

Washington University in St. Louis

## Washington University Open Scholarship

---

All Computer Science and Engineering  
Research

Computer Science and Engineering

---

Report Number: WUCS-94-12

1994-01-01

### Rationales and Argument Moves

R. P. Loui and Jeff Norman

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



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

---

#### Recommended Citation

Loui, R. P. and Norman, Jeff, "Rationales and Argument Moves" Report Number: WUCS-94-12 (1994). *All Computer Science and Engineering Research*.

[https://openscholarship.wustl.edu/cse\\_research/334](https://openscholarship.wustl.edu/cse_research/334)

Department of Computer Science & Engineering - Washington University in St. Louis  
Campus Box 1045 - St. Louis, MO - 63130 - ph: (314) 935-6160.

**Rationales and Argument Moves**

**R. P. Loui and Jeff Norman**

**WUCS-94-12**

**May 1994**

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

*Submitted to AI and Law 5/94.*



## Rationales and Argument Moves

R. P. Loui  
Washington University  
St. Louis

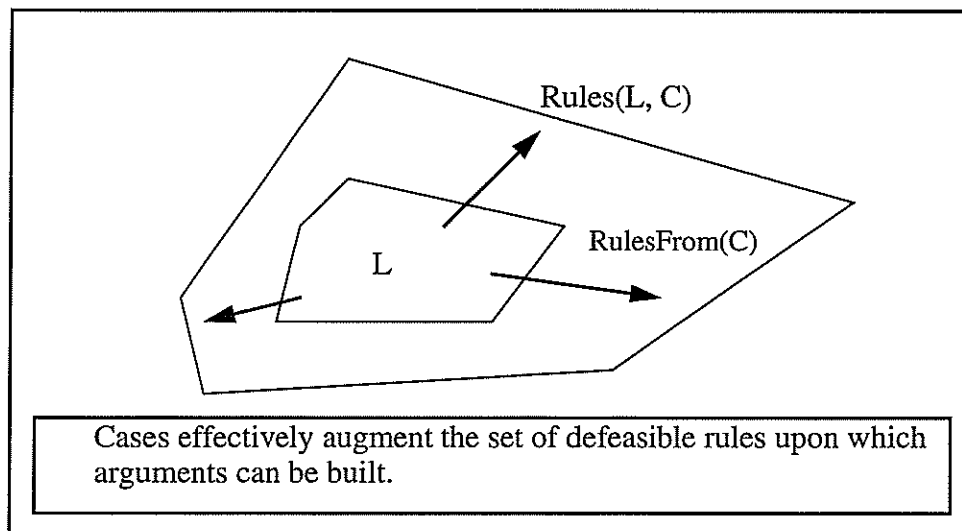
Jeff Norman  
Seyfarth, Shaw, Geraldson, & Fairweather  
Chicago

### I. Procedure and Rationale.

In disputation, claims are supported by arguments, which refer to rules, cases, and evidence. Sometimes, the rationales of rules and the rationales for decisions in cases appear in disputation as well.

Rationales appear in procedural contexts such as debate, dialectical inquiry, and legal argument. In purely declarative contexts, wherein logic usually is defined, the effect of rationales on inference is more mysterious.

In many models of argument, arguments are produced by chaining rules from a set of defeasible rules,  $L$ , and are grounded in a set of undisputed premises, or evidence,  $E$ . Case-based reasoning supplements the set of rules by permitting rules to be formulated from cases. There is thus a set,  $C$  of cases, the effect of which is to augment  $L$ ,  $\text{Rules}(L, C) = L \cup \text{RulesFrom}(C)$ .



This is one framework in which to formalize legal and policy reasoning, a framework inherited from the logicians, suitably altered by researchers in philosophical logic and artificial intelligence. There is considerable agreement on the way that cases augment rules, as described in the Rissland-Ashley work [Rissland-Ashley87].

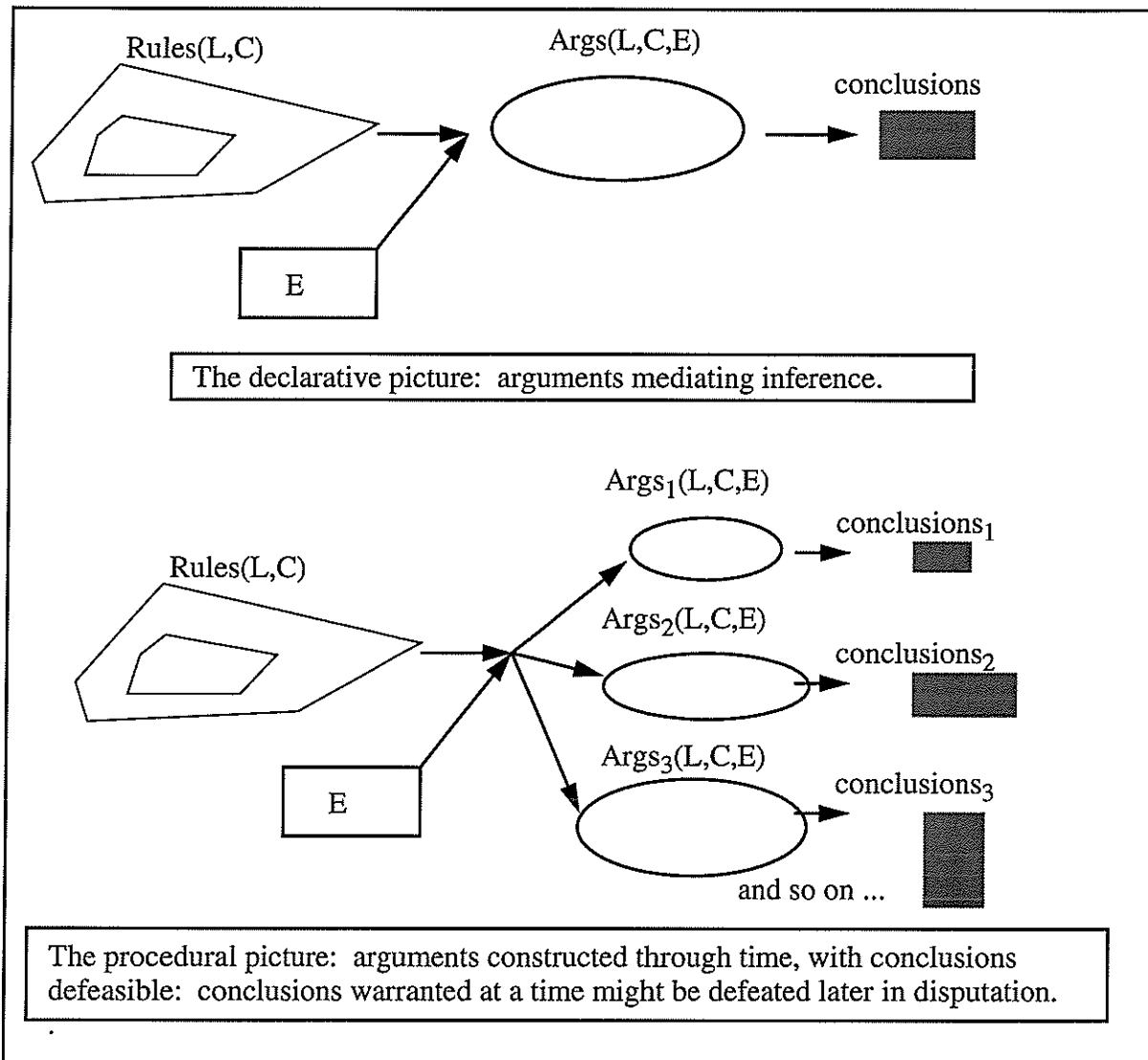
Declaratively (i.e., from a traditional logical point of view), the interest is in the complete set of arguments that can be constructed upon L, C, and E:  $\text{Args}(L, C, E)$ . A conclusion must be warranted with respect to the full complement of arguments: if an effective counterargument is possible, then an argument fails to warrant its conclusion, whether anyone ever finds that counterargument. Alchourron and Bulygin [Alchourron-Bulygin71], Soeteman [Soeteman89], and Simari and Loui [Simari-Loui92] are examples of declarative accounts of legal reasoning, though the former works do not explicitly construct arguments.

Procedurally (i.e., computationally), parties to a dialogue consume limited resources, producing a succession of sets  $\text{Args}_i(L, C, E)$ , corresponding to the various stages  $i=1, 2, \dots, n$  of the dispute. Warrant is defined with respect to a particular set of arguments at a particular stage,  $\text{Args}_i(L, C, E)$ . The appropriate stage is determined by a termination rule. Sometimes it is known in advance; sometimes it is a bound that is revealed when it is reached; sometimes an  $i$  is appropriate merely because  $\text{Args}_i$  has a quiescence, e.g., because one party can make no useful response. The protocol for inquiry may allow parties to introduce objects significant to the dispute, other than arguments. For example, challenges might be allowed in the dialogue. Rescher [Rescher77], Alexy [Alexy89], Loui [Loui92], Vreeswijk [Vreeswijk93], and Gordon [Gordon93] are examples of this kind of work; Skalak and Rissland [Skalak-Rissland91], Pollock [Pollock92], Prakken [Prakken93], and can be interpreted as this kind of work, too.

A rationale for a rule is a structure that contains relevant information about the reason for the rule's adoption. A rationale for a case, similarly, is relevant additional information about the decision reached. Following Toulmin, one can refer to the rationale as the "backing" of the "warrant," where Toulmin's warrants are essentially the rules in L. In the language of legal philosophy, *ratio decidendum* and *ratio legis* are rationales. Recent authors have referred to "principle" or "purpose." The essential question is what are the appropriate forms of rationales and how they can be brought to bear in argument.

There are many kinds of rationales; rules are adopted on various grounds. Defeasible rules can be adopted through game-theoretic or inductive reasoning, on the whim of authority, or by arbitrary convention. A defeasible rule that is a policy, which might be a rule in a legal domain, often balances competing interests. Sometimes the particular balance struck is part of the grounds for the rule's adoption. Sometimes the rule strikes a balance in a principled way, but as often it has the right politic or expedience. Sometimes the grounds for adopting a policy are decision-theoretic. Decisions in precedent-setting cases can be made on similar grounds.

