

Report Number: WUCS-92-43

1992-10-01

# Process and Policy: Resource-Bounded Non-Demonstrative Reasoning

Authors: Ronald P. Loui

This paper investigates the appropriateness of formal dialectics as a basis for non-monotonic and defeasible reasoning that takes computational limits seriously. Rules that can come into conflict should be regarded as policies, which are inputs to deliberative processes. Dialectical protocols are appropriate for such deliberations when resources are bounded and search is serial.

AI, it is claimed here, is now perfectly positioned to correct many misconceptions about reasoning that have resulted from mathematical logic's enormous success in this century: among them (1) that all reasons are demonstrative, (2) that rational belief is constrained, not constructed, (3) that process and disputation are not essential to reasoning. AI mainly provides new impetus to formalize that alternative (but older) conception of reasoning, and AI provides mechanisms with which to create compelling formalism that describes the control of processes.

The technical contributions here are: the partial justification of dialectic based on controlling search; the observation that non-monotonic reasoning can be subsumed under certain kinds of dialectics; the portrayal of inference in knowledge based on policy reasoning; the review... **Read complete abstract on page 2.**

Follow this and additional works at: [http://openscholarship.wustl.edu/cse\\_research](http://openscholarship.wustl.edu/cse_research)



Part of the [Computer Engineering Commons](#), and the [Computer Sciences Commons](#)

---

## Recommended Citation

Loui, Ronald P., "Process and Policy: Resource-Bounded Non-Demonstrative Reasoning" Report Number: WUCS-92-43 (1992). *All Computer Science and Engineering Research*.  
[http://openscholarship.wustl.edu/cse\\_research/605](http://openscholarship.wustl.edu/cse_research/605)

---

# Process and Policy: Resource-Bounded Non-Demonstrative Reasoning

## **Complete Abstract:**

This paper investigates the appropriateness of formal dialectics as a basis for non-monotonic and defeasible reasoning that takes computational limits seriously. Rules that can come into conflict should be regarded as policies, which are inputs to deliberative processes. Dialectical protocols are appropriate for such deliberations when resources are bounded and search is serial.

AI, it is claimed here, is now perfectly positioned to correct many misconceptions about reasoning that have resulted from mathematical logic's enormous success in this century: among them (1) that all reasons are demonstrative, (2) that rational belief is constrained, not constructed, (3) that process and disputation are not essential to reasoning. AI mainly provides new impetus to formalize that alternative (but older) conception of reasoning, and AI provides mechanisms with which to create compelling formalism that describes the control of processes.

The technical contributions here are: the partial justification of dialectic based on controlling search; the observation that non-monotonic reasoning can be subsumed under certain kinds of dialectics; the portrayal of inference in knowledge based on policy reasoning; the review of logics of dialogue and proposed extensions; and the pre-formal and initial formal discussion of aspects and variations of dialectical systems with non-demonstrative reasons.

**Process and Policy: Resource-Bounded Non-  
Demonstrative Reasoning**

**R. P. Loui**

**WUCS-92-43**

**October, 1992**

**Department of Computer Science  
Washington University  
Campus Box 1045  
One Brookings Drive  
St. Louis MO 63130-4899**



# Process and Policy: Resource-Bounded Non-Demonstrative Reasoning

R. P. Loui  
Department of Computer Science  
Washington University  
St. Louis, MO 63130

## Abstract.

This paper investigates the appropriateness of formal dialectics as a basis for non-monotonic and defeasible reasoning that takes computational limits seriously. Rules that can come into conflict should be regarded as policies, which are inputs to deliberative processes. Dialectical protocols are appropriate for such deliberations when resources are bounded and search is serial.

AI, it is claimed here, is now perfectly positioned to correct many misconceptions about reasoning that have resulted from mathematical logic's enormous success in this century: among them, (1) that all reasons are demonstrative, (2) that rational belief is constrained, not constructed, (3) that process and disputation are not essential to reasoning. AI mainly provides new impetus to formalize the alternative (but older) conception of reasoning, and AI provides mechanisms with which to create compelling formalism that describes the control of processes.

The technical contributions here are: the partial justification of dialectic based on controlling search; the observation that non-monotonic reasoning can be subsumed under certain kinds of dialectics; the portrayal of inference in knowledge bases as policy reasoning; the review of logics of dialogue and proposed extensions; and the pre-formal and initial formal discussion of aspects and variations of dialectical systems with non-demonstrative reasons.

This work was supported under NSF R9008012.

I am grateful for the invitations to talk at Linköping, Amsterdam, Illinois, SLU, TRI, SIGART, and Rochester, which helped me organize this material.

Jon Doyle and Michael Loui, Gerard Vreeswijk, Patrick Doherty, and John-Jules Charles Meyer have been especially supportive. I thank Peter Ladkin for listening to these ideas at KR89 Toronto.

## 1. Arguments and Demonstration.

Sometimes there is proof; mostly, there are arguments. A proof demonstrates its claim once and for all. An argument, in contrast, produces warrant for its claim only when it is effective, in a context of counterarguments, rebuttal, and further counterargument. It produces warrant as a result of a process which subjects the claim to dispute.

This distinction between argument and proof is historically upheld: there is demonstrative reasoning and there is non-demonstrative reasoning.

Mathematical proof is the model of demonstration. Demonstrative reasoning has been the basis of mathematically satisfying formal languages. For non-demonstrative reasoning, formal linguistic systems have been developing more slowly in this century.

Major historical figures have written about this distinction: from Aristotle, who introduced the study of both; to Thomas Aquinas, whose discussions of argument occurred during a time when, reversing current thought, non-demonstrative reasoning dwarfed demonstrative reasoning as a topic of intellection; to the unsuccessful admonitions of John Maynard Keynes against Russell and Whitehead as they were conscripted under Hilbert's reductionist mathematicism; to Karl Popper's attempt to distinguish respectable study of disputation from Hegelian dialectical metaphysics; to Alonzo Church's infamous claim of dialectic's intrinsic incompatibility with formalism.

For a discussion of Aristotle and dialectics, see Hogan. For a discussion of Keynes, see Carabelli. For Aquinas, see Byrne. Church and Popper references are in the bibliography.

AI's interest in this decade in non-monotonic reasoning forces a return to the non-demonstrative paradigm. Formalists in AI, especially those of the knowledge representation milieu, have not yet acknowledged that their work connects to a longer tradition. Applied AI research in law, discourse, and uncertainty and have shown better understanding and better scholarship on this matter.

## 2. Reasons and Policies.

Reasoning begins not only with claims, but also with reasons. Chaining reasons forms arguments for further claims.

Some take the relation of one claim being reason for another claim to be primitive; others feel that reasons can be derived from other reasons, or warranted by argument, particularly statistical argument.

In the long tradition of logical thought, the reason relation has been thought to have many nuances. This century's logical mainstream departs from the longer intellectual tradition at Rus-

sell; he insisted on a narrow formulation of reasons uniformly as mere (disjunctive) claims. Inductive logicians distinguish themselves by their conviction that all reasons derive from probabilistic considerations. Assimilating reasons as subjunctive conditionals has been popular recently. All three are reductionist programmes that construe reasons too narrowly.

The defeasible reason, as studied principally by John Pollock in the past three decades, is a serious alternative to the narrow and reductionist views of reasons. Where Pollock has focused on the logic of defeasible reasons, [74] [87] [92], the present discussion focuses on the process in which those reasons participate.

Represented knowledge can include knowledge about how to conduct reasoning about that knowledge. Reasons are that kind of knowledge; a kind of meta-knowledge. Representing object-level knowledge has been well-studied. Representing certain transformations of knowledge has also been well-studied, but under the assumption that the only transformations are those that rewrite object-level knowledge. “Inference rules” define, solely as a function of asserted sentences, a set of sentences as theorems, which are the results of transformations. Legitimate transformations are only those “sound” or “truth-preserving” transformations that rewrite the meanings of what is already asserted. The transformations define a shorthand system in which sentences are succinct ways of writing all of their entailments. This is true of inductive entailments and non-monotonic entailments, as well as deductive entailments.

A representer who conveys knowledge as “p” and “if p then q” also conveys as knowledge “q”, implicitly. This is what is meant by using both sentences in a propositional language, *PC*. Both Jon Doyle [89] and I [91] have recently written at length about this view of implicit knowledge (the phrase is Levesque’s, but used differently by him) and its alternatives (a more comprehensive work is Gottlieb [68]). Philosophers have called “non-ampliative” all inferences that merely rewrite meanings of sentences. Any consequence relation that is a function only of sentences can be viewed as merely providing meanings for those sentences in a system of representation; thus, the consequences are non-ampliative.

The ampliative alternative is to construct knowledge using meta-knowledge, and to let a process mediate explicit sentences, explicit reasons, and their implicit (or constructible) consequences. Such a process would utilize reason relations between claims as if they were knowledge about how the process may proceed.

AI already widely recognizes procedural meta-knowledge. But it has not yet recognized reasons as one kind of procedural meta-knowledge. Default rules, non-monotonic rules, expert system rules, and inheritance links are knowledge about how to construct knowledge: they are more like heuristics in a theorem prover than they are like assertions that are input to the theorem prover.

Reasons are policies. They are policies for constructing arguments.

Sometimes the grounds for adopting a policy can be represented: for example, the game-theoretic justification of the policy “castle early in chess” might be representable as a cache of analyzed positions; the statistical arguments that justify a policy “arguments for something being a bird” can be extended to arguments for that something being able to fly” might be represented as the summary of sampling with the description of sampling procedure.

Most policies are adopted simply on the grounds that they come from authority: legislative bodies, for example, can create policy by fiat. This is the situation with rule-based programs and knowledge-representers, or with inheritance systems and their authors. Grounds for policies are not represented. If for some policies, grounds are represented, while for others, they are not, this forces policy-based reasoning. Policy reasoning is the common denominator for processing reasons adopted on different grounds. Conflict among policies forces deliberations upon the disputes. Ideally, policies do not conflict and can be compiled into decision tables; but in reality policies are twisted and piecemeal, have multiple sources, and are inconsistent: it is often unclear what are their collective pronouncements. Because policies are defeasible rules, their individual pronouncements are relevant only inasmuch as they contribute to collective pronouncements.

This is exactly the observation that defeasible reasoning is non-local, non-modular, or holistic; this holism separates (syntactic) non-monotonic systems from modular inferential systems.

Policies do not share many of the behaviors of demonstrative reasons. Many non-monotonic reasoning research programmes have formed expectations for the behavior of defeasible reasons based on behaviors of deductive or probabilistic entailment relations; researchers even take these expectations as uncriticized axioms of their non-demonstrative systems. These aims are regrettable. Even a probabilistic account of policies is misdirected since not all policies have probabilistic grounds.

Though Pearl [91] for example, adequately explains why behaviors motivated by other kinds of reasoning, such as axioms from  $\mathcal{E}$ -semantics, may still be interesting as guides in choice of convenient communication convention: "The quest for probabilistic semantics is motivated by the assumption that the conventions of discourse are not totally arbitrary, but rather, respect certain universal norms of coherence, norms that reflect the empirical origins of these conventions."

Consider the following pair of policies pertaining to motor vehicle accidents:

*being rear-ended is reason for not being liable;*

*being rear-ended after an abrupt elective stop is reason for being liable;*

these do not contrapose, do not survive logical strengthening of the referent (the first of the objects related; the 'left-side'), nor weakening of the relatum (the second of the objects related; the 'right-side'). Arguments that make use of a policy might allow weakening of that policy's claims (thus, deriving a weaker argument from an argument), but altering the policy itself by weakening relata (thus, deriving a policy from a policy) violates the spirit of a policy. These policies can also be embedded in a system of rules wherein it strains intuition to admit reasoning by cases, or to create new policies by chaining, as if the reason relation were transitive (Loui [87]).

### 3. Meaning and Process.

Part of the meaning of a policy is the policy-maker's understanding that the policy will be used *eristically*, that is, in processes of disputation where arguments are based on policies.



That policies can conflict is clear: so clear, that the dialectician's song of "synthesis from conflicting thesis and antithesis" is banal. More interesting is that disputations based on policy often warrant conclusions because of the particular way in which the disputation occurred. Repeating the disputation under similar conditions is not guaranteed to produce the same outcome. Nevertheless, the conclusion is warranted. Indeed it is odd to say that we accept an outcome even though we would have accepted the opposite, had it occurred as the outcome instead.

This non-determinism

It is not an indeterminate result, nor is it necessarily probabilistic; however, the rule for acceptance differs from that for non-deterministic machines.

is the main difference between ampliative inference and mere expansion of shorthand. Consequences depend on non-deterministic choices that are input to the process, and on the protocol for the process. The function that maps represented knowledge to constructed inferences not only depends on the entire set of claims, but also must be indexed by a particular process: its individuating features, such as particular non-deterministic choices, and the context of the process, such as resource bounds that were imposed and the regimen distributing those resources.

Another difference is non-monotonicity in computation: had the process continued, a different result might have emerged. Note that this non-monotonicity is not a property of syntax: syntactically non-monotonic systems

defined in the familiar way in terms of non-monotonic growth of theorems with respect to axioms: i.e., there does not exist  $S, T$ , s.t.,  $\text{Thm}(S \cup T)$  does not contain  $\text{Thm}(S)$ .

do not necessarily define results for partial computations. It is non-monotonic dynamics of partial computations that deserves attention: the phenomenon that more computation could cause retraction of conclusions; mere non-monotonicity of syntax is less interesting.

Inference is constructed through process. Policies do not demonstrate conclusions; they are ingredients to a process that warrants conclusions. In general there is no ideal process,

Perhaps we should say there is no ideal tokening of the process type.

no correct outcome of the process. There is sometimes a natural termination: for instance, when the set of arguments is finite and can be exhausted, or when protocol leads to deadlock. The outcome at a natural point of process-termination (which still may be a non-deterministic outcome) may be desirable. But this is because we prefer processes that terminate at natural points of termination over those that terminate ungraciously; we prefer unbounded processes over bounded ones (not necessarily because they produce ideal results). We need not insist on an ideal; not every process computes an approximation of some ideal.

Constructing belief non-deterministically is anathema to mathematical logic. The idea, however, has been successful elsewhere. In mathematical statistics, for example, the dominant view is that hypotheses undergo testing. What makes a statistical hypothesis acceptable is the process by which it is conceived and tested, not just the relation it bears to other statistical assertions. The testing could have returned a different answer, and we are bound (at least until further hear-

ing) to the outcome no matter what it is. This Neyman-Pearson view has dominated the Bayesian alternative. Data are most important to acceptability, but not the sole arbiters.

Since Popper, Lakatos, and Kuhn, the constructivist view of scientific theory formation is mainstream, not the Carnapian view based on theory's probabilistic relation to data. In decision theory, the works of Simon and of Shafer and Tversky use constructivism to address Savage's and von Neumann-Morgenstern's non-constructivist shortcomings. Heuristic approaches to optimization, including the outputs of connection networks that attempt to minimize energy functions, but settle in local minima, are almost always constructive.

Not all constructions are heuristic. Some properties, especially social properties, are defined to hold by construction. What makes a deserving Supreme Court Justice is the positive outcome in a confirmation process. What makes political mandate is the fairness of the election. What makes a championship sports team is its prevailing in the playoffs. Most footprints in the sand are not the right size and shape, but are caused by feet and subjected to the right erosive processes.

There is nothing odd about having both a constructivist and a constraint-based view of what makes something what it is. What makes a belief rational, in one sense, is the relation it bears to other beliefs. This is a constraint imposed on claims of rationality. But a constructive conception is also possible. What makes beliefs rational could be the way in which they are constructed: they are the outcomes of the right kind of deliberative process.

The distinction is not merely one of statics versus dynamics. Axiomatic theories of belief-revision and axiomatic theories of time-attitudes toward preference seek to impose constraints on the dynamics of belief, but do not thereby advocate a constructive view. What marks constructivism is the non-determinism (and to a lesser extent, the non-monotonicity) of the construction.

Sometimes it is possible to have both constructive and constraint-based grounds for a claim. For example, at the completion of a Solovay and Strassen [77] probabilistic primality test, the conclusion has both been vindicated by a test of the right kind and is highly probable given the data. In that case, the propriety of the test has probabilistic origins.

