REASON BASED LOGIC
A LOGIC THAT DEALS WITH RULES AND REASONS

J.C. Hage, H.B. Verheij, A.R. Lodder

University of Limburg
Department of Metajuridica
P.O. Box 616
6200 MD Maastricht
tel. +31 43 8830 53 / 48 / 05
email: {jaap.hage, bart.verheij, arno.lodder}@metajur.rulimburg.nl

ABSTRACT

This paper proposes a logic that is specially designed to deal with rules and the reasons based on them. Although the logic is inspired by characteristics of legal reasoning, it aspires to be suitable for dealing with other kinds of rule guided reasoning as well.

It is characteristic for this logic that rules are treated as individuals rather than as conditionals, and that application of a rule only leads to a reason for the rule's conclusion. Reasons may have to be weighed against other reasons. The paper describes respectively the motivation behind the logic, a formalization of the logic, some examples of its application, and a dialogical extension.

1. INTRODUCTION

Reasoning with rules is so much different from reasoning with propositions, that it is worth the trouble to develop a logic that is specially designed to deal with rules. This view is backed up by a theory about logic which holds that a logic should model valid reasoning as it actually occurs. The task of logic according to this view, is not only to provide criteria for the validity of arguments, but also to offer a reasoning mechanism that is as faithful as possible to valid reasoning steps as they occur in human practice.

This paper provides three different approaches to Reason Based Logic (RBL), a logic that deals with rules and reasons. RBL comes in both a monological and a dialogical variant. The former focuses on valid inference, while the latter takes the modelling of actual human reasoning as its starting point. The paper starts with an overview of the philosophical considerations on which RBL is based. Second follows a formal description of the monological version of RBL, followed by some examples of its application. Finally, we present an informal description of the dialogical version of RBL, together with the motivation for having a dialogical version of the logic next to a monological one.

2. THE PARADIGM OF CLASSICAL LOGIC

In theory, logics can be developed as uninterpreted calculi. Whether the result can be used as a logic for a particular domain depends on whether the calculus can be given an interpretation which makes it adequate as a logic for the domain in question. Actually, most logics are developed within the paradigm set by first order predicate logic. This paradigm holds that theories formulated in a logical language are characterized by sets of statements, and that inference consists of making explicit the information contained in a theory.

The language of a logic consists of logical constants on the one hand, and signs for predicates (sets) and individuals on the other hand. The validity of arguments and of formulas depends solely on logical form, and this form is exhaustively determined by the logical constants of the language. Moreover, the logical constants are to be characterized by constraints on the models of logical theories. For instance, the conjunction is characterized by the constraint that in every logically possible world a conjunction has the value 'true' if and only if both conjuncts have the value 'true'. Modal operators are analogously characterized by constraints on sets of possible worlds (frames). Together, these assumptions approach logic as a tool to transport truth values (in a wide sense) among propositions, and this approach is reflected in the demand that a logical language is completed with a semantics that specifies the truth values of legal formulas.

Although this paradigm is challenged by some nonmonotonic logics (e.g. Reiter's default logic), it still holds sway and influences the way in which logicians define the subject of their work. As a consequence, reasoning with rules is usually dealt with as reasoning with some kind of universal conditionals, where the unusual characteristics of rules are taken into account by specifying unusual truth conditions of these conditionals. For instance, Stalnaker proposes an analysis of conditionals which makes them true material implications in worlds which differ minimally from the actual world [Stalnaker 1968]. We think that this analogy between rules and conditional statements is inadequate, and that rules are better compared to laws of nature which underlie forces that can interact with each other. This contribution describes a logic for rules that is based on this metaphor.

3. OF RULES AND REASONS

A rule is valid (or accepted, or used) if and only if the facts of the type identified by the rule's condition are (considered to be) reasons for facts of the type identified in the rule's conclusion. For instance, the validity of the legal rule that thieves are punishable comes down to it that the generic fact that somebody is a thief is a reason for the generic fact that this person is (legally) punishable. The acceptation of the rule of inference that thieves may be taken to be armed comes down to that the generic fact that somebody is a thief is (considered to be) a reason to assume that this person is armed.

3.1. Kinds of rules and types of rules

Reasons can be subdivided according to the *kinds* of their conclusions, and the *types* of their modes of existence. A rule's conclusion can be classificatory, anankastic, deontic, or epistemic[1]. Examples are respectively: Who takes away another person's property is a thief, Thieves are punishable (there is a possibility to punish), Thieves ought to be punished, and Who comes running out of a robbed bank under suspect circumstances, is probably a thief.

According to their modes of existence, three types of rules can be distinguished. First there are rules based on explicit decisions. Most legal rules belong to this type of decision based rules. Second there are social rules, which exist in their being used within a social group. Moral rules, most rules of language, and many anankastic rules belong to this type. Finally there are what might be called 'personal rules'. Somebody uses a personal rule if (s)he is disposed to consider facts of a certain kind as reasons for facts of another kind. Personal rules may coincide with decision based rules and with social rules. In fact, social rules are personal rules used by sufficiently many members of a group.