The rationales for rules and cases considered here constitute a subset of those that can be imagined. They might generally be classified as compilation rationales. These rationales might be described fully within the vocabulary of a system of argument. Excluded, for instance, are decision-theoretic rationales, which require the introduction of additional information representing utilities. The present aim is to discuss how dialectical moves force reversion to the forms from which rules were compiled, and how the the model might thereby be made more comprehensive, how the repertoire



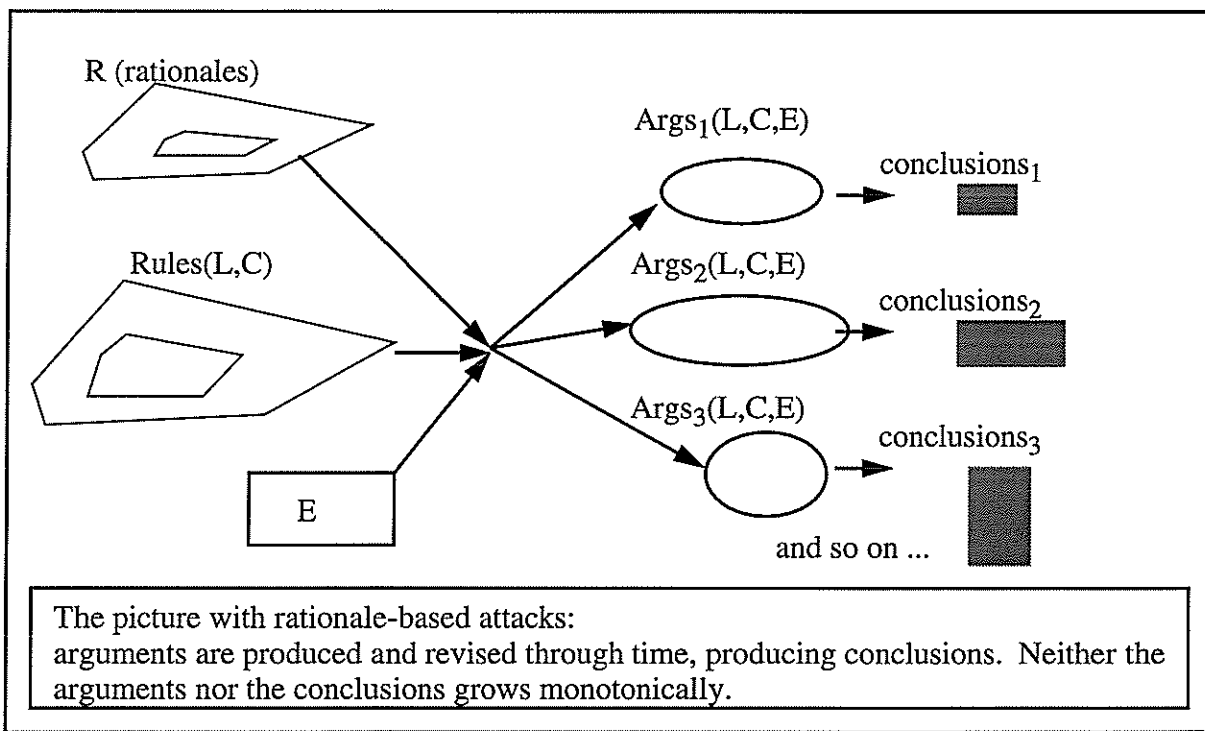
of formally modeled moves might be increased.

Rationales resulting from compilation are compilation rationales. The compilation rationales to be discussed are as follows. For rules: adopt a rule because it compresses a line of reasoning (compression, or c-rationale); adopt a rule because it specializes a general principle (specialization, or s-rationale); adopt a rule because it expresses a regularity that fits a set of cases (fit, or f-rationale). For cases: form an opinion despite competing factors by resolving the competition (resolution, or r-rationale); form an opinion because the set of arguments in the recorded disputation of the case, which may be incomplete, warrants the opinion (disputation, or d-rationale).

To attack a well-formed argument, with rationales unmentioned, there are already a couple of alternatives. Attack some subpart of the argument or provide a separate argument for a contrary conclusion. If the protocol allows meta-argument, attack the sufficiency of the argument at the present point in the dialogue. Attack the legitimacy of the move at the point at which it occurs. That is, show that it does not meet the player's burdens, or provide reasons for claiming that it does

not defeat the argument that it is said to defeat, or otherwise impugn its place or station. There is yet another class of moves: moves that make use of rationales.

Attack a rule used in the argument by citing its rationale and claiming that the rule's scope is exceeded in this application. Or attack a rule that derives from a case, a case-based argument, by citing the rationale for the decision in the case and arguing that the decision is flawed. A new argument that should have been considered has been discovered. Or argue that the decision provides no guidance in the current situation. An argument that can now be made, in the current situation, would have effectively countered the argument that led to the decision in the case. Or argue that a rule is a particular form of a more general rule, or a modern restatement of an older rule, a derivation from broader principle, or a version of a separate rule; then restate the rule in its original form; then attack the new rule, the strength of which has now been altered under considerations of *lex specialis*, *lex posterior*, or *lex superior*.



The purpose of this paper is to sketch a formal account of what data would be required to make such moves, and how those moves affect the state of a disputation.

## II. Informal Examples of Compilation Rationales.

First, consider each of the rationales in simple quasi-legal examples.

### 1. *compression*, or *c-rationale*:

As a defeasible rule, “vehicles used for private transportation are not allowed in the park.” Meanwhile, defeasibly, “vehicles are normally for private transportation.” There is therefore the

two-step argument, for disallowing vehicles in the park. Adopt the policy, with the rationale of compression, that as a rule, “no vehicles in the park.”

To attack an argument using a rule with a c-rationale, restate the argument in an uncompressed form. The resulting uncompressed argument will be more susceptible to attack. Uncompressed arguments present more points to counterattack. They are also not as direct, so they are more easily defeated in some syntactic accounts of defeat among arguments.

Attack by arguing that emergency vehicles are not used for private transportation. This does not affect the argument using the compressed rule, but it devastates the uncompressed argument.

Compression tends not to be fragile because rules are typically compressed when the argument being compressed is resistant to counterargument. It would be folly to compress long arguments, for instance, into one-step rules. Arguments constructed thereupon would always be challenged because so many challenges of the uncompressed form will be possible.

## 2. *specialization, or s-rationale:*

As a principle, “tranquil public spaces should be preserved.” “Parks are tranquil public spaces.” “Disallowing vehicles in a space is a kind of preserving tranquility.” There is currently an argument for a particular park, defeasibly, that it is a public space, and defeasibly, that it should be undisturbed. There is currently no argument that such a park should have vehicles disallowed from being in it. There may be other ways to preserve. The problem is that so far, the contrapositive of the third rule is missing: “In order to preserve a tranquil space, disallow vehicles.” But the principle is entreating in such a way that a policy-maker might still adopt as a rule, “no vehicles in the park.”

Usually, such a rule is adopted in the presence of counterarguments. For instance, “barring vehicles increases commuting effort.” In such instances, the rationale for the rule is more advanced. It is a balance of competing interests, which might actually be an f-rationale, d-rationale, or r-rationale. The s-rationale here is more naive. One way to preserve is to disallow vehicles; one way of meeting the demands of an imperative is to implement in a particular way. Disallowing vehicles implements a kind of preservation. This kind of reasoning could be subtle. Here, a simpler approach is taken. There are often hidden assumptions or background rules. The missing contraposition might be implicit: “defeasibly, preservation entails disallowing vehicles.”

Attack, again, by restating the argument so that the rule in question has its antecedent expressed more generally: “insofar as public spaces should be preserved, no vehicles allowed in public spaces.” A counterargument based on a rule, “vehicles are allowed on public roads in public spaces,” will now suffice as a reply (assuming that the vehicle in question can be argued to have been on a public road). Prior to citing the rationale, such a counterargument might be ineffective because it is less specific: it refers to public spaces, not to parks. Formally, this will be equivalent to attacks on c-rationales, except that there is usually a dominant rule or principle in the argument that is compressed.

Attack, too, by identifying a different way of meeting the demands of the guiding rule, by suggesting a different implementation. This attack corresponds to attacking a rule adopted without full consideration of counterarguments. It will be formally equivalent to attacks using r- or d-



rationales.

Since principles can be implemented in numerous ways, rules based on s-rationales are fragile. They eventually will be challenged and replaced with rules or cases with based on f-, d-, or r-rationales.

### 3. *fit, or f-rationale:*

Adopt a rule because it expresses a regularity among cases. Cases can be actual or hypothetical, but there should be widespread agreement over the way they are decided.

Rule-adoption based on f-rationale is related to theory-formation in scientific reasoning. As in scientific theory-formation, coherence and simplicity are important. Unlike scientific theory-formation, errors of fit are not tolerable. Though errors are not tolerated, there is a way to eliminate errors of legal fit that is not available to scientific theorizing. Errors of legal fit can be eliminated without altering the offending legal rule. A more specific legal rule that deals with exceptions can simply be added to the body of rules. This option is not possible for scientific theory.

The policy “no vehicles in the park” might allegedly fit the cases:

(case1) *disallowed*: a private automobile driving through the park;

(case2) *disallowed*: an antique automobile parked in the park;

(case3) *disallowed*: a golf cart driven into the park;

(case4) *allowed*: a pedestrian strolling in the park.

An argument using a rule with f-rationale can be attacked by first proposing a new rule that also successfully distinguishes the recorded cases of allowed parking from cases of disallowed parking. Then the attack continues by (a) noting that the new rule no longer applies to the current fact situation (or at least, that the argument that it applies has not been given), or by (b) noting that the reformulated rule applies but is not as specific as originally suggested. In the latter situation, the argument is susceptible to attack by counterarguments that would have been considered less specific on the earlier formulation of the rule. Once again, the defeat relations among arguments can be altered by citing rationale.

Reformulating rules corresponds to arguing over which theory has the best fit. We are unwilling at the moment to postulate the conditions under which a “policy theory” fits better than a competing policy theory. Philosophy of science, similarly, still cannot articulate conditions precisely for one scientific theory fitting better than another. Thus, competing formulations of policies to fit cases merely interfere with each other; neither can defeat the other until coherence and simplicity and other criteria of fit have been addressed.

An argument using a rule with an f-rationale may also be attacked by adding or deleting cases. For example, to attack the rationale of this rule, add a case, such as

(case5) *allowed*: a vehicle delivering a tree to be planted in the park;

or delete a case e.g., claim that (case3) is incorrectly decided. Or propose a different regularity consistent with the cases: “vehicles, excluding those performing essential park or public functions, are not allowed in the park”.

Since it is easy to theorize about cases with sets of defeasible rules that allegedly express regularity, rules with f-rationale will be robust only if the question of fit is addressed in detail.

#### *4. resolution, or r-rationale:*

Suppose the case of the hurried government official who detours his automobile through the park raised two arguments that interfered with each other without either defeating the other. “No vehicles in the park.” “Public officials on official business have use of public spaces.” The decision to disallow such detours sets a precedent upon which a variety of future arguments can be constructed. The rationale of the decision, of the precedent, is that it resolves the competing factors: preservation of public space, versus availability of public resources for government business.

To attack an argument relying on this precedent, cite the rationale. Then claim that the current case introduces additional interests which, when weighed, could alter the balance against the prior decision. For instance, suppose a subsequent case also introduces the element of time. In the current case, the official drove through the park when the park was closed. In the prior case, the time of day was Saturday noon. Now arguments based on time of day, government business, and preservation of public spaces must be weighed.

Without citing the rationale and attacking the grounds for the precedent, an argument for allowing the new detour based on time of day could still be made. The argument based on time and the argument based on vehicles in the park are of incomparable specificity. Hence, their disagreement might force a resolution in this case. Discount the old precedent and the new resolution must start anew, weighing time of day and government business against the ban on vehicles in the park. Leave the rationale of prior decision unmentioned, and the current case will compare time of day against an established precedent based both on government business and the ban on vehicles in the park. The difference is subtle.

Because new cases almost invariably introduce significant new factors unweighed in the earlier decision, precedents with r-rationales tend to be fragile. But as noted, the stakes are not high when an r-rationale case is discounted by this kind of attack.

R-rationales are special kinds of d-rationales. So there are other attacks on r-rationales that are described as attacks on d-rationales.

#### *5. disputation, or d-rationale:*

Disputation that informs decision in a case can be easier than the weighing of indeterminate factors. A decision might simply be mandated by superior argument, where the adjudicator played no major role. Appeal to such a case in future argument is susceptible to review of the recorded disputation of the case. The arguments that were persuasive in the prior case may not be persuasive in the present case and its new context.

When the record of the disputation that led to decision can be recalled, the decision has a d-rationale. The use of such a precedent depends on whether the result of disputation cannot be significantly corrupted when the same arguments are applied again. Sometimes there are new arguments that can be made that would change the decision or the clarity of the decision. A different attack on a rule with d-rationale occurs when past disputation was incomplete. Not all arguments that could have been advanced were advanced in the prior disputation.