In a trite sense, all beliefs are the result of some metaphysical non-determinism: the world could have differed from what it was. A rational agent could have come to believe  $p$  instead of  $\neg p$ , by whatever means of belief formation. Furthermore, this belief is revisable, which is a kind of non-monotonicity. Is this construction? We are interested only in constructions that are based upon a fixed set of claims. Even for a fixed set of background beliefs, there is disagreement about inference. Constraint-based, non-ampliative, rewriting, demonstrative, Russellian inference gives one answer: only one construction is permissible; while constructivist, ampliative, dewriting, non-demonstrative, Keynesian inference gives another answer: many constructions are permissible.

#### **4. Dialectic and Protocol.**

The importance of process is widely recognized in many intellectual fields. Computer science has the additional responsibility of specifying the process.

One of the right kinds of process for constructing rational belief is dialectic.

Dialectic refers to one form of disputation in which a serializable resource is distributed so that one party's use of that resource is informed by the result of the other party's (or parties') prior use of resource. That is, disputation that is dialectical involves response. The serialized resource is typically either search for arguments or time for presentation of arguments, but could also be the adjudicator's sequential consideration of arguments.

Dialectics can be immediate-response or not. If response is immediate, then satisfaction of some condition, such as the changing of adjudicator's current opinion, or the mere presentation of an argument, causes a switch in control of resource. Most evolved systems of disputation are dialectical, but not immediate-response.

I doubt that the protocol of a disputation must be dialectical in order for the outcome to be rational, constructed belief. Still, dialectical protocols satisfy certain desiderata having to do with effectiveness and fairness of the process.

Lobbying protocols are inappropriate. Policies conflict and lead to conflicting arguments. Producing arguments for one side of a dispute is too easy when reasons abound. Goal-directed search for arguments must admit search-targets from both sides.

Consider the example of negation-as-failure in logic programming or default reasoning, where search for counterargument is restricted to demonstrative counterargument. Suppose two default rules (Reiter [80]):

1.  $\frac{A : B}{B}$  ; and
2.  $\frac{A : \neg B}{\neg B}$  .

Given  $A$ , try to use the first rule. In order non-demonstratively to conclude  $B$ , show demonstratively that  $B$  is consistent. This is undecidable if  $B$  is a sentence of a sufficiently expressive language, or computationally expensive in any case; so to show this, try to prove  $\neg B$  and fail. Fail by consuming all resources on the attempted proof. Therefore,  $B$ . Note that the opposing rule would not be considered. A much better strategy would be to take

1.  $A$  to be reason for  $B$ ; also
2.  $A$  to be reason for  $\neg B$ .

Given  $A$ , pro advances the argument for  $B$ .

If time does not remain at this point, the same answer,  $B$ ,  
might be allowed, but for very different reasons.

If time remains, con advances the argument for  $\neg B$ . If time still remains, one side tries to resolve the dispute in its favor. This last search is the one that fails because of exhausted resource.

This is a bit unfair as a portrayal of resource-bounded default reasoning. The form of the default rule requires resource-exhaustion in order to advance an argument. The attempt to prove  $\neg B$  is an

invitation of counterargument, with which we agree; the problem is that counterargument must be demonstrative. Surely a better implementation awaits the clever deployer of resource-bounded defaults. Still, the negation-as-failure example illustrates how non-monotonic reasoning, in its early forms, ignores dialectical ideas, resource distribution, and the considerations of fairness under resource-bounded construction that lead to dialectical ideas.

An even more extreme example of lobbying protocol starts with the latter pair of defeasible rules. The argument for  $B$  is advanced. Then search continues to target the construction of arguments for  $B$  until resources are exhausted. The results of deliberation based on such a protocol are clearly unacceptable.

The fairer the protocol of a disputation, and the better the strategic play, and the more effective the expenditure of resource, the better warranted the outcome. Dialectic ensures two properties that bear on fairness and effectiveness:

1. when a side is losing, it gets resources;

this is fair in at least one sense;

2. when an opinion results, maximum resources were spent attempting criticism (and failing);

this is effective.

Nicholas Rescher (whose monograph is the most sensible disquisition to date on the subject) justified dialectic by claiming that knowledge is social, hence the formation of knowledge must meet social standards.

The root issue in probative rationality is that of “building up a good case” to enlist the conviction of one’s fellows. Accordingly, probative standards are person-indifferent; they are inherently public and communal. ... The very conception of duly validated knowledge claims relates to publicly established and interpersonally operative standards. [p. 60]

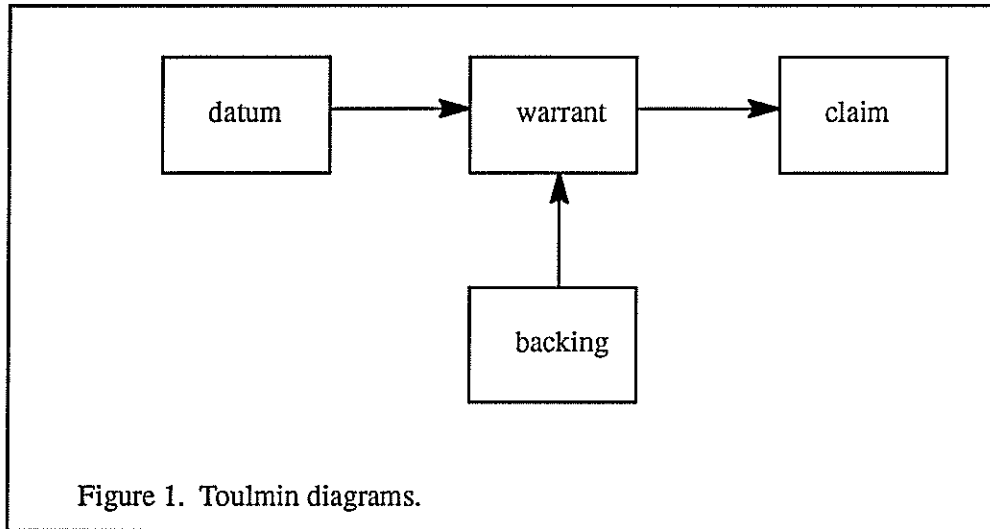
A different justification is possible, by referring to fairness, effectiveness, and the need to adjudicate resource-bounded disputes (this is closely related to Simon’s call for procedural rationality, heuristics, and satisficing). Fairness demands maximal opportunity for response, and effectiveness demands maximal information about what are the aims of response. Fair protocol and effective advocacy are required of a dispute for rational believers to accede to a process’s outcome.

When resources are unbounded and arguments exhaustible (which is impossible if arguments are infinite), or when search is completely parallelizable and strategies for adjudication can cope with unserialized argumentation, then perhaps dialectic is unnecessary even in disputational deliberations. There are other criteria of fairness, such as equivalent expenditure of resource, which then dominate the selection of protocol.

## 5. Existing Work.

Work related to the logic of disputation is considerable, though most of it is not widely known, even among philosophical logicians. In AI there are also a few precedents. Literatures of rhetoric, legal philosophy, and debate are all relevant, but not formal in a useful way.

The diagram of Stephen Toulmin [58] is currently popular among those who must pictorially



represent non-demonstrative arguments (e.g., Marshall et al. [91]). Toulmin's form for an argument has four parts: claim, warrant (the non-demonstrative reason which allows the claim), datum (the evidence needed to use the reason), and backing (the grounds for the reason) (see figure 1). In AI's Tweety example, the claim is that Tweety flies, the warrant is that birds fly, the datum is that Tweety is a bird, and the backing is some statement of the grounds for the non-demonstrative reason: presumably some statistical claim about flying's frequency among birds. Counter-argument can attack any of these elements. Toulmin diagrams can be chained for visual effect and are currently investigated mainly for user-interface applications (Marshall et al. [91] and Cavalli-Sforza et al. [92]). Toulmin's essay is of historical interest and suggests that proper formalism will be found in modal deontic logic, which was fashionable at the time.

A problem with Toulmin's essay is that little guidance is given regarding the distinction between datum and backing. Readers of non-monotonicity's literature know data to be nodes and warrants to be links; backings, then, are the equally mysterious justifications for links. As a form in which to present arguments pictorially, Toulmin form is minimalist: macro forms that are structured collections of Toulmin schemata could be defined for particular styles of non-demonstrative argument, such as analogy, inference to the best explanation, decision under risk, and so forth, which would perhaps be more cogent.

Lorenz "strip" diagrams were invented in a subsequent German dissertation [61]. These diagrams consist of a two-column table with one column labeled "proponent" and the other labeled "opponent," though labels "attacker" and "defender" are also used (see figure 2). Play alternates between the two sides, and the symbols that can be written as legitimate responses are regulated by rules that provide little choice. Winning positions are formally defined. The effect of attacks, responses, and winning is essentially the building of a partial semantic tableaux, or the

Opponent	Proponent
	$[(a \supset b) \supset b] \supset [(b \supset a) \supset a]$
$(a \supset b) \supset b$	
	$(b \supset a) \supset a$
$b \supset a$	
	$a \supset b$
$a$	
	$a$

Figure 2. A Lorenz strip diagram in which the opponent fails to show that the proponent's first locution is not a tautology.

traversal of one path in a refutation tree. The rules for responses are basically rules for a game that would determine whether a sentence is a tautology, except that the underlying logic is intuitionist.

A large body of work follows Paul Lorenzen and Jaakko Hintikka's use of Lorenz's strip diagram for game-theoretic semantics, noting that they are especially well-suited for intuitionist logic (see Krabbe [85] for example). Lorenz however demonstrates a broader conception of this diagram's use when he suggests in a later paper [73] that the dialogue-game could be limited to a finite number of attacks, or a finite number of plays, thus suggesting resource bounds.

A large literature also surrounds Jim MacKenzie's augmentation of Lorenz games to include locutions from Charles Hamblin's theory of dialogues (see Walton [85] for a review). The main additions are the use of interrogatives (hence, its affinities to erotetic logic) and the concept of information-store or commitments (see figure 3). Systems in this paradigm have developed quite formally. MacKenzie [90] and Walton [89] are examples. Hamblin's system purports to model exchanges of information through dialogue. MacKenzie's systems appear to revive medieval styles of dialectical reasoning: a player might adopt inconsistent or unwarranted commitments, especially because of limited inferential capacity, and interrogation seeks to put the set of commitments in order. Walton's games are particularly suited to modeling informal fallacious reasoning, such as question-begging or ad hominem arguments. All of these systems are interesting inasmuch as they are formalizations of two-player protocols over formal logical languages.

Speaker and Locution	Commitments
A: Is it the case that P?	A:
B: P	B: A: P
A: Q	B: P A: P, Q
B: $P \supset R$	B: P, Q A: P, Q, $P \supset R$
A: $R \supset \neg Q$	B: P, Q, $P \supset R$ A: P, Q, $P \supset R$ , $R \supset \neg Q$
B: Resolve {P, Q, $P \supset R$ , $R \supset \neg Q$ }	B: P, Q, $P \supset R$ , $R \supset \neg Q$ A: P, Q, $P \supset R$ , $R \supset \neg Q$
A: Withdraw P	B: P, Q, $P \supset R$ , $R \supset \neg Q$ A: Q, $P \supset R$ , $R \supset \neg Q$
B: $R \supset S$	B: P, Q, $P \supset R$ , $R \supset \neg Q$ A: Q, $P \supset R$ , $R \supset \neg Q$ , $R \supset S$
A: Resolve {P, Q, $P \supset R$ , $R \supset \neg Q$ }	B: P, Q, $P \supset R$ , $R \supset \neg Q$ , $R \supset S$
B: Withdraw $R \supset \neg Q$	A: Q, $P \supset R$ , $R \supset \neg Q$ , $R \supset S$
	B: P, Q, $P \supset R$ , $R \supset S$

Figure 3. Part of one play of a MacKenzie dialogue game.

MacKenzie's system is the closest in spirit to the present work. The medieval "obligation games" sought to land opponents in contradiction mainly by asking questions that required answers that created commitments. Commitments are added to a player's store because the player must answer questions raised by the opponent. When an inconsistency arises, the player may be asked to resolve the inconsistency by withdrawing an earlier commitment. Certain logical forms and inferences cannot be withdrawn. The games rely on restricted forms of implication, and consistency, each qualified by the adjective "immediate." The underlying logic of immediate implication and immediate inconsistency need not include anything more than modus ponens for conditionals.

Logics of dialogue do not however employ non-demonstrative reasons. When inconsistency arises, a player may retreat by withdrawing claims or conditionals. The missing option is that a player might retain all claims and conditionals and resolve *prima facie* inconsistency by considering some (chains of) conditionals to be defeated. Argument systems, inheritance, and non-monotonic systems also take claims and conditionals to be largely shared, which logics of dialogue do not. Still, the relationship between logics of dialogue and what AI is attempting in non-demonstrative reasoning has many facets deserving further attention.

Investigations of logics of dialogue demonstrate considerable scholarship in pre-Fregean logical history. Much of their polemical writing against the direction of logic in this century could equally be applied here, in making room for formal non-demonstrative reasoning.

By far, the most relevant work is the monograph of Nicholas Rescher [77], which follows his monograph on plausible reasoning [76]. The later work has more focus and formalism, but has been heretofore largely neglected compared to the earlier work.

The formal system Rescher develops begs many questions. But it is undoubtedly the most elegant system to date; his essay has a clarity and purposiveness that could make it last for centuries. His discussion is complete. Although he dedicates the essay to famous dialecticians, it is the book that every dialectician wishes he had written.

Rescher says he seeks:

... a disputational approach to inquiry, ... [to pose] the prospect of utilizing disputation and rational controversy as a model for issues in the theory of knowledge. [Preface]

... The prime aims ... are to exhibit the sociocommunal roots of the foundations of rationality, ... to illuminate the communal and controversy-oriented aspects of rational argumentation and inquiry — *scientific* inquiry in particular. [p. xiii]

Rescher's system augments Lorenz strip diagrams mainly by adding non-demonstrative reasons and informally discussing termination based on plausibility (see figure 4).

proponent	opponent
!P	$\neg P/Q \ \& \ \dagger Q$
! $\neg Q$	$Q/R \ \& \ \dagger R$
! $\neg R$	$R/S \ \& \ \dagger S$
$!\neg R/(S \ \& \ T) \ \& \ !(S \ \& \ T)$	$Q/U \ \& \ \dagger U$
$!\neg Q/(U \ \& \ V) \ \& \ !(U \ \& \ V)$	

Figure 4. A Rescher dispute.

I have found no further development of Rescher (though it is occasionally referenced). The dialogical logic community could make no sense of reasons which relate sentences and the sen-



tences for which they are reasons. Walton's dissertation recoiled at "provisoed assertion," since (paraphrasing Walton [84]) it is not transitive, it does not detach, its reflexivity and symmetry properties are unknown, and not everything follows from inconsistency. It is supposed to be read "P generally (or usually or ordinarily) obtains provided that Q," or "P obtains, other things being equal, When Q does," or "Q constitutes prima facie evidence for P." It is not a matter of probabilities, but a matter of how things go normally or as a rule. This eerily recollects all of the later readings of default rules in AI's effort to understand non-demonstrative reasons.

One might wonder: with all these properties the slash doesn't have, what properties does it have? The fact that no substantive answer is given indicates at least that the slash is something of a stranger on the logical scene. Perhaps that is why Rescher "for simplicity" supposes that in a formal disputation, moves of the form  $P/Q$  are always "correct," i.e., beyond dispute or challenge. This supposition is bound to be unsettling, not least to logicians. [p. 102]

Hegselmann has similar frustrations,

Rescher has executed an abandonment of logical laws in dialectics that needs clarification. So it should be, for example, that ex-falso-quodlibet must be abandoned, since it might be reasonable to accept  $P/Q$  and  $Q$  as well as  $\neg P/(Q \wedge R)$  and  $Q \wedge R$ , without thus accepting anything at all. What follows from the abandonment of ex-falso-quodlibet, tertium-non-datur, the law of double-negation, and the law of non-contradiction (one should be obliged to abandon the last three laws as well) cannot be judged in detail, because Rescher did not characterize the resulting "dialectical logic" beyond the stated non-validity of the aforementioned theorems.

