumentation, and given the fact that new concepts of logic and inference have been developed, any discussion in which logical matters are involved, should take these new concepts into account. Conversely, AI has much to learn about modern argumentation theory. Regarding the work in defeasible reasoning, we certainly can welcome ideas taken from dialectics, debate and the legal field. However, at the same time there is a little disappointment as well, since research in this field commonly seems to ignore much of the ideas and results of modern argumentation theory. As we stated already, this is regrettable since a worked out account of argument seems an excellent candidate for supplying a conceptual model underlying the formalism. Nevertheless, current research will show whether one will succeed in developing sufficiently rich concepts of argument or whether one uses the notion mainly metaphorically. A divergence in concepts of argument seems undesirable for any theory of reasoning.

References

Barth, E.M. and E.C.W. Krabbe (1982). *From Axiom to Dialogue: a Philosophical Study of Logics and Argumentation*, New York: Walter de Gruyter.

Beach, L.R. and T.R. Mitchell (1987). 'Image theory: principles, goals and plans in decision making'. In: *Acta Psychologica*, 66, pp. 201-220.

Brewka, G. (1991). *Non-monotonic Reasoning. Logical foundations of Common-sense.* Cambridge: Cambridge Press.

Eemeren, F.H. van and R. Grootendorst (1984). *Speech Acts in Argumentative Discussions; a theoretical model for the analysis of discussions directed towards solving conflicts of opinion*, Dordrecht: Foris Publications.

Eemeren, F.H. van and R. Grootendorst (1992). *Argumentation, Communication and Fallacies; a Pragma-Dialectical Perspective.* Hilsdale, N.J.: Lawrence Erlbaum.

Gordon, T. (1993). '"The Pleadings Game", Formalizing Procedural Justice'. In: *Proceedings ICAIL-1993.* Amsterdam: Compute-Law Institute VU, pp. 10-20.

Groarke, L. (1992). 'In Defense of Deductivism: Replying to Govier'. In: F.H. van Eemeren e.a. (eds.). *Argumentation Illuminated.* Dordrecht: SICSAT, pp. 113-122.

Hage, J.C. (1992). 'A dialogical model of legal reasoning'. In: C. Gruetters e.a. (eds.). *Legal Knowledge Based Systems,* JURIX 1992. Lelystad: Koninklijke Vermande, pp. 135-145.

Hamblin, C.L. (1970). *Fallacies.* London: Methuen.

Janis, I.L. and L. Mann (1977). *Decision Making: a psychological analysis of conflict, choice and commitment.* New York: The Free Press.

Johnson, R. H. (1987). 'Logic Naturalized; Recovering a Tradition'. In: F.H. van Eemeren, R. Grootendorst (eds.). *Argumentation: Across the Lines of Discipline.* Dordrecht: Foris Publications, pp. 47-56.

Kahneman, D. and A. Tversky (1982). 'Prospect Theory: an analysis of decision under risk'. In: *Econometrica*, 47, pp. 263-291.

Kahneman, D. and A. Tversky (1982). *Judgement under Uncertainty. Heuristics and Biases.* Cambridge: Cambridge University Press.

Konoligue, K. (1988). 'Defeasible Argumentation in Reasoning about Events'. In: *Methodologies for Intelligent Systems* 3, pp. 380-390.

Loui, R.P. (1987). 'Defeat among Arguments: A System of defeasible Inference'. In: *Computational Intelligence* 3(2), pp. 100-106.

Loui, R. P. (1990). 'Ampliative Inference, Computation and Dialectic'. In: J.L. Pollock (ed.), *AI and Philosophy.* Cambridge, Mass.: MIT Press.

Loui, R. P. (1991). 'Argument and Belief: Where we stand in Keynesian Tradition'. In: *Minds and Machines* 1, pp. 357-365.

Luckasziwicz, W. (1990). *Non-monotonic Logic; Formalisation of Commonsense.* Ellis Horwood.

McCarthy, J. (1980). 'Circumscription - A form of non-monotonic reasoning'. In: *Artificial Intelligence* 13, pp. 27-39.

Montgomery, H. (1989). 'From cognition to action: the search for a dominant search in decision-making'. In: Montgomery, H. and O. Svenson (1989). *Process and structure in Decision Making.* Chicester (UK): Wiley.

Nute, D. (1988). 'Defeasible Reasoning and Decision Support Systems'. In: *Decision Support Systems* 4(1). Amsterdam: North Holland, pp. 97-110.

Pollock, J.L. (1987). 'Defeasible Reasoning'. In: *Cognitive Science* 11, pp. 481-518.

Scriven M. (1987). 'Probative Logic: review and preview'. In: F.H. van Eemeren, R. Grootendorst (eds). *Argumentation: Across the Lines of Discipline.* Dordrecht: Foris Publications, pp. 7-33.

Simari, G.R. and R.P. Loui. (1992). 'A Mathematical treatment of Defeasible Reasoning and its Implementation'. In: *Artificial Intelligence* 53, pp. 125-157.

Vreeswijk, G.A.W. (1993). *Studies in defeasible Argumentation*, doctoral dissertation, Amsterdam.

Walton, D. (1989). 'Reasoned use of expertise in argumentation'. In: *Argumentation* 3.

Woods, John, (1989). 'The Necessity of Formalism in Informal Logic'. In: *Argumentation* 3, pp. 149-167.

Embedding logic in communication: lessons from the logic classroom

Keith Stenning[1],
Human Communication Research Centre,
Edinburgh University

Abstract

In this paper, a fresh approach is proposed to the tension between language as a formal structure and language as a social practice. Learning elementary logic can improve reasoning skills, and therefore to characterise what it is that is being learnt in elementary logic classes can help to characterise what argumentation is, and what relation it bears to logical theory.

1. Introducing logic to society

At the time at which argumentation theory parted company with logic under the impetus of such authors as Toulmin and Hamblin in the 1950s, logic was seen as a source of authority for arguments. Its rejection by those interested in the empirical study of argument was based on the assumption that this was indeed logic's role. Their argument that symbolic logic had little to contribute to the empirical study of argumentation was based on the thesis that logic could not contribute authority.

At that time, symbolic logic was still strongly rooted in the study of the foundations of mathematics, and had been couched in artificial languages in order to divorce it from the natural language discourses of such paradigmatic domains of argumentation as philosophy and the law. This period coincided with the heyday of analytical philosophy which saw the analysis of vernacular natural language *use* as its main source of data and insight, and which opposed the analysis of natural languages by the new formalisms of logic. When we come to reassess the potential relations between logic and argumentation in the current environment we operate in a transformed intellectual space.

Chomsky's and Montague's demonstrations that logics and automata theory could produce insightful analysis of the structure of natural languages was a returning full cycle of the 'formal' revolution. What started as an escape from the vernacular for

philosophical practice, became a tool of a linguistic theory of the vernacular itself. Meanwhile, logicians have developed their art in the direction of the proliferation of logics to a point where the idea of *the* logic as a source of authority is quite foreign to the discipline. Selection of a tool appropriately designed for the job in hand from an expandable kit of tools fashioned during practical computational experience, is different metaphor for the discipline than that of the laws descending on tablets of stone as the origin of all authority.

Now that we are so much clearer about the formal structure of natural languages and of the multifarious possibilities of logic, we might expect to be all set for a productive collaboration between logic and the empirical study of argumentation. Certainly, if the characterisation of the *languages* of argument could contribute to the understanding of argumentation, then all should be set fair. However, on a different view, it never was clarification of the *languages* of argument that was required, but clarification of the *discourse* of argumentation. Languages, whether natural or artificial, are embedded in a pragmatics of use whenever language is observed in action. And it is only discourse that ever presents itself as the data for a theory of communication. This was the observation that lead Wittgenstein, Austen and their followers to ditch formalism and argue that it contributed little or nothing to the understanding of language use. The tension between seeing language as a set of sentences and seeing language as a social practice is just as real today, despite our much greater sophistication about the formal structure of natural languages.

The goal of the present paper is to propose the usefulness of approaching this problem afresh from the perspective of an account of the phenomenon of learning logic. The very notion of 'learning logic' is paradoxical, approached from the perspective of the classical conceptualisation of logic as the origin of authority and the basis of the 'laws of thought'---and yet mundanely referential. Students take logic courses, some of them taught by members of this symposium, and do indeed sometimes succeed in them. There is even good evidence, contrary to much received psychological wisdom, that such classes can improve reasoning skills beyond those practised in class (Stenning, Cox & Oberlander (1995a)). We propose that to characterise what it is that is being learnt in elementary logic classes can help to characterise what argumentation is, and what relation it bears to logical theory.

What we take to be the classical account of the place of logic in a theory of communication makes a fundamental distinction between what we will call *exposition* and *derivation*. These are fundamentally bases for *transitions* between sentences of a discourse. One is the basis of transitions between successive assumptions appearing on the left of the logician's turnstile ($\vdash$): the other is the basis for the transitions between the sentences of the derivation which arrives at the conclusion on the right of the turnstile.