It is important to notice that the existence of any type of rule is *a matter of social reality*, and can therefore not be derived from truth values of the rule's conditions and conclusions. For this reason, the logic which will be presented considers rules as individuals, and statements about their validity as propositions of the object language.

---

1 A sentence is said to be anankastic if it expresses a necessity or a possibility (an alethic modality). The term anankastic derives from Von Wright 1963. Deontic sentences are concerned with what is obliged, permitted, or forbidden. They are *not* concerned with what is ideally the case as ideal world semantics might want us to belief. The ideal and the prescribed are not necessarily the same. Epistemic sentences are concerned with what is certain, probable, and - in a non-anankastic sense - possible. Cf. White 1975, p. 78f.

The distinctions between the kinds of rule conclusions and between the types of rule existence are orthogonal to each other, thus creating twelve possible categories of rules. Most, but not all, of these categories actually occur.

3.2. Rules of inference and constitutive rules

The four kinds of rules can be divided in the two main categories of constitutive rules and rules of inference. The *rules of inference* coincide with the kind of epistemic rules. They are characterized by the phenomenon that the truth of the rule's conclusion (what may be inferred) is independent of the truth of the rule's conditions. For instance, whether somebody has robbed a bank is not (even partly) determined by the facts that this person came running out of a bank, with a mask on and brandishing a gun. Yet, these facts make it reasonable to *assume* that this person robbed the bank.

In the case of *constitutive rules*, the facts that correspond to the rule's conditions make it the case that the facts of the rule's conclusion are the case. Taking somebody else's property *makes* one into a thief. Being a thief *makes* one punishable, and that one ought to be punished.

The distinction between rules of inference and constitutive rules may be summarized by saying that the former underlie reasons to assume their conclusions, while the latter underlie reasons why their conclusions are the case (true). Adapting some terminology of Searle's, we can say that rules of inference have a word to world fit, while constitutive rules have a world to word fit [Searle 1979, p. 3-5] [2].

In the remainder of this paper our discussion will be confined to constitutive rules.

4. REASONS AS INTERACTING FORCES

If there is a constitutive reason for a conclusion, this does not guarantee the truth of that conclusion. Neither does a reason against a conclusion guarantee the falsity of that conclusion. Whether a fact is constituted depends on all available reasons. This is particularly clear in deontic cases: What I ought to do depends on all available reasons for (and against) behavior. The same is, however, also true for classifications and for anankastic conclusions. There can both be reasons why this object is to be classified as a table and reasons why it should not be classified so. Some facts contribute to the possibility to travel to Mars, while other facts tend to make it impossible. The actual conclusion about a classification, or about a possibility depends on the interaction of all available reasons.

The relation between rules and reasons on the one hand and the conclusions based on them on the other hand, is comparable to respectively the relation between laws of nature, forces and the outcome of their interaction. Just as it is impossible to predict what will happen if only one force is known, it is impossible to draw a conclusion from only one reason, unless, of course, this reason is the only available one.

A force on a body is one of the factors that determine the body's acceleration. A law of nature describes what will happen to a body as far as only this particular force is taken into account. Similarly a constitutive reason is one of the factors that determine a particular piece of reality, and a constitutive rule describes reality as far as this reason is concerned. For example, a person is punishable as far as the fact that he is a thief is concerned. The corresponding rule that thieves are punishable 'describes' reality as far as it is determined by persons being thieves.

Laws of nature are to be distinguished from hypothetical descriptions of the effects of forces. Hypothetical descriptions indicate what will happen as a result of all interacting forces, if the facts are this or that. These descriptions are statements which aim to be true. The laws, on the contrary, are formula's to compute forces, and do not aim to describe what happens to a body on which the force acts. This does not exclude that the law correctly describes the body's behavior if only one force is relevant.

---

[2] Cf. also Putnam's internal realism (Putnam 1976 and 1981), and Goodman's view about world making (Goodman 1978).

Similarly, rules do *not describe* a relation in reality between its conditions and its conclusion. A rule only indicates how the facts of its conditions contribute to the facts of its conclusion. This does not exclude that the rule gives a true description of reality in case there is only one relevant reason.

4.1. Decision based rules

Decision based rules resemble hypothetical descriptions much more than other rules. The purpose of making a decision is to weigh all reasons for and against something and store the result (the contents of the decision) for future use. The decision is now a reason (for instance, to act) and this reason comes *in the place* of the reasons on which it is based. Because it replaces these previous reasons, the reason based on the decision need not to be weighed against these other reasons[3].

For instance, the legislator weighs economical against environmental interests in making ecological legislation. This legislation tells us what count as reasons in discussions about whether some industrial pollution is allowed. The interests underlying the statute do not count anymore; they have been replaced by the legislation.

If all relevant reasons were taken into account while making the decision, the decision need not be weighed against other reasons. As a consequence, rules based on decisions seem to describe a connection in reality. They are very much like the hypothetical descriptions of what will happen as a result of all interacting forces.