The author's translation is rough: Klärungsbedürftig sind Reschers Ausführungen ueber in der Dialektik aufzugebende logische Gesetze. So soll z.B. das ex-falso-quodlibet aufgegeben werden muessen, weil es durchaus moeglich und vernuenftig sein koenne, zugleich  $P/Q$  und  $Q$  und  $\neg P/(Q \wedge R)$  und  $Q \wedge R$  zu akzeptieren, ohne deshalb beliebig zu akzeptieren. Welche Folgen sich aus der Aufgabe des ex-falso-quodlibet, des tertium-non-datur, des Gesetzes der doppelten Negation und des Nicht-Widerspruchs (auch diese letzten drei Gesetze sollen angegeben werden muessen) insgesamt ergeben, laesst sich nicht im einzelnen beurteilen, weil Rescher die sich dann ergebende "dialectical logic" ueber die behauptete Nicht-gueltigkeit der genannten Theoreme hinaus nicht charakterisiert. [pp.42-43]

But a non-demonstrative reason whose behavior is characterized by the role it plays in a system of disputation is perfectly familiar to AI. Rescher himself gives "birds fly" as an example of "linkage that gives rise to provisoed assertion." Rescher says these assertions are ceteris paribus rules, "reasonably safe presumption rather than an airtight guarantee." [p.7] "It is sensible to suppose that P once Q is given." Rescher even notes the ampliative character of resource-bounded non-demonstrative reasoning, compared to non-ampliative expansions of shorthands (see Loui [91]):

If the only evidential move available were logical entailment, rather than this weaker, essentially ampliative stroke-relationship, then the very reason for being of disputation would be undermined. For in a strictly deductive argument, the conclusion can-

not be epistemically weaker than its weakest premiss. This would preclude any prospect of building up a case for an epistemically frail conclusion from relatively firm premisses (just as in inductive reasoning), and exactly this is one of the key aims of disputation. [p. 7]

Rescher was aware of the emerging popularity of defeasible reasons in epistemology, citing Chisholm (and apparently knowing of Pollock; p. 34 and p. 93 of Rescher), and could have appealed to their work to justify provisoed assertion. However, he takes the existence of non-demonstrative reasons to be unproblematic. He appears not to have been aware of the emerging popularity of non-demonstrative reasons for representing knowledge in AI rule-based and inheritance systems (Doyle [80], Reiter [80]).

Technical aspects of Rescher's system are interesting: non-monotonicity, search, termination, and even defeat by specificity all appear prominently in the text without special concern for their novelty in formal reasoning systems. This was one year before AI began mulling over non-monotonicity. It took a decade for that community to recognize specificity, and perhaps here is a first call for non-monotonicity's researchers to study search and termination.

Rescher's system has technical limitations from AI's current stance on resource-bounded protocols for non-demonstrative reasoning. Mainly, Rescher's protocol allows only a single reason to be given at each stage, together with a commitment to the reason's antecedent. Thus, arguments at any point are not grounded in evidence, and in fact there is no shared basis of evidence. Furthermore, specificity compares only the logical strength of reasons' antecedents (as in Nute's LDR [88]), rather than comparing whole lines of argument. So something like a directness defeater, a shortcutting of a chain of defeasible rules, which is present in most argument systems evolving from inheritance or non-monotonic reasoning, is unavailable to Rescher. Rescher, finally, has little to say about termination, which he regards as externally mediated by measures of plausibility on claims; and Rescher has no formal role for an adjudicator who intervenes or contributes to the dialogue.

Herbert Simon forebode the concepts presented in this paper. Simon's articulation of the distinction between substantive and procedural rationality was effective and celebrated; his analogy to chess-playing heuristics was apt. However, Simon's examples are limited to economic rationality (decision-making) despite his broad intentions; the present remarks aim squarely at logic. Also, Simon insists on conceiving of an optimum against which satisficing is defined; he insists on an objective world interpreted by a mathematical model that is an unfortunate but computationally necessary approximation. He resists claiming strongly that some things are defined solely by process, with no reference to an ideal. Says Simon (all quotations from [82]):

Behavior is procedurally rational when it is the outcome of appropriate deliberation. Its procedural rationality depends on the process that generated it. [p. 426]

The shift from theories of substantive rationality to theories of procedural rationality requires a basic shift in scientific style, from an emphasis on deductive reasoning within a tight system of axioms to an emphasis on detailed empirical exploration of complex algorithms of thought. [p. 442]

... Economics has largely been preoccupied with the *results* of rational choice rather than the *process* of choice. Yet as economic analysis acquires a broader concern with the dynamics of choice under uncertainty, it will become more and more essential to consider choice processes. In the past twenty years, there have been important advances in our understanding of procedural rationality, particularly as a result of research in artificial intelligence ... [p. 446]

Procedural rationality takes on importance ... in those situations where the “real world” out there cannot be equated with the world as perceived and calculated by the economic agent. Procedural rationality is the rationality of a person for whom computation is the scarce resource ... [p. 470]

I believe that formalism for non-demonstrative argument, especially non-demonstrative argument about decision-making under risk, is the right development for those who take seriously Simon’s ambitions.

Pollock’s defeasible reasoning is essentially disputational because reasons can conflict. His impressive technical developments have not attended much to processes; in part, this must be because there was plenty to say about how reasons rebut and undercut each other. Pollock’s framework and non-technical comments make room for process: defeasible reasons make more sense as inputs to a constructive process than as stylistic representation of non-constructive commitment.

Pollock [92] considers arbitrary bounds on some of his specifications, particularly, the determination of ultimate warrant, as might arise from limited computation. But Pollock, too, does not pursue ramifications of bounds: particularly, the effects of protocol on the construction of belief. Pollock recently writes about recursively enumerable warrant, but does not appear to make the same claims made here. Specifically, he is interested in correctness at the limit, and in probabilistic guarantees that “interrupted” computations will give the same answer as the limit computation. In fact, Pollock does not necessarily view defeasible reasons as policies for constructing belief under resource bounds. He appears unwilling to advocate a particular position on non-monotonicity arising from computation (what I have called internal temporal non-monotonicity), or on non-determinism.

In AI, Doyle has recently made about representation and belief many of the same claims made here [89]. Among them, Doyle agrees that belief is constructed from “manifest representation,” and that sometimes choice is involved.

Doyle also sees belief construction as “an activity, one step in the more general process of reasoning in time.” However, he sees the choice as a matter of preference, to be attacked as a problem in decision theory. If this choice is prescribed by the meanings of epistemic utilities (*a la* Levi [80]) or preferences (*a la* Wellman-Doyle [91]), then Doyle and I disagree over the ampliative nature of belief construction. This is because a represented preference would force a particular belief to be constructed solely in virtue of the meanings of the preferences applied to the meanings of preference sentences. There would not necessarily be non-determinism nor necessarily non-monotonicity in computation. However, Doyle is attuned to the tradition of constructivism in decision theory (esp. the writings of March [86] and Simon [82]), so the disagreement

may be merely a matter of emphasis. In fact, there is an established tradition of dialectic in decision theory (Mitroff and Mason [82], Churchman [71], Sabre [91]).

More interesting are passages of Doyle's doctoral dissertation [80], which Doyle himself later describes as being about "controlling action through dialectical deliberation." [91]

Doyle even has a section with the explicit title: [Section]  
3.11 Defeasible Reasons and Dialectical Argumentation  
[p. 112]

This dissertation was not well-received, as evidenced by its lack of citation. My conclusion is that Doyle's TMS/RMS was misunderstood for a decade. Legitimate developments of the TMS as a system of dialectic went unappreciated. The Fregean grip on the non-monotonic reasoning community is to be blamed.

Says Doyle:

... Deliberations make use of policies. Policies are used in reasoned deliberations to indicate new options to consider and to give reasons for or against the options. [p. 22]

Rational thought is the process of constructing reasons for attitudes. To say that some attitude is rational is to say that there is some acceptable reason for holding that attitude. Rational thought is a process of finding such acceptable reasons. (Doyle's footnote: this thesis allows as rational thought inferences involving random choices. For example, we might count as an acceptable reason "I couldn't think of anything else to do, so I flipped a coin.") [p. 90]

The non-monotonic logic developed by McDermott and myself also appears to suffer from this confusion: ... writing inference rules as implications; but I was never happy with this. [p. 92]

The policy's putative effect may fail to be realized because the policy's application ... may be defeated ... The deliberation procedure reflects on each new reason to find further policies relevant to the new reason. ... Of course, these defeating reasons can in turn be defeated. Finally, the deliberation procedure reflects on the entire set of reasons to decide whether to make a decision ... [p. 164]

Explicit, non-monotonic reasons form the basis of ... representation. These are used in defeasible reasons in ... deliberation, ... dialectical argumentation ... [p. 215]

The remaining work in AI exhibits less conviction than Doyle's.

The step-logic of Elgot-Drapkin, Miller, and Perlis [91] is vaguely related to the present paper because it views reasoning as a process through time that is not deterministic and not monotonic. But in their systems, when reasons fail to demonstrate, the non-demonstrative behavior is epiphenomenal, not fundamental.

Loui [87] considers justification with respect to a restricted set of arguments, but does not consider dialectic explicitly. Baker and Ginsberg's report on computing prioritized circumscription [89] is essentially an example of the kind of disputational process discussed here. The authors,

however, did not agree on the import of the dialectical process, venturing only to suggest that it was a useful strategy of implementation. These authors subordinate the dialectical aspects of their investigation to the affinities it has to circumscriptive and default approaches to non-monotonic reasoning. Similar implementational studies that employ dialectical protocols in automated argument systems can be found in Simari-Loui [92], and in Loui-Norman-Merrill-Olson-Costello-Stiefvater [92]. Both begin with no dialectical aims and are forced to define dialectical processes.

There is a newly proposed system of Vreeswijk [91], which begins by adopting exactly the position taken here. Unfortunately, Vreeswijk takes as desideratum an equivalence theorem which equates declarative and procedural approaches. This guarantees the work's reception in the existing non-monotonic reasoning community, but evades the important philosophical point.

As mentioned earlier, there is work in AI on non-demonstrative argument and process in the communities of natural language, education, and legal reasoning that is less formal. The recent symposium on argument and belief (see Alvarado [91]) abstracts much of the relevant work. Alvarado [90] and Cavalli-Sforza and Moore [92] are excellent examples of the semi-formal analysis of argument that can be found in these AI communities.

## 6. Non-Monotonic Reasoning and Other Disputations.

### *Elements.*

Disputations can have multiple *parties*. Mostly they have two parties, *pro* and *con*. Sometimes there is a non-trivial *adjudicator* role. When a distinction is desired, some parties are *advocates*, and some are not. Advocates incline to certain claims that are in dispute. There must be at least two advocates for there to be a dispute. Parties that have strategic roles, roles that include choice, are *players*.

We are mostly interested in two-advocate disputations with and without adjudicators as parties or players. What distinguishes adjudicators as parties versus adjudication that is a syntactic feature of the system is that an adjudicator, as a party, can engage in dialogue. What distinguishes adjudicator players from parties that are non-players is that their dialogue moves involve choice.

Parties can cooperate about various things. Mostly, they are cooperative about reaching a deeply reasoned opinion (which precludes filibustering), but non-cooperative (adversarial) about what that opinion should be. The interaction of these attitudes complicates strategy.

As Doyle notes, a society of minds or social choice problem is a multi-party dispute, and adopting a position on a dispute is like exhibiting a preference [89]. But we are interested in disputes where social choice is based primarily on the reasons given for preferences rather than the relative social standing of the parties. Though social properties such as authority and privilege do figure in the conduct and result of disputation, the emphasis here on the production of reasons separates this investigation from work on social choice.

Sometimes a single agent acts so as to provide both parties to a dispute. Rescher calls these “unilateral dialectics,” which is surely a non-sequitur. What is more interesting is the degree of *shared basis* and *shared resource* between the parties. An executive committee can have many individuals, but be so regulated that resources are shared among advocates and agreement is substantial over what claims are foundational and what reasons are recognized. Meanwhile, a single person, even with no mental disorder, might be capable of providing each party to the dispute and be able to partition resources. A single person also might have sufficient skepticism that all claims and reasons can be challenged and subject to dispute.

In non-monotonic reasoning, or argument systems that have been proposed for non-monotonic reasoning, the two parties share both foundational evidence and reasons as a basis. In Toulmin, reasons can be disputed as well as evidential claims. In Rescher, all evidential claims can be disputed, but reasons cannot. Some claims must be shared, however, or else a result of disputation could never be reached, according to Rescher. Another way of saying this is that rules for termination often depend on reaching claims that are shared. Note that sharing in this sense has to do with the right or willingness to dispute, not with the freedom of one advocate to use what may be used by another. Considerations of fairness are what allow equal access to a single corpus of undisputed evidential claims and reasons.

In non-monotonic reasoning, the critical resource is time for search and the right to select or schedule targets for goal-directed search. In legal and policy forums, the critical resource is time

for presentation of arguments. Communication of arguments in non-monotonic automated reasoners is trivial, once arguments have been constructed: argument presentation may require no more than pushing a pointer on a stack. Consequently, the protocols and rules for adjudicating automated reasoning disputes may differ markedly from existing social rules for adjudicating policy disputes.

Most significant is the availability of lookahead for arguments and its effect on player strategy. When players can perform private search (where other players are not automatically informed of results of search), there can be lookahead. Particular moves have greater significance to termination and adjudication when there is lookahead than when strategy is uninformed. Players can be penalized for poor moves and rewarded for good moves. Players can be exhorted to advance their best arguments when time is limited. This is impossible if arguments are advanced immediately when discovered. When the first admissible argument that is found by search is automatically adopted as the player's move, as would naturally be the case in an immediate-response implementation of non-monotonic reasoning, arguments are not necessarily presented best-first: the strength of the first argument cannot be taken to indicate the strength of the ensemble of arguments for the claim.

#### *Protocol.*

In non-monotonic reasoning, protocol is not formally studied because resource-bounds are not taken seriously. Search strategy matters only when resources do not suffice to exhaust the arguments. AI has considered search strategy to be a detail of implementation.

Arguments, however, are likely to be infinite. Arguments proliferate because of syntactic variation; but even if there is a canonical form, which bars multiplicity due to syntactic variation, arguments can be inexhaustible. Rule-bases upon which arguments can be constructed only by chaining are the exception, not the norm. Classes of decision-theoretic argument (where trees can be extended indefinitely) and best-explanation argument (where more candidate explanations or theories always can be put forth) are especially likely to be infinite. Analogical argument and statistical argument can usually be produced in combinatorial numbers. All argument can potentially ascend to unbounded meta-levels, although the requisite meta-knowledge and meta-language is likely to be sparse.

More importantly, even a finite set of arguments may not be exhausted by search. This is of special concern when connections between arguments can involve non-recursive or just plain impracticable computations, such as determining logical entailment or consistency.

Reorientation in the knowledge representation community is needed to include the formal study of protocol as a part of the study of non-monotonic reasoning. Policies produce warrant only through a process of disputation. It is inappropriate to study the forms of policy assertion without also studying the processes in which they will be involved. Policies without processes have no meaning. It does not help to study their transformational behavior or the mathematical structure of their presumably compositional form. For policies, protocol is semantics.

Protocol has two parts: *locution* and *control of locution*. *Termination*, an aspect of controlling locution, is treated separately due to its importance.

In non-monotonic reasoning, the appropriate locutions are declarative. Logics of dialogue use a broader class of speech acts, and there may well be good use for simple requests and queries; these could for instance support strategic aims regarding adjudication, control, and termination, such as calling bluffs and focusing dispute. However, the full repertoire of rhetorical maneuvers used for persuasion is beyond the present scope; much of what has been studied as persuasive dialogue detracts from the idea that dispute is a model of rational inquiry.

Locutions pertain not only to the substance of the dispute, but also to the protocol. Arguments can advance claims about how to define or alter the protocol. Arguments can seek to establish conditions relevant to protocol, such as defeat relations among arguments, adequacy of responses, and termination. These are rightly considered meta-arguments. The more effective the definitions of aspects of protocol, such as when an argument defeats another, when response is adequate, and when termination is appropriate, the less the need to ascend to meta-argument. Frequently, agreement over protocol is a part of shared basis.

