# 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 $\varphi$ is derivable from a theory $\Sigma$, where $\Sigma$ represents a set of premises, then $\varphi$ is also derivable from every superset of $\Sigma$. 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 cooperativity 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, unambiguous 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 $\Delta$ 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 $\Delta$. 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 $\Sigma$ can be reduced to a monotonic inference based on a superset of $\Sigma$.

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 parameters 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 advantages. 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: *Artifficial 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.