This phenomenon may make us loose sight of the exceptional character of this situation. Normally rules only underlie reasons which must be weighed against other reasons; only in special circumstances decision based rules generate reasons which exclude other reasons and which need not be weighed. However, even reasons from decision based rules need to be weighed against reasons which were not taken into account while deciding on the rule[4].

5. A FORMALIZATION OF MONOLOGICAL REASON BASED LOGIC

Reason Based Logic (RBL) is an extension of first order predicate logic which is especially designed to deal with the peculiarities of rules and reasons as described above. Derivation of constituted facts goes in two steps. In the first (iterative) step, all reasons for and against the potential conclusions are collected. In the second step weighing knowledge is employed to determine what follows from the found reasons.

RBL comes in a monological and in a dialogical variant. For many purposes the monological variant suffices. Some kinds of arguments can, however, not well be modelled in the monological variant. For instance the phenomenon that some reasons (although by themselves perfectly in order) may not be used in drawing a conclusion is better modelled in the dialogical version of the logic[5].

The following sections describe the monological version of RBL. The basic idea is to consider the condition of a rule merely as a *reason* for the rule's conclusion. Other reasons considering that conclusion, originating from other rules, can influence the actual derivation. In this way the resolution of rule conflicts can be represented as the result of weighing the reasons concerning some statement.

Only an *applicable* rule gives rise to a reason. By means of the notion of rule applicability, a natural rule priority mechanism using rule names can be incorporated in RBL. By this mechanism reasons can be excluded in the presence of particular knowledge. As a result, the resolution of a conflict between rules can be divided in two parts. First, priority knowledge can

---

3 The present account of decisions, of rules, and of reasons based on decisions is inspired by the account of decisions, exclusionary reasons, and mandatory rules, provided by Raz 1975, 1978a and 1978b.

4 Cf. Hage 1992.

5 An example of this phenomenon would be that some criminal evidence may not be used, because it was obtained illegally.

simplify a conflict by making rules inapplicable. Second, the conflict is solved by knowledge on the relative weight of the reasons originating from the remaining applicable rules.

The following two sections define respectively RBL-theories, and the inference rules of RBL.

5.1. RBL-Theories: facts, rules, weighing results

Definition 1. The language L_{RBL} of Reason Based Logic is the language L_{FOPL} of First Order Predicate Logic, extended with the unary predicate symbols applicable, condition_satisfied and condition_denied, and the unary function symbols cs and cd. An *RBL-formula* is a formula of L_{RBL}. An *RBL-proposition* is a closed RBL-formula.

The role of the extra predicate and function symbols will be explained later.

Definition 2. An *RBL-rule* has the form

$r : \varphi \Rightarrow \lambda$

where r is a function symbol, φ an RBL-formula, and λ an *RBL-literal*, i.e., an RBL-formula that is also a literal. A rule is called *closed*, if neither φ, nor λ contain free variables; otherwise *open*. If $\lambda = \alpha$ for an atom α, we call φ a *possible reason for* α; if $\lambda = \neg\alpha$, φ is a *possible reason against* α. The term $r(x)$, where x stands for the free variables in the rule, is called the *name* of the rule.

Rules are the reason-constituting objects in Reason Based Logic. If a rule is *applicable* its condition is a reason for its conclusion.[6] This constitution of a reason is called the *application* of the rule. Whether the conclusion of a rule follows, is not solely determined by its application. A conclusion follows on the basis of knowledge on the relative weight of all reasons concerning a statement that arise through the application of applicable rules.

Definition 3. A *weighing result* has the form

(α, P, C, v)

where α is an *RBL-atom* (that is an RBL-formula that is also an atom), P and C are finite sets of formulas, and v equals either *pos* or *neg*. A weighing result is called *closed*, if neither α, nor any formula in P or C contains free variables; otherwise *open*. Let R be a set of rules. We call a set of weighing results W_R a *weighing relation for R*, if it has the following properties:

(1) $\forall\, (\alpha, P, C, v) \in W_R$: $P \subseteq P_R(\alpha)$ and $C \subseteq C_R(\alpha)$.

Here $P_R(\alpha)$ and $C_R(\alpha)$ are respectively the set of possible reasons for α (stemming from rules in R), and the set of possible reasons against α.

(2) For all non-empty sets $P \subseteq P_R(\alpha)$ and $C \subseteq C_R(\alpha)$,

(i) $(\alpha, P, \emptyset, pos) \in W_R$, $(\alpha, \emptyset, C, neg) \in W_R$.

(ii) $(\alpha, P, \emptyset, neg) \notin W_R$, $(\alpha, \emptyset, C, pos) \notin W_R$.

The *minimal weighing relation for R*, denoted $W^{\min}{}_R$, is the intersection of all weighing relations for R. The *maximal weighing relation for R*, denoted $W^{\max}{}_R$, is the union of all weighing relations for R.