An important distinction between Rescher's protocol and search protocols that implement argument systems in AI is that locutions in Rescher are one-step arguments. Rescher presumes no evidential basis in which to ground arguments, so argument chains could extend arbitrarily. Rescher preempts arbitrary-length locutions by simply restricting locutions to single claims and (possibly) a single reason.

Refer to restrictions on locutions that are legitimate moves as *locution-restrictions*, and refer to minimal requirements as *locution-obligations*. Locution-restrictions help define control of the process, which usually means control of resources. Restrictions prevent unbridled consumption of resource. Locution-obligations are helpful in defining termination; unmet obligations cause a dispute to be lost. Note that there may be no obligation at some points in a disputation, but for the process to be dialectical, there must at some point be an obligation.

Protocol assigns *burdens* to parties. Rescher identifies burden that arises from "presumption," an inclination of opinion toward one advocate or another. If the presumption is pro, then the burden-to-advance-argument is on con. Presumption occurs throughout non-monotonic reasoning when failure is taken to be negation.

Additional burdens can be assigned.

There can be a burden-to-respond-with-defeat. A lesser burden is the burden-to-respond-with-interference. Each defines a locution-obligation. Lack of either burden allows subsequent locution that does not directly respond to the other advocate's argument. Reinstatement potentially occurs when the defeater of an argument is in turn attacked. The burden-to-reinstate-with-defeat, as opposed to reinstating simply with new interference, gives rise to an ambiguity-propagating system.

This distinction arises in the "clash of intuitions" over inheritance systems, Touretzky et al. [87].

Without the burden, the system is ambiguity-blocking. Unlike in the inheritance systems, this burden can be assigned to one side and not the other. In general, assignment of burdens may be not symmetric.



Burden can be assigned regarding most aspects of protocol. There could be a burden-to-argue-for-termination, a burden-to-argue-defeat, or more generally, a burden-to-argue-sufficiency-of-response. These burdens make sense when their conditions are not syntactically determined trivially; it makes no sense to assign a burden to argue for termination if determining termination conditions is syntactic and immediate.

Pragmatics defines which burdens are assigned. A dispute in which one side carries heavy burdens compared to the other may be desirable from a strategic point of view. It might be contracted in advance that if the unburdened advocate cannot win such a dispute, then the adjudicator should find in the opponent's favor.

### *Termination.*

AI is familiar with issues of termination because similar issues arise in heuristic search. However, unlike heuristics, critical appraisal of termination rules for disputes sometimes cannot refer to a correct answer. The process itself determines what is the "correct" application of policies and "correct" adjudication of conflicts among policies. Some protocols for processes are acceptable and others are not, but not directly because of the prospect or probability of arriving at some right outcome. Protocols may satisfy desiderata that refer to outcome: if there is at least one final state in which all arguments have been advanced, and in all such final states, a certain opinion is mandated, then we may aspire to find a protocol that gives game-theoretic or weak probabilistic assurances of reaching that opinion with earlier termination. We can do no better in appraising protocols in general.

Sometimes disputation just ends; termination befalls. Sometimes the allocation of time is known in advance; time's expiration is scheduled; time for dispute expires. Sometimes disputation is terminated because a condition arises, other than expenditure of resource, such as a winning position. Immediate-response dialectic is incompatible with scheduled time; the latter suggests *stages* of a disputation. Protocols for actual debates even restrict the kinds of arguments that can be advanced in each stage. All three kinds of termination (*befalling*, *expiry*, and *satisfaction* of non-resource conditions) are relevant to non-monotonic reasoning.

A state of the disputation yields descriptions at various levels. The *current opinion*, favoring an advocate or favoring none, is the most succinct description. A more detailed description is formed by the list of undefeated arguments, together with the kinds of responses to each that would suffice to change opinion. A record of stakes that have been won and lost during previous moves, a *score*, can be kept. Finally, there are more detailed levels of description that refer to the actual sequence and content of argumentation. Fixing exact language for describing state is crucial for disputations that permit negotiation on protocol, or meta-argument about adjudication and termination.

Current opinion may be visible to players or hidden during disputation. When dialectic is immediate-response, opinion must be visible; what makes a response adequate and forces a change in control, after all, is an argument's changing current opinion. In non-monotonic reasoning, opinion should be visible, except when computation of opinion is expensive.

Adjudicators can be aggressive. They can intervene to set conditions for termination or focus a dispute. One way to intervene is to declare deadlock on a hopelessly contentious issue and re-

quire that advocates substantiate their claims without reliance on the issue. Another intervention is to declare that termination is imminent and to create burdens for one player or for both players which must be satisfied in order to prevent termination or to change opinion. Adjudicator intervention is most useful when dialectic seems not to be moving forward, or when players' interests as advocates of an opinion unduly outweigh the common interest in deep and effective reasoning.

Part of the shared basis includes agreement over what is a winning position. In Rescher, winning positions are defined by plausibility of reasons, and plausibility is a metric supplied independently. In non-monotonic reasoning, a winning position is implicit in the rules for defeat among arguments. Whether a position is winning for an advocate can be argued, especially by analogy to canonical disputations.

### *Strategy.*

In current approaches to non-monotonic reasoning, there is no strategy because without independent means for evaluating contentiousness of claims, with no lookahead and no rating the prospect of goal-directed search's success, choices are blind.

There is a place for strategy. When choices are informed, the play of players can be improved by strategy. Improvement of choice is not easily defined, but information about choice carries game-theoretic implications. For instance, if points of counterargument are ordered by prospect of successful counterargument, and resources are limited, the dominating strategy is to target the point of counterargument with greatest prospect of success, all other things being equal. When the play of each advocate can be bettered, it can result in a process, the results of which are more widely accepted. When asked to accept the result of a deliberative process, an agent is better justified in believing the result of a process in which moves were well-played, as long as no unfairness results from different quality of play among players. A manager forming belief based on disputes in committee, as well as an audience watching debate, is more justified in believing the pronouncement resulting from a well-played competition. If information can be brought to bear on choices, then strategy is a legitimate part of disputations that support rational inquiry.

Termination can be based on score-keeping, which counts the number of ineffective counterarguments, the number of unproductive searches, the number of challenges met, and so on. When there is no strategic play, termination based on score is essentially termination due to a resource bound. When there is strategic play, scores are attributable not only to the inherent strength of an advocate's position, but also to the quality of play.

Thus, policy-makers make policy with the understanding that their policies' application will be holistic, heuristic, non-deterministic, and sometimes even influenced by strategy and tactics.

## 7. Skeletal Model.

### Definitions.

A locution record is a finite sequence  $\Gamma = \langle l_1, \dots, l_k \rangle$  of locutions where each locution  $l_i$  is a 3-tuple  $\langle p_i, s_i, r_i \rangle$ ;

each  $p_i$  is a party, from a set  $P$  of parties (some of which are players);

each  $s_i$  is either a well-formed sentence in a language  $L$  (which may actually be better thought to be a set of languages, such as an object language and its metalanguage), or is a structure meeting some well-formedness conditions, such as being an argument for a sentence in  $L$ ;

and where  $r_i$  is a description of resources, such as a  $d$ -vector, a marker of resources consumed, of  $d$  different resource types.

This structure induces the functions  $party(i)(\Gamma)$ ,  $sentence(i)(\Gamma)$ ,  $resources(i)(\Gamma)$ , and  $sentences(i)(\Gamma)$ , which collects locutions up to and including  $sentence(i)(\Gamma)$ ,  $\{sentence(j)(\Gamma) : \text{for some } j \leq i\}$ . Without the index  $i$ ,  $sentences(\Gamma) = sentences(length(\Gamma))(\Gamma)$ . The overloading of function name is too convenient not to adopt, and possible to distinguish from form and context.

A locution record also induces a record of moves,  $\langle m_1, \dots, m_k \rangle = M$ , a sequence, each element of which is itself a sequence of consecutive locutions by a single player in  $\Gamma$ :

each  $m_i$  is  $\langle l_{1(i)}, \dots, l_{k(i)} \rangle$ , as above;

the functions  $party(j)(m)$ ,  $sentence(j)(m)$ , and  $resources(j)(m)$  are induced (deliberately overloading the symbols), as well as  $party(m)$ , and  $resources(m)$ , where the latter is  $resources$  applied to the last locution in  $m$ ; and the function  $sentences(p)(i)$ , which is the set of all sentences by  $p$  up to and including move  $i$ ,  $\{sentence(j)(m_h) : \text{for some } h \leq i, l_j \text{ occurs in } m_h\}$ ;

$\langle l_u, l_v \rangle$  are adjacent in any  $m_i$  iff  $party(l_u) = party(l_v)$ ; and  $\langle l_u, l_v \rangle$  are adjacent in  $\Gamma$  iff for some  $\langle m_s, m_t \rangle$  in  $M$ ,  $l_u$  is the last locution of  $m_s$  and  $l_v$  is the first locution of  $m_t$ .

A disputation record,  $D$ , is a structure from which a locution record and a record of moves can be derived,  $\Gamma(D)$  and  $M(D)$ .

### Example.

For a resource-bounded defeasible reasoner, such as the one described in Simari-Loui [92],

$P = \{\text{pro, con}\}$  and  $p_i$  switches every other locution: moves are 2-tuples.

A move  $m_i$  is always a wff of  $L$  paired with an argument from  $\Omega_{K,\Delta}$ , the set of arguments constructible from evidence  $K$  (a subset of  $L$ ) and defeasible reasons  $\Delta$  (a subset of  $L \times L$ ). The argument,  $l_{2(i)}$ , is an argument for the sentence,  $l_{1(i)}$ .

$\Omega_{K,\Delta}$  is a part of the shared basis; they do not argue over whether sentences or structures are well-formed.

In the implementation described in Loui-Norman-Merrill-Stiefvater-Costello-Olson [92],  $d=2$ , so  $\text{codomain}(\text{resources}) = \mathcal{R}^2$ , because only two resources affect protocol: number of resolvents and number of attempted resolution steps.

Since sentences are complex (an argument is itself a 2-tuple, whose first element is the main claim of the argument and whose second element is the set of reasons used to derive the claim from  $K$ ), we prefer to depict locution records as numbered sequences. The locution record for a familiar dispute is cogently rendered:

1. pro	files(opus)	<0, 0>
2. pro	<flies(opus), {<bird(opus), flies(opus)>}>	<2, 30>
3. con	¬flies(opus)	<2, 30>
4. con	<¬flies(opus), {<penguin(opus), ¬flies(opus)>}>	<6, 50>

and might be the two-move result of disputing *flies(opus)* in which 6 resolvent clauses were created in the search for arguments that attempted 50 resolutions.

#### Example.

In a Rescher dispute, again

$P = \{\text{pro}, \text{con}\}$ , where moves are either 1-tuples, claims, or 2-tuples, a claim in  $L$  and a reason for the claim, a member of  $L^2$ .

Rescher distinguishes between sentences pro and con:  $L$  is actually modal, with the two operators “!” and “†”, for categorical and cautious assertion, respectively. However, since all and only pro’s assertions are categorical, and all and only con’s assertions are cautious, the distinction is superfluous. Adjudication may be sensitive to the modality of the assertion, but this need not be represented linguistically since it is trivially reconstructed.

Resources are formally irrelevant to Rescher, so  $d = 0$  and  $\text{codomain}(\text{resources}) = \{\langle \rangle\}$ .

1. pro	P	<>
2. con	¬P/Q	<>
3. con	Q	<>
4. pro	P/(Q ∧ R)	<>
5. pro	Q ∧ R	<>

is the locution record of a basic Rescher dispute.

#### Definitions.

A disputation protocol consists of the following functions, all of which may be partial:

because the functions might be defined inductively only on disputation records that the protocol itself could generate; since they are partial, it is easier to give their domains by example than to name the full sets, for example, of all disputation records.

Where  $D$  is a disputation record and  $r$  is a description of resources, let a disputation stage be the pair  $s = \langle D, r \rangle$ :

1.  $current\_opinion(s) \in P \cup \{none\}$ ; opinion currently favors a party, or favors none;
2.  $party\_to\_move(s) \in P \cup \{none\}$ ; it is always some party's turn to move unless there is natural termination;
3.  $move\_options(s)$  is a set of moves, i.e., a set of sequences of locutions;
4.  $player\_information(s)$  is a structure that may include a restriction of  $D$  that permits subsequences of  $\Gamma(D)$  and  $M(D)$  to be derived; it may also include information about the plausibility of members of  $L$ , and the results of search.

These functions do not appear to restrict protocols much, but note that already a choice has been made regarding non-determinism. In multi-party (more than two-party) protocols, the order of participation is pre-determined since  $party\_to\_move$  is a function. Non-determinism can be simulated, though, by parties choosing to make empty locutions non-deterministically.

When there is a resource bound but these protocol functions are insensitive to that bound, the bound is unannounced (or hidden) and “time expires” without warning.

These functions serve essentially the same roles as “rules for dialogue,” “rules for commitment,” and “rules for termination” in logics of dialogue. Mackenzie's dialogues, for example, are formalized by ignoring resources, taking disputation records to be locution records, and writing constraints on locution records that define the disputation protocol. We must complicate the model, since our interest in dialogue arises from considerations of search. So we require that protocol can be sensitive to resources consumed within a player's turn, and that players can have information about search for argument, not just production of argument.

The function,  $move\_options$ , represents both the locution-obligations and the locution-restrictions. For any member of  $move\_options$ , a sequence of locutions, obligations quantify existentially over the sequence, while restrictions quantify universally, in their normal forms. For example, a protocol has timed moves just in case, as a restriction, for all  $m_1, m_2$ , adjacent in a record of moves,  $|resources(m_2) - resources(m_1)| < k$ . A protocol restricts moves to a single claim and its argument just in case  $party(i)(\Gamma) \neq party(i+2)(\Gamma)$ , and if  $sentence(i) \in L$ , then  $sentence(i+1)$  is an argument for  $sentence(i)$ . A protocol prohibits repeated arguments just in case  $\langle p, s, r_1 \rangle$  and  $\langle p, s, r_2 \rangle$  do not both appear in a locution record (this restriction makes sense when paired with the restriction that arguments cannot be introduced in locutions if the record already contains an effective rebuttal). A protocol obligates novel argument instead of rebuttal just in case, for all disputation stages  $s$ , if  $party\_to\_move(s) = pro$ , and  $current\_opinion(s) \neq pro$ , then every member of  $move\_option(s)$  contains as some locution the (re-)assertion of  $sentence(0)$ , the original claim under dispute.

Most locution-obligations and restrictions are complicated and refer to both the theory of argumentation on which the disputation is based, and the criteria for adjudicatory preferences, as captured in  $current\_opinion$ .

#### Definition.

At a stage,  $s = \langle D, r \rangle$ , the  $party\_to\_move$  is a player if  $move\_options$  contains more than one element.

At a stage where *move\_options(s)* is empty, the disputation is at natural termination; if *party\_to\_move = p*, then *p* is a loser. Unlike logics of dialogue which define winning and losing in terms of *party\_to\_move* at termination, we use the function *current\_opinion* to track whether *sentence(0)* has been established (or its negation, or neither). This recognizes that many disputations may end in draws, in which neither party was able to establish the point of contention. Some protocols shift burden in such a way that there is always a winner; in such protocols, being the *party\_to\_move* can indicate loss.

Note that we are beginning to drop a function's arguments  
if they are obvious from context.

A protocol is dialectical if at all stages, *player\_information* is the entire disputation record. A protocol is an immediate-response dialectic if it is dialectical and at all stages, *party\_to\_move* is not equal to *current\_opinion*.

A strongly alternating two-party immediate response dialectic is one in which, as a restriction, all moves are pairs of locutions: a claim in *L* and an argument for that claim; and as an obligation, for any  $s = \langle d, r \rangle$ ,  $current\_opinion(s) \neq current\_opinion(s')$ , where  $s'$  is any stage formed by concatenating a member,  $m \in move\_options(s)$  onto the record of moves  $M(d)$ , forming a new disputation record, and pairing it with *resources(m)*.

Current opinion is hidden if at some stage, *player\_information* is not sufficient to calculate *current\_opinion*; otherwise, current opinion is visible (either because the protocol is dialectical, or because *current\_opinion* can be constructed from a function on a smaller domain).