Suppose that in the case of an ambulance parked in the park while its crew was off-duty, the parking was decided to be disallowed. One side argued that the ambulance was permitted to park because it was an emergency vehicle. The adjudicator held, on the other hand, that despite being an emergency vehicle, parking was disallowed for emergency vehicles that were not prepared to respond to an emergency. Suppose the arguments were of such a form that the case could be decided on syntactic grounds. Consider a subsequent case of an off-duty ambulance that remained prepared to respond. The persuasiveness of the precedent depends on whether the decision's rationale, a d-rationale, is exposed. According to precedent, "off-duty ambulances are not allowed in the park." The opposition successfully defeats this argument by recalling that an ability to respond to emergency was crucial to the past decision. Imagine that the ambulance's ability to respond can be established in the present case. The precedent is no longer persuasive.

### **III. Formal Requirements.**

The behaviors of the various rationales are sufficiently different that some complexity is required in order properly to model them.

The c- and s-rationales for rules, and the d-rationale, for cases, are appropriately modeled by substituting less compact arguments (or suites of arguments) for the original argument. Disputation can then continue after the unpacking.

The f-rationale and r-rationale require an entirely different mechanism. When arguments using rules or cases with these rationales are attacked, the dispute is about the outcomes of other disputes. A mechanism for meta-argument is required, and the simplest adequate devices are used here.

We present first the formalism for rationales that can be treated wholly at the object level, then augment the formalism to handle the meta-level disputes. For a formal system to treat all five kinds of compilation rationales, it would have to integrate the mechanisms for unpacking and meta-argument. In the interest of clarity, this is not done here.

#### **III.1. Object-Level Disputation.**

Formally, the record of a disputation includes a sequence of moves,  $M_n = \langle m_0, m_1, \dots, m_n \rangle$ . Each move may consist of several things: a claim, an argument for the claim, perhaps an argument for the sufficiency of the pair at this point in the disputation.

A simple protocol for dispute is a two-player immediate-response dialectic (several alternate protocols are described in [Loui92]):

- i. there are two advocates, who are players, and an adjudicator who has no moves;
- ii. moves alternate between players;
- iii. each move must alter current opinion, which is determinable syntactically, and which is either pro, con, or none;
- iv. all arguments must be well-formed: i.e., members of  $\text{Args}(L, C, E)$ ;
- v. arguments do not occur repeatedly in the record: i.e., once an argument is introduced, it need not and cannot be introduced again.

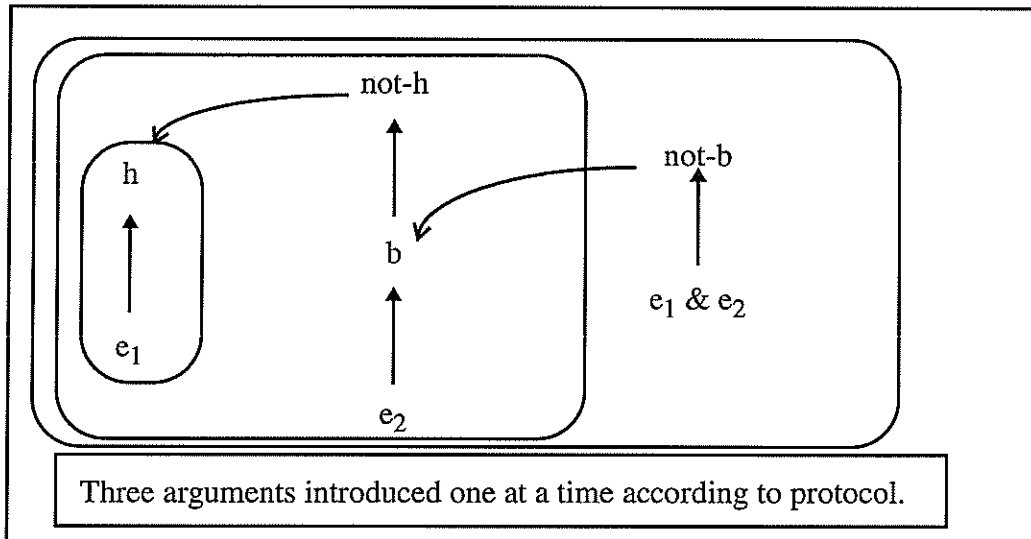
The internal structure of the argument,  $a_i$ , in each move,  $m_i$ , has been the topic of most recent research on argument (e.g., [Sartor93b], [Prakken93], [Simari-Loui92], [Loui-Norman93a]). An argument is a structure,  $\langle \{l_1, l_2, \dots\}, h \rangle$ , where each  $l_j$  is in  $\text{Rules}(L, C)$ , and  $h$  is the main claim of the argument. Other authors discuss minimality and connectedness properties at length: that is, how the  $l_j$ 's and  $h$  relate to one another.

The sufficiency of the move depends crucially on the defeat relations among arguments introduced in  $M_n$ .

As an example, the following dispute proceeds according to a reasonable dialectical protocol:

- |   |               |
|---|---------------|
| 1. pro: $\{e_1 \dashv\vdash h\}, h$   | opinion: pro  |
| 2. con: $\{e_2 \dashv\vdash b; b \dashv\vdash \text{not-}h\}, \text{not-}h$ | opinion: none |
| 3. pro: $\{e_1 \ \& \ e_2 \dashv\vdash \text{not-}b\}, \text{not-}b$ ,      | opinion: pro  |

establishing  $h$ .



### III.1.1. Preliminaries.

In procedural models, citing a rationale is a prelude to an attack. It is a necessary part of what

makes the subsequent argument relevant and sufficient. It alters the context so that the subsequent argument is a legitimate move by the player. A rationale, cited, permits  $a_{i+1}$  to respond to  $a_i$ . Without citing the rationale,  $a_{i+1}$  is not a legitimate successor of  $a_i$ .

In declarative models, there are at least two ways to regard rationales. Citing rationales may alter L or C, restating the formulation of rules and cases to which a player is willing to agree. In this way, the appropriate basis for warrant is  $\text{Args}(L', C', E)$  instead of  $\text{Args}(L, C, E)$ . The second way is to suppose that the rationale, cited, changes the relations of defeat among the arguments, defeat, the binary relation on  $\text{Args}(L, C, E)$  that is defined by the underlying system of argument. Contemporary systems regard argument  $a_1$  as defeating argument  $a_2$  on syntactic criteria, or according to an externally supplied ordering,  $\text{extord}(\text{Args}(L, C, E))$ . The latter might arise from orderings on its constituents:  $\text{extord}(L)$ ,  $\text{extord}(C)$ ,  $\text{extord}(E)$ . With rationales, the cited rationales, R, becomes a new index in the determination of defeat. What was the mapping:

$$\text{extord}(L), \text{extord}(C), \text{extord}(E) \rightarrow \text{extord}(\text{Args}(L, C, E)) \rightarrow \text{defeat};$$

becomes the mapping:

$$\text{extord}(L), \text{extord}(C), \text{extord}(E), R \rightarrow \text{extord}(\text{Args}(L, C, E)) \rightarrow \text{defeat}.$$

If R is static, then R is only important to those who study these mappings. This is why rationales do not have much importance in declarative settings. Taking R to be dynamic, as a sequence  $R_i$ , one may as well return to a procedural model. We will just assume a procedural model of argument.

Let  $R(l)$  and  $R(c)$  respectively be all citable rationales for a defeasible rule,  $l$  in L, and a precedent-setting case,  $c$  in C. The form of a rationale depends on the kind of rationale, as elaborated below. Practically, the most difficult part of arguing with rationales will be discovering and formulating the set  $R(l)$ , for a rule  $l$ , or  $R(c)$ , for a case,  $c$ . Even when an authoritative decision is explained, and even when legislative deliberation is well preserved, formal expression of rationales as R is problematic. The assumption here is the usual assumption among formalists: just as evidence, E, wording of rules, L, and documentation of cases, C, can be expressed in formal language, so too can rationales, R. To address the problem more fully would be more ambitious than what is usually attempted by formalists.

The intention is that  $R(l)$  will be a singleton set for most rules (likewise for  $R(c)$ ). There could be multiple citable rationales, and the difference could matter to the course of a particular dispute. There could also be no citable rationale.

If  $R(c)$  or  $R(l)$  is not a singleton for the relevant case or rule, then the parties can dispute the appropriate rationale. Since we suppose no information that would allow adjudication of such a dispute, the simplest reasonable rule will be adopted. If a rationale is raised for a rule or case in an argument, and there is a multiplicity of possible rationales, then it suffices that the player who makes use of the original argument can make progress under one of the many rationales. For example, suppose pro has the burden to establish  $h$ . Pro produces an argument A for  $h$ , using rule  $l$ . If  $l$  has multiple rationales in  $R(l)$ , let them be named  $\text{rat}_1(l)$ ,  $\text{rat}_2(l)$ , etc.; if they are of different types, name them  $c\text{-rat}_1(l)$ ,  $d\text{-rat}_2(l)$ , etc. In the suggested protocol it suffices for pro if  $h$  can be established after unpacking pro's choice of the permitted rationales.

Call the aforementioned the *User's Prerogative Assumption*. Since the procedural rules that

implement this assumption merely complicate the presentation in an uninteresting way, we shall assume that all  $R(c)$  and  $R(l)$  are always singletons for a rule  $l$  or case  $c$ ; this is the simplest way of making the assumption of user's prerogative.

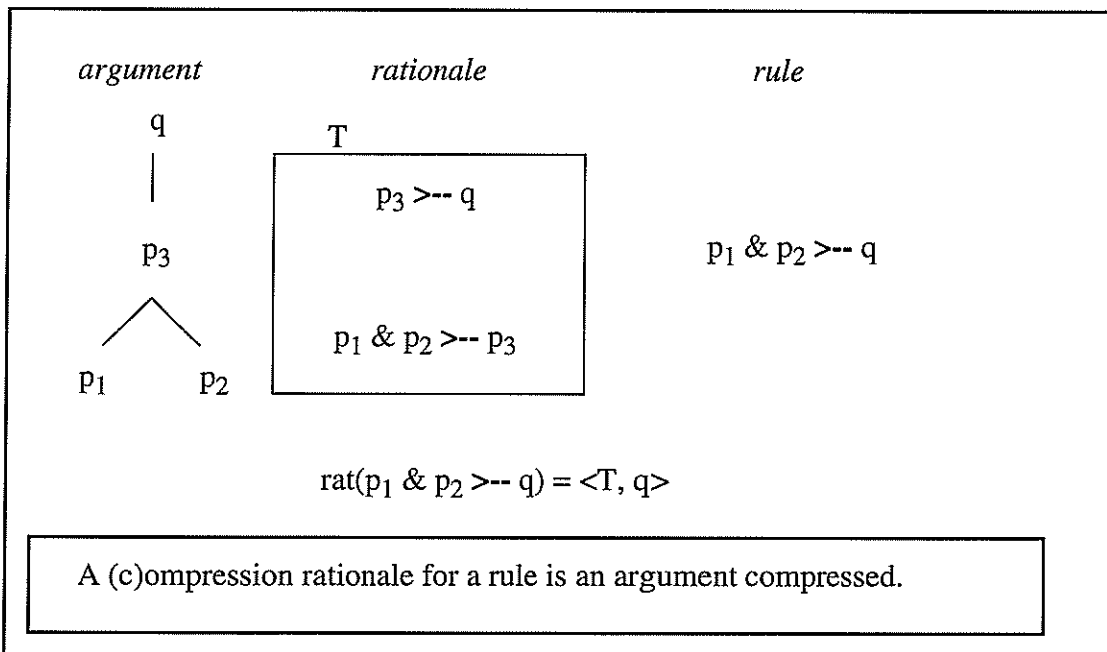
### III.1.2. Structures.