Logic has next to nothing to say about the former transitions. The assumptions on the left of the turnstile are conceived of as an unstructured set of sentences, and so their ordering conveys no information. The burgeoning pscyhological and computation-

al literature on natural language processing is about how discourses have to be structured to achieve the kind of interpretation which is represented by the unstructured logical idealisation. Logic is ostensibly all about the transitions of derivation toward the conclusion on the right. However much concerns of derivation dominate logic, and however little logic has to say about the structuring of the premisses of an argument, this conceptualisation is still an implicit theory of the role of logic in communication. Derivation is about the re-representation of information. Logic says nothing about what set of assumptions we should reason from. It only speaks of what re-representations of these premisses will make no further assumptions. Logic does not even indicate what to do about ones' assumptions when derivation reveals inconsistency---it only indicates that something must be done. We propose that any theory of communication (and, *a fortiori*, of argument) that does not embrace this fundamental distinction between the bases of transitions between sentences will prove an inadequate theory of communication.

Naturally occurring discourse may consist of complex alternations of these different types of transition between sentences. In developing a set of mutually agreed assumptions which express shared knowledge, belief or hypothesis, participants may have to explore each others' interpretations by deriving consequences from them. The appearance of contradiction through derivation may reveal the necessity for reformulation of assumptions. But the fact that naturally occurring discourses interleave exposition and derivation makes the need of a theory of communication for the fundamental distinction all the greater.

The most obvious deficiency of logic as the basis for a theory of communication its *asocial* character. To base a theory of communication on logic we have to explain how the categories of exposition and derivation map onto social categories of participant. It is not enough to merely note that each sentence can be assigned to a speaker/author and that an intended audience/reader can be identified. How does the distinction between exposition and derivation map onto these same participants? The central concept appears to be that of *authority* for information. Let us begin by illustrating with the simplest form of exposition---story-telling with two participants. What constitutes story-telling is that one participant knows about some situation which the other does not, and the former undertakes to communicate this situation by uttering a sequence of sentences designed so that the other can construct the intended model of them. In this simplest case of exposition, authority for information is completely asymmetrical, and this asymmetry is acknowledged by the participants---one knows and the other does not. Story-telling is a paradigm example of what we call exposition.

This account is, of course, severely idealised. The situation with real story-telling can be quite complex. Ignorance on the part of the listener may be just a pretence. As most parents discover, breakdowns in this pretence may lead to challenges on the part of the listener to the authority of the storyteller---"the story doesn't go like that", says the listener.

Another example of exposition with a socially anomalous authority pattern is the discourse of examination. The teacher knows the answer, and the pupil knows that the teacher knows the answer, but still the pupil's task is to exposit for the teacher *as if she didn't know the answer* so that the teacher can judge whether the pupil knows the answer. But these anomalies merely serve to highlight the all important fact that asymmetry of authority is the ruling condition of exposition---these are anomalies that emphasise the rule.

Neither is the important distinction the one between fictitious and factitious storytelling. Recounting a true narrative to an audience who was not witness to its events assumes just the same asymmetry of authority as a fictional story. Setting out on an expository discourse with an audience who *does* already know what they are being told will commonly lead to the question 'Why are you telling me this?', or 'Don't you know I was there?', thus underlining that the required asymmetry of authority does not exist.

The one place where the expository speaker's authority can be challenged by his audience is when the hearer deems that contradiction occurs, but this challenge is a derivational challenge---'but you said Fido was a cat, and now you say she is a dog'. Such inconsistencies may be resolved in many ways---there are two Fidos or whatever ---but resolved they must be if exposition is to continue. The goal of exposition is to arrive at a shared set of assumptions, and when inconsistency threatens a hearer's interpretation, then derivational challenge is essential.

Derivations have a quite different social structure from expositions. Here the participants have equal authority. They fully share a set of mutual assumptions (either from knowledge, belief or hypothesis), and they seek to reformulate them by derivation. Here symmetry of authority is the order of the day. It will not do for one party to assert a sentence as deriving from another on the basis of superior *authority*. Appeals to superior authority from knowledge can only be appeals to change the assumptions---an essentially expository move.

Using patterns of equal and unequal authority to define derivation and exposition suggests an obvious corollary. For unequal authority, at least two participants are required---for equal authority one will suffice. Derivation can be either a social or an isolated activity, but expositing to ourselves is... well, the first sign of madness.

In understanding where *arguments* belong in these mappings of the logical onto the social, it is important to observe that discourses vary on two dimensions which are to be carefully distinguished. Besides being made of expository and derivational transitions, discourses can be cooperative or adversarial. The latter categories apply to the participants' overall interests in the achievement of goals.

Story-telling is the simplest form of exposition because the participants typically are cooperating to 'transfer' the intended model from the speaker to the listener. The listener does not have preconceived ideas about the story to be told. Only contradiction is unacceptable, and that because it threatens the hearer's new-found interpretation---

the very object of the communication. But clearly, not all participants of expositions are so indifferent to the particularities of the resulting set of assumptions.

How would this framework apply to the typically adversarial communications of, say, a court of law? To say that defence and prosecution are engaged in a cooperative activity seems initially bizarre. But as an over all analysis of the communication in court this seems fundamentally right. Defence and prosecution seek to arrive at a set of propositions believed by the court. The defence wants this set to include that the defendant did not commit the crime---the prosecution that he did. But both must subsume the 'facts of the case' and be internally consistent. It is of critical importance which facts are 'agreed' and which 'contested'. A great deal of cooperation must serve as background for the focus of adversarial combat.

How does the discourse divide between exposition and derivation? Like most actually occurring discourse, intertwining of these two kinds of transitions is intimate. The authority relationships are far more complex than in our paradigm example in which one participant knows all and the other knows nothing, and this status is known and accepted by both. Neither defence nor prosecution, nor witnesses have an overall special status with regard to their authority for their statements. But each of their expository statements is made on their asymmetrical authority with regard to that statement. A large part of the manoeverings consists of attempts to bolster or impugn the authority of speakers for their propositions. Testimony may be bolstered or impugned by assessments of status, character, self-interest, or expertise---all ingredients which go to determine authority for information. But equally derivation plays an important role in the maintainance of consistency. If testimony is contradictory, then repairing this contradiction requires adjustments to the interpretation of expositions, perhaps by changes in judgement of character or expertise or whatever. Derivation enters mainly in reformulating propositions to reveal consistency or inconsistency, or in integrating them. Its operation makes no appeal to authority.

Legal discourse may be usefully compared with mathematical discourse. Two mathematicians engaged in proving or disproving a conjecture may take sides in much the same way as the counsels for defense and prosecution. They may elaborate the assumptions which they make. Or they may make deductions from their assumptions. Lakatos (1963) illustrates how these modes are interwoven in the development of a mathematical topic---the 'proof' of Euclid's conjecture about the function relating the numbers of faces and edges of polygons.[2] Expositional transitions---changes in assumptions---cannot be supported by derivation (though the need for change may be motivated by the derivation of inconsistencies). Changes of assumptions have to be motivated by the range of factors that affect mathematical authority. Here issues of the style of the school, past masters' practice, or mathematical elegance are ingredients of authority.

---

[1]   History illustrates that these issues may be as partisan, and as socially motivated, as in a court of law.

While mathematical discourse may differ from that of the courtroom in the rigidity of participants' comitments to outcome, and in the sort of considerations which bear on the changing of assumptions, it is similar in that it makes a clear distinction between the considerations for changing assumptions, and the process of reformulating them in derivation.

The concept of *argument* in common parlance appears to attach to two distinctions. One between adversarial and non-adversarial discourse goals, and the other between exposition and derivation as bases for transitions from sentence to sentence. We have argued that both of these distinctions are fundamentally socially grounded, and that it is the concept of authority for information is the key to understanding both.

Our destination is an analysis of what is taught (and learnt) in elementary logic classes. So far we have argued that understanding the relation between logic and argumentation requires the socialisation of logic through the concept of authority for information. But much more than this will be required to resolve the paradox of logic learning. Logic focusses on the discourse of derivation, and derivation plays its role in communication through its part in the maintenance of consistency of interpretation. But surely students are capable of linguistic communication (and argument, for that matter), before they get to university, and logic classes?

Our proposals about what is learnt in logic classes will assume the answer to this rhetorical question is positive. Part of the resolution must lie in the question what distinguishes such institutionalised professional activities as courtroom persuasion and mathematical proof from the story-telling with which our exposition of exposition began? Professional practices rely on interpretation of experience through document alone, and this is what introduces formality. But first let us take a closer look at one theory of the relation between logic and communication in the next section, and at some data from the classroom in the next.

2. Grice's account of logic and communication

Grice's (1975) *The logic of conversation* is one of the few systematic attempts to embed logic in a theory of linguistic communication. Grice's seminal idea was that by embedding a classical view of logic within a suitable theory of the particpants' communication goals, he could provide a pragmatic theory which would accommodate apparent discrepancies between 'natural' and 'technical' interpretations of English constructions, notably quantifiers.