When search is goal-directed, a player can set a search target. Descriptions of states of search for this goal are not a part of the disputation stage, but may be part of computations that players perform in choosing their moves, which are part of the stage. This is a deliberate choice, with the consequence that change of control can be forced only on the basis of consumed resource, not on how poorly a search is going. If a search space is exhausted, the protocol as defined here cannot force change of control; however, a cooperative player will make the appropriate locution upon exhausting the space, or set another search target, as appropriate. Since descriptions of intermediate states of players' computations of actions are likely to be as varied as their strategies, the alternative design choice is formally unattractive (as we stand, at least we can restrict the domains if not the ranges of the protocol functions). Moreover, a protocol influenced by the particular way a player chooses a move seems improper, even if that choice makes use of a shared data structure, such as a theorem-prover's clause stack.

#### Example One: AND-OR searching for arguments.

The protocol in Simari-Loui [92] and in Loui-Norman-Merrill-Stiefvater-Costello-Olson [92] is a fairly standard approach, where arguments that ground in the (shared) evidence and that change current opinion are required in each move. *pro* and *con* take turns introducing complete arguments until time expires or no response can be found. Opinion must switch after each move, and may take on any of three values: *pro*, *con*, and *none*, according to existing rules of support with respect to a possibly incomplete set of arguments. The main choices for players are what should be the targets of search among the various points of counterargument. This protocol generates arguments in much the same way that an AND-OR search strategy would generate arguments.

*current\_opinion* for the null disputation record is *none*. Thereafter, *current\_opinion* is *pro* iff  $s_0$ , the main point under dispute, is justified among all of the arguments presented,  $arguments(D) = sentences(\Gamma(D)) \cap \Omega_{K\Delta}$ . A sentence is justified just in case it has an ultimately supporting argument (the idea comes from Pollock [87] was altered in Simari-Loui [92], Prakken [92], and Loui et al. [92]); roughly:

All arguments can *interfere* and *support* at *level*(0). At *level*( $n+1$ ), an argument can support if no argument that can interfere at *level*( $n$ ) interferes with it. At *level*( $n+1$ ), an argument can interfere if no argument that can interfere at *level*( $n$ ) defeats it. An argument is *ultimately supporting* if for some  $k$ , it can support at all levels greater than  $k$ ; likewise, for ultimately interfering. An argument *interferes* with another if the claim of the first is contrary to what can be derived with the second; a special case of interference is *disagreement*, when the main claims of arguments are contraries. An argument *defeats* another if it interferes and is more specific than the subargument with which it disagrees. Rules for *specificity* are based on the existence of a non-trivial asymmetric activator, along the lines of Poole [85]. An *asymmetric activator* is a contingent sentence that, together with necessary evidence and the combined rules of the theories being compared, allows a consistent defeasible derivation of the conclusion of one argument, but does not do the same for the other argument. *Non-triviality* requires that all top rules of an argument be used by an asymmetric activator. A rule is *top* if the antecedents of all other rules in the argument can be derived without using the alleged top rule. Separation of *contingent* from *necessary* sentences is arbitrary and defines the background against which defeasible rules are adopted. *Consistent defeasible derivation* is entailment in the underlying logic augmented by (modus ponens for defeasible rules) detachment of defeasible rule consequents when their respective antecedents can be derived. See Loui et al. [92] for more detail and a pruning lemma that reduces the search (from infinite to two) for asymmetric activators for any pairwise specificity comparison.

For example, letting “ $\succ$ ” be an infix symbol for the defeasible reason relation, and letting all and only  $e_i$  be evidence, consider the arguments:

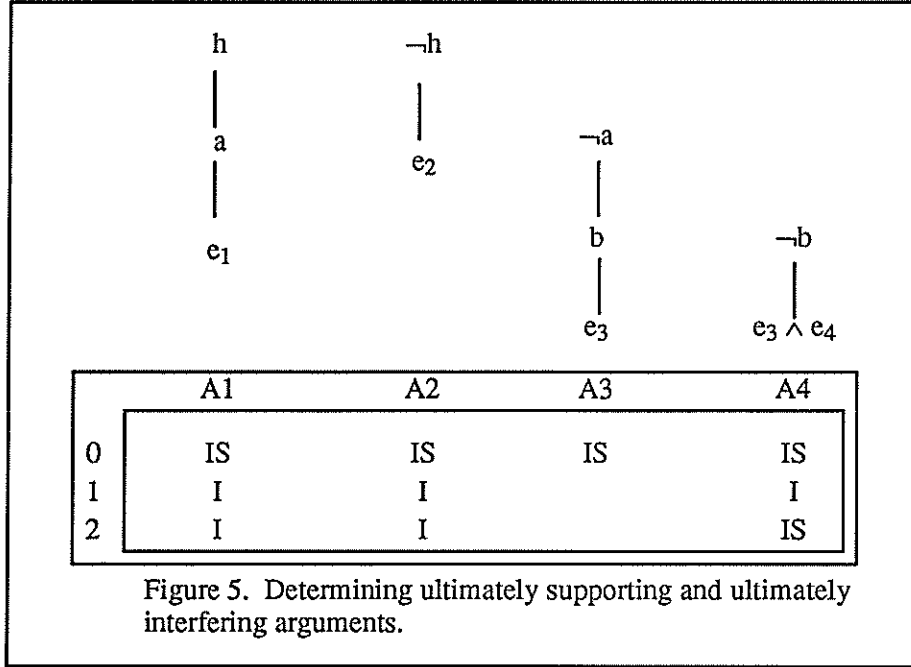
A1:  $e_1 \succ a; a \succ h;$

A2:  $e_2 \succ \neg h;$

A3:  $e_3 \succ b; b \succ \neg a;$

A4:  $e_3 \wedge e_4 \succ \neg b;$

Then A1 and A2 disagree; A3 counterargues A1 at  $a$ ; A4 counterargues A3 at  $b$ ; and A4 defeats A3, thus reinstating A1 against A3; note that A2 is undefeated, so it will remain ultimately interfering with A1. At *level*(0), all arguments are interfering and supporting. At *level*(1), no arguments are supporting, A1, A2, and A4 are interfering; at *level*(2), A4 is supporting, and A1, A2, and A4 are interfering. All subsequent levels are the same. So A4 and only A4 is ultimately supporting.



Also, *current\_opinion* is *con* iff the negation of  $s_0$  is justified among all of the arguments presented. Since moves are always two locutions (claim and argument for claim), *party\_to\_move* can be determined by the length of  $\Gamma(D)$ . However, play is resource-bounded, so *party\_to\_move* = *none* if  $r - \text{resources}(m)$  exceeds  $\langle \text{max}_1, \text{max}_2 \rangle$  in either dimension, where  $m$  is the last move in  $M(D)$  by the player who is not *party\_to\_move*. In a move, all locutions report the same resource: the cumulative search time required to identify an argument that is a sufficient response.

At any time, *player\_information* consists of: the entirety of  $D$ ,  $r$ , and a list of targets for which search was unsuccessful in this turn and in past turns.

But since success is defined by finding an argument that can affect *current\_opinion*, and since the last move by the opposing player affects this, it may not be necessary to know all previous failures.

This protocol is a strongly alternating two-party immediate response dialectic. Hence, *move\_options*( $D, r$ ) is at any time the set of arguments such that concatenating any one of them onto  $M(D)$  causes *current\_opinion* to change. Note that a non-move by the player is to make no locution while resources are exhausted beyond bound. This causes a change in the state of disputation (disputation terminates with *current\_opinion* unchanged), but does not alter the disputation record.

From our knowledge of defeat and justification, we can state *move\_options*( $D$ ) more precisely.

Case 1. *current\_opinion* = *none*. *party\_to\_move* can introduce a new argument for which there is no interference, for some important claim, or can rebut (by producing a defeating subargument) some argument that alone interferes with some important argument of its own.



Note that since moves contain a single argument, this protocol thus precludes a move by a party to rebut all arguments, if there are more than one, by producing multiple arguments. However, a single argument for the conjunction of several points, if it has the right structure, can have the same effect.

Let an *establishable point* be any sentence such that *current\_opinion* would change if *party\_to\_move* could make the sentence ultimately justified with respect to *arguments(D')*, where *D'* is the disputation record after the party's move.

$s_0$  is the basic *establishable point* for *pro*; its negation is the basic establishable point for *con*. Let  $s$  be any establishable point for *party\_to\_move*. Let  $a$  be any argument for  $s$  and let *live\_opposition(a)* be the subset of *arguments* that can ultimately interfere with  $a$ . Let *targets(a)* =  $\{t_i\}$  be any set of sentences such that for each argument  $o$ , a member of *live\_opposition(a)*, there is in *targets(a)* some  $t_i$  that is contrary to some point derivable from  $o$ : i.e., the weakest *targets(a)* collects for each argument in *live\_opposition(a)* a disjunction of negations of interesting subpoints: defeasible rule antecedents or consequents are interesting subpoints. Then the conjunction of members of *targets(a)* is also an establishable point for *party\_to\_move*.

Finally, consider any establishable point paired with an argument for the establishable point, where that argument is legitimate, i.e., a member of  $\Omega_{K\Delta}$ , and that it has no interferer in *arguments* that can ultimately interfere; this pair will be a member of *move\_options*.

Case 2. *current\_opinion* = *opposing\_party*; which is *pro* if *party\_to\_move* is *con*, and vice versa. *party\_to\_move* can make any of the moves available in Case 1, and can also: introduce a new argument for which there is no defeat, for some important claim, or can rebut (by producing an interfering subargument) some argument that alone ultimately justifies some important point of *opposing\_party*.

*sentence(0)* is the basic *sufficient point* for *pro*; its negation is the basic sufficient point for *con*. For  $s$  a sufficient point for *party\_to\_move* and  $a$  an argument for  $s$ , *established\_opposition(a)* is the subset of *arguments*, each of which can ultimately defeat  $a$ . *targets(a)* is as before, some set that includes at least one member contrary to some point derivable in each argument in *established\_opposition(a)*. Then the conjunction of members of such a set is also a sufficient point for *party\_to\_move*.

Consider any sufficient point paired with an argument for the point, where that argument is legitimate, i.e., a member of  $\Omega_{K\Delta}$ , and that it has no defeater in *arguments* that can ultimately defeat; this pair will be a member of *move\_options*, as will any pair constructed as above, with an establishable point.

For example, let  $e_i$  be (shared) evidence; the following depicts each player's two-locution move: claim, argument for claim, and a running total of resources consumed:

1,2. PRO: $h$	$e_1 \succ a$	80
	$a \succ b$	
	$b \wedge e_2 \succ h.$	

*current\_opinion* is *pro*;

*move\_options* are for *con*:

1. any argument for  $\neg h$  that uses something other than just  $e_1$ , or just  $e_2$ , or uses both through  $a$  and  $b$  but in a less direct way, e.g.,

$$\begin{aligned} e_2 &\succ c \\ c \wedge e_1 &\succ \neg h \end{aligned}$$

would be o.k., but

$$\begin{aligned} e_1 &\succ a \\ a &\succ b \\ b &\succ c \\ c \wedge e_2 &\succ \neg h; \end{aligned}$$

would not be;

2. any argument for some  $s$  that entails  $\neg a \vee \neg b \vee \neg h$ , e.g.

$$\begin{aligned} e_3 &\succ c \\ c &\succ \neg a. \end{aligned}$$

3,4. CON:  $\neg a$

$$\begin{aligned} e_3 &\succ c \\ c &\succ \neg a. \end{aligned}$$

60

*current\_opinion* is *none*;

*move\_options* are for *pro*:

1. any new argument for  $a$  that defeats *con*'s argument for  $\neg a$ ; e.g.

$$e_3 \succ a;$$

or

$$\begin{aligned} e_3 &\succ c \\ c \wedge e_4 &\succ h; \end{aligned}$$

2. any argument for  $\neg c$  that defeats *con*'s argument for  $c$ ; e.g.

$$e_3 \wedge e_4 \succ \neg c;$$

3. any new argument for  $h$  that avoids use of  $a$ ; e.g.

$$e_4 \succ h.$$

5,6. PRO:  $h$

$$e_4 \succ h.$$

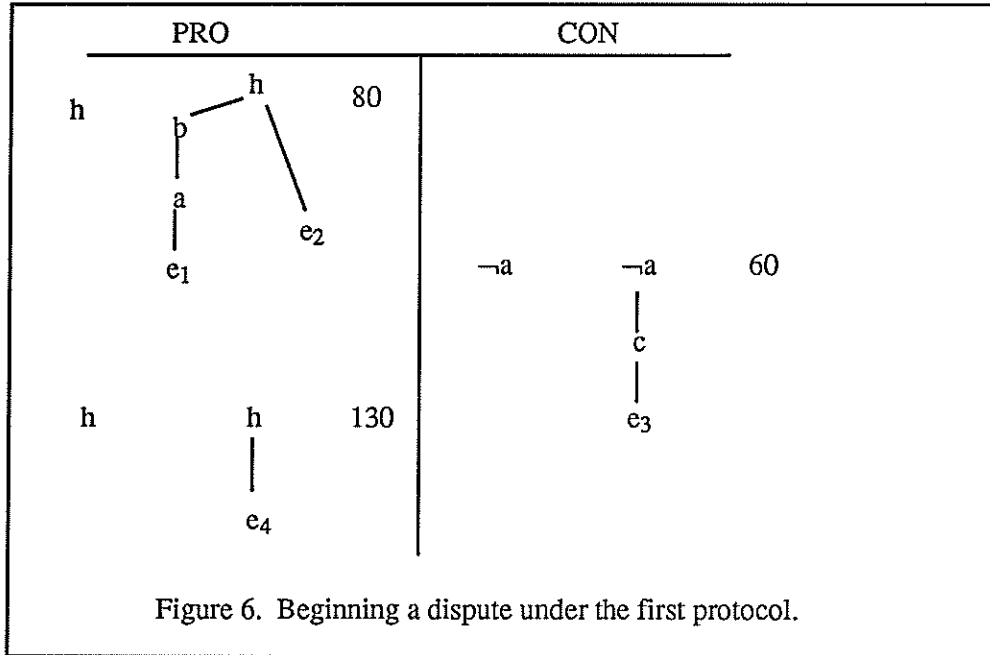
130

*current\_opinion* is *pro*.

... and so on until either resources are exhausted or arguments are exhausted.

### Example Two: Dialectic in Chaining.

A different protocol is given in Loui-Chen [92] for the same conception of argument, support, and defeat. In the last protocol, dialectical control is exerted every time an argument is introduced. Here, dialectic occurs every time a potentially adequate backward chaining step is taken while arguments are being constructed. This protocol is close to the computation currently performed in a quasi-legal reasoner that (among other things) peruses decisions in cases (precedents) to find relevant rules for argument. It is an interesting protocol because it allows dialectic to guide search. Instead of allowing backward chaining to finish, a weak step may be questioned immediately by beginning the search for counterargument. The choices for players not only include targets for search, but also whether to continue even after a promising reason



has been found (or else stop and allow control to pass to the opponent), and whether to allow the opponent to continue (or else to start counterarguing immediately).

There are strategies for assigning burdens, but we may assume that *pro* carries the burden-to-establish, the burden-to-open, and the burden-to-reinstate-with-defeat.

If resources are not exceeded, *current\_opinion* is always *pro* if *player\_to\_move* is *con*; *current\_opinion* is *none* if *player\_to\_move* is *pro*. *con*'s only objective is to thwart *pro* in this game. Moves are not limited to two locutions, as in the last protocol, so control is not rigidly structured. In fact, players can elect to make as many or as few locutions as they choose, subject to a locution-restriction regarding resources, and a locution-obligation that requires sufficiency of response. As before, the resource-field of a locution reports amount of computation spent, and is not elective.

The resource restriction is simple: at no time can *pro*'s consumption exceed *con*'s by more than some constant *k*. This is also a termination condition: when *player\_to\_move* is *pro* and any locution that *pro* could make would violate this resource restriction, the dispute ceases with *pro* failing to establish the main point, and *current\_opinion* switching to *none*. If *player\_to\_move* had been *con* and resources expired, *current\_opinion* would terminate *pro*.