Weighing results represent the knowledge on the relative weight of reasons. The weighing result

(α, P, C, v)

informally means that weighing the reasons in P, pleading for α, and those in C, pleading against α has the outcome v. Note that property (2) describes the weighing knowledge in case all reasons point in one direction, and there is no conflict.

Definition 4. A triple $\boldsymbol{T} = (F, R, W_R)$ is an *RBL-theory*, if F is a set of RBL-propositions, R a set of RBL-rules (with names not of the form cs(r) or cd(r)), and W_R a weighing relation for

---

[6] Hage (1993) proposes an extra mechanism to model *exclusionary reasons*. If such a reason arises no weighing knowledge is needed. It simply excludes the application of a rule, whatever reasons there are for applying it.

R. The elements of *F* are called the *facts* of the theory ***T***. An RBL-theory is called *closed*, if both *R* and W_R are closed; otherwise *open*.

To simplify the definitions this formalization of Reason Based Logic does not allow reasoning with rules and weighing results. The sets *R* and W_R can be interpreted as shorthand for propositions of the form

valid(⌜$r : \varphi \Rightarrow \lambda$⌝)

weighing_result(⌜α⌝, ⌜P⌝, ⌜C⌝, v)

for elements $r : \varphi \Rightarrow \lambda \in R$ and $(\alpha, P, C, v) \in W_R$, in some way encoded as terms of the logical language by a suitable function λx.⌜x⌝.

This anticipates a natural extension of Reason Based Logic, in which reasoning with rules and weighing results is incorporated in the formalism, as proposed by Hage (1993).

Definition 5. Let $\boldsymbol{T} = (F, R, W_R)$ be a (closed) RBL-theory. $\boldsymbol{T}^* = (F^*, R^*, W_R^*)$, the *associated theory of* ***T***, is defined as follows:

$F^* := F \cup \{ \varphi \leftrightarrow \text{applicable}(\text{cs}(r)) \mid r : \varphi \Rightarrow \lambda \in R \}$
$\quad \cup \{ \neg\varphi \leftrightarrow \text{applicable}(\text{cd}(r)) \mid r : \varphi \Rightarrow \lambda \in R \}$

$R^* := R \cup \{ \text{cs}(r)\text{: condition\_satisfied}(r) \Rightarrow \text{applicable}(r) \mid r : \varphi \Rightarrow \lambda \in R \}$
$\quad \cup \{ \text{cd}(r)\text{: condition\_denied}(r) \Rightarrow \neg\text{applicable}(r) \mid r : \varphi \Rightarrow \lambda \in R \}$

$W_R^* := W^{\min}{}_{R^*} \cup W_R$.

The rules with names of the form cs(*r*) and cd(*r*) mentioned in definition 4 play a special role. By definition 5 certain axiomatic knowledge is added to a theory. For each rule in a theory two rules, with names cs(*r*) and cd(*r*), are added. They define default reasons for applicability and non-applicability of a rule, namely the satisfaction and the denial of its condition. Note that W_R^* is the smallest weighing relation for R^* containing W_R.

Recall that the applicability of a rule just means that it constitutes a reason for its conclusion. Thus, if r: $\varphi \Rightarrow \alpha$ is a rule, and $\varphi \leftrightarrow$ applicable(cs(*r*)) is RBL-derivable this will simply mean that condition_satisfied(*r*) is a reason for applicable(*r*), if and only if the condition φ of *r* holds.

5.2. An additional rule of inference

To avoid unnecessary attention to technicalities, theories are implicitly assumed to be closed in the following definitions. Measures involving Skolemized versions of a theory should be taken to deal with theories that contain open rules and weighing results.

Definition 6. Let $\boldsymbol{T} = (F, R, W_R)$ be a (closed) RBL-theory, and α an RBL-atom. We define the following sets:

$R^+(\alpha) := \{ \rho \mid \rho = r : \varphi \Rightarrow \alpha \in R \}$

$R^-(\alpha) := \{ \rho \mid \rho = r : \varphi \Rightarrow \neg\alpha \in R \}$

$R(\alpha) := R^+(\alpha) \cup R^-(\alpha)$

Definition 7. The relation $\vdash^\circ_{RBL}$ is the smallest relation, with the following three properties. Let $\boldsymbol{T} = (F, R, W_R)$ be a (closed) RBL-theory, and $\boldsymbol{T}^*$ its associated theory.

(1) For an RBL-proposition φ, $\boldsymbol{T}^* \vdash^\circ_{RBL} \varphi$, if $F \vdash_{FOPL} \varphi$. Here $\vdash_{FOPL}$ denotes the deduction relation in First Order Predicate Logic.

(2) Let φ and ψ be RBL-propositions. If $\boldsymbol{T}^* \vdash^\circ_{RBL} \varphi$, and $\boldsymbol{T}^* \vdash^\circ_{RBL} \varphi \to \psi$, then $\boldsymbol{T}^* \vdash^\circ_{RBL} \psi$.