One particular phenomenon which motivated Grice was the 'misinterpretation' of quantifers revealed when students come to learn elementary logic. It is often noted by logic teachers, and well supported in the psychological literature (e.g. Newstead (1989)), that students will commonly claim that from *Some A are B* it follows that *Some A are not B*. Grice noted that one can make good sense of these interpretations if students assume that the author of the premiss is cooperative and seeks to be maximally informative. According to this account, the student reasons: "If all A are

B, then the speaker would have said so, rather than saying just *Some A are B,* so some A must be non-B."

Grice does not make our distinction between exposition and derivation. The term *conversation* in his title connects logic with the phenomena of communication, and one way of interpreting his title suggests that Grice's intention was to give, in our terms, an account of expository discourse. This would be a conflation of conversation with one of its conspicuous functions. Nevertheless the title successfully conveys what was in its time the radical proposal---that logic required connection to communicative phenomena other than deductive proof.

But it is quite clear that Grice's account of his students' errors does not stand up. Grice's explanation of the implicature from *Some A are B* to *Some A are not B* is couched in terms of his maxim of quantity: *Give as much information as is informative.* Gazdar's 1979 formalization of this inference shows quite clearly what is missing. From Grice's maxim, and the assertion that some A are B, it follows that the speaker does not know that all are B. The reason that it does not follow that some A are not B is simply that the speaker may not know whether this is true or false. The maxim of quantity enjoins informativeness, but it is balanced by the maxim of quality which enjoins authority for information asserted. The student's inference that all A are B is highly revealing of what we will call *the assumption of omniscience.*[3]

Where could this ascription of omniscience to the source of the premiss come from? An obvious suggestion can be couched in terms of the student assimilating the new 'game' of deduction to an expository discourse with asymmetrical authority (not unlike the story-telling we chose as the simplest expository situation above). If this explanation is along the right lines, Grice's theory will require augmenting with a typology of kinds of discourse (or perhaps kinds of discourse transition) in order for its maxims to be sensitive to the different possible purposes of discourse.

But before we explore the possibilities of repair for Grice's approach, we want to stand back and ask rather more systematically what peoples' interpretations of quantifiers are like before they undergo elementary logic teaching.

3. Towards a model of what is learnt in logic class

We draw here on data collected in a study of first year social science faculty undergraduates' judgements about the logical relations between English quantifiers. These students were taking an introductory psychology class and few of them had had any formal instruction in logic. A few may have seen some 'set diagrams' in learning mathematics. The study is reported in greater detail in Stenning, Cox & Oberlander (1995b).

---

[3] We intend omniscience with regard to the relations between sentences expressible in terms of the quantifiers we consider---clearly true omniscience is considerably stronger.

We elicited interpretations of elementary quantifiers from students by asking them to make judgements about what followed from sentences containing each of the quantifiers in simple syllogistic premisses: *Some A are B; Some A are not B; All A are B;* and *No A are B*. Given a sentence of the form *Quantifier A B* as premiss, students were asked to say for each of eight sentences, whether they must be true if the premiss is true, whether they must be false if the premiss is true, or whether their value was undetermined by the truth of the premiss---abbreviated henceforth as *true, false* and *can't tell* responses. The eight sentences were generated by combining the four quantifiers with the matrices *AB* and *BA* respectively. 124 Edinburgh University first year undergraduate students contributed their judgements.

Our aim here is to stand back from the focus on explaining specific patterns of interpretation (such as the pattern of implicature from *some* to *some not* discussed above) to ask what the overall pattern of interpretation is like. With 3^8 possible responses for each of four premiss quantifiers, this is a rather complex data set and we will not attempt a comprehensive account here. But there are some striking patterns of consistency across quantifiers which are suggestive of underlying factors. The first pattern of responses which struck us was that there was a substantial group of students who *always* responded *can't tell* to any query sentence that reversed the premiss' subject/predicate structure. So not only would these students correctly respond 'can't tell' when presented with *"All A are B. Are All B A?"*, but they would incorrectly respond the same way to *"Some A are B. Are Some B A?"*. In fact there was a strongly bi-modal distribution for these reverse subject/predicate questions, regardless of quantifier. Either subjects nearly always responded *can't tell* or they rarely did, with few doing both equally often.

What is striking about this pattern of responses, both for those who always respond *can't tell* and those who never respond so, is that it suggests that responses are determined by the arrangement of the terms *A* and *B* between premiss and conclusion in the questions, rather than by the nature of their quantifiers. Pursuing this suggestion, there are four quadrants of the data relevant. We examined the patterns of response where A and B stayed in the same position, compared with those where they were changed in their premiss subject/predicate positions in the conclusions. And we examined questions where the *correct* answer was *can't tell*, as compared with questions where the correct answer was determinate: *true* or *false*.

The bi-modality of the distribution of scores in these four quadrants was strong, just as it had been in the first portion of the data which had first drawn our attention to the phenomenon. Students either tended to respond *can't tell* or to respond determinately, across all the questions of a group defined by the subject/predicate relations between premiss and conclusion. To make these patterns of response vivid we will call students with the tendency to respond *can't tell* to questions which should be answered *true* or *false* 'hesitant', and those with the opposite tendency to respond determinately when they should respond *can't tell* 'rash'.

The obvious next question is how these tendencies are sensitive to the structure of the groups of questions defined by subject/predicate relations between premiss and conclusion. We will call *"Q AB. Does it follow that Quantifier AB?"* questions[4], *in-place* questions, and *"Q AB. Does it follow that Q BA?"* questions, *out-of-place* questions. It is logically possible to be hesitant or rash on either kind of question, meaning that each student has a score on four kinds of questions. As mentioned before, there is a strong tendency for these scores to be bi-modally distributed in this population. If we choose a criterion, and collapse the data onto binary scales, the possible space of responses is just four-dimensional.

Examination of data reveals that in fact only three of these dimensions differentiate students. There are effectively no students who are hesitant on in-place questions. This leaves three binary dimensions for classifying students---a cubic space structures the data. It further turns out that four of the eight corners of this space account for 94% of the students, with no other corner accounting for more than 4%.

Since each of the dimensions reflect correct/incorrect responding, we might expect these four groups to consist of those who make no errors at all, and the three groups each constituted by a single error-type. But in fact this is not what we see. A substantial group of students *do* make no errors at all (17%). There is a substantial group of students who are just rash on in-place questions (23%). But both of the other substantial groups makes two kinds of error. The largest group (35%) are rash on both in-place and out-of-place questions. The fourth group of students (20%) are simultaneously rash on in-place questions and hesitant on out-of-place questions. If rashness and hesitancy were some sort of 'temperamental response tendencies' this is not the pattern we would expect at all---it would amount to a change of personality in response to a change of grammar. The overwhelming generalisation in the data is that it is the relation between subject/predicate structures in premiss and conclusion which determines patterns of response.

Since there are four main patterns of response among these students, no single all encompassing explanation for inference behaviour can explain the patterns. We cannot tell from this data whether these distinct patterns are stages along a path of learning which a single student may pass through on the road to enlightenment. Or whether they are alternative starting points which will converge by different routes during logical teaching. We have only just begun to investigate these same students' inferential behaviour, and this may reveal something about possible alternatives.

How do the traditional questions about errors of interpretation in deductive reasoning look when visited afresh in the light of this data? Just as with Grice's implicatures, fallacies are generally considered in isolation from each other. For example, because data on *illicit conversion* of *all* are obtained from reasoning data, they have not been examined against the pattern of interpretation of other quantifiers. Asking why a student might believe that *All B are A* follows from *All A are B*

---

[4]   Where *Q* is a variable ranging over the four possible quantifiers.

provokes a quite different line of thought than asking why a student always replies either true or false to any question about inverted subject and predicate.

Another casualty of the methodology of studying 'fallacies' has been the lack of attention to sins of omission. In our data, many students respond *can't tell* to the question "Does *Some B are A* follow from *Some A are B?*". This is surely just as striking as the sins of comission that constitute the fallacies. Furthermore, omission and comission are closely related errors when viewed at the level of abstraction which cuts across the quantifiers. We observe the student responds the same way to all such inverted questions, and that the response which avoids the fallacy of deriving *All B are A* from *All A are B* can be described as the very same response that leads to the failure to conclude *Some B are A* from *Some A are B*.

A common response among the logically initiated to this failure to 'convert' *some* is incredulity---it is hard to see what could lead to such a response. Yet these are highly intelligent students who are behaving anything but randomly, despite the tedium of a large number of irritating questions. So what is the systematic basis that leads to these behaviours? Why should these patterns be driven by subject/predicate structure rather than by the nature of their quantifiers?