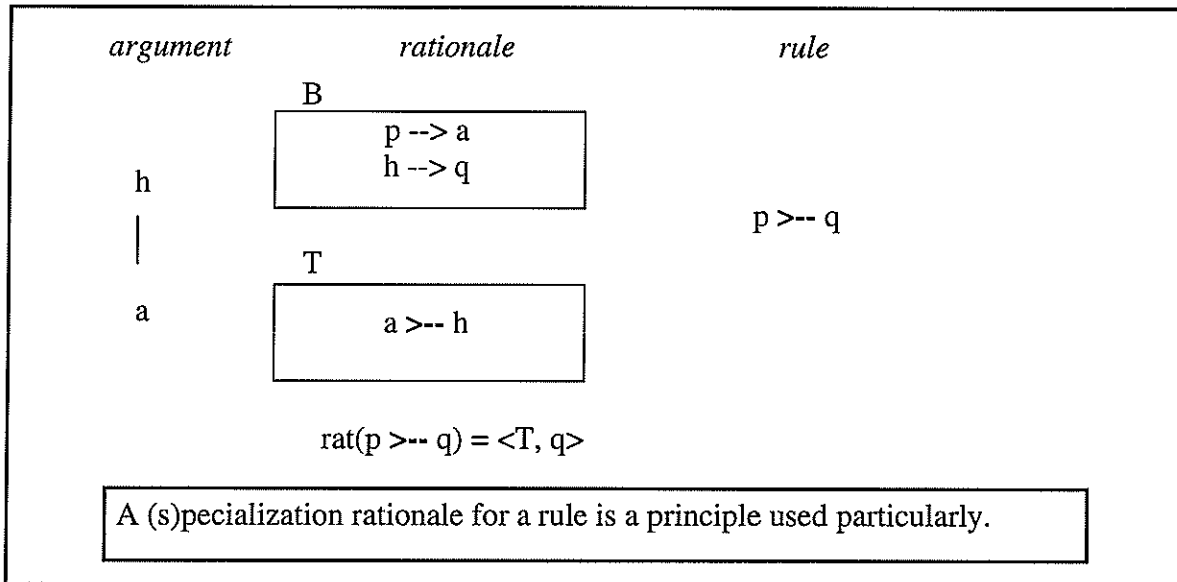
For rules of the form “ $p$  is defeasible reason for  $q$ ”, rationales take the following forms.

A  $c$ -rationale is typically an argument, with premises  $p$  and conclusion  $q$ . There may be a background,  $B$ , against which this argument was made. The intention in adopting the rule is that usually  $B$  will be present in contexts in which the rule is applied. A  $c$ -rationale may thus be an argument from  $p$  and  $B$  to  $q$ . Such an argument is as above, a 2-tuple:  $c\text{-rat}(l) = \langle T, h \rangle$ , of rules  $T$  and conclusion  $h$ . Roughly speaking, this argument would be in the set of well-formed arguments  $\text{Args}(T, \{\}, \{p\} \cup B)$ .

An  $s$ -rationale is formally identical to a  $c$ -rationale except that the argument takes a particular form. In a  $c$ -rationale, several rules are used in the argument, which is compressed.  $T$  is not a singleton set. An  $s$ -rationale uses a single guiding rule. It uses supplementary rules in  $L$  and/or  $B$ .

A  $d$ -rationale of a case  $c$  is a set of arguments,  $d\text{-rat}(c)$ , that has a certain structure. These arguments are typically not from  $\text{Args}(L, C, E)$  but are from a different set,  $\text{Args}(L', C', E')$ . The applicability of the case to the current situation depends crucially on whether the arguments in  $d\text{-rat}(c)$  are in both of these sets, not just in  $\text{Args}(L', C', E')$ . For a  $d$ -rationale,  $d\text{-rat}(c)$  is a set of arguments that warrants the decision of the case.

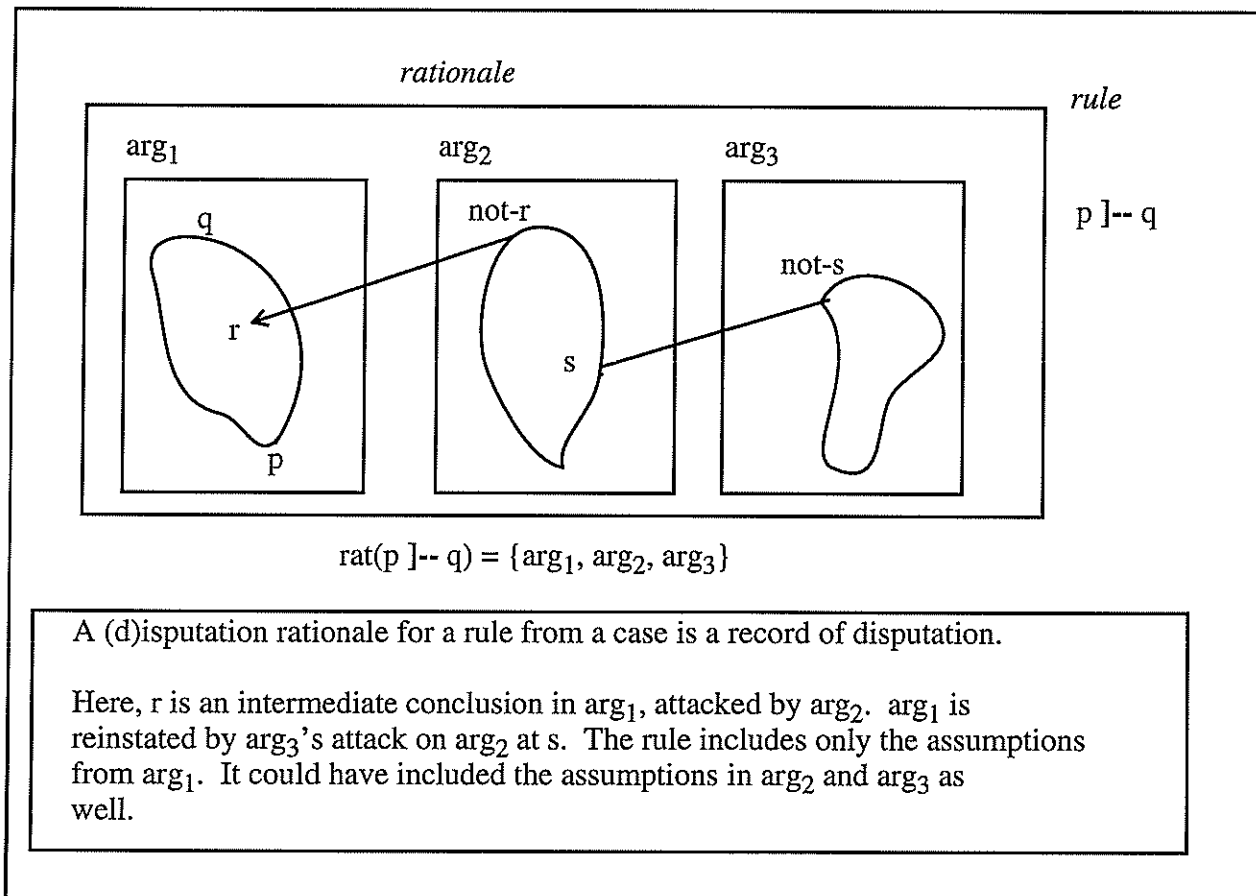




### III.1.3. Attacks

Let  $\text{ArgRec}_n$  be the recognized arguments at a stage  $n$  in the sequence of moves. In a simple disputation where moves do not refer to rationales and all arguments are well-formed members of  $\text{Args}(L, C, E)$ ,  $\text{ArgRec}_n$  is simply the union of all arguments introduced in all moves,  $m_1, \dots, m_n$ . In disputations where well-formedness of an argument can be called into question (which includes disputations that allow rationale-based attacks on arguments),  $\text{ArgRec}_n$  expands and contracts as the disputation proceeds.

$\text{ArgRec}_n$  includes both defeated and undefeated arguments. It is not the “arguments in force” concept that Vreeswijk defines [Vreeswijk93], which is useful in analyzing warranted conclusions. It is preliminary to determining which arguments are in force. An argument that was introduced in the dialogue but has subsequently been excluded from  $\text{ArgRec}$  cannot even be considered for its properties of defeat relative to the other arguments.



ArgRec is defined inductively: an argument in  $m_0$  is in  $\text{ArgRec}_0$ . If  $m_i$  does not contain a statement of rationale, then for an argument  $a_i$  in  $m_i$ ,  $\text{ArgRec}_i$  contains  $\text{ArgRec}_{(i-1)} \cup a_i$ . If  $m_i$  contains a statement of a rationale, then this is either a rationale-based attack on prior argument, or else a reinstatement of an argument that has suffered a rationale-based attack.

A statement of a rationale in a move requires the naming of the argument that is attacked or defended. It also requires the naming of the rule or case, the rationale of which is at issue. Let  $@_i$  indicate that  $i$  is a move that uses a rationale. For such a move, let  $a_i$  be the argument concerned, and let  $l_i$  or  $c_i$  be the rule or case, as appropriate, in question. Let  $r_i$  be the rationale, whether it is a c-, s-, or d-rationale. Let  $\text{player}_i$  be the player who moved in move  $i$ .

Trivial well-formedness requires that the attacked argument has occurred, that the rule occurs in the argument, and that the rationale is recognized:

$a_i$  must be in  $\text{ArgRec}_k$  for some  $k < i$  ( $k$  need not be  $i-1$  since the defense against rationale-based attacks must be possible);

$l_i$  or  $c_i$  must be a part of  $a_i$ ;



$r_i$  must be in  $R(l_i)$ , or  $R(c_i)$ .

$r_i$  and  $a_i$  may have occurred together in a previous move,  $j < i$ , where both  $@_i$  and  $@_j$ . This seems to permit repetition of moves, but  $m_i$  still might not be the same as  $m_j$ . Repetition will be prevented by a requirement that moves be effective.

The same rule or case can repeatedly be the target of rationale-based attacks if it occurs in distinct arguments. This is simpler than calling all arguments based on a rule or case into question at once. Call this the *One-At-A-Time Protocol*. Also note that both players might be relying on the same argument,  $a_i$ . Let the first player who attacks a rule,  $l$ , in an argument,  $a$ , be  $\text{attacker}(l, a)$ ; the opposing player is  $\text{user}(l, a)$ ; likewise for attacking rules extracted from cases.

Consider that  $@_i$  and that  $i$  is a first attack. To be a first attack,  $a_i$  and whichever of  $l_i$  or  $c_i$  is relevant, do not already occur for some  $j < i$ . Note  $\text{attacker}(l_i, a_i) = \text{player}_i$ .

Apparently, the immediate effect of the attack is to remove  $a_i$  from  $\text{ArgRec}_i$ . That is, apparently,  $\text{ArgRec}_i = \text{ArgRec}_{(i-1)} - \{a_i\}$ . But the move might modify the form of the argument removing  $a_i$  and substituting a revised argument in its stead. So apparently,  $\text{ArgRec}_i = \text{ArgRec}_{(i-1)} - \{a_i\} \cup \{\text{argrev}\}$ , where  $\text{argrev}$  is a modified form of  $a_i$ .

For a  $c$ -rationale or  $s$ -rationale, which is an argument  $\langle T, h \rangle$ ,  $\text{argrev}$  is just the argument that uses  $\langle T, h \rangle$  as a subargument instead of the rule,  $l_i$ , that is attacked. So if  $\langle T_0, h_0 \rangle = a_i$ , then  $\text{argrev}$  is the argument in which a subargument has been substituted for the compressed rule:  $\langle T_0 \cup T - \{l_i\}, h_0 \rangle$ . This substitution will be denoted:  $\text{subst}(\langle T_0, h_0 \rangle, l_i, T)$  if indeed it is an argument (satisfies minimality and consistency conditions). Otherwise, the statement of the rationale destroys  $a_i$  and there is no revised argument added in its place.

This completes the inductive definition of  $\text{ArgRec}_i$ .