Sufficient response is more interesting. Informally, instead of requiring whole arguments grounded in evidence to be produced in each move, only one reason need be given (although more reasons are permitted). This reason must eventually take part in an argument, grounded in evidence, that helps support the main point of contention. However, completion of the argument is not required in any turn; it suffices that this be merely possible, given what reasons have so far been introduced, and given knowledge about possible completions. If a player chooses not to

complete an argument, but merely to cite its possibility, then the opposing player may immediately investigate possible lines of counterargument, without even waiting for completion of the argument. The opponent could also challenge, asking that the partial argument be brought closer to completion (or challenge all the way to completion) before responding. However, there are implications regarding the accounting of resources when there is search in response to a challenge.

The shared basis includes evidence  $K$ , and reasons,  $\Delta$ , as before. The language  $L$  is a restricted propositional language with negation only for literals and with only the 2-connective for conjunction (no disjunction or conditional). The main point of contention is  $h$ .

The sentence of a locution is either a reason,  $\delta \in \Delta$ ; or a challenge of an antecedent of some reason, e.g., “challenge  $p$ ” for the reason  $p \supset q$ . If there is a challenge, then it ends a move. *player\_information*( $D, r$ ) consists of the disputation record  $D$ , and  $r$ , together with a list of all rules so far inspected by either player, and the rules that remain uninspected, indexed by their consequents. In response to a challenge, the challenged player has resources  $k'$  in which to find a sufficient response. If successful, the search is conducted for free and its time is not added to cumulative search time. If unsuccessful with the bounded resource, not only is a response to the challenge still required, but the entire search time, including the  $k'$ , is charged to the responding player.

The definition of *move\_options* is completed by giving a precise rule for sufficient response. After any move by *pro*, *pro* must be capable of engaging in dialogue establishing  $h$ ; after any move by *con*, *con* must be able to prevent *pro* from engaging in such dialogue. This *sufficiency dialogue* could be part of the disputation record, but we exclude it for simplicity. The effect of such dialogue, however, is important. Let *Reasons* be the set of reasons thus far introduced by either player, i.e.,  $\Delta \cap \text{sentences}(\Gamma(D))$ . *pro* must be able to show that  $h$  is *live*.

A literal is *live* for a player if (a) it is a evidence; or (b) it is a literal for which there is no reason in *Reasons* for its negation, and for which there are no challenges in this move (this is a promissory note for future search); or (c) a *tree of potential argument* can be cited for the literal which is *undefeated* for the player.

A *tree of potential argument* for a literal (henceforth, an argument) is a collection of reasons that can be organized into a tree by: (i) taking the literal to be root; (ii) taking the antecedents of the unique reason for this literal to be its children; (iii) checking that no literal appears with its negation in this tree; (iv) checking that the leaves of this tree are each live. An argument is *undefeated* for the player if the negation of a literal in the tree cannot be cited to have potential argument for the opposing player where the argument cited has *appropriate strength*.

An argument *has appropriate strength* against its counterargument if: (a) the argument in question is for *con*, and this argument for *con* is not *less specific* than the argument for *pro*; or (b) the argument in question is for *pro* and this argument for *pro* is *more specific* than the argument for *con*. Specificity depends on *activation*: an argument for a literal is activated by a set of literals if the set of literals contains a cutset (not including the root) of the argument.

An argument is *more specific* than another if there is a set of literal that can be cited that activates the lesser argument, but does not activate the greater argument.

In addition to termination because *pro* has used too much resource, natural termination occurs when either player can make no sufficient response.

For example, let  $a_i$  be literals that are consequents of reasons and  $e_i$  be evidence.  $k = 15$ ;  $k' = 5$ .  $h$  is  $a_0$ . Literals in antecedents of defeasible rules are assumed conjoined.

1. PRO:  $\{a_1, a_2, a_3\} \succ \neg a_0$  3

Note that this argument is not (yet) grounded in evidence. *current\_opinion* is *pro*; *move\_options* for *con* include: *challenge*  $a_1$ ,  $a_2$ , or  $a_3$ ; cite a reason for  $\neg a_1$ ,  $\neg a_2$ , or  $\neg a_3$ ; or cite a reason for  $\neg a_0$ .

2. CON: *challenge*  $a_1$  0

*current\_opinion* is *none*; *move\_options* for *pro* forces giving a reason for  $a_1$ , or else giving a new reason for  $a_0$ . Suppose *pro* does not meet the challenge in 5 units of search, and finding a response requires 7 units, retreating from the first reason, and giving a new reason for  $h$ :

3. PRO:  $\{a_3, a_4, e_1\} \succ \neg a_0$  10

This argument also is not fully grounded in evidence. Note that since the challenge was not met, search was charged to *pro*. *current\_opinion* is *pro*; *move\_options* for *con* include: *challenge*  $a_2$  or  $a_4$ ; cite a reason for  $\neg a_3$  or  $\neg a_4$ ; or cite a reason for  $\neg a_0$ .

4. CON:  $\{a_5\} \succ \neg a_3$  4

*Con*, too is not obliged to ground in evidence immediately. *current\_opinion* is *none*; *move\_options* for *pro* include challenging  $a_5$ , giving a reason for  $\neg a_5$ , giving a new reason for  $a_3$ , or else giving a new reason for  $a_0$ .

5. PRO: *challenge*  $a_5$  10

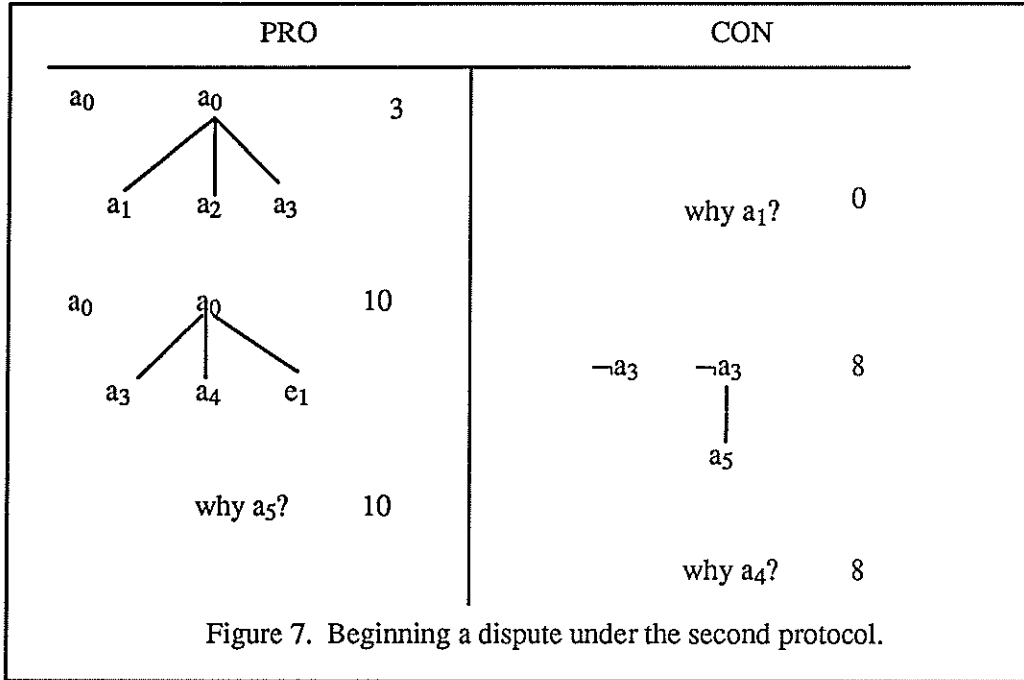
If *con* cannot meet this challenge, or if he chooses not to try, then he must give an alternative attack *pro*'s reason in locution 3:

6. CON: *challenge*  $a_4$  8

... and so on. Since challenges do not expend the resources of the challenger, the apparently winning player can always be forced to ground arguments on which he eventually relies in evidence, through challenges by the opposing player.

### Example Three: Bundling Claims.

This is the third protocol that is based on the same conception of argument, support, and defeat. This protocol was invented to study strategic aspects of disputation. Like the first protocol, this one requires that an argument, if given, be a complete argument grounded in evidence  $K$ . Unlike that protocol, more than one claim can be made in a single move, and not all claims need be made with argument; in fact, only a single argument is allowed in a move, though the number of claims is unrestricted (claims must be relevant). Like the second protocol, the opponent can challenge claims for which arguments are not complete (in this protocol, because they are not given; in the last protocol, because they are not grounded in evidence).



When a claim is made, it is bundled with the other claims in that move. As any claim is won or lost, so too are all of the other claims bundled with it that have not yet been argued. As claims are won and lost in this manner, points are awarded. A lopsided score can cause early termination. This models the rhetorical maneuver that is often made: “I reject  $a$ ,  $b$ , and  $c$ ; but even granting all of them, rejecting  $d$  suffices, and here is my argument for  $\neg d$ .” This is a posturing which attempts to focus debate on  $d$ , presumably because of limited time, and asks that  $\neg a$ ,  $\neg b$ , and  $\neg c$  be conceded with  $\neg d$ , if indeed  $\neg d$  can be established. The opposition can accept the challenge to focus on  $d$ , or can instead focus on  $a$ ,  $b$ , or  $c$ . Doing any of the latter, and establishing for example  $a$ , would be damaging to the player who postured.

Why would a player attempt such a maneuver when rejection of  $d$  suffices, and when rejection of  $a$ ,  $b$ , and  $c$  could be attempted at a later time if needed? Because no matter which of  $a$ ,  $b$ ,  $c$ , or  $d$  is the focus of dispute, victory on any one is dramatic if the opponent chose the focus of the dispute. Presumably this affects adjudication, or in our case, can force premature termination. The potential costs for such a move are that the opposing player can choose where now to focus the dialogue. So this protocol allows risky moves to be made by a player.

*current\_opinion* is *pro* or *con*, and starts at *con*. *player\_to\_move* begins *pro* and switches between *pro* and *con* as each meets the obligation of changing *current\_opinion*, or otherwise meets locution obligations. If resources have not expired nor early termination occurred, *current\_opinion(D)* is *pro* if a distinguished sentence,  $s_0$ , the main point of the dispute, is justified with respect to *arguments(D)*. Otherwise, it is *con*. *player\_information* includes the entire disputation record together with the *current\_score*, which is detailed below. Players also have available information about relative implausibility or vulnerability of various claims, which players might want to use to inform their strategic choices.

Locution sequences include claims and arguments for claims, as in the first protocol, and also include challenges, as in the second protocol. Unlike the first protocol, where a claim had to be paired with an argument for the claim, claims need not appear with arguments for them. A move is restricted to at most one argument, its claim, and a number of additional claims, so long as each claim is *new* and *relevant*.

To define *newness* and *relevance* of claims, *arguments* and *sentences*( $\Gamma(D)$ ) will again be used. A claim is *new* if it has not yet occurred, i.e., if *sentences*( $\Gamma(D)$ ) does not contain the claim. A claim is *initially unargued* if no argument appears with it in the move in which it is first introduced.

*Current\_opinion* is defined in terms of *arguments* as in the first protocol. In fact, *current\_opinion\_on(s)* will be defined for every sentence *s* that is claimed initially unargued, then challenged. This opinion is *p* if *p* is first to claim *s*, and there is an ultimately supporting argument for *s* among *arguments*; it is the opposite of *p* if there is an ultimately supporting argument for the negation of *s*. A claim is *relevant* if, were it to be established (i.e., were there to be an ultimately supporting argument for it, which in this system is equivalent to supposing that the claim in question were background contingent evidence), *current\_opinion* or *current\_opinion\_on(s)* for any *s* would change. This condition can also be stated in terms of sufficient responses since we have already fixed the rules for defeat and justification; however, we will not restate this condition here.

As before, *move\_options* is defined in terms of changes to *current\_opinion*. If *party\_to\_move* is *p*, then a move, *m*, can contain locutions by *p* at acceptable resource levels such that:

1) *current\_opinion* or *current\_opinion\_on(s)* for any *s* changes by concatenating *m* onto  $M(D)$ , and taking the resources of the resulting disputation stage to be *resources(m)*; and *m* contains only relevant claims and possibly arguments for those claims; and any unargued claim must be new;

2) *m* contains at least one new claim, unargued and relevant; and *m* contains only relevant claims and possibly arguments for those claims; and any initially unargued claim must be new;

3) *m* contains just a challenge to an initially unargued claim made by the opposing player; or

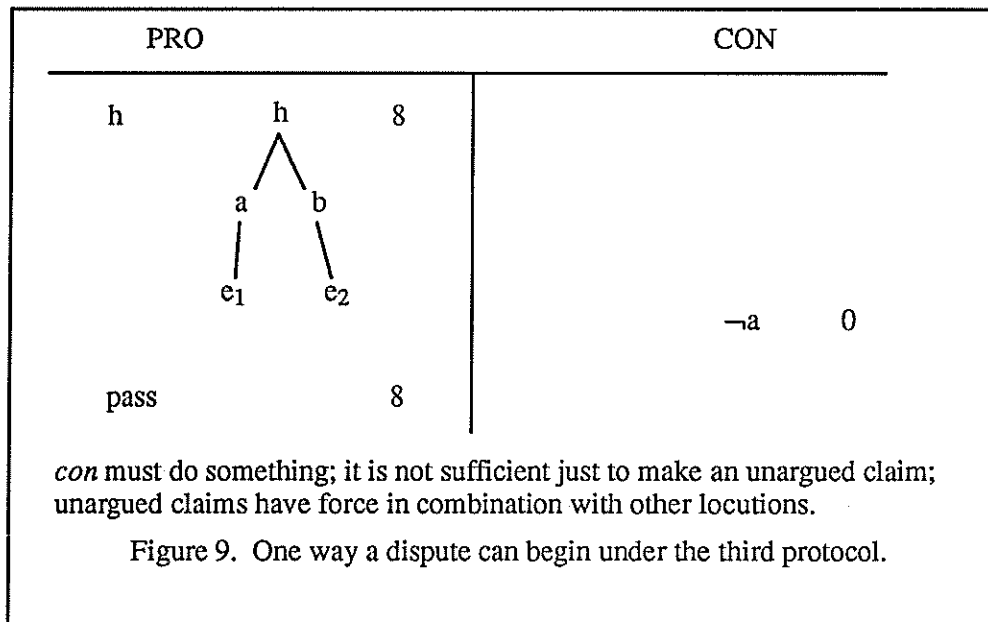
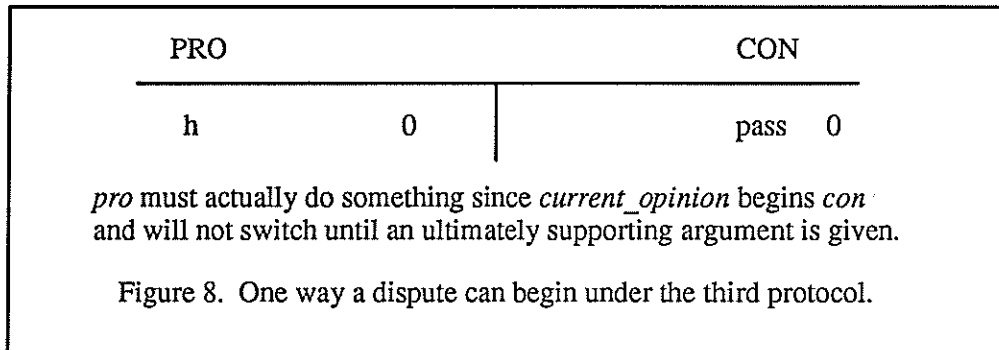
4) *current\_opinion* favors *p* and *m* contains the single locution, with a null sentence, which is interpreted as a *pass* (this can be used to shift the burden to go forward; when an initially unargued claim has been lost by *p*, but *current\_opinion* still favors *p*).

Unlike MacKenzie dialogue games, failure to challenge an unargued claim is not concession of the claim. Unchallenged, initially unargued claims contribute to early termination. They are bundled with the other claims in the move in which they occur, and under certain conditions, these bundles may have value.

Let *s* be an initially unargued claim made by *p* in a move *m* that contained *n* new claims. The *value* of *s* is *n*. At any disputation stage, the *current\_score* for *p* is obtained by summing the value of *s* for every challenged, initially unargued claim *s* for which *current\_opinion\_on(s)* = *p*. By design choice, unchallenged initially unargued claims do not contribute directly to score. Also, claims bundled together in a single move can each contribute directly to score if chal-

lenged and supported, and they may not even contribute to the same player's score. Finally, score draws from claims that may have been made in response to arguments upon which a player no longer even relies, again, by design choice (mostly for simplicity).