Natural language structures and exposition

The structuring of natural language sentences into subject and predicate is one important syntactic device which expresses what has come to be called 'information packaging' (see e.g. Vallduvi (1992)). In these simple sentences, neutrally intoned, the subject forms the 'ground' and the 'topic' whereas the predicate forms the 'focus' and the 'comment'. The subject tells the hearer what the statement is about in terms already recognizable to the hearer, and therefore 'where' the new information has to be 'attached' to existing knowledge. The predicate is the 'focus' which provides the new information to be added to the hearer's knowledge base. Many other syntactic (as well as prosodic) devices play a part in this structuring.[5] The point here is that these categories only make sense in terms of expository discourse---the discourse of informing.

Formal logical languages obliterate subject/predicate distinctions in favour of structures of function and variable. This transformation is accompanied by a change of discourse in which the distinction between ground and focus is obliterated. The discourse of derivation does not recognise these categories since all information in the assumptions is common ground. So to teach elementary logic is to teach that reference

---

[5]   The linguistics literature understandably concentrates on the complex interplay between prosodic, syntactic and lexical resources for expressing information packaging, and the differences between languages in these patterns. It also elaborates the distinctions made, for example, between focus and comment. But in the extremely limited range of sentences at play in these experiments, the only determination of packaging is by subject/predicate structuring. See Vallduví & Engdahl (to appear) for an overview of the crosslinguistic realisation of information packaging.

and attribution give way to the interpretation of both subject and predicate terms as denoting sets, and quantifiers as relations between sets. This new perspective might be called the *extensionalist stance*. It is quite foreign to an understanding of natural language restricted to its use in exposition.

Let us take a simple example to illustrate the effect of information perspective on logical analysis---the syllogism: *All competitors were running. Some children were running.* This is conventionally analysed into the logical form:

$$\forall\, y,\ competitor\ y \rightarrow running\ y.\ \exists\, x,\ x\ a\ child\ \&\ running\ x$$

The relation between three sets of individuals (those who are children, competitors, and those running) are partially determined by the quantifier relations.

Now consider two ways which the second premiss of this syllogism might function in an *expository* discourse. These two interpretations can be distinguished by the positioning of contrastive stress. Reduced stress on 'some' is noted as 'sm'. Capitals indicate contrastive stress.

1. Sm children were running
2. SOME children were running

Sentence 1 would be used to introduce into the discourse the new information that there were children in the domain of the story and that they were running. This reduced quantifier functions like the indefinite article. Sentence 2 would assume as already commonly known information that there were children, and assert the new information that some of them were running. What implications would these two different interpretations have for the question whether it is true that all children were running?

In the case of 1, we are inclined to say that some running children have been introduced into the story, and, at least until we hear about some more children, these are the only ones in the story, and all of *these ones* are running. The narrator is in authority and the story is unfolding in this way, and there may or may not be more children to come, but that is definitely not the issue at this point. In the case of 2, it is quite clear that some of the children which we already knew about before this part of the story are *not* running, and this information *is* the focus of the current assertion. So treating this syllogism as exposition, and assigning two different prosodies to one premiss gives opposite but determinate conclusions.

What implications would these two different interpretations have for questions about the relations between the children and the competitors? In an expository discourse, it would be anomalous to use these two logically independent terms unless it was assumed that the hearer knew what relation they had before the assertion. The most likely interpretation without further context is perhaps that the children are distinct from the competitors---no children are competitors. If not, then the statement is

experienced as redundant. Though given a context in which the fact that young people were running was very striking and against expectation, and perhaps needed observing to be a consequence of the previous universal, then the second statement might be understood as indicating that some children were among the competitors. Again, if there are children *not* among the competitors, then they may or may not turn up later in the story, but until they do, they so to speak 'do not exist for the story'. Again, authority relations between participants in an expository discourse, together with the purpose of exposition, conspire to fully determine the relations between the sets of entities in the intended domain (see Stenning 1978). Speakers may actually have to explicitly state their ignorance of these relations if they are to avoid generating assumptions about them.

Conclusions of the argument?

Perhaps, at this point, a condensed exposition of our argument would be of service.

Logic regarded as a theory of communication makes its fundamental distinction the one between exposition (the introduction of new assumptions) and derivation (the derivation of conclusions). Socially these are distinguished by their patterns of authority.

Argument is commonly distinguished as being adversarial discourse. Whether discourse is cooperative or adversarial in participants' immediate goals, our distinction between exposition and derivation allows us to see that at a more fundamental level there has to be a cooperative element in all communication. Exposition and derivation play distinctive roles in the process of communication.

Grice's program for pragmatics couched in terms of general purpose principles such as informativeness fails in its explanation of some implicatures because it fails to capture inferences hearers base on the omniscience of expositors. This omniscience derives from the authority relations between speaker and hearer.

An empirical approach to the question how students' naive interpretations of quantifiers differ from the deductively normative interpretations seeks general patterns of interpretation rather than explanations of specific fallacies of inference. The program of analysis of such data sketched here presents the view that what distinguishes their interpretations is their stance toward the information-packaging structures of natural language. They tend to assimilate logical problems such as syllogisms to their well-honed model of expository discourse.

Standing back from particular fallacies of reasoning, and from particular theories of their origins, and examining the global pattern of interpretation of quantifiers reveals links between what appear in the literature of rhetoric as quite independent fallacies. It also reveals failures of omission of inferences, and connects their occurrence to the comissions we label as fallacies. Finally it reveals that groups of students exhibit highly distinctive patterns of omission and comission which cannot be subsumed under uniform tendencies to err in reasoning. The force shaping the field of these several

distinctive patterns is uniformly the information packaging structures of natural languages. Learning elementary logic is learning the 'extensionalist stance' which un-packages propositions.

Many students appear to begin their task at a point at which they assimilate the interpretational task set them in the experiment described above to the discourse of exposition. For some this leads them to interpret the sentences presented as part of an exposition with an omniscient source, and so to full determine the relations between their terms in the way we illustrated by our distinctively intoned syllogism. Others appear to appreciate that subject/predicate is problematic in this new game of pure deduction, but are not yet able to differentiate between the cases in which sub-ject/predicate maps onto antecedent/consequent, and the cases where it is merely redundant information packging left over from its expositional origins. Perhaps these problems lead to the hesitancy in which subject/predicate changes mean that all bets are off. Far more analysis is required before we can understand the several patterns we observe.

Reviewing the paper's argument in compression emphasises one puzzle that we have barely touched on. These students are perfectly capable of argument well before they hit Logic 101. If they were not able at derivation they would not be competant expositors, nor would they reach Logic 101. On these grounds alone it is too easy to dismiss the approach proposed here, and indeed the idea that learning logic has any impact on general reasoning abilities. Strong currents in the psychology of the last 40 years have tended towards a kind of naturalism with regard to reasoning. Ever since Piaget's logicism, the tide has been flowing towards 'contextualised', 'ecological', or 'situated' accounts of human reasoning which reject logic as playing any part either theoretical or didactic.

Important insights have flowed from this tide. Piaget's identification of reasoning with the *language* rather than the *discourse* of reasoning lead him to logicise early childhood. Certainly much of human reasoning is extremely efficient in its exploitation of contextual simplifications. But this tide has tended to lead to us ignoring the possibly confined, but nevertheless highly distinctive reasoning that is formal and decontextualised---the kind of reasoning logic is intended to formalise and logic teaching to disseminate. As Sylvia Scribner's work (e.g. 1977) did so much to show, westernised formal schooling leads to the capability for a relatively decontextualised stance toward language and discourse. The ability to reason about situations experienced solely through representations, and defined by those representations, is catalysed by teaching focussed on the business of derivation (whether in mathematics, language teaching or logic). This cognitive exercise has an intimately social side--- learning how authority is established and consequences derived from language (and other representations) requires a culture with a social organisation that is willing to go beyond eye-witness experience as evidence.

The professional discourses mentioned earlier---law and mathematics---are examples where elaborate practices make explicit the rituals of exposition and derivation. The

embedding of logic---both the exposition on the left of its turnstile, and the derivation that takes us to the right---in these social discourses give us a way into understanding formal thought. In all cases, formality strips off the information packaging which controls the rapid and felicitous comprehension of cooperatively told stories, and replaces it with the explicit procedures for discourses with more complex authority patterns. The nature of formal thought is as worthy a goal for psychology as any--- certainly we can understand remarkably little of the particular culture which we have inherited without an understanding of it. Our civilisation of formalisation may be fraught with its discontents, but its what we have been landed with and is in need of our attention.

References

Gazdar, G. (1979). *Pragmatics, Implicature, Presupposition, and Logical Form.* London: Academic Press.

Grice, H.P. (1975). *Logic and Conversation.* London: Academic Press.

Lakatos, I. (1963). 'Proofs and refutations'. *British Journal for the Philosophy of Science,* 14. Also as Worral & Zahar (eds.) Cambridge: Cambridge University Press 1976.