(3) Let α be an RBL-atom. Suppose that we have

$\forall\, r : \varphi \Rightarrow \lambda \in R(\alpha)$: $\boldsymbol{T}^* \vdash^\circ_{RBL} \text{applicable}(r) \vee \boldsymbol{T}^* \vdash^\circ_{RBL} \neg\text{applicable}(r)$,

Define the sets

$P := \{ \varphi \mid \exists\, r : \varphi \Rightarrow \alpha \in R^+(\alpha)\text{: } \boldsymbol{T}^* \vdash^\circ_{RBL} \text{applicable}(r) \}$,

$C := \{ \psi \mid \exists\, s : \psi \Rightarrow \neg\alpha \in R^-(\alpha)\text{: } \boldsymbol{T}^* \vdash^\circ_{RBL} \text{applicable}(s) \}$.

Then:

(i) If (α, P, C, *pos*) $\in W_R$, then $\boldsymbol{T}^* \vdash^\circ_{RBL} \alpha$.

(ii) If (α, P, C, *neg*) $\in W_R$, then $\boldsymbol{T}^* \vdash^\circ_{RBL} \neg\alpha$.

In this situation the elements of P are called *reasons for* α, those of C *reasons against* α.

Let $\boldsymbol{T} = (F, R, W_R)$ be a (closed) RBL-theory, and φ an RBL-proposition. We define $\boldsymbol{T} \vdash_{RBL} \varphi$, if and only if $\boldsymbol{T}^* \vdash^\circ_{RBL} \varphi$.

The RBL-derivability relation $\vdash_{RBL}$ is defined by means of the relation $\vdash^\circ_{RBL}$ in order to add the axiomatic knowledge of definition 5. By (1) all tautologies of First Order Predicate Logic are RBL-derivable, hence property (2), Modus Ponens, suffices to include all FOPL-derivations in the derivability relation of Reason Based Logic. Property (3) is the essence of Reason Based Logic. It encompasses the main points, that applicable rules constitute reasons, and that a conclusion on the basis of reasons only follows if the relative weight of the reasons is known.

6. EXAMPLES

6.1. Application of a single rule

The simplest case is how the conclusion of a rule follows from its condition, if no interfering rules are available. This involves several steps that are worked out in detail.

We will consider the bird Tweety, and the rule that, if Tweety is a bird, she can fly. Let the theory $\boldsymbol{T}_1 = (F_1, R_1, W_1)$ be defined by

F_1 := { bird(Tweety) },

R_1 := { f(Tweety): bird(Tweety) ⇒ fly(Tweety) },

$W_1 := W^{\min}{}_{R_1}$.

The associated theory of $\boldsymbol{T}_1$, $\boldsymbol{T}_1^* = (F_1^*, R_1^*, W_1^*)$, is

F_1^* := F_1 ∪ { bird(Tweety) ↔ applicable(cs(f(Tweety))),
¬bird(Tweety) ↔ applicable(cd(f(Tweety))) },

R_1^* := R_1 ∪ { cs(f(Tweety)): condition_satisfied(f(Tweety)) ⇒ applicable(f(Tweety)),
cd(f(Tweety)): condition_denied(f(Tweety)) ⇒ ¬applicable(f(Tweety)) },

$W_1^* := W^{\min}{}_{R_1^*}$.

We want to derive fly(Tweety), using the reason bird(Tweety). The rule with name f(Tweety) constitutes this reason, if it is applicable, i.e., if applicable(f(Tweety)) is derivable. By (1) in definition 7 we have

$\boldsymbol{T}_1^* \vdash^\circ_{RBL}$ bird(Tweety)

$\boldsymbol{T}_1^* \vdash^\circ_{RBL}$ bird(Tweety) ↔ applicable(cs(f(Tweety)))

Now applying (2) in definition 7, we find

$\boldsymbol{T}_1^* \vdash^\circ_{RBL}$ applicable(cs(f(Tweety)))

Analogously we conclude

$\boldsymbol{T}_1^* \vdash^\circ_{RBL}$ ¬applicable(cd(f(Tweety)))

Thus condition_satisfied(f(Tweety)) is the only reason concerning applicable(f(Tweety)). Applying (3) in definition 7 to applicable(f(Tweety)), using the weighing result

(applicable(f(Tweety)), { condition_satisfied(f(Tweety)) }, ∅, *pos*) $\in W_1^*$,

yields $\boldsymbol{T}_1^* \vdash^\circ_{RBL}$ applicable(f(Tweety)). There are no other rules concerning fly(Tweety), so this time applying (3) in definition 7 to fly(Tweety), and using

(fly(Tweety), { bird(Tweety) }, ∅, *pos*) $\in W_1 \subseteq W_1^*$,

we find $\boldsymbol{T}_1^* \vdash^\circ_{RBL}$ fly(Tweety). By definition 8 we finally arrive at $\boldsymbol{T}_1 \vdash_{RBL}$ fly(Tweety).