Finally, having defined  $\text{ArgRec}_i$  in terms of  $m_i$  and  $\text{ArgRec}_{(i-1)}$ , we require the move  $m_i$  to be effective: it must alter current opinion so that  $\text{player}_i$ 's position is improved. When  $\text{argrev}$  is part of  $m_i$ ,  $m_i$  must sometimes contain a response to the revised argument as well.

For a  $d$ -rationale,  $r_i$  is a set of arguments.  $m_i$  must contain a new argument,  $\text{attack}_i$ , which attacks  $\text{argrev}$  and changes the current opinion,  $\text{dec}_i$ . In a  $d$ -rationale, the decision of the case,  $\text{dec}(c_i)$ , is warranted among the arguments recorded for the case,  $r_i$ . That is, enough arguments were recorded to support the decision outright. So  $\text{attack}_i$  must be such that  $r_i \cup \{\text{attack}_i\}$ , taken as a set of arguments, would not have warranted the old decision outright,  $\text{dec}(c_i)$ .

Responses to these simple kinds of rationale-based attacks are just moves that augment  $\text{ArgRec}$ .

For example, for a  $d$ -rationale, all of the arguments recorded of the past case have been entered into  $\text{ArgRec}$ . Among them was an argument,  $\text{argw}$ , that warranted the decision of the case. Also included is the opponent's  $\text{argrev}$ , which has somehow changed the status of  $\text{argw}$ . To respond substantively, simply give an argument that restores  $\text{argw}$ 's ability to support its conclusion again. Or simply retreat from the attacked precedent-based argument and provide a new, unrelated argument to support the old decision.

### III.1.4. Simple Symbolic Examples.

*Example. A dialogue without a rationale.*

1. pro:  $\text{arg}_1 = \langle \{ b \text{>--} a \}, b! \text{ } \therefore \text{ } a \rangle$

$\text{ArgRec}_1 = \{ \text{arg}_1 \}; \text{dec}_1 = \text{pro.}$

2. con:  $\text{arg}_2 = \langle \{ c \text{>--} d, d \ \& \ b \text{>--} \text{not-}a \}, c! b! \text{ } \therefore \text{ } \text{not-}a \rangle$

$\text{ArgRec}_2 = \{ \text{arg}_1, \text{arg}_2 \}; \text{dec}_2 = \text{none.}$

3. pro:  $\text{arg}_3 = \langle \{ c \ \& \ e \text{>--} \text{not-}d \}, c! e! \text{ } \therefore \text{ } \text{not-}d \rangle$

$\text{ArgRec}_3 = \{ \text{arg}_1, \text{arg}_2, \text{arg}_3 \}; \text{dec}_3 = \text{pro.}$

Here, arguments are 2-tuples (a set of rules paired with a conclusion). We have embellished the second argument. Instead of just giving the conclusion of the argument, we give the evidence on which the argument rests (marked with an exclamation point) and write “ $\therefore$ ” to separate the conclusion. Rules are given in their sentential form,  $a \text{>--} b$  instead of using the pair  $\langle a, b \rangle$ , and quotes are suppressed. Note that we use “=” to give a name to a structure. This is not to be confused with an assertion.

At move 4, it is con’s turn to move and opinion favors pro:  $\text{arg}_3$  defeats  $\text{arg}_2$ , thus reinstating  $\text{arg}_1$ . If con cannot move at this point, pro wins.

*Example. A c-rationale.*

1. pro:  $\text{arg}_1 = \langle \{ b \text{>--} a \}, b! \text{ } \therefore \text{ } a \rangle$

$\text{ArgRec}_1 = \{ \text{arg}_1 \}; \text{dec}_1 = \text{pro.}$

2. con:

2.1.  $\text{c-rat}_1(\{ b \text{>--} a \}) = \langle \{ b \text{>--} c, c \text{>--} a \}, b! \text{ } \therefore \text{ } a \rangle$

2.2.  $\text{Attack} \langle b \text{>--} a, \text{arg}_1 \rangle$

2.3.  $\text{arg}_2 = \langle \{ b \ \& \ d \text{>--} \text{not-}c \}, b! d! \text{ } \therefore \text{ } \text{not-}c \rangle$

Con states a rationale for a rule, which presumably is the legitimate member of  $R(\{ b \text{>--} a \})$ . Con must state the argument being attacked. The effect is to replace the compressed argument with the uncompressed, but this is still not a sufficient response for con. So con provides  $\text{arg}_2$ , an attack on the uncompressed argument, the *raison d’etre* for stating the rule’s rationale. @<sub>2</sub>. Note  $\text{pro} = \text{user}(b \text{>--} a, \text{arg}_1)$ .  $\text{con} = \text{attacker}(b \text{>--} a, \text{arg}_1)$ .

$\text{ArgRec}_2 = \{ \text{arg}_1 \text{rev}, \text{arg}_2 \}; \text{dec}_2 = \text{none.}$

where

$\text{arg}_1\text{rev} = \text{subst}(\text{arg}_1, b \succ\text{--} a, \text{c-rat}_1(b \succ\text{--} a))$

3. pro:  $\text{arg}_3 = \langle \{ b \& d \& e \succ\text{--} c, c \succ\text{--} a \}, b! d! e! \therefore a \rangle$

Pro provides a new argument for a,  $\text{arg}_3$ , that is undefeated in  $\text{ArgRec}_3$ . The new argument for c is specific enough to defeat  $\text{arg}_2$ , which would reinstate  $\text{arg}_1$ , but  $\text{arg}_1$  has been flushed from the  $\text{ArgRec}$  in favor of  $\text{arg}_1\text{rev}$ .

$\text{ArgRec}_3 = \text{ArgRec}_2 \cup \{ \text{arg}_3 \}; \text{dec}_3 = \text{pro}.$

*Example. An s-rationale.*

1. pro:  $\text{arg}_1 = \langle \{ b \succ\text{--} a \}, b! \therefore a \rangle$

$\text{ArgRec}_1 = \{ \text{arg}_1 \}; \text{dec}_1 = \text{pro}.$

2. con:

2.1.  $\text{s-rat}_1(b \succ\text{--} a) = \langle \{ d \succ\text{--} e \}, b! (d \vee \text{not-}b)! (a \vee \text{not-}e)! \therefore a \rangle$

2.2. Attack  $\langle b \succ\text{--} a, \text{arg}_1 \rangle$

2.3.  $\text{arg}_2 = \langle \{ d \& f \succ\text{--} \text{not-}e \}, d! f! \therefore \text{not-}c \rangle$

$\text{ArgRec}_2 = \{ \text{arg}_1\text{rev}, \text{arg}_2 \}; \text{dec}_2 = \text{none}.$

where

$\text{arg}_1\text{rev} = \text{subst}(\text{arg}_1, b \succ\text{--} a, \text{s-rat}_1(b \succ\text{--} a))$

Here,  $\text{arg}_2$  defeats  $\text{arg}_1\text{rev}$  because it is more specific. But  $\text{arg}_2$  does not defeat  $\text{arg}_1$  on specificity.

3. pro:  $\text{arg}_3 = \langle \{ d \& f \& g \succ\text{--} e \}, b! (d \vee \text{not-}b)! f! g! (a \vee \text{not-}e)! \therefore a \rangle$

$\text{ArgRec}_3 = \text{ArgRec}_2 \cup \{ \text{arg}_3 \}; \text{dec}_3 = \text{pro}.$

Notice how similar this example is to the example of a c-rationale and recovery from a c-rationale-based attack.

*Example. A d-rationale.*

1. pro:  $\text{arg}_1 = \langle \{ b \& c ]\text{--} a \}, b! c! \therefore a \rangle$

$\text{ArgRec}_1 = \{ \text{arg}_1 \}; \text{dec}_1 = \text{pro}.$

Here, ]-- is used instead of >-- to indicate that this rule is extracted from a case.

2. con:

2.1.  $\text{d-rat}_1(b \& c ]\text{--} a) = \{ \text{da}_1, \text{da}_2, \text{da}_3 \}$

where

$$da_1 = \langle \{ b \>-- a \}, b! \therefore a \rangle,$$
$$da_2 = \langle \{ c \>-- d, d \>-- \text{not-}a \}, c! \therefore \text{not-}a \rangle$$
$$da_3 = \langle \{ b \>-- f, f \& c \>-- \text{not-}d \}, b! c! \therefore \text{not-}d \rangle$$

In  $d\text{-rat}_1(b \& c ]-- a)$ , the argument for not-d ( $da_3$ ) defeats the argument for not-a ( $da_2$ ), thus reinstating the argument for a ( $da_1$ ). Con attacks, now, by showing that the argument for not-d is defeated in the present context, where more evidence is available.

2.2. Attack  $\langle b \& c ]-- a, arg_1 \rangle$

2.3.  $arg_2 = \langle \{ b \& g \>-- \text{not-}f \}, b! g! \therefore \text{not-}f \rangle$

$ArgRec_2 = \{ da_1, da_2, da_3, arg_2 \}; dec_2 = \text{none}.$

3. pro:  $arg_3 = \langle \{ g \& c \>-- \text{not-}d \}, g! c! \therefore \text{not-}d \rangle$

$ArgRec_3 = ArgRec_2 \cup \{ arg_3 \}; dec_3 = \text{pro}.$

This move suffices for pro because  $arg_3$  now defeats  $da_2$ , thus reinstating  $da_1$ . Current opinion is for pro, to decide a, because  $da_1$  is now in the  $ArgRec$ . Pro is lucky; the same new condition, g, that allows a distinction from the precedent also allows a new argument that replaces essentially what was lost.

### III.2. Meta-Level Disputation.

For both the r-rationale and the f-rationale, disputation about disputation is required. This requires meta-claims, i.e., claims about claims, and meta-arguments about meta-claims. There are serious semantic difficulties for the formalist who takes defeasible reasoning to the meta-level.

Technically, the semantic ascent to the meta-language is usually preferable to the demotion of the reason relation “>--” from a 2-place relation in the metalanguage to an iterable 2-connective in the language. As Quine asks, “How do you hope to establish semantic connections between object level and meta-level?” [Quine90] A declarative approach to meta-disputation would indeed require serious study. A procedural approach would be one way to side-step the difficulty. Earlier writers who have reached this point, notably Toulmin and Rescher, have ignored this technical problem. Pollock, meanwhile, can be viewed as a first step in the procedural approach.

Here, we assume that a system of declarative argument and meta-argument exists. The system determines a few things: whether meta-arguments are undefeated and supporting, hence, whether their meta-claims warranted; whether the arguments that make use of the meta-claims are undefeated and supporting, hence, whether their object-level claims are warranted. This assumption is made for two reasons: (a) to define the envisioned procedural system would be a separate study not crucial to the picture of rationales that is drawn here; and (b) the use of meta-argument here provides requirements for those who would attempt to define such a system.

### III.2.1. Preliminaries.

For an r-rationale, the crucial assertion is an assertion about the relation among arguments, a meta-assertion. Arguments that were maximal, that is, undefeated, in the cited case should continue to be maximal in the current case in order for the past decision to have the proper influence.

For an f-rationale, two kinds of meta-assertion are crucial. They are that a case should have a certain decision, and that a rule fits a set of cases.

There are a few kinds of attacks on rules with f-rationales that we can model and a few that we cannot model. Even for the few that we do model, we must introduce negated assertions of a sentence being a reason for another sentence, e.g.  $\text{not}(b \text{ >-- } a)$ .

An r-rationale is similar to a d-rationale. It is at least a set of recorded arguments from the prior dispute. For an r-rationale of a case  $c$ ,  $\text{r-rat}(c)$  is a set of undefeated arguments that interfere with each other. It is this interfering set of arguments that the decision of the case resolves.  $\text{r-rat}(c)$  also contains the implicit argument that when those arguments are maximal, interfering, and unresolved, a certain decision is appropriate. An assumption here is that the maximality of a set of arguments is assertible without dispute (or what is nearly the same, that maximality can be determined easily). If the defeat relations among arguments are clear, then this assumption is reasonable.