Newstead, S.E. (1989). 'Interpretational errors in syllogistic reasoning'. *Journal of Memory and Language,* 28, pp. 78-91.

Scribner (1977). 'Modes of thinking and ways of speaking: culture and logic considered'. In P.N. Johnson-Laird, P. C. Wason (eds.) *Thinking.* Cambridge: Cambridge University Press.

Stenning, K. (1978). 'Anaphora as an approach to pragmatics'. In: M. Halle, Bresnan, J. and Miller, G. A. (eds.) *Linguistic Theory and Psychological Reality.* Cambridge, Mass.: MIT Press.

Stenning, K., Cox, R. & Oberlander, J. (1995a). 'Contrasting the cognitive effects of graphical and sentential logic teaching: reasoning, representation and individual differences'. *Language and Cognitive Processes,* 10 (3/4), pp. 333-354.

Stenning, K., Cox, R. & Oberlander, J. (1995b). 'Attitudes to logical independence: traits in quantifier interpretation'. *Proceedings of Seventeenth Meeting of the Cognitive Science Society,* Pittsburgh, 1995, pp. 742-747.

Vallduví, E., (1992). *The informational component.* New York: Garland.

Vallduví, E. & E. Engdahl, to appear. 'The linguistic realisation of information packaging'. *Linguistics.*

Functioning and teachings of adaptive logics

Diderik Batens,
University of Gent

Abstract

This paper concerns some formal systems, viz. adaptive logics, that display a specific flexibility in the meanings of logical terms. Both the flexibility that occurs within the systems and the question as to how we may arrive at such systems is discussed. Both, it is argued, are relevant to bridging the gap between logic and argumentation.

1. Aim of this paper

In the present paper, I report on some formal systems, viz. adaptive logics, that display a specific flexibility in the meanings of logical terms. I shall discuss both the flexibility that occurs within the systems and the question as to how we may arrive at such systems. Both, I maintain, are relevant to bridging the gap between logic and argumentation.

I shall start by examining an opposition that underlies the alleged opposition between logic and argumentation and is more fundamental. Neglecting the underlying opposition may result in a misguided approach to the relation between logic and argumentation. I shall show how the underlying opposition brings us to the problem of the flexibility of terms, including logical terms. At that point I start my story on adaptive logics.

I set out for the modest task to discuss one of the many aspects of the logic-argumentation opposition. It is not all-embracing and perhaps not even central. Still, it seems to me that it is fundamental in that the opposition cannot be overcome if one does not crack this nut.

2. Monologism and Plurilogism

Monologism is the doctrine of the one true logic. Few logicians maintain that the true logic is available right now, but many believe that it exists and may eventually be fully described. It may contain a wide variety of so-called logical terms: connectives, quantifiers, modalities, etc. All of them, however, should belong to one single system. As seen from monologism, the interpretation of a logical term, say, in a natural language, concerns the mapping of this term on a term *of* the logical system. An alternative is *plurilogism*. It allows not only for a variety of logical terms, but also for

a variety of mutually incompatible logical systems. According to this view, there is room for contextual meaning in a very extreme sense (for some clarification and consequences, see my (1985), (1992a) and (1992b)). Here language is flexible in a totally different sense. The flexibility does not concern the interpretation of a term with respect to a formal system, but the choice of a formal system from a variety of alternatives. Even this description is not fully accurate, because the appropriate alternative may not be known; so, the interpretation of a term may require one to devise a new formal system. According to plurilogism, we do not only use the same terms to express different meanings (in natural language, in thinking and communicating), but we jump from one logical universe to another.

The majority view among formal logicians always was monologism. Not that everyone agreed about the true logic. Many classical logicians even opposed extending classical logic with non-extensional logical terms. Relevance logicians, especially Anderson and Belnap, and students of them, argued at length that some classical inferences are not correct (in the one true logic). Australasian paraconsistent logicians (Routley/Sylvan, Priest, etc.), argued that even most relevant logics are too rich and too classical to be sensible candidates for the true logic.

Most people from the argumentation tradition adhere to plurilogism (although they hardly ever put the matter in these terms); so do a variety of linguists and many literature theorists (who face concrete problems of text interpretation).

Quite often the opposition is underestimated. No one doubts that there is a variety of (technically respectable) logical systems. Actually, a battle had to be fought to arrive even at this meagre agreement. First intuitionistic logic and modal logic became tolerated. Later followed relevant logics and paraconsistent logics, both having a much harder time. This tolerance, however, is merely passive: philosophical fights about who is right have almost stopped, some results are carried over from one approach to the other, but the philosophical disputes are not settled. The parties consider each other as technically interesting curiosities. If monologism is correct, however, then either the relevance view or the classical view is mistaken; and similarly for other comparisons. (Just as one cannot at the same time be a classicist and an intuitionist with respect to mathematics.) The situation is even worse: there is not and cannot be an agreement on what the precise distinction between rival logics is; the metalanguages available to the several parties result in different descriptions of the same logics (see my (1990)).

Natural language may be approached from both the monologicist and the plurilogicist tradition. In my view, the present most impressive attempt to do so from the first tradition is actually a very heterodox one: the approach defended in Graham Priest's (1987). It is impressive because it tackles the presumably most fundamental theoretical problem of natural languages, viz. self-reference and the semantic paradoxes, as well as the paradoxes and limitative theorems (Gödel, Church, Löb, etc.) of classical mathematics. The result is a coherent system that contains its own metatheory in a way comparable to natural language. Priest's actual logical system is rather poor, but it may almost as easily be supplemented with new logical terms as classical

logic. It has one major theoretical drawback: it does not enable us to adequately express that we *reject* a statement — this is argued at length in my (1990).

Logicians and argumentation theorists start at opposite ends of a continuum. At the bottom, there is the exactness and accuracy, but also poverty and rigidity of formal languages. At the top, there are the rich and relevant contributions from argumentation theorists, that leave room for flexibility and interpretation, but usually lack formal strictness and mathematical accuracy.

Precisely because both sides seem to come closer to each other, a warning is needed: the opposition between monologism and plurilogism should be taken seriously. The continuous enrichment of a monolithic logical system will not enable us to solve the problems dealt with by argumentation theorists. If monologism is correct, the (many times implicit but nevertheless) standard doctrine behind argumentation is simply mistaken. That is quite possible, of course, but it seems to me that the present evidence points in the opposite direction. I expand a bit on this in the next section.

3. Some problems for monologism

According to monologism, the logical terms that occur in natural languages may be ambiguous and perhaps even vague, but the underlying logical concepts, the meanings of the disambiguated terms, are stable and fixed. At the conscious level, these meanings may be discovered (in the literal sense) and this may require time and analysis. At the unconscious level, the true logic must always have been there, not (only) in some platonic heaven, but in people's minds, even if they had no theory about it.

Outside of the domain of logic, a similar point of view would sound completely outdated. At least to anyone who has some knowledge about the evolution of natural languages. Also to anyone who has some insights in the history of the sciences, especially where active thinking is concerned (creative understanding, creative problem solving, etc.). New concepts are created, not 'discovered'. They are the products of the evolution of our theories about the world, much more than of the world itself. They contain heavy interpretations, many of which turn out to be completely on the wrong track with respect to reality (as it is seen in later periods). Given that this is the common view for non-logical terms, what is so special about logical terms, that their meanings should be stable and fixed, even within the human mind (at the unconscious level)?

The reason cannot be that logical terms are not 'referring'. For neither are mathematical terms, and we all agree (i) that there is a multiplicity of mutually exclusive systems in many mathematical domains, and (ii) that the choice of a mathematical system for some empirical theory is an empirical matter (think about geometry and relativity theory). The latter point is the central one, of course. It entails that the multiplicity of mutually exclusive systems is not merely a technical matter;

only our continuing interaction with the world will enable us to decide which of the alternatives is appropriate to shape our theories about specific domains.

I know of one sensible argument for this exceptional position of logical constants: they would be fixed in our hardware. This sounds like a good argument, but is it true? There is no evidence to support the thesis that the logical terms from natural languages would be genetically determined. There is even no evidence for the thesis that those terms would be the same in all natural languages and in all subcultures. Quite to the contrary, we witness children mastering, from a certain age on, the way in which the (wide) variety of those terms is employed in their environments. Even as grown-ups we are able to master the logical terms from mutually exclusive logical systems (and to actually use them in proofs *about* logical systems). So, how could we even arrive at hypotheses that determine which meanings are genetically fixed and which are not? Apart from all this, it is much more plausible, by present evidence, that genetically fixed matters are quite remote from anything like logical terms.

4. Flexible meanings in formal systems

It takes only a small step from plurilogism to a position that may be summarized as follows. The terms occurring in formal logical systems are sharp, univocal and static, whereas those occurring in natural languages are vague, ambiguous and flexible. Formal logical systems might be fit for mathematics and science, especially handbook science; perhaps it might even be fit for other finished theories as well. But we need natural languages and argumentation in those domains and/or situations in which the exactness of formal systems cannot be reached or should be purposively avoided. Among the latter are the recently disclosed domains where creativity plays a role, both in and outside the sciences. Similar positions are advocated by Perelman and Olbrechts-Tyteca (1958) and Perelman (1968) and by many later argumentation theorists.