6.2. Weighing reasons

The next imaginary example is taken from the field of law. It illustrates the weighing of more than two reasons. Pat, who is sixteen year old, is coming to trial for shoplifting. Normally she would be punished. But because it's her first offense, the judge does not convict her. The seventeen year old Bob, who attacked and injured somebody in a fight, got away with a warning, because it was his first offense. But John, sixteen years of age, was punished, even though it was his first encounter with the law: while stealing from a shop, he got involved in a fight and injured a customer, who was trying to stop him. This was considered too serious to let him get away with it.

In this case three rules are involved in a conflict. Pairwise priorities do suffice in Pat's and Bob's case (because in their cases only two rules are actually applicable), but in John's case the relative weight of the reasons provided by all three rules is needed to solve the conflict. This type of reasoning can be modeled easily in Reason Based Logic.

Define the theory $\boldsymbol{T}_2 = (F_2, R_2, W_2)$ by

F_2 := { steal(Pat) $\wedge$ $\neg$injure(Pat) $\wedge$ minor(Pat),
$\neg$steal(Bob) $\wedge$ injure(Bob) $\wedge$ minor(Bob),
steal(John) $\wedge$ injure(John) $\wedge$ minor(John) },

R_2 := { p_1: steal(x) $\Rightarrow$ punish(x),
p_2: injure(x) $\Rightarrow$ punish(x),
p_3: minor(x) $\Rightarrow$ $\neg$punish(x) },

W_2 := $W^{\circ}_{R_2}$ $\cup$ { (punish(x), { steal(x) }, { minor(x) }, *neg*),
(punish(x), { injure(x) }, { minor(x) }, *neg*),
(punish(x), { steal(x), injure(x) }, { minor(x) }, *pos*) }.

In this definition the convention is broken that theories should be closed. Each element of R_2 and W_2 can be read as shorthand for its three instances, one for Pat, Bob and John. It is then straightforward to conclude the following:

$\boldsymbol{T}_2 \vdash_{RBL}$ applicable(p_1(Pat)) $\wedge$ $\neg$applicable(p_2(Pat)) $\wedge$ applicable(p_3(Pat)),
$\boldsymbol{T}_2 \vdash_{RBL}$ $\neg$applicable(p_1(Bob)) $\wedge$ applicable(p_2(Bob)) $\wedge$ applicable(p_3(Bob)),
$\boldsymbol{T}_2 \vdash_{RBL}$ applicable(p_1(John)) $\wedge$ applicable(p_2(John)) $\wedge$ applicable(p_3(John)).

The denial of the conditions of rules p_2(Pat) and p_1(Bob) leads to their inapplicability. Using the weighing results

(punish(Pat), { steal(Pat) }, { minor(Pat) }, *neg*),
(punish(Bob), { injure(Bob) }, { minor(Bob) }, *neg*),
(punish(John), { steal(John), injure(John) }, { minor(John) }, *pos*),

we have the following outcome:

$\boldsymbol{T}_2 \vdash_{RBL}$ $\neg$punish(Pat) $\wedge$ $\neg$punish(Bob) $\wedge$ punish(John).

6.3. Exceptions to rules

The next example shows how exceptions to rules can be modeled in the formalism. Suppose we include the following (open) rule and weighing result in a theory.

exc: exception(r) $\Rightarrow$ $\neg$applicable(r)

(applicable(r), { condition_satisfied(r) }, { exception(r) }, neg)

The meaning of the weighing result is that the reason for the applicability of a rule, namely that its condition is satisfied, is outweighed by the reason against applicability, namely that there is an exception. If this scheme is included in a theory it becomes very simple to model rules and (undercutting) exceptions. For example, consider the general rule that an offense leads to punishment. An exception to this rule occurs, if the offense has happened too long ago. But then again, there is an exception to this latter rule, which states that war crimes can never become outdated. The following rules lead to the expected conclusions.

r_1: offense ⇒ punish
r_2: out_of_date ⇒ exception(r_1)
r_3: war_crime ⇒ exception(r_2).

7. DIALOGICAL REASON BASED LOGIC

Traditionally, logic focuses on the relations between sets of sentences or propositions. Real arguments, however, are not sequences of sentences which stand in a particular relation to each other. Arguments are a kind of speech *acts*, in which sentences are used to convince some audience of the truth, validity or acceptability of a statement or rule. The audience can be considered to be the other party in a dialogue in which both parties can make dialogue moves. The determination of an argument's validity in the traditional way can be compared to the determination which party *ought* to win the dialogue.

The advantage of taking a dialogical approach to arguments is that some pragmatic aspects of arguments, which follow from the process-character of dialogues, can easily be taken into account. For instance, in a monological logic, it is not so easy to account for phenomena as the division of the burden of proof, and that a valid argument should sometimes not be considered in drawing the conclusion. Yet, these phenomena play an important role in actual dialogues, and often they will determine which dialogue party wins the dialogue. To be able to account for these peculiarities of *actual* arguments (as opposed to sequences of sentences), we are developing a dialogical version of RBL, which will also serve as the basis of an Intelligent Tutoring System. The following contains a first approach to this dialogical version.