Termination occurs when a player is unable to make a legitimate move within acceptable resource limits, or when *current\_opinion* favors *p* and *current\_score* is lopsided in favor of *p*. Lopsidedness, like resource limitation, is determined by pragmatics. The next figures demonstrate different ways a disputation can begin under this protocol.

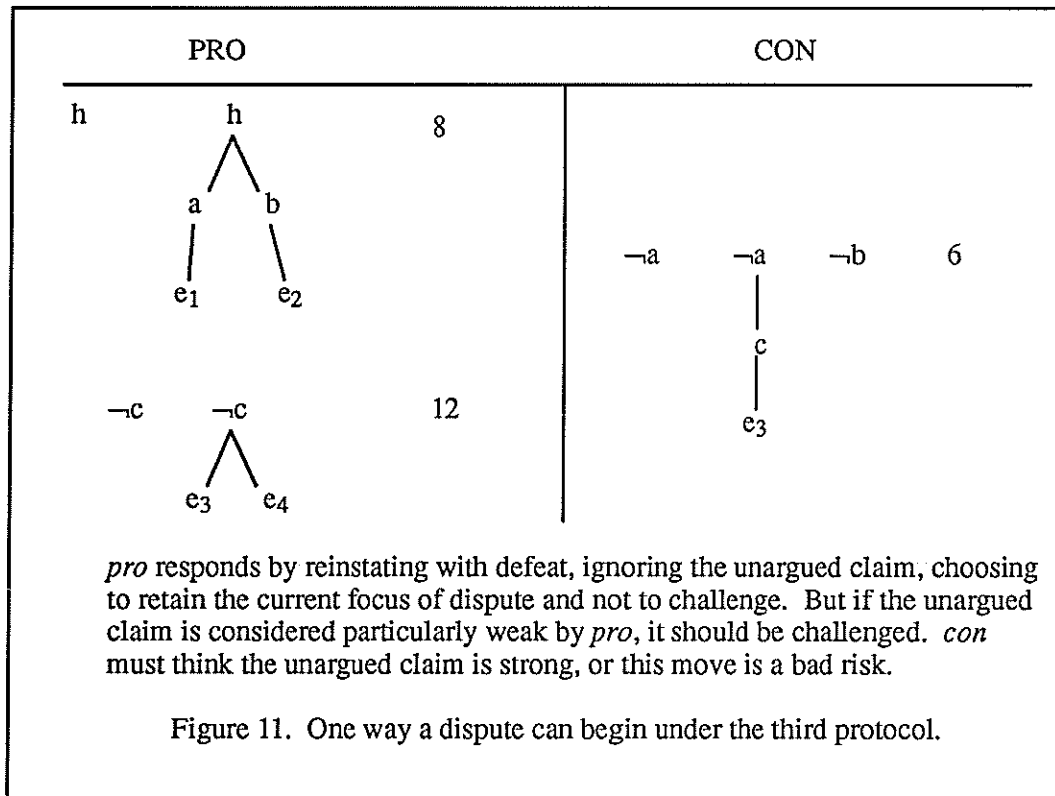
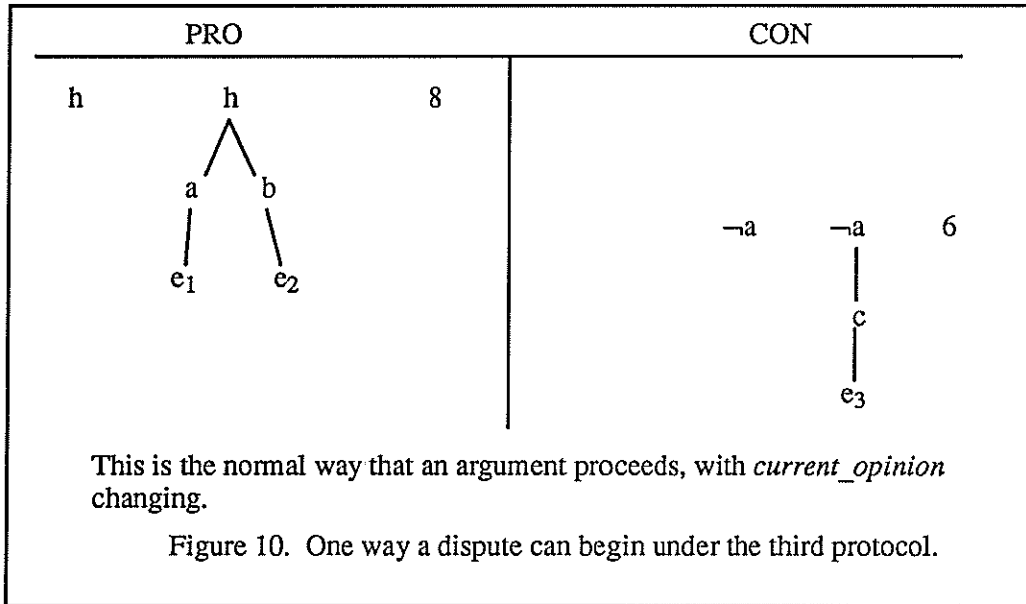


#### Example Four: Players Responsible for Meta-Arguments.

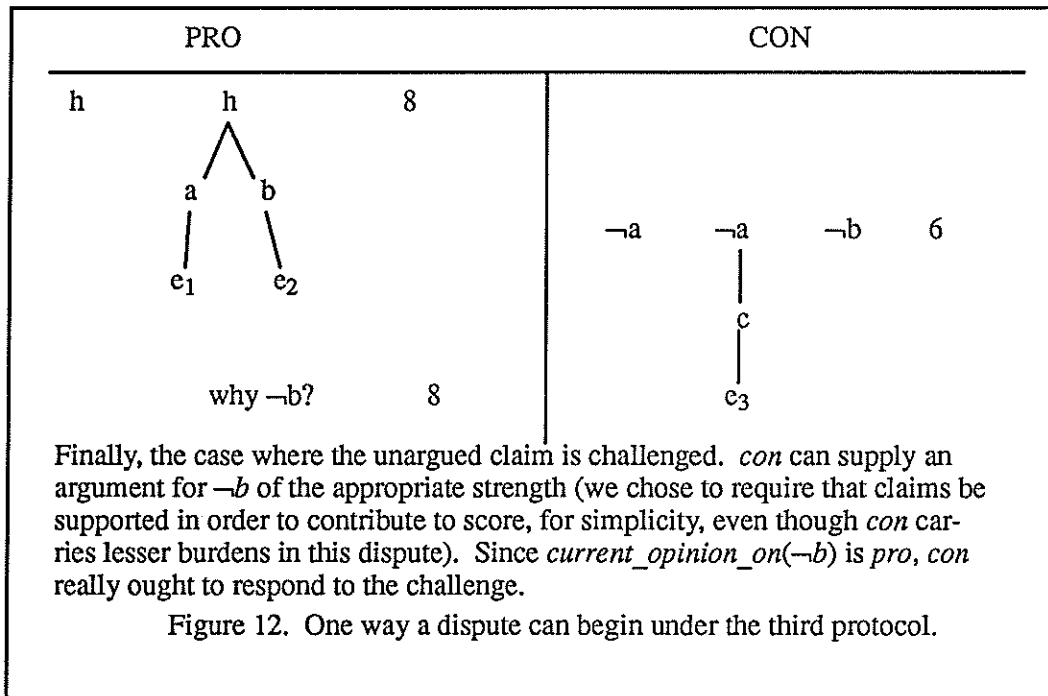
Finally, a protocol may require defeat relations between arguments to be argued by the parties.

The locutions of the two parties thus include meta-claims and meta-arguments in addition to claims and arguments. All object-level arguments are named when introduced. Meta-claims about the arguments use these argument names. Meta-claims are supported by meta-arguments of a special kind so that there is no regress to meta-meta-claims about meta-arguments. The





meta-claims are that either an argument defeats another or that it is undefeated by another, therefore free to interfere, unless some other argument defeats it. Meta-claims concern asymmetric activation, which is important under the current conception of defeat; but this is just an example:

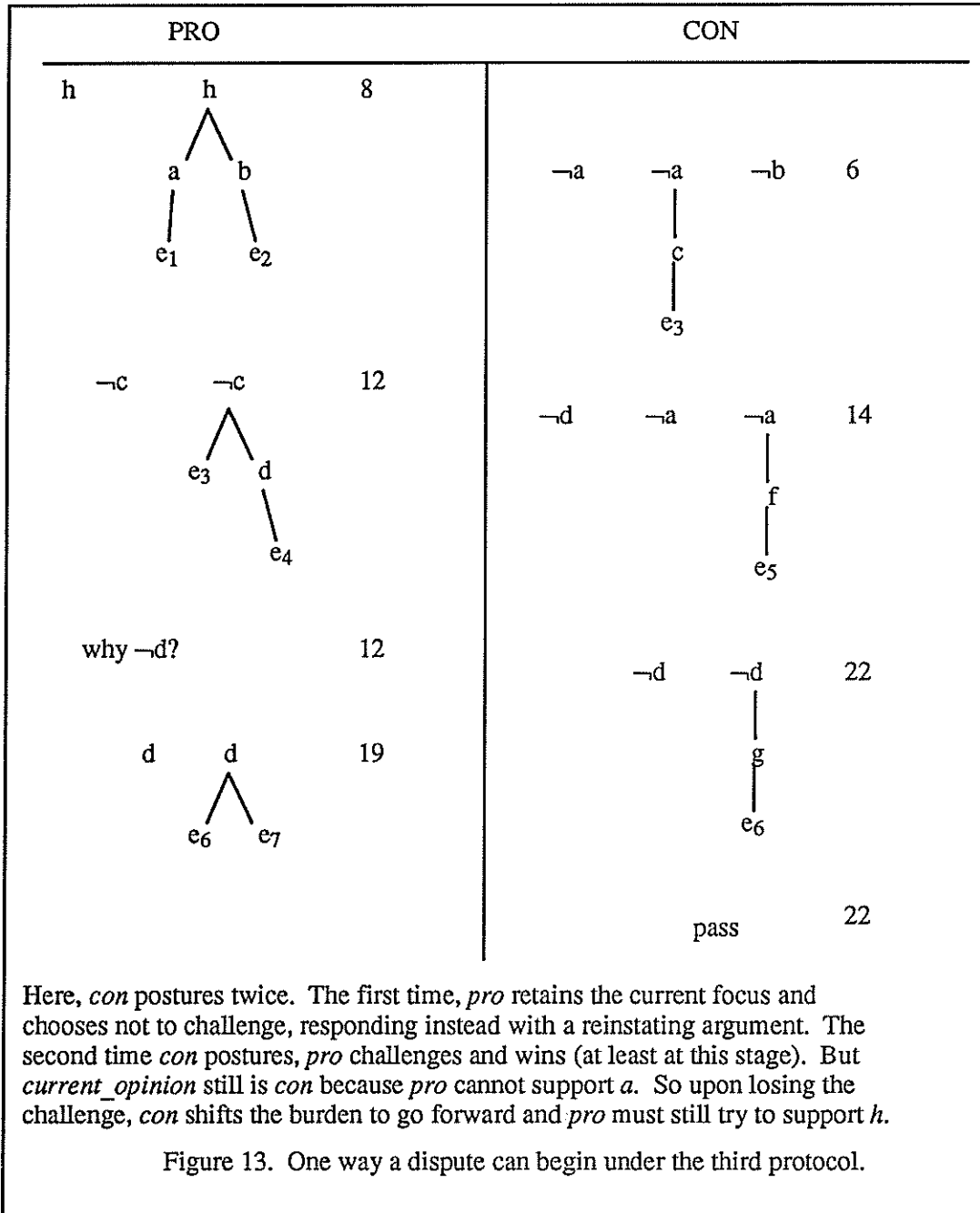


a different conception of defeat among arguments would suggest different kinds of meta-claims.

Requiring parties to supply their own arguments for defeat (hence, for specificity, hence for asymmetric activation) is useful since the test for specificity can be computationally expensive. Technically, the computation requires determining non-activation of one theory as well as activation of another, so it is undecidable. Since the idea of asymmetric activation is itself dialectical, it makes sense to make its determination a part of the dialogue.

Locutions may contain a single object-level claim, always paired with an argument for the claim, which is named. A claim always occurs with a meta-claim, which essentially claims that this object-level claim is an adequate response, unless it is the first argument for the main point of the dispute,  $s_0$ , in which case the adequacy is clear. Meta-claims can sometimes be made without object-level claims. Meta-claims always occur with meta-arguments for the meta-claim.

For simplicity, assume that the most recent meta-claim among those that conflict is always the best among them. To guarantee that this makes sense requires stating some conditions on meta-arguments that will not be stated here. However, this is technically possible since entailment claims are assumed to be correct when made, but claims of non-entailment can be in error because they are the result of negation-as-failure proof attempts. Meta-arguments about asymmetric activation consist of an entailment claim and a non-entailment claim. A sequence of disagreeing meta-claims occurs, for example, when: first, an argument A1 is claimed to bear no specificity relation to another A2, because there are no asymmetric activators (discovered) for either; then, one party claims to find an asymmetric activator, a sentence that activates A1, but does not (appear to) activate A2; then, the opposing party finds the entailment relation that had

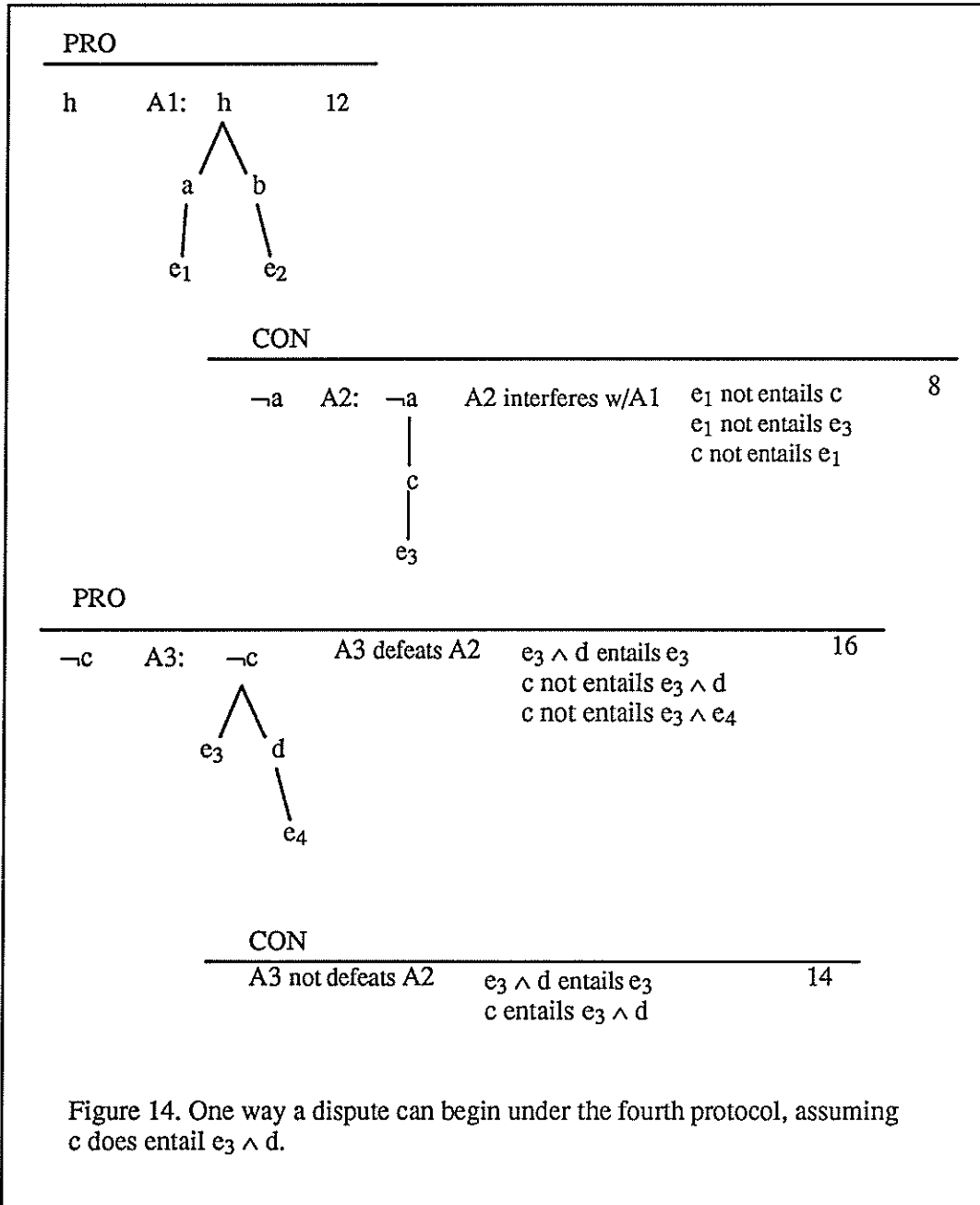


been overlooked, so the putative asymmetric activator is actually a symmetric activator, activating both arguments; then, another asymmetric activator is claimed to be found, and so forth.