As small and obvious as this step may be, it is mistaken. There is no reason why we might not devise formal systems the terms of which have flexible meanings. Consider even programmed computers. There is no reason why these would not be able to use terms in a flexible way, and, where communicating with each other, to trace the meanings of the terms used by the other party. To refuse the label "formal system" for the underlying logical systems would be just a matter of fiat.

Formal results that allow for flexible meanings are quite meagre. This holds even for those parts of artificial intelligence that concern discovery and creativity. Simon's BACON, for example, merely adds new terms that are devisable from available ones by simple algorithmic means. Contrary to Simon's claims, there is hardly any relation with real historical creative processes, as was convincingly shown by Fernand Hallyn (1993). For this reason, I think that the results I report below are valuable to the present discussion.

Before I continue, there is a point that I need to stress again. A realistic approach based on plurilogism not only requires that the intended tool is capable of mapping some term, say from a natural language, to one of the logical terms of a given formal system. It would not even be sufficient that it did so with respect to a given set of mutually exclusive formal systems. The central issue is that new meanings for logical terms may originate, for example while one is trying to get a grasp on some domain, and that such terms may occur in communication (hence that the addressee should be able to detect them).

Work on the interpretation of texts is relevant in this respect, but quite remote from formal systems. I shall start at the other end of the continuum, from results that were arrived at for other reasons. The results are poor, but they are relevant to our problem and they allow for some generalization.

5. Adaptive logics: the problem

In the next section, I consider examples of two kinds of adaptive logics: inconsistency-adaptive and incompleteness-adaptive ones. Most results on (my preferred) inconsistency-adaptive logics have been published or are in print — see my (1989)[1], (1986) and (199+a), which contains the most comprehensive technical description of the (predicative) logics. Some results on decision methods have still to be written up, and so do all results on incompleteness-adaptive logics. But let us start with a general characterization of an adaptive logic.

Consider a theory $\langle \Gamma, L \rangle$, where Γ is the set of axioms and L the underlying logic. L will contain several presuppositions about the domain described. For example, classical logic presupposes that the domain (as approached by observational and operational, or other criteria) is consistent (that the criteria do not, for some A, lead to both A and ~A). Sometimes Γ will violate some of these presuppositions, in which case we shall say that Γ has *abnormal* properties (with respect to the intended underlying logic). For example, where L is classical logic, the consequences of Γ may turn out to be inconsistent or to assert incompleteness (by way of non-logical theorems of the form ~(A∨~A)). If the abnormal properties cannot be readily removed, or if we have to reason about $\langle \Gamma, L \rangle$ in order to improve this theory, then neither L nor a monotonic weakening of L will do — see my (1989) for the inconsistency-adaptive case, other cases being analogous.

Here *adaptive* logics come in. They localize the abnormal properties of the theory, safeguard the theory for triviality by preventing specific rules of L from being applied to abnormal consequences of Γ, but behave exactly like L in all other cases.

The easiest way to understand how all this proceeds, is to realize that an adaptive logic 'oscillates' between the original logic L and a fragment Lf of L that differs from

---

[1]   The semantics presented in 1989 (written around 1981) should be forgotten as soon as possible in view of the extremely clarifying 1986 semantics.

L in not sanctioning the abnormal properties involved. If the abnormal property displayed by Γ is inconsistency, *Lf* will allow for inconsistencies (will not lead from inconsistency to triviality); if the abnormal property is (negation-)incompleteness *Lf* will allow for incompleteness (by not having such theorems as A∨~A or such rules as A⊃B, ~A⊃B / B). That the adaptive logic *La* oscillates between *L* and *Lf* may now be characterized intuitively, but somewhat inaccurately, by saying that *La* allows for the application of the (i.e. all) rules of *L*, except for applications to sets of consequences of Γ for which it is derivable from Γ that they display abnormal properties. This formulation is inaccurate because the "derivable" is not specified. The correct specification is somewhat complicated but, on closer inspection, turns out extremely intuitive.

I hope that the previous paragraph clarifies that an adaptive logic *localizes* the abnormal properties. At the syntactic level, a rule operates on finite sets of consequences of Γ (as in the case of any other logic); if a rule presupposes that a (specific) abnormal property is not involved, then it will be applicable or not applicable according as it is or is not derivable from Γ that the formulas included in the set have the abnormal property. To phrase it differently, the adaptive logic prevents that abnormal properties of specific consequences of Γ result in a trivial consequence set, but it does not restrict the rules of *L* in as far as they are applied to consequences of Γ that do not display abnormal properties. If applied to a normal theory, nothing has to be restricted and the adaptive logic *La* leads to exactly the same set of consequences as *L* itself.

Another way to look upon adaptive logics is to say that they interpret the premises as maximally normal. *L* presupposes normality. *Lf* gives up this presupposition (for some form of normality), thus heavily restricting on the set of consequences of Γ. *La* takes into account that Γ is abnormal at specific points, but goes on presupposing normality elsewhere, thus leading to a set of consequences that is a real subset of the *L*-consequence set iff the latter is trivial[2], but is in general a real superset of the *Lf*-consequence set.

It should be stressed that the adaptive character of the logics does not rely on any inventiveness (or even any intervention) on the part of whoever applies them: applying the adaptive logic leads to correct, although not necessarily interesting, results. Also, adaptive logics, at least, those I report upon below, have a nice and intuitive semantics that is directed precisely at maximizing normality.

Adaptive logics are non-monotonic (if $\Gamma \cup \{A\}$ is more abnormal than Γ, some B derivable from the latter need not be derivable from the former). Some adaptive logics, e.g., the examples I shall discuss, are decidable at the propositional level and exactly as undecidable as classical logic at the predicative level.

To end this section, I record some facts. Adaptive logics differ from the kind of logics usually labelled "non-monotonic logics" because of two (related) properties: (i)

---

[2]   In the present paper, the trivial set of sentences is the set of all formulas. This convention is handier here than the usual convention that calls a set trivial iff all formulas are derivable from it.

they do not involve any non-logical preferences and (ii) they do not rule out the abnormal properties. In a sense, they form the purely logical basis for some (the usual) non-monotonic logics: they localize the problems but do not resolve them. This result is established in Batens (199+b) in that a circumscription logic is reconstructed by (i) an inconsistency-adaptive logic (that 'minimizes' the inconsistent consequences), (ii) a purely logical mechanism, defined in terms of transformation rules, that connects a set of consistent models to the set of inconsistent models, and (iii) a (non-logical) preferential mechanism that selects the preferred set of models from the set of consistent models.

Some aspects of Nicholas Rescher's famous mechanism — Rescher (1968) and several later publications — are somewhat similar to adaptive logics. The main difference is that, at the syntactic level, Rescher's mechanism operates in terms of sets of premises, whereas adaptive logics operate in terms of deductive proofs. As a result, Rescher's mechanism is extremely dependent on the *formulation* of the set of premises; for example, {p, ~p, q, ...} and {p, ~p&q, ...} determine different sets of 'weak consequences'. For some applications this dependency is suitable, for others adaptive logics are preferable.

In his (1991) Graham Priest invokes adaptive logics to an end that is completely different from the one I originally intended, but proves very interesting from his philosophical stand. Priest is a (monologistic) dialetheist for whom the true logic is a paraconsistent (and relevant) one. He agrees, however, that in many situations we are justified in presupposing consistency. He goes on to show that, if his preferred paraconsistent logic *LP* (from Priest (1987)) is turned into an adaptive logic LP^m by assuming consistency "until and unless proven otherwise,"[3] then LP^m recaptures all classical reasoning where it is sensible (according to his so qualified dialetheist view).

6. Two adaptive logics: semantics

Although the syntax of adaptive logics is both more impressive and more realistic (with respect to actual revisionist thinking) than the semantics, I start with the latter because it is simple and intuitive. But first to the two forms of abnormality.

A simple paraconsistent logic is obtained by giving up the consistency requirement from classical logic *CL*. We keep binary connectives, quantifiers and identity unchanged but weaken negation to the completeness requirement (if $v(A) = 0$, then $v(\sim A) = 1$), dropping its converse, which is the consistency requirement. Let us call this logic *PIL*. Actually, *PIL* has an infinite number of paraconsistent extensions, obtained by adding such requirements as $v(\sim\sim A) = v(A)$, and some of these are maximally paraconsistent (have *CL* as their only non-trivial extension). Each of these

---

[3] This phrase is appealing but only accurate if it is not derivable from the premises that some formulas are connected with respect to their inconsistent behaviour. For example, p and q are so connected if $(p\&\sim p) \lor (q\&\sim q)$ is derivable from the premises, but neither disjunct is. In this case exactly one of the contradictions is true in each model of the premises.

results in different inconsistency-adaptive logics, but astonishing as it might seem, *adaptive* logics based in *PIL* seems preferable for most applications. A simple paracomplete logic *POL* is obtained by giving up the completeness-requirement instead of the consistency requirement. *POL* has an infinite number of paracomplete extensions, etc.