7.1. Basic ideas

Dialogical RBL specifies the rights and duties of two parties that have a dialogue. A dialogue starts if a party, in its role as *proponent,* advances (claims) a thesis. The advanced thesis is the subject of the dialogue. The other party, in its role as *opponent,* has three different ways to react: he can agree to the thesis, he can deny the thesis and he can question the thesis. The party who advances a thesis is always the proponent of this particular thesis and has the burden of proof for it; the other party is the opponent of this thesis. The roles can change during a dialogue.

All dialogues together are called the *dialogue game.* Each dialogue takes place at a certain level: the main dialogue has level 1; sub-dialogues have higher level-numbers. A sub-dialogue originates, when an advanced thesis is denied or questioned. So if a thesis is advanced at level n, after questioning by the opponent a dialogue starts at level n+1. This n+1 dialogue is a sub-dialogue of the level n dialogue. A sub-dialogue of this n+1 dialogue is also a sub-dialogue of the level n dialogue. In fact every dialogue at a level n+k (k>0) that originates from a level n dialogue, is a sub-dialogue of the level n dialogue.

If in a dialogue statements are made on level n, these statements are both arguments for the level n-thesis, and theses for the level n+1-dialogue. The thesis of the level 1-dialogue is the main thesis. Each dialogue is defined by the thesis it is about. A dialogue ends as soon as its thesis has been decided. The structure of dialogues is recursive.

7.2. Example dialogue

The following dialogue is to give an impression of the course of a dialogue.

	Bert	Ernie
1.	0> You should be punished with 15 years imprisonment	? (questions the thesis)
2.	1> You've shot the deputy	A (Agreement)
3.	1> The killing was on purpose	A

4.	1> If you kill someone on purpose, you should be sentenced to 15 years imprisonment	?
5.	2> Section 287 Dutch Penal Code lays down rule: move 4	A4
6.	1> Apply(rule: move 4; conditions: move 2, 3)]	2> not [Apply(rule: move 4; conditions: move 2, 3)]
7.	?	3> The sheriff was threatened by the deputy
8.	A	3> I killed to defend the sheriff
9.	A	3> If you kill someone to defend a person who is threatened, you are not punishable.
10.	?	4> section 41 Dutch Penal Code lays down rule: move 9
11.	withdraw (1)	

This example dialogue game starts when Bert advances the thesis that Ernie should be sent to prison for fifteen years. Ernie questions this rather strong claim. Bert begins his defense by successively mentioning two facts, to which Ernie agrees (moves 2 - 3). Ernie still is not convinced about the main thesis, so in move 4 Bert gives a rule. Ernie wants to question the rule, so Bert names the source of the rule. Ernie, presumably after reading the section mentioned by Bert, agrees to the rule (move 5). Because the rule was claimed in a former move, Ernie specifies his agreement by giving the number of the move in which the rule was claimed.

As a next step, Bert wants to apply the rule. Ernie denies that the rule should be applied and has to defend his negation. In doing so he provides two facts and a rule (moves 7 - 9). Bert wants to know the source of the rule and after reading the section, he realizes that Ernie is not punishable. In his last move Bert is not waiting for Ernie to apply the rule, but he withdraws his first claim immediately, knowing that "You should be punished with 15 years imprisonment" is not going to hold. With the withdrawal of the first claim the dialogue game ends; Bert has failed in defending the main thesis claimed in the first move.

7.3. Sentences

The theses are sentences that belong to either one of three kinds: factual statements, rules, and statements concerning the application of rules. Factual statements present some factual state of affairs. If one claims a fact, one claims the truth of that fact. Facts are for instance, "He shot the deputy", or "He is a murderer". In claiming a rule one claims the validity of a rule. An example of a rule is "If you kill someone, then you are a murderer". Since in RBL rules are not applied automatically if their conditions are satisfied, there are statements that explicitly indicate that a rule is to be applied.

7.4. Dialogue moves

There are four moves the players have disposal of: claim(sentence), accept(sentence), question(sentence) and withdraw(sentence). If a party claims a sentence, he believes the sentence to be true. If a party accepts a sentence, he agrees to a sentence claimed by his opponent. If a party wants a claim to be defended, he can question the sentence of that claim. Finally, a party withdraws a statement if he no longer thinks that the sentence he put forward is defensible.

The move in which the application of a rule is claimed is a special one. The application of a rule may only be claimed if both parties are committed to the factual statements that match with the condition of a rule and if both are committed to the rule. The consequence of a claim about the application of a rule is that the other party is committed to the conclusion of the rule, unless he has arguments why the rule should not be applied. The opponent cannot question a rule's application.

7.5. Commitment stores

The theses a dialogue party becomes committed to during a dialogue, are part of this person's *personal commitment store* (CS). This CS is empty when a dialogue starts, and is filled during the dialogue. Every time a party claims a thesis, or agrees to a claimed thesis, this thesis is added to his CS. If a player no longer feels bound to a sentence that he claimed earlier, he can withdraw this sentence only if it concerns an open claim.