*current\_opinion* is determined by following the Pollock-motivated labeling scheme that determines which arguments are ultimately supporting; however, there is a restriction: the defeat relations that are recognized during labeling are only those that have been claimed in the disputation record. In fact, only the best (latest among disagreeing) defeat claims (actually meat-claims) are recognized. *current\_opinion* begins *con*, with *player\_to\_move pro*, then both alternate as a sufficient response is made, that is, as moves are taken from *move\_options*.

*move\_options* consists of appropriately resource-restricted locutions that would alter *current\_opinion*. *current\_opinion* is *pro* just in case there is ultimate support for  $s_0$  under the restricted notion of ultimate support. Note that a sufficient response can consist of either a claim, argument, meta-claim, and meta-argument, or just the latter pair.

Termination occurs, as before, through resource expenditure, or movement into a winning position, where there are no *move\_options* for *player\_to\_move*. An example is contained in the next figure.



## 8. Prospects.

Disputation is a complex thing.

The formal theory of rational belief for demonstrative argument, mathematical logic, is not simple, nor particularly elegant, despite the beauty beheld by some. Still, the pre-formal and semi-formal investigations here have unsatisfying complexity and point in unsettling directions. We must be certain that the complexity and compass are unavoidable.

We begin with the assumption that there are claims, and there are reasons relating one claim to another. When the reasons are policies for constructing arguments, and this is most of the time, the constructions are non-demonstrative because policies can conflict. Constructivism and resource limitation require us to acknowledge non-determinism of outcome and non-monotonicity in the computation of outcome. Hence, there may be only weak guarantees about how the outcome of the process may be regarded as "correct." Yet we must avoid protocols too simple to produce what we would regard as rational belief. Fair and effective protocols are desired. Crap shoots are disdained. Under a protocol, choice may be unavoidable, and choice can often be attributed to a party. Choice raises the question of whether some choices are better than others, and the question of how to exclude choices that do not further rational inquiry.

We begin wanting only to compute entailments for a database expressed in what seemed to be and what experience tells us is a convenient language. The linguistic artifice of non-demonstrative inference merely reflects the non-demonstrative reasoning we use; this must be the explanation of why such language is convenient. Its convenience belies its historical resistance to formalization.

We begin looking for a calculus for non-demonstrative entailments. We end fixated on the design of formal games. The games' rules depend on a conception of argument defeat and claim support too new to have had proper critical appraisal. Guidance on the selection of an appropriate game does not exist. Further, we can expect even the best future guidance to be indeterminate, permitting many possibilities. This expectation is the result of both negative experience with the *a priori* justification of demonstrative systems, and experience with protocols for disputation in the world.

In fact, this investigation merely begins the formal invention and analysis of protocols. Important possibilities remain unexplored. Adjudicators intervene. Protocol is negotiated. Real adjudicators attend to pragmatics, to the ramifications of their decisions, to utilities and disutilities. Whether they ought to do this, whether doing so jeopardizes *stare decisis* or some other principle dearly held, is within the scope of an investigation of this kind. This is broad work not for the timid.

History provides a mandate and artificial intelligence provides a need. Aristotle cleaved non-demonstrative from demonstrative reasoning. Keynes signaled that failure to formalize the non-demonstrative in favor of the demonstrative would run us afoul. Pollock, Simon, Rescher, and Doyle have pointed to where formalism can today be found. Whether adequate, cogent, formal tools can be developed, or at least how quickly, seems now to be a matter of effort. It would be a mistake, however, to suppose that much of the work is done.

## 9. Bibliography

- Ackrill, J.L., ed. *New Aristotle Reader*, Princeton, 1987.
- Alexy, R. *A Theory of Legal Argumentation*, Clarendon, 1989.
- Alvarado, S., ed. *Proc. AAAI Spring Symposium on Argument and Belief*, 1991.
- Alvarado, S. *Understanding Editorial Text*, Kluwer, Dordrecht, 1990.
- Ashley, K. "Defining salience in case-based arguments," *Proc. IJCAI 1989*.
- Ashley, K. and V. Aleven. "Generating dialectical examples automatically," *Proc. AAAI 1992*.
- Baker, A. and M. Ginsberg. "A theorem-prover for prioritized circumscription," *Proc. IJCAI 1989*.
- Barth, E.M. and Krabbe, E.C.W. *From Axiom to Dialogue*, de Gruyter, Berlin, 1982.
- Barth, E.M. and Martens, J.L., eds. *Argumentation*, Benjamins, Amsterdam, 1982.
- Barth, E.M. "A new field," *Synthese* 63, pp. 375–388, 1985.
- Becker, W. and Essler, W., eds. *Konzepte der Dialektik*, Klostermann, 1981.
- Brown, R. "The burden of proof," *American Philosophical Quarterly* 7, pp. 74–82, 1970.
- Bulygin, E. "Norms and logic," *Law and Philosophy* 4, pp. 145–163, 1985.
- Byrne, Edmund. *Probability and Opinion: A Study in The Medieval Presuppositions of Post-Medieval Theories of Probability*, Nijhoff, Netherlands, 1968.
- Carabelli, A. *On Keynes's Method*, St. Martin's Press, 1981.
- Cavalli-Sforza, V. and Moore, J. "Collaborating on arguments and explanations," *Proc. AAAI Spring Symposium on Producing Cooperative Explanations 1992*.
- Cavalli-Sforza, V., Lesgold, A., and Weiner, A. "Strategies for contributing to collaborative arguments," *Proc. Cognitive Science 1992*.
- Church, Alonzo and Kline, G. "Review of A. Philipov's logic and dialectic in the Soviet Union," *Journal of Symbolic Logic* 18, pp. 272–273, 1958.
- Churchman, C. *The Design of Inquiring Systems*, Basic Books, 1971.
- Cullison, A. "The logic of legal rules and obligation," *Connecticut Law Review* 13, pp. 215–261, 1981.
- Doyle, J. "A truth maintenance system," *Artificial Intelligence* 12, 1980.
- Doyle, J. *A Model for Deliberation, Action, and Introspection*, doctoral dissertation, MIT, 1980.

- Doyle, J. "Constructive belief and rational representation," *Computational Intelligence* 5, 1989.
- Doyle, J. "The foundations of psychology: a logico-computational inquiry into the concept of mind," in R. Cummins and J. Pollock, eds., *Philosophy and AI*, MIT, 1991.
- Dworkin, R. *A Matter of Principle*, Harvard, 1985.
- Ehninger, D. and Brockriede, W. *Decision by Debate*, Harper and Row, 1978.
- Elgot-Drapkin, J., Miller, M., and D. Perlis. "Memory, reason, and time: the step-logic approach," in R. Cummins and J. Pollock, eds., *Philosophy and AI*, MIT Press, 1991.
- Flowers, M., McGuire, R., and Birnbaum, L. "Adversary arguments and the logic of personal attacks," in W. Lehnert and M. Ringle, eds., *Strategies for Natural Language Processing*, Erlbaum, 1982.
- Gottlieb, Gidon. *The Logic of Choice: An Investigation of The Concepts of Rule and Rationality*, Allen-Unwin, London, 1968.
- Hall, C.G. "On the logic of the 'internal logic' of law," *Cambrian Law Review* 15, pp. 31–38, 1984.
- Hamblin, Charles "Mathematical models of dialogue," *Theoria* 37, pp. 130-155, 1971.
- Hart, H.L.A. *The Concept of Law*, Oxford, 1961.
- Haynes, D. "The language and logic of law," *Univ. Miami Law Review* 35, pp. 183–254, 1981.
- Hegselmann, Rainer. *Formale Dialektik*, Meiner, Hamburg, 1985.
- Hogan, J. "The Dialectics of Aristotle," *Philosophical Studies* 5, pp. 3–21, 1955.
- Kahane, Howard. *Logic and Contemporary Rhetoric*, Wadsworth, Belmont, 1971.
- Kneale, W. and Kneale, M. *Development of Logic*, Oxford, 1962.
- Konolige, K. "Defeasible argumentation in reasoning about events," *Methodologies for Intelligent Systems* 3, 1988.
- Krabbe, Erik. "The adequacy of material dialogue-games," *Notre Dame Journal of Formal Logic* 19, pp. 321-330, 1975.
- Krabbe, Erik. "Formal systems of dialogue rules," *Synthese* 63, pp. 295-328, 1985.
- Levesque, H. "A logic of implicit and explicit belief," *Proc. AAAI 1984*.
- Levi, I. *Enterprise of Knowledge*, MIT Press, 1980.
- Loui, R. "Defeat among arguments," *Computational Intelligence* 4, 1987.

- Loui, R. "Theory and computation of uncertain inference and decision," dissertation, University of Rochester, 1987.
- Loui, R. "Dialectic, computation, and ampliative inference," in *Philosophy and AI*, R. Cummins and J. Pollock, eds., MIT Press, 1991.
- Loui, R. "Argument and belief: where we stand in the Keynesian tradition," *Minds and Machines 1*, 1991.
- Loui, R., Norman, J., Merrill, A., Olson, J., Costello, A., and K. Stiefvater. "Computing specificity," submitted to *Computational Intelligence*, 1992.
- Loui, R. and W. Chen. "Rules for arguing with cards," Technical Report, Washington University Department of Computer Science, 1992.
- Loui, R. "How a formal theory can be normative: implementation versus interpretation," *Journal of Philosophy*, 1993 (to appear).
- Lorenz, Kuno. *Arithmetik und Logik als Spiele*, dissertation, Kiel, 1961.
- Lorenz, Kuno. "Rules versus theorems," *Journal of Philosophical Logic* 2, pp. 352-369, 1973.
- Lorenzen, Paul. "Dialogical foundation of logical calculi," in *Constructive Philosophy*, (collected papers), U. Mass., Amherst, 1987.
- MacEwan, Elias. *The Essentials of Argumentation*, Heath, Boston, 1900.
- MacKenzie, J.D. "The dialectics of logic," *Logique et Analyse* 94, pp. 159-177, 1981.
- MacKenzie, J. "Four dialogue systems," *Studia Logica* 44, pp. 567-583, 1990.
- Marshall, C., Halasz, F., Rogers, R., and W. Janssen, Jr. "Aquanet: a hypertext tool to hold your knowledge in place," *Proc. Hypertext 1991*.
- March, J. "Bounded rationality," in J. Elster, ed., *Rational Choice*, NYU Press, 1986.
- Mason, R. "A dialectical approach to strategic planning," *Management Science* 15, 1969.
- Meyer, Michel. *From Logic To Rhetoric*, Benjamins, Amsterdam, 1986.
- Mitroff, I. "A Brunswik lens model of dialectical inquiring systems," *Theory and Decision* 5, 1974.
- Mitroff, I. and R. Mason. "On the structure of dialectical reasoning in the social and policy sciences," *Theory and Decision* 14, 1982.
- Mitroff, I., R. Mason, and Barabba, V. "Policy as argument: a logic for ill-structured decision problems," *Management Science* 12, 1982.
- Newman, S. and C. Marshall. "Pushing Toulmin too far: learning from an argument representation scheme," Xerox PARC Technical Report SSL-92-45, 1992.



- Nute, D. "Defeasible reasoning and decision support systems," *Decision Support Systems* 4, 1988.
- Pearl, J. "Probabilistic semantics for nonmonotonic reasoning," in R. Cummins and J. Pollock, eds., *Philosophy and AI*, MIT Press, 1991.
- Pollock, J. *Knowledge and Justification*, Princeton, 1974.
- Pollock, J. "A theory of moral reasoning," *Ethics* 96, pp. 506–523, 1986.
- Pollock, J. "Defeasible reasoning," *Cognitive Science* 11, 1987.
- Pollock, J. *Nomic Probability and The Foundations of Induction*, Oxford U. Press, 1990.
- Pollock, J. "How to reason defeasibly," *Artificial Intelligence* 57, 1992.
- Poole, D. "On the comparison of theories: preferring the most specific explanation," *Proc. IJCAI* 1985.
- Poole, D. "A logical framework for default reasoning," *Artificial Intelligence* 36, 1988.
- Poole, D. "Dialectics and specificity: conditioning in logic-based hypothetical reasoning," *Proc. Workshop on Nonmonotonic Reasoning*, 1990.
- Popper, K. *Conjectures and Refutations*, Routledge, 1963.
- Prakken, H. "A tool in modelling disagreement in law: preferring the most specific argument," *Proc. Intl. Conference on AI and Law*, 1991.
- Prakken, H. "An argumentation framework in default logic," working paper, Computer/Law Institute, Free University Amsterdam, 1992.
- Rescher, N. *Plausible Reasoning*, van Gorcum, Amsterdam, 1976.
- Rescher, Nicholas. *Dialectics*, SUNY Albany, 1977.
- Reiter, R. "A logic for default reasoning," *Artificial Intelligence* 13, 1980.
- Rissland, E. "Artificial intelligence and law," *Yale Law Journal* 99, pp. 1957–1981, 1990.
- Routley, Richard. "Dialectical logic, semantics, and meta-mathematics," *Erkenntnis* 14, pp. 301–331, 1979.
- Sabre, R. M. "An alternative logical framework for dialectical reasoning in the social and policy sciences," *Theory and Decision* 30, 1991.
- Schlink, B. "On a principle of contradiction in normative logic and jurisprudence," *Theory and Decision* 2, pp. 35–48, 1971.
- Shafer, G. and A. Tversky. "Languages and designs for probability judgement," *Cognitive Science* 9, 1985.

- Simon, H. *Models of Bounded Rationality*, v. II, (collected papers), MIT Press, 1982.
- Simari, G. and R. Loui. "A mathematical treatment of defeasible reasoning and its implementation," *Artificial Intelligence* 53, 1992.
- Solovay, E. and V. Strassen. "A fast Monte-Carlo test for primality," *SIAM Journal on Computing* 6, 1977.
- Toulmin, Stephen. *The Uses of Argument*, Cambridge, 1958.
- Touretzky, D. *Mathematics of Inheritance Systems*, Pitman, 1986.
- Touretzky, D., Horty, J., and Thomason, R. "A clash of intuitions," *Proc. IJCAI* 1987.
- van Benthem, J.F.A.K. "What is dialectical logic?" *Erkenntnis* 14, pp. 333-347, 1979.
- Vreeswijk, G. "Defeasible dialectics," technical report IR-265, Faculteit der Wiskunde en Informatica, Vrije Universiteit Amsterdam, 1991.
- Vreeswijk, G. "Abstract argumentation systems," *Proc. First World Conference on Fundamentals of Artificial Intelligence*, 1991.
- Vreeswijk, G. "The feasibility of defeat in defeasible reasoning," *Proc. International Conference on Knowledge Representation and Reasoning*, 1991.
- Walton, Douglas. *Logical Dialogue-Games and Fallacies*, Lanham, New York, 1984.
- Walton, Douglas. "New directions in the logic of dialogue," *Synthese* 63, pp. 259-274, 1985.
- Walton, Douglas. *Question-Reply Argumentation*, Greenwood, New York, 1989.
- Wellman, M. and Doyle, J. "Preferential semantics for goals," *Proc. AAAI* 1991.