Consider the set of *PIL*-models.[4] All *CL*-models are *PIL*-models (and all *PIL*-models that do not contain any inconsistency are *CL*-models). Consider some set Γ of formulas and some formula A. A is a *CL*-consequence of Γ iff A is true in all classical models of Γ. A is a *PIL*-consequence of Γ iff A is true in all paraconsistent models of Γ. Clearly, as any set of premises has more paraconsistent models than classical models, its set of paraconsistent consequences will in general be a subset of its set of classical consequences.[5]

Where M is a *PIL*-model, let K(M) be the set of contradictions (formulas of the form A&~A) occurring in M.[6] There are at least two strategies to select maximally normal models from the *PIL*-models of some set of premises Γ. The first selects models on the basis of *reliability*. The idea is that, if (p&~p)∨(q&~q) is a (*PIL*-)consequence of the premises and neither p&~p nor q&~q is, then both p and q are considered unreliable. This leads to the inconsistency-adaptive logic *APIL1*. The second strategy, which is less cautious, proceeds by *minimizing abnormality*. A *PIL*-model M of Γ is selected if and only if there is no *PIL*-model M' of Γ such that K(M') ⊂ K(M). In other words, there are no models of Γ that are strictly less inconsistent than M. This strategy leads to the inconsistency-adaptive logic *APIL2*. The choice of a logic will obviously depend on the appropriateness of the strategy in a specific situation.

Given all this, we define: A is an *APIL1*-consequence of Γ iff it is true in all *APIL1*-models of Γ; in other words, in all *PIL*-models in which only unreliable formulas behave inconsistently. Similarly, A is an *APIL2*-consequence of Γ iff it is true in all *PIL*-models of Γ that are not more inconsistent than is required by Γ. If Γ is consistent, both logics select exactly the classical models of Γ. If it is not, they select no classical model, but in general (i.e., unless Γ is trivial) they select a subset of the *PIL*-models of Γ — and the *APIL2* models form a subset of the *APIL1* models. In the former case, the inconsistency-adaptive consequence set will be identical to the classical consequence set; in the latter case, the inconsistency-adaptive consequence set will be in general (i.e., whenever Γ is not trivial) a subset of the *CL*-consequence set

---

[4]   To keep things simple, let us consider a model as an ω-complete set of formulas throughout this paper.

[5]   There is only one exception, viz. when the set of premises is trivial itself.

[6]   If ω-incomplete models are included, the treatment becomes quite a bit more difficult — see Batens 199+a.

(the trivial set), but will be a superset of the *PIL*-consequence set (that takes a *larger* set of models into account).[7]

Although results were readily within reach, I stopped working, some ten years ago, on adaptive logics based on *POL* because I did not find sensible applications. Recently, a discussion with Diederik Aerts, who is doing advanced research in quantum physics, suddenly made me see the light: an incompleteness-adaptive logic (the philosophical rationale of which I evidently cannot discuss here) *might* spare us the awkward properties of quantum logics. The *APOL*-systems are nice counterparts of the *APIL*-systems, the two strategies now being defined with respect to formulas of the form $\sim(A \vee \sim A)$. Again, if Γ is complete, *APOL*-systems define the same consequence set as *CL*, whereas they define a poorer consequence set, but one richer than the *POL*-consequence set, in the opposite case.

7. Two adaptive logics: a glimpse on the syntax

The proof theory of *APIL*-systems and *APOL*-systems is most interesting and (in a specific sense) realistic. The proof procedure is dynamic (or revisionist) in that we start from the supposition that all formulas derivable from Γ are consistent (respectively negation-complete) unless and until proven otherwise — but compare note 3. It turns out indeed that the articulation of a proof procedure leaves us no other way than to rely on the formulas *that actually occur in the proof.* This is realistic in the sense of conforming to what happens in our 'natural' thinking: *to revise our view according as our understanding improves.* However, the proof procedure is still deterministic in the sense that, if we proceed sensibly, we shall eventually arrive at a result that may be defined statically. In other words, even if different people set off in different directions from the same set of premises Γ, they will all end up at the same fixed point.

The technical details are complicated but intuitive. A central feature is that formulas may be *connected* with respect to their inconsistent — respectively incomplete — behaviour. This will be expressed by formulas of the form $(A_1 \& \sim A_1) \vee \ldots \vee (A_n \& \sim A_n)$ — respectively $\sim(A_1 \vee \sim A_1) \vee \ldots \vee \sim(A_n \vee \sim A_n))$ — occurring in the proof in the absence of sub-disjunctions of them. Also, the aforementioned "proceeding sensibly" is a bit tricky, but strictly definable. And then, all this is simple at the propositional level, where everything is effectively decidable, but at the predicative level the usual lack of an algorithm for derivability interferes, and forbids even a general algorithm for "proceeding sensibly".

There is more tricky stuff, like the notion of a theorem. Both *APIL*-systems and *APOL*-systems have exactly the same set of theorems as *CL* if a theorem is defined by derivability from the empty set. However, their theorems reduce to those of *PIL* and

---

POL respectively, if theorems are defined as formulas derivable from any set of formulas.

Although all this is too technical to be continued here, I hope I made it clear that the notion of a formal system may have some very unexpected properties. All this is at the level of 'logic', even formal logic in the strict sense; no external preferences are involved, no non-logical terms, not even relevance requirements on either connectives or derivability.

8. Adaptive logics and argumentation

The main point I wanted to make is that it is quite possible for a formal logic (defined with respect to a formal language) to deal with *flexible* meanings of *logical* terms. Of course, some aim need to be determined in some way or other. In adaptive logics, it is determined by the specific points at which a theory displays abnormal properties of the given kind. The flexibility displayed by this procedure has some generality already. This shows at least a certain similarity with such problems as discovering new meanings from the interaction with some domain, or grasping unknown meanings hidden in a text. In the latter case, there clearly is not an intended logic (as the L in our couple $\langle \Gamma, L \rangle$). Yet, a preferred logic is given by assumptions deriving from the pragmatic context.

Of course, it would be nicer if these pragmatic assumptions themselves might be incorporated within the formal machinery. But then I set myself a modest task, which was to report on a machinery — an exact and formal one for that matter — that, given those pragmatic assumptions, proves able to deal with flexible meanings of logical terms.

The force of a logical machinery manifests itself in its applications. Fortunately, some applications of inconsistency-adaptive logics have been tried out, and with success. The most impressive one concerns the reconstruction of a creative discovery process from the history of thermodynamics. In her 199+, Joke Meheus considers the case of Clausius who forged a consistent theory from the inconsistent set comprising Carnot's thermodynamics as well as Joule's principle on the conversion of work to heat (and back) and a set of experimental results (mainly obtained by Joule). By relying on concrete passages from Clausius's text, she convincingly shows that the process may be reconstructed in terms of adaptive logics; and not in terms of classical logic, Rescher's aforementioned mechanism, or (the usual) non-monotonic logics (that are directed at handling rules with exceptions). The reconstruction is especially interesting because, in the presence of inconsistent premises, Clausius nevertheless applies a Reductio ad Absurdum (concludes to the falsehood of a supposition by showing that it leads to an inconsistency), and this application is indeed justified in view of the inconsistency-adaptive logic.

During the process by which Clausius transformed the inconsistent set of theories and data into his consistent theory, the meaning of "heat" (and many other terms)

changed drastically. A rather impressive aspect of the Meheus' reconstruction is that this change does not require any special treatment. The following hypothesis seems plausible in view of the reconstruction. Relying on non-logical preferences[8], Clausius arrives at his theory by stepwise eliminating (halves of) inconsistencies from the consequence set of the premises. But as some of the eliminated statements pertain directly to the meanings of the (non-logical) terms, the latter are modified at once. The elimination of other statements indirectly modifies those meanings. Roughly, this happens because of the connection between the meaning of terms and the accepted statements in which they occur (needless to say, this should be further specified[9]). This aspect has not been sufficiently studied. But if the analysis is roughly correct, *some* changes in the meanings of non-logical terms may be understood as consequences of the elimination of inconsistencies. In other words, the explicit flexibility of the meaning of negation, as it occurs in inconsistency-adaptive logics, is sufficient, *in such cases*, to understand the (implicit) flexibility of non-logical terms.

All I said up to here was related to a given logical frame, viz. a specific adaptive logic determined by a given maximal logic (*CL*) and a hypothesized minimal logic (*PIL*, respectively *POL*). Let me now turn to the question whether it is within the reach of algorithmic means to devise a minimal logic and next an adaptive logic in view of a specific abnormality problem. If we succeed in establishing this, we take another, rather remarkable, step from the stability of logical systems to the flexibility of argumentative procedures.