If you are committed to a thesis, you are limited in the moves that you can perform. First, if your opponent claims a thesis you are committed to, you may not question it. Second, you may not claim a thesis, when you are committed to the opposite. Finally, if you are committed to a thesis, you may not agree to its opposite.

A claimed sentence that is in the CS of only one of the parties is called an open claim. An open claim occurs if a claimed sentence is questioned and the dialogue about the sentence has not been decided yet.

8. THE RULES OF THE DIALOGUE

The dialogues are defined by what in logics of dialogue are called rules for commitment, rules for dialogue and rules for termination [Loui 1992]. Rules for commitment describe the situations in which parties become committed to theses and the removal of commitment. The rules of dialogue describe the way the parties can use the *dialogue moves.* Finally, the rules of termination are about the situations in which dialogues end.

The rules described are a minimal set of rules. In the future more elements of RBL have to be incorporated, like counter-arguments and arguments using weighing knowledge. The commitment stores of a party P is indicated as CS_P. The commitment stores of the two parties a and b are similarly represented by CS_a and CS_b.

commitment rules

rule 1. A party is committed to the sentences in his CS.

rule 2. A claimed sentence is added to the CS of the claiming party.

rule 3. An accepted sentence is added to the CS of the accepting party.

rule 4. A withdrawn sentence is removed from the CS of the withdrawing party.

rule 5. If the dialogue ends because a sentence was accepted, all open claims of the dialog's opponent are withdrawn and all open claims of the dialog's proponent are added to the opponent's CS.

rule 6. If the dialogue ends because a sentence was withdrawn, all open claims of the dialog's proponent are withdrawn and all open claims of the dialog's opponent are added to the proponent's CS.

dialogue rules

rule 7. A dialogue about sentence S starts by claim(S). There is no other way to start a dialogue.

rule 8. It is possible to make a claim as the first move of a dialogue game, direct after another claim, direct after the acceptance of another claim, and direct after a question. These possibilities are subject to the following constraints:

a. if a claim is direct after another claim, it should claim the negation of the contents of that other claim;

b. a party P can only claim a sentence S, if neither S nor the negation of S is in CS_P;

c. a claim to the effect that a rule applies is only possible if the statement that this rule is valid and the sentences corresponding to the rule's conditions are in both party's CS's.

There are no other possibilities to make a claim.

rule 9. A party can accept a sentence S if and only if both S is not in his CS, and S is in his opponent's CS.

rule 10. A party can question the sentence S as a reaction to Claim(S), unless this party is committed to S, or S states that a particular rule applies.

rule 11. A party can withdraw a sentence S if and only if both S is in his CS, and S is not in his opponent's CS.

termination rules

rule 12. A dialogue about the sentence S ends if and only if either:

a. S is in CS_a and in CS_b;

b. S is neither in CS_a, nor in CS_b.

rule 13. If a dialogue ends, all of its sub-dialogues end.

REFERENCES

Goodman, N. (1978). *Ways of worldmaking*, Hackett Publishing Company, Indianapolis and Cambridge.

Hage, J.C. (1992). Redeneren met de strekking van rechtsregels. (Reasoning with the purpose of rules) (H. de Swaan Arons, H. Koppelaar and E.J. H. Kerckhoffs eds.) *Proceedings NAIC '92*, Delft, p. 43-52.

Hage, J.C. (1993). Monological Reason Based Logic. *Proceedings of the fourth International Conference on Law and Artificial Intelligence*, ACM-press, Amsterdam, p. 30-39.

Loui, R.P. (1992). *Process and Policy: Resource-Bounded Non-Demonstrative Reasoning.* Wash. U. Comp. Sci, WUCS-92-43, Submitted to Cognitive Science, 10/92.

Putnam, H. (1976). Realism and Reason. H. Putnam, *Meaning and the Moral Sciences*. Routledge and Kegan Paul, Boston e.a. 1978, p. 123-138.

Putnam, H. (1981). *Reason, Truth and History*. Cambridge University Press, Cambridge e.a.

Raz, J. (1975). *Practical Reason and Norms*, Hutchinson, London.

Raz, J. (ed.) (1978a). *Practical Reasoning*. Oxford University Press, Oxford.

Raz, J. (1978b). Introduction. J. Raz (ed.) *Practical Reasoning*. Oxford University Press, Oxford, p. 1-17.

Searle, J. (1983). A taxonomy of Illocutionary Acts, *Expression and Meaning*, Cambridge University Press, Cambridge e.a., p. 1-29.

Stalnaker, R.C. (1968). A Theory of Conditionals, *Studies in Logical Theory*, (N. Rescher ed.), Blackwell, Oxford.

White, A.R. (1975). *Modal Thinking*. Basil Blackwell, Oxford.

Wright, G.H. Von (1963). *Norm and Action*, Routledge and Kegan Paul, London.

This research was partly financed by the Foundation for Knowledge-based Systems (SKBS).