An argument so far has been a set of defeasible rules from  $L$ , and a conclusion derivable from those rules using the accepted (incontrovertible) evidential claims. Henceforth, the set of rules can contain rules from  $L$  and meta-rules. These meta-rules relate a specific kind of meta-claim, that a certain set of arguments is maximal or not, to a different kind of meta-claim, a defeasible rule. For example, instead of allowing just

$$a \text{ >-- } b$$

we allow as well

$$\text{maximal}(S) \text{ >-- } (a \text{ >-- } b).$$

The first occurrence of “>--” above is a relation in the meta-meta-language, and the second occurrence of “>--” is the relation in the meta-language used throughout the preceding sections. Derivations are now allowed to make use of an argument’s rules and meta-rules (arguments may use both kinds of “>--”). The mathematical demands on the theory of argument imposed here are serious and not to be ignored. Fortunately, implementations have shown that progress can be made with systems defined in this way, e.g., [Sartor93b], despite the mathematical and philosophical conundrums.

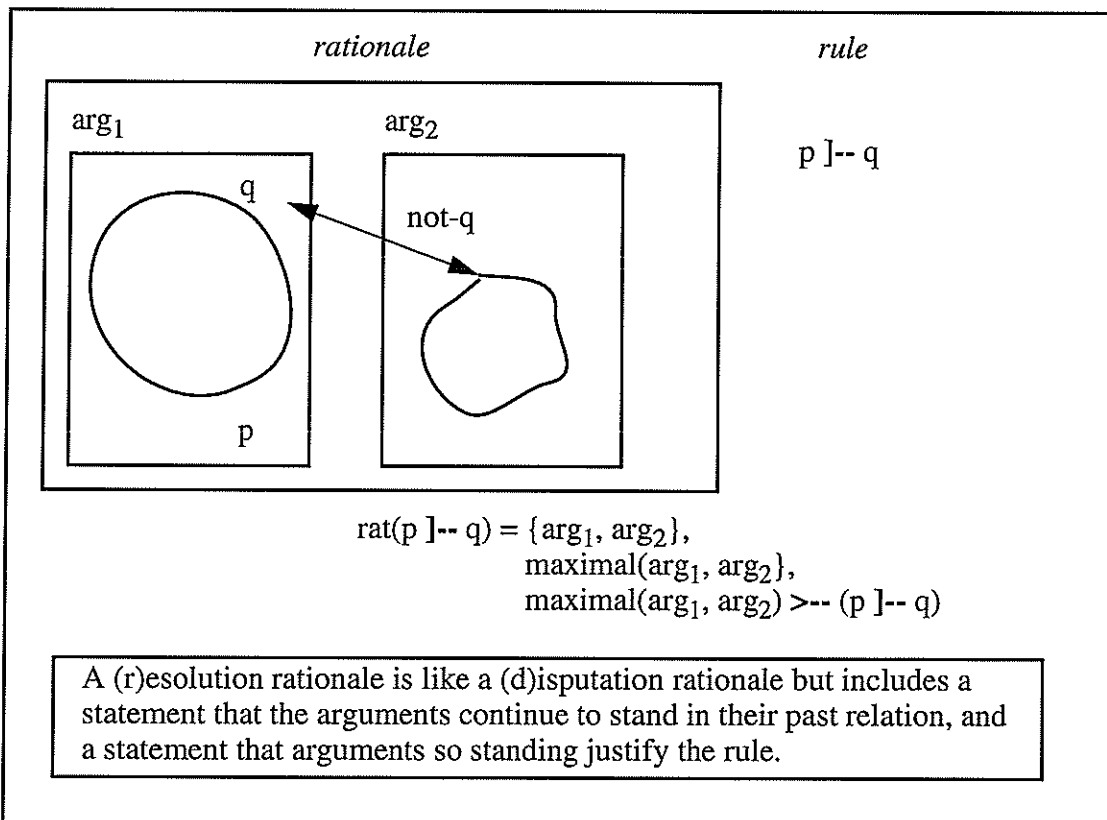
$\text{Maximal}(S)$  asserts of a set of arguments,  $S = \{A_1, A_2, \dots, A_n\}$  that each of those arguments contained in  $S$  can be made in the current setting and is undefeated. If  $S$  is such that  $\text{maximal}(S)$ , then  $S$  must be contained in  $\text{Args}_i(L, C, E)$ . The predicate,  $\text{maximal}$ , is thus dependent on all of the indexes:  $i$ ,  $L$ ,  $C$ , and  $E$ . Since  $i$  is the most important index here, write  $\text{maximal}_i(S)$  when it is important to distinguish the stage at which maximality is claimed.

This is not the final formulation for r-rationales. As HYPO makes clear, a precedent can be used

in the presence of additional arguments pro. If a case decides that argument  $A_1$  for  $h$  outweighs argument  $A_2$  for  $\text{not-}h$ , then if  $A_1$ ,  $A_2$ , and an additional argument for  $h$ ,  $A_3$ , can be made in a new setting, the decision of the prior case is still relevant. The precedent can also be used in a context where there are fewer arguments con than those that figured in the past resolution. Hence, it need not be required that exactly those maximal arguments in the past case be maximal in the present case.

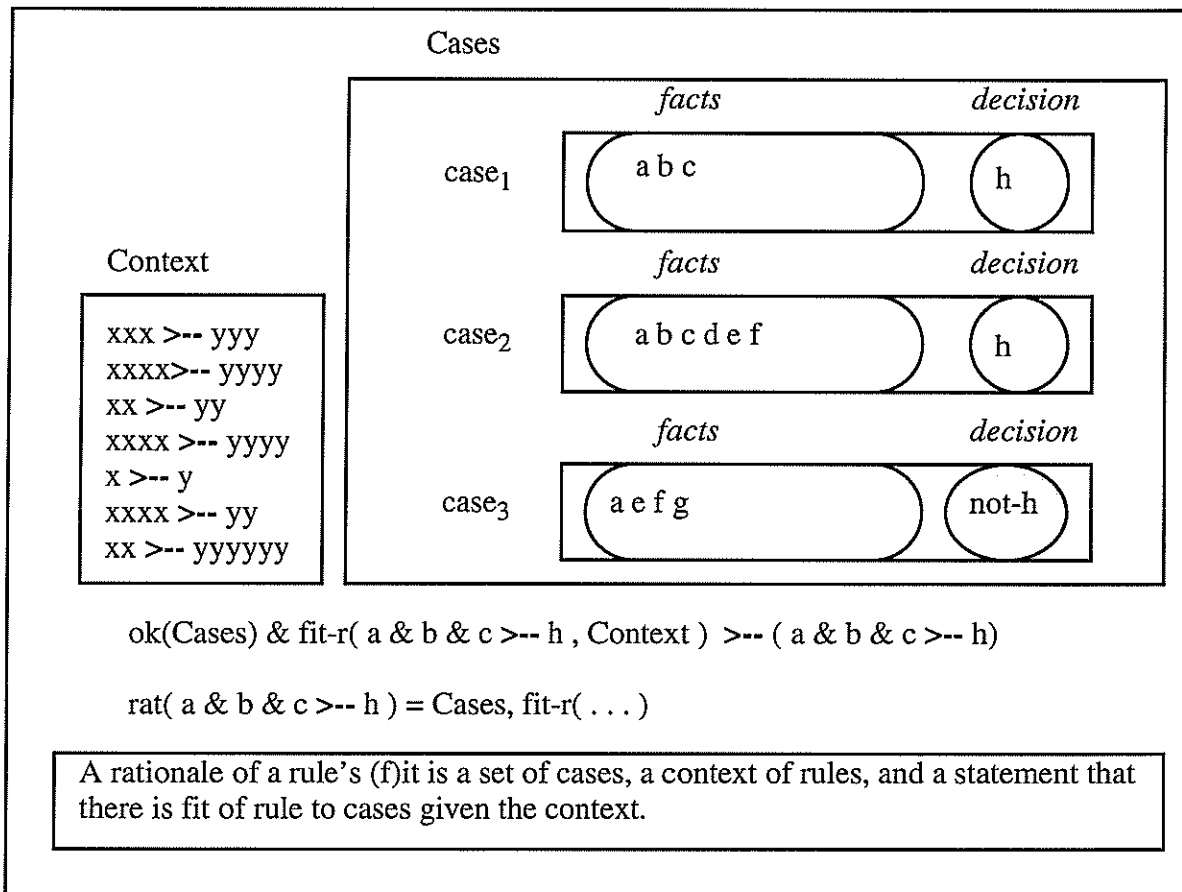
Instead, of the predicate  $\text{maximal}_i(S)$ , suppose there are two predicates:  $\text{minpro-maximal}_i(S_1)$  and  $\text{maxcon-maximal}_i(S_2)$ . The arguments in each of the sets  $S_1$  and  $S_2$  were weighed in the prior decision, and arguments in  $S_1$  taken together were more persuasive than those in  $S_2$ . If at stage  $i$ , at least the arguments in  $S_1$  are maximal, and if no more than the arguments in  $S_2$  for con are maximal, then there is defeasible reason for using the case (or more precisely, the rule extracted from the case). If the maximality of the arguments in  $S_1$  cannot be maintained at stage  $i$ , then the precedent cannot be used. If there are additional arguments con, which were not weighed previously, then the case is again inapplicable.

A formal account of determining when an argument should be considered additional pro or additional con is owed. As in HYPO, this determination is trivial if the decision of the case is  $h$  and all arguments in  $S_1$  and  $S_2$  are for  $h$  and  $\text{not-}h$ , respectively. Determining whether an argument is pro or con can be inductively defined in an obvious way, and will not be done here. There will be arguments that support both pro and con arguments, and these most naturally contribute to  $\text{minpro-maximal}_i$  and be excluded from  $\text{maxcon-maximal}_i$ .



An f-rationale is a set of cases, not all of which might be from C. Let  $HC \cup C$  be the actual cases augmented by hypothetical cases. HC are cases that did not actually occur, but are cases upon which there is agreement of the hypothetical decisions given the hypothetical facts. Assume that the assertion that a case and its decision are agreeable to all parties is incontrovertible. Write  $ok(case_1)$  if  $case_1$  is so agreeable. Once again, we do not seek to model every kind of dispute. Here, we choose not to model disputes over whether a case should have a certain outcome. Some of these disputes could be modeled easily, some would require a complicated recursion, and some are not obviously subject to reason at all.

Instead of introducing a language in which cases are described, suppose simply that there is a relation  $fit-s(l, c)$ , which holds when a defeasible rule fits a case c. For example, the rule  $l_1 =$  "if  $f_1$  and  $f_2$ , then h" fits those cases wherein  $f_1, f_2$ , and h hold. It also fits those cases wherein h does not hold, but at least one of  $f_1$  or  $f_2$  also does not hold. The rule does not fit a case wherein  $f_1, f_2$ , and not-h are all true. This is fit *simpliciter*. A more robust concept of a defeasible rule's fit to a case takes account of context.



In the context of a set of other defeasible rules, LCont, l might fit a case c even if it does not fit *simpliciter*. This is because LCont may provide through more specific rules an explanation of why c does not conform to l, why it is not the case that  $fit-s(l, c)$ . So a rule may fit, *robustus*,  $fit-r(l, c, LCont)$  when  $fit-s(l, c)$  or when there is a more specific rule  $l'$ , in LCont which does fit c. The

whole apparatus of arguments which permits chaining of rules could be applied here to explain why  $l$  does not fit  $c$ , *simpliciter*, but is allowed not to fit by a superior chain of rules in LCont. Under the criteria of syntactic superiority advanced in [Simari-Loui92], [Prakken93], and [Loui-Norman93a], no chain of rules in LCont will be superior to  $l$  unless a single rule in LCont is superior to  $l$ . But under other systems, a recursion of disputation is possible here, and woe be to the formalizer who seeks to model this recursion.