9. Devising adaptive logics

We shall learn more about flexible meanings by turning to the question as to how we arrive at an adaptive logic. (I shall try to stay as close as possible to the facts: the available systems, how I arrived at them, and how, by trying to generalize, other problems may be handled.)

We start from a problem: a theory showing some abnormal properties. This first of all presupposes a notion of normality (a mapping of the (logical) terms of some text to the terms of some given formal system). We may safely consider this to be determined by pragmatic considerations external to the formal task under discussion. Next we need a *criterion* for abnormality. Triviality clearly is a good indicator, but at first sight it seems too narrow. That a theory turns out non-trivial under some interpretation, does not seem to warrant the correctness of the interpretation. Yet, some reflection leads to a rather startling conclusion: triviality is sufficient as a criterion. Suppose that our interpretation of a text entails that its author X subscribes to a view

---

[8] Mainly: favouring data over theoretical statements and favouring some principles over others — the preferred ones derive from his world-view, which, however, is modified itself as an effect of his analysis.

[9] The main missing aspects are the role of *interpretations* and the (empirically supported) fact that only 'pieces' of the meanings of terms play a role in specific thought episodes.

A, whereas we assumed him or her to reject A. If we take other interpretations to be possible, we shall not consider our interpretation a sufficient reason to reject the assumption that X rejects A. Taking "X accepts A" and "X rejects A" to be strong negations of each other (presumably the only sensibly assumption) we arrive at a strong contradiction and hence at triviality. Or consider a case in which our interpretation leads to the conclusion that X utters a truism (like in interpreting "If it rains, it rains" as a statement of the form "p→p"), whereas we assumed that X was not uttering a truism; here again we end up with triviality. In other words, if we take our assumptions serious, triviality is a sufficient criterion for localizing problems.[10]

As a second step, we might identify the inference rules that lead to the problem. This offers important but confusing information. Indeed, the result will highly depend on the actual inference rules, whereas we know these to be exchangeable. In other words, the gathered information will be highly contingent on arbitrary choices.

Identifying combinations of derivable inference rules that lead to the problem sounds better, but again there is a difficulty: whenever an arbitrary statement A is derivable from some (obviously finite) set of premises Γ, there is (in general) an infinite number of sets of derivable rules of inference (even if we rule out supersets of other sets) by which A may be derived from Γ.

In order to arrive at a diagnosis of the problem, we need a theory about the *elements* of the meanings of the logical terms ("meaning-elements" for short). For example, with respect to classical logic, we standardly consider the meaning of implication to consist of three elements: if v(A) = 0, then v(A⊃B) = 1; if v(B) = 1, then v(A⊃B) = 1; if v(A) = 1 and v(B) = 0, then v(A⊃B) = 0. Please remark that these are not the standard semantic clauses, but rather a way of summarizing them (that mainly derives from standard metatheoretical proofs about classical logic).

Given a theory about the meaning-elements of the logical terms, the information about derivable rules of inference becomes relevant. We may now study which rules depend on which meaning-elements. Once this is accomplished, we know which of the meaning-elements leads to the problem. More correctly, we will have arrived at a set of sets of meaning-elements and we know that restrictions should be imposed on one of these sets.

If some set of meaning-elements is a subset of another, we leave out the superset. There are two reasons to do so. The first lies with the notion of normality: for a start we try to keep as close as possible to the standard interpretation defining normality. The second reason is a bit disappointing: if we do not introduce this restriction, then each meaning-element will occur in some of the aforementioned sets.[11] This does not entail that we might not have a good reason, later on, to get further away from the

---

[10] I am not arguing that a theory about the assumptions involved in the interpretation of texts is irrelevant, but rather that it need not be incorporated in the formal machinery for devising adaptive logics.

[11] Consider any meaning-element of some logical term. If A is derivable from Γ, then there is a derivation of A from Γ in which is applied a rule relying on the considered meaning-element.

standard interpretation defining normality (for example, doing so might lead to greater systematicity in the interpretation of the logical terms). But this again will be a consideration that is external to the formal machinery under discussion (compare note 10).

The third step consists in choosing the set of meaning-elements that we shall impose restrictions upon. If the abnormal property occurs systematically and frequently, there is only one such set (and nothing to choose). Whenever there is a choice to be made, we shall again have to rely upon assumptions external to our logical machinery. We might pick the set containing the least elements, or a set containing meaning-elements of one term only. Even these, however, are to be considered as extra-logical assumptions.

The fourth and final step consists in devising the adaptive logic. The extra-logical element here is that we have to choose a strategy (minimizing abnormality, or reliability — but more might be discovered later). Once this is done, devising the adaptive logic is completely straightforward. We first define a (monotonic) subsystem Lf of the original logic L by dropping the meaning-elements of the chosen set. Then we define abnormal properties of models; these are typically properties of Lf-models that do not occur in the L-models.[12] Here are some examples: the presence of a formula of the form A&~A, the absence of a formula of the form A∨~A, the absence of both A and B in the presence of A∨B, etc. The adaptive logic La is arrived at by defining the maximally normal models in view of the chosen strategy. This procedure is clearly algorithmic.[13]

All this may seem somewhat theoretical. However, some nice and ready problems may function as a test for the procedure sketched. Simple one's, like the lottery paradox, and complex ones like the paradox of Curry and Moh Shaw Kwei. Neither of these has anything to do with negation.

Before leaving the matter, I want to stress the important role played by the theory about the meaning-elements of the logical terms. If this theory is based on the standard semantics of classical logic, the meaning of each connective reduces to two or three elements only. If it is based on, say, the Routley-Meyer semantics for relevant logics, the meaning of each connective (of classical logic) consists of a host of (independent) elements — see, e.g., Routley (1982). In the latter case, deviations from normality will be much smaller, which is not necessarily preferable. The importance of philosophical theories should be stressed in this connection. Each of these semantic systems, as most others, leads to theories that *isolate* the meanings of logical terms. Again, there is a philosophical position behind this. The position may be justified, but it is neither unimportant nor straightforward (as many logicians seem to presume). Frege did some

---

[12]   By the definition of Lf, all L-models are Lf-models, but not conversely.

[13]   It is still an open problem whether there is an algorithm for devising proof methods that are dynamic in the sense explained in section 7.

excellent thinking for classical logic; after him, we got too much technique and bad metaphysics.

References

Apostel, L. (1971). 'Assertion logic and argumentation.' *Philosophy and Rhetoric* 4, pp. 92-110.

Batens, D. (1985). 'Meaning, acceptance, and dialectics.' In: J. C. Pitt (ed.), *Change and Progress in Modern Science*. Dordrecht: Reidel, pp. 333-360.

Batens, D. (1986). 'Dialectical dynamics within formal logics.' *Logique et Analyse* 114, pp. 161-173.

Batens, D. (1989). 'Dynamic dialectical logics.' In: G. Priest, R. Routley & J. Norman (eds.), *Paraconsistent Logic*. München: Philosophia Verlag, pp. 187-217.

Batens, D. (1990). 'Against global paraconsistency.' *Studies in Soviet Thought* 37, pp. 27-47.

Batens, D. (1992a). 'Do we need a hierarchical model of science?' In: J. Earman (ed.), *Inference, Explanation, and Other Frustrations. Essays in the Philosophy of Science*. Berkeley-Los Angeles-Oxford: University of California Press, pp. 199-215.

Batens, D. (1992b). *Menselijke Kennis. Pleidooi voor een bruikbare rationaliteit*. Garant: Leuven-Apeldoorn.

Batens, D. (199+a). 'Inconsistency-Adaptive Logics.' In: E. Orłowska (ed.) *Logic at Work. Essays Dedicated to the Memory of Helena Rasiowa*. Dordrecht: Kluwer.

Batens, D. (199+b). 'Inconsistency-Adaptive Logics and the Foundation of Non-Monotonic Logic.' *Logique et Analyse*.

Hallyn, F. (1993). 'La troisième loi de Kepler et la 'psychologie de la découverte'.' *Archives internationales d'histoire des sciences*. vol. 43, pp. 247-257.

Meheus, J. (199+). 'Adaptive logic in scientific discovery: the case of Clausius.' *Logique et Analyse*.

Perelman, Ch. (1968). *Logique et argumentation*. Bruxelles: Presses Universitaires de Bruxelles.

Perelman, Ch. & L. Olbrechts-Tyteca (1958). *Traité de l'argumentation*. Paris: PUF (many later editions).

Priest, G. (1987). *In Contradiction. A Study of the Transconsistent*. Dordrecht: Nijhoff.

Priest, G. (1991). 'Minimally inconsistent LP.' *Studia Logica* 50, pp. 321-331.

Rescher, N. (1968). *Hypothetical Reasoning*. Amsterdam: North-Holland.

Routley, R. (1982). *Relevant Logics and their Rivals*. vol. 1, Atascadero, Ca., Ridgeview.

Index