Arguments of fit based on a larger number of cases are considered better arguments. Likewise, arguments of fit based on a larger context are better arguments. Fit is assumed incontrovertible.

### III.2.2. Structures.

Formally, an  $r$ -rationale is a set of arguments that occurred in the case, a true assertion that all of those arguments are maximal with respect to one another (none is defeated), and a meta-reason that connects the maximality of certain arguments to the rule putatively extracted from the case. The arguments are partitioned into minpro and maxcon, as discussed above. Technically all that are required are the two sets of arguments,  $S_1$  and  $S_2$ , that occurred in the resolution of the case. The rest can be recovered from context. Making the other parts explicit adds clarity.  $r\text{-rat}(c) = \langle$

1.  $S = S_1 \cup S_2$ ,
  2. minpro-maximal( $S_1$ ), maxcon-maximal( $S_2$ ),
  3. minpro-maximal( $S_1$ ) & maxcon-maximal( $S_2$ )  $\>-- l$
- $\rangle$ .

The claim of maximality must be true, and the arguments in  $S$  must be properly relevant to  $c$ . To define proper relevance, consider the rule extracted from the case to be:  $l = l_1 \>-- l_2$ . There must be an argument in  $S$  for  $l_2$ . We do not insist that all arguments in  $S$  be possible with just  $l_1$  as evidence. Nor must there be some argument in  $S$  that requires all of  $l_1$  to be made. Much more weakly,  $l_1$  must be relevant to  $S$ ; it should contain no superfluous literals.

An  $f$ -rationale is a set of cases together with a claim that the rule fits (*robustus*) that set of cases with respect to a context.  $\langle HC \cup C, LCont \rangle$  is the  $f$ -rationale for a rule  $l$ , if and only if for every  $c$  in  $HC \cup C$ ,  $\text{fit-}r(l, c, LCont)$ . LCont must be a subset of  $L$ .

### III.2.3. Attacks.

To attack an argument using a rule with  $r$ -rationale, cite the rationale and name the argument and rule attacked. Then substitute a revised argument into  $\text{ArgRec}_i$  that makes explicit the connection between the maximal arguments resolved by the case and the rule extracted therefrom. The modified argument is just like the earlier argument except that a meta-rule has been put in place of the object-level rule. The meta-rule makes explicit the origin of the object-level rule. This allows subsequent attack of the statement that  $S$  is maximal.

To attack the statement minpro-maximal( $S$ ), provide a new argument or set of arguments,  $S'$ , that



are arguments that can be made in the current case. Then claim  $\text{not}(\text{minpro-maximal}(S))$  as a result of the presumed system of determining defeat among arguments. If  $S'$  contains an argument that defeats an argument in  $S$ , for example, then not every member of  $S$  will be maximal (some member will now be defeated). Similarly, to attack  $\text{maxcon-maximal}(S)$ , provide a new argument or set of arguments,  $S'$  that can be made in the current case. Then claim  $\text{not}(\text{maxcon-maximal}(S))$  because  $S'$  contains a new maximal argument con.

To attack an argument using an f-rationale, state the f-rationale, then alter either the cases or the context with respect to which fit is claimed. We have assumed that claims of fit are incontrovertible, since we are providing no theory of how this is determined. Adding a case that the rule does not fit is the usual attack.

### III.2.4. Simple Symbolic Examples.

*Example. An r-rationale with maxcon-maximal attacked.*

1. pro:  $\text{arg}_1 = \langle \{ b \ \& \ c \} \text{--} a \rangle, b! \ c! \ \therefore \ a \rangle$

$\text{ArgRec}_1 = \{ \text{arg}_1 \}; \text{dec}_1 = \text{pro}.$

2. con:

- 2.1.  $\text{r-rat}_1(\{ b \ \& \ c \} \text{--} a) =$   
 $\langle \{ da_1, da_2 \}, \text{minpro-max}(\{ da_1 \}) \ \& \ \text{maxcon-max}(\{ da_2 \}),$   
 $\text{minpro-max}(\{ da_1 \}) \ \& \ \text{maxcon-max}(\{ da_2 \}) \text{--} ( b \ \& \ c \} \text{--} a) \rangle$

where

- $$da_1 = \langle \{ b \} \text{--} a \rangle, b! \ \therefore \ a \rangle,$$
- $$da_2 = \langle \{ c \} \text{--} d, d \text{--} \text{not-}a \rangle, c! \ \therefore \ \text{not-}a \rangle$$

- 2.2. Attack  $\langle b \ \& \ c \} \text{--} a, \text{arg}_1 \rangle$

- 2.3.  $\text{arg}_1\text{rev} = \langle$   
 $\{ \text{minpro-max}(\{ da_1 \}) \ \& \ \text{maxcon-max}(\{ da_2 \}) \text{--} ( b \ \& \ c \} \text{--} a) \},$   
 $b! \ c! \ \text{minpro-max}(\{ da_1 \}) \ \& \ \text{maxcon-max}(\{ da_2 \})! \ \therefore \ a \rangle$

- 2.4.  $\text{arg}_2 = \langle \{ e \} \text{--} f, c \ \& \ f \text{--} \text{not-}d \rangle, e! \ c! \ f! \ \therefore \ \text{not-}d \rangle$

- 2.5.  $\text{not}(\text{maxcon-max}(\{ da_2 \}))$

$\text{ArgRec}_2 = \{ \text{arg}_1\text{rev}, da_1, da_2, \text{arg}_2 \}; \text{dec}_2 = \text{none}.$

Note that among  $da_1$  and  $da_2$ , neither is defeating. The opinion in this case is significant because it creates a preference that cannot be found in  $\text{extord}(\text{Args}(L, \{ \}, E))$  or in defeat, that is, among the syntactically determinable or externally supplied orderings among arguments. The adjudicator chose between competing arguments that could not be decided purely on form. Con revises pro's argument  $\text{arg}_1$  by producing  $\text{arg}_1\text{rev}$  which exposes the meta-reasoning. Then  $\text{arg}_2$  is produced, which is an additional argument con, and is maximal. So 2.5 is assertible,  $\text{not}(\text{maxcon-max}(\{ da_2 \}))$ , which is one of the requirements for using the precedent.

*Example. An r-rationale with minpro-maximal attacked.*

1. pro:  $\text{arg}_1 = \langle \{ b \ \& \ c \ ]\!-\!-\ a \}, b! \ c! \ \therefore \ a \rangle$

$\text{ArgRec}_1 = \{ \text{arg}_1 \}; \text{dec}_1 = \text{pro}.$

2. con:

2.1.  $\text{r-rat}_1( b \ \& \ c \ ]\!-\!-\ a ) =$   
 $\langle \{ \text{da}_1, \text{da}_2 \}, \text{minpro-max}(\{ \text{da}_1 \}) \ \& \ \text{maxcon-max}(\{ \text{da}_2 \}),$   
 $\text{minpro-max}(\{ \text{da}_1 \}) \ \& \ \text{maxcon-max}(\{ \text{da}_2 \}) \ ]\!-\!-\ ( b \ \& \ c \ ]\!-\!-\ a ) \rangle$

where

$\text{da}_1 = \langle \{ b \ ]\!-\!-\ g, g \ ]\!-\!-\ a \}, b! \ \therefore \ a \rangle,$

$\text{da}_2 = \langle \{ c \ ]\!-\!-\ d, d \ ]\!-\!-\ \text{not-a} \}, c! \ \therefore \ \text{not-a} \rangle$

2.2. Attack  $\langle b \ \& \ c \ ]\!-\!-\ a, \text{arg}_1 \rangle$

2.3.  $\text{arg}_1\text{rev} = \langle$   
 $\{ \text{minpro-max}(\{ \text{da}_1 \}) \ \& \ \text{maxcon-max}(\{ \text{da}_2 \}) \ ]\!-\!-\ ( b \ \& \ c \ ]\!-\!-\ a ) \},$   
 $b! \ c! \ \text{minpro-max}(\{ \text{da}_1 \}) \ \& \ \text{maxcon-max}(\{ \text{da}_2 \})! \ \therefore \ a \rangle$

2.4.  $\text{arg}_2 = \langle \{ b \ \& \ e \ ]\!-\!-\ \text{not-g} \}, b! \ e! \ \therefore \ \text{not-g} \rangle$

2.5.  $\text{not}(\text{minpro-max}(\{ \text{da}_1 \}))$

$\text{ArgRec}_2 = \{ \text{arg}_1\text{rev}, \text{da}_1, \text{da}_2, \text{arg}_2 \}; \text{dec}_2 = \text{none}.$

Again, among  $\text{da}_1$  and  $\text{da}_2$ , neither is defeating, so the opinion in this case is significant. Again, the adjudicator chose between competing arguments that could not be decided purely on form. Con revises pro's argument  $\text{arg}_1$  by producing  $\text{arg}_1\text{rev}$  which exposes the meta-reasoning as before. This time,  $\text{arg}_2$  is produced, which defeats an essential part of the case, one of the minpro-maximal arguments (in fact, the only minpro-maximal argument),  $\text{da}_1$ . So 2.5 is assertible,  $\text{not}(\text{minpro-max}(\{ \text{da}_1 \}))$ , which again is a requirements for using the precedent.

*Example. An f-rationale attacked for fit and redeemed in larger context.*

1. pro:  $\text{arg}_1 = \langle \{ b \ ]\!-\!-\ a \}, b! \ \therefore \ a \rangle$

$\text{ArgRec}_1 = \{ \text{arg}_1 \}; \text{dec}_1 = \text{pro}.$

2. con:

2.1.  $\text{f-rat}_2( b \ ]\!-\!-\ a ) = \{ \text{Cases}_1, \text{fit-r}( b \ ]\!-\!-\ a, \text{Cases}_1, \text{Context}_1 ) \}$

where

$\text{Cases}_1 = \{ \text{case}_1, \text{case}_2 \}$

$\text{case}_1 = \langle \{ b, d \}, a \rangle$

$\text{case}_2 = \langle \{ \text{not-b} \}, \text{not-a} \rangle$

$\text{Context}_1 = \{ \}$

2.2. Attack  $\langle b \ ]\!-\!-\ a, \text{arg}_1 \rangle$

- 2.3.  $\text{arg}_1\text{rev} = \langle$   
 $\{ \text{ok}(\text{Cases}_1) \ \& \ \text{fit-r}(b \succ\!\!\succ a, \text{Cases}_1, \text{Context}_1) \succ\!\!\succ (b \succ\!\!\succ a) \},$   
 $\text{ok}(\text{Cases}_1)! \ \text{fit-r}(b \succ\!\!\succ a, \text{Cases}_1, \text{Context}_1)! \ b! \ \therefore \ a \rangle$
- 2.4.  $\text{arg}_2 = \langle$   
 $\{ \text{ok}(\text{Cases}_2) \ \& \ \text{not-fit-r}(b \succ\!\!\succ a, \text{Cases}_2, \text{Context}_1) \succ\!\!\succ \text{not}(b \succ\!\!\succ a) \},$   
 $\text{ok}(\text{Cases}_2)! \ \text{fit-r}(b \succ\!\!\succ a, \text{Cases}_2, \text{Context}_1)! \ \therefore \ \text{not}(b \succ\!\!\succ a) \rangle$   
 where  
 $\text{Cases}_2 = \{ \text{case}_1, \text{case}_2, \text{case}_3 \}$   
 $\text{case}_3 = \langle b, c, \text{not-}a \rangle$

$\text{ArgRec}_2 = \{ \text{arg}_1\text{rev}, \text{arg}_2 \}; \text{dec}_2 = \text{none}.$

Con has added a case,  $\text{case}_3$ , which the rule,  $b \succ\!\!\succ a$ , does not fit in  $\text{Context}_1$ . Pro responds by exhibiting a larger context which accounts for the alleged deviation.

3. pro:  $\text{arg}_3 = \langle$   
 $\{ \text{ok}(\text{Cases}_2) \ \& \ \text{fit-r}(b \succ\!\!\succ a, \text{Cases}_1, \text{Context}_2) \succ\!\!\succ (b \succ\!\!\succ a) \},$   
 $\text{ok}(\text{Cases}_2)! \ \text{fit-r}(b \succ\!\!\succ a, \text{Cases}_2, \text{Context}_2)! \ b! \ \therefore \ a \rangle$   
 where  
 $\text{Context}_2 = \{ b \ \& \ c \succ\!\!\succ \text{not-}a \}$

$\text{ArgRec}_3 = \{ \text{arg}_1\text{rev}, \text{arg}_2, \text{arg}_3 \}; \text{dec}_3 = \text{pro}.$

*Example. Deepening an r-rationale-based attack.*

1. pro:  $\text{arg}_1 = \langle \{ b \ \& \ c \ ]\!\!\!- a \}, b! \ c! \ \therefore \ a \rangle$

$\text{ArgRec}_1 = \{ \text{arg}_1 \}; \text{dec}_1 = \text{pro}.$

2. con:

- 2.1.  $\text{r-rat}_1(b \ \& \ c \ ]\!\!\!- a) =$   
 $\langle \{ \text{da}_1, \text{da}_2 \}, \text{minpro-max}(\{ \text{da}_1 \}) \ \& \ \text{maxcon-max}(\{ \text{da}_2 \}),$   
 $\text{minpro-max}(\{ \text{da}_1 \}) \ \& \ \text{maxcon-max}(\{ \text{da}_2 \}) \succ\!\!\succ (b \ \& \ c \ ]\!\!\!- a) \rangle$

where

- $\text{da}_1 = \langle \{ b \succ\!\!\succ g, g \succ\!\!\succ a \}, b! \ \therefore \ a \rangle,$   
 $\text{da}_2 = \langle \{ c \succ\!\!\succ d, d \succ\!\!\succ \text{not-}a \}, c! \ \therefore \ \text{not-}a \rangle$

2.2. Attack  $\langle b \ \& \ c \ ]\!\!\!- a, \text{arg}_1 \rangle$

- 2.3.  $\text{arg}_1\text{rev} = \langle$   
 $\{ \text{minpro-max}(\{ \text{da}_1 \}) \ \& \ \text{maxcon-max}(\{ \text{da}_2 \}) \succ\!\!\succ (b \ \& \ c \ ]\!\!\!- a) \},$   
 $b! \ c! \ \text{minpro-max}(\{ \text{da}_1 \}) \ \& \ \text{maxcon-max}(\{ \text{da}_2 \})! \ \therefore \ a \rangle$

2.4.  $\text{arg}_2 = \langle \{ b \ \& \ e \succ\!\!\succ \text{not-}g \}, b! \ e! \ \therefore \ \text{not-}g \rangle$

2.5.  $\text{not}(\text{minpro-max}(\{ \text{da}_1 \}))$

$\text{ArgRec}_2 = \{ \text{arg}_1\text{rev}, \text{da}_1, \text{da}_2, \text{arg}_2 \}; \text{dec}_2 = \text{none}.$

This is as before:  $arg_2$  is produced, which defeats an argument that must be minpro-maximal. Now pro responds by attacking  $arg_2$  and ejecting it from  $ArgRec_3$ .

3. pro:

3.1.  $arg_3 = \langle \{ b \ \& \ e \ \& \ f \} \rightarrow g \rangle, b! \ e! \ f! \ \therefore \ g$

3.2.  $minpro-max(\{ da_1 \})$

$ArgRec_3 = \{ arg_1rev, da_1, da_2, arg_2, arg_3 \}; dec_3 = pro.$

Note that the argument pro in the prior case,  $da_1$ , can be strengthened in the current situation: instead of supporting  $g$  with merely  $b$ , support  $g$  with all of  $b \ \& \ e \ \& \ f$ . This alone does not suffice to establish  $a$ , however. Pro needs to cite the case in which  $b$ , a prima facie reason for  $a$ , was decided to be more important than  $c$ , a prima facie reason for not- $a$ . Among the object-level arguments in  $ArgRec_3$ ,  $\{ da_1, da_2, arg_2, \text{ and } arg_3 \}$ ,  $a$  is not warranted. The important point is that  $da_1$  was judged more important than  $da_2$  in the past, and  $arg_3$  prevents  $arg_2$  from disrupting the past's bearing on the present.

#### IV. Related Work and Conclusions.

A *ratio decidendum* is an essential part of a case. To cite a case without attention to the *ratio* is to make analogies on surface similarities. Casually determined similarities may not actually be relevant to the past decision or its subsequent use. Similar remarks could be made about the application of rules without regard to the principles on which the rules are based: without regard for *ratio legis*.

There may be some latitude in the use of cases. In some protocols, whoever cites a case must produce the rationale. In other protocols, it is the responsibility of the opposition to raise any problem regarding rationale. Here, the assumption is that *ratio decidendi* need not always be given when a case is cited as part of an argument. The assumption is that the use of the case is appropriate. The benefits of such an assumption are like the benefits of any protocol that permits simpler argument moves. Unless the use of the case or the use of the rule is disputed, the assumption is that the use is appropriate. The same is true for rules. What is important is a mechanism in the protocol for permitting an objection to the use of a rule or case, when a party to the dispute elects to do so.

Rationales, especially r-rationales and d-rationales, illuminate why Ashley can make distinctions among cases [Ashley89]. In our work, rules are extracted from cases. In any rule derived in this way, the sharing of important properties with the precedent is reason for sharing the decision of the precedent. Counterarguments are possible only by finding reasons for the opposite decision. Ashley and our improved analysis of r-rationales permit more sophisticated counterarguing. The opposition can simply cite properties that the present case does not share with the precedent. This kind of distinction becomes possible when the r-rationale or d-rationale of the case is available for examination.

Berman and Hafner recently discussed the telology of rules [Berman-Hafner93]. The rationales of

the rules they discuss are more complicated than the kinds discussed here. Their rationales involve a fundamentally different kind of reasoning: the reasoning about compromises in policy-making. The present work is less ambitious in scope and more ambitious in formality. It remains to be seen how many rationales can be expressed as compilation rationales of the five kinds explored here. Future work must surely be directed at representing the most important rationales in a particular legal domain. We do not expect that the full range of principles discussed for instance in Hart [Hart61], Dworkin [Dworkin85], or Peczenik [Peczenik89] could be accommodated with compilation rationales. The rationales chosen for investigation here are just the ones most amenable to treatment in the existing model of argument and disputation. Branting [Branting93] is a formal analysis of rationales that does not rely on argument systems.

Prakken briefly discusses the possibility of modeling principle and purpose [Prakken93]. He follows Gardner [Gardner87] by taking most of the principle and purpose to be reflected in the matching of past cases to present case: i.e., in the extraction of rules from cases. “In this way it is possible to account for the defeasibility of legal rules caused by principle and purpose without having to complicate the formal model too much.” Of course, the matching of cases must be performed on the essential aspects of the case. It was precisely the delineation of the essential from the inessential that led us to investigate explicit representation of rationales. Ashley [Ashley89], for example, assumes that only the important aspects of cases are formalized in the first place.

Prakken, noting agreement with Berman and Hafner, notes that reasoning with rationales will be unavoidably meta-level reasoning. Applying the model will determine whether the added representational power is worth the complication of the model to include meta-reasoning.

Formality plays an important role in some of AI’s interpretations of legal reasoning. Once the importance of rationales is allowed, the questions arise: what kinds of rationales? represented in what way? introduced according to what protocol? General studies of argument, such as Toulmin’s, have provided room for criticising the backing of a rule, the grounds on which it was adopted. But no prior work known to us has attempted to present details of this knowledge, their structures, and the processes in which they participate. To future automatizers of rule-based and case-based legal reasoning, we hope studies of this kind will be useful.

### **Acknowledgements.**

We are grateful for the meta-arguing of Violetta Cavalli-Sforza; also, to Kevin Ashley and Mark Mittelman for their comments, and the opportunity to present this work at U. Pittsburgh LRDC.

### **References.**

Alchourron, C. and E. Bulygin. *Normative Systems*, Springer-Verlag, 1971.

Alexy, R. *A Theory of Legal Argumentation*, (translation), Oxford, 1989.

Ashley, K. “Defining salience in case-based arguments,” *Proc. IJCAI*, 1989.

Ashley, K. and V. Aleven. “Toward an intelligent tutoring system for teaching law students to

- argue with cases,” *Proc. ACM ICAIL*, 1991.
- Bayles, M. *Procedural Justice*, Kluwer, 1990.
- Berman D. and C. Hafner. “Representing teleological structure in case-based legal reasoning,” *Proc. ICAIL*, 1993.
- Branting, K. “A reduction-graph model of ratio decidendi,” *Proc. ICAIL*, 1993.
- Dworkin, R. *A Matter of Principle*, Harvard, 1985.
- Gardner, A. *An Artificial Intelligence Approach to Legal Reasoning*, MIT Press, 1987.
- Gordon, T. *The Pleadings Game*, doctoral dissertation, University of Darmstadt, 1993.
- Hart, H.L.A. “The ascription of rights and responsibilities,” in A. Flew, ed., *Logic and Language*, Oxford, 1951.
- Hart, H.L.A. *Concept of Law*, Oxford, 1961.
- Loui, R. “Analogical reasoning, defeasible reasoning, and the reference class,” *Proc. Knowledge Representation and Reasoning*, 1989.
- Loui, R. “Process and policy,” Washington University Computer Science Technical Report 92-43, 1992.
- Loui, R., J. Norman, A. Merrill, K. Stiefvater, and A. Costello. “Computing specificity,” Washington University Computer Science Technical Report 93-03, 1993a.
- Loui, R., J. Norman, A. Merrill, and J. Olson. “A design for reasoning with policies, precedents, and rationales,” *Proc. ICAIL*, 1993b.
- Peczenik, A. *On Law and Reason*, Kluwer, 1989.
- Pollock, J. “How to reason defeasibly,” *Artificial Intelligence* 57, 1992.
- Prakken, H. *Logical Tools for Modelling Legal Argument*, doctoral dissertation, Free University Amsterdam, 1993.
- Quine, W. V. O. personal communication, 1990.
- Rescher, N. *Dialectics*, SUNY Buffalo, 1977.
- Rissland, E. and K. Ashley. “A case-based system for trade secrets law,” *Proc. ACM ICAIL*, 1987.
- Rissland, E. and D. Skalak. “Interpreting statutory predicates,” *Proc. ACM ICAIL*, 1989.
- Sartor, G. *Artificial Intelligence in Law*, Tano, 1993a.
- Sartor, G. “A simple computational model for non-monotonic and adversarial legal reasoning,” *Proc. ICAIL*, 1993b.

Simari, G. and R. Loui. "A mathematical treatment of defeasible reasoning and its implementation," *Artificial Intelligence* 53, 1992.

Skalak, D. and E. Rissland. "Argument moves in a rule-guided domain," *Proc. ACM ICAIL*, 1991.

Soeteman, A. *Logic in Law*, Kluwer, 1989.

Toulmin, S. *The Uses of Argument*, Cambridge, 1958.

Vreeswijk, G. *Studies in Defeasible Argumentation*, doctoral dissertation, Free University Amsterdam, 1993.