Artificial Intelligence and Law **6:** 3–26, 1998.

An Integrated View on Rules and Principles ⋆

BART VERHEIJ
Department of Metajuridica, Universiteit Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands
email: bart.verheij@metajur.unimaas.nl, internet: http://www.metajur.unimaas.nl/~bart/

JAAP C. HAGE
Department of Metajuridica, Universiteit Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands
email: jaap.hage@metajur.unimaas.nl

H. JAAP VAN DEN HERIK
Department of Law and Informatics, Leiden University, P.O. Box 9520, 2300 RA Leiden, The Netherlands
email: herik@cs.unimaas.nl

Abstract. In the law, it is generally acknowledged that there are intuitive differences between reasoning with rules and reasoning with principles. For instance, a rule seems to lead directly to its conclusion if its condition is satisfied, while a principle seems to lead merely to a reason for its conclusion. However, the implications of these intuitive differences for the logical status of rules and principles remain controversial.

A radical opinion has been put forward by Dworkin (1978). The intuitive differences led him to argue for a strict logical distinction between rules and principles. Ever since, there has been a controversy whether the intuitive differences between rules and principles require a strict logical distinction between the two. For instance, Soeteman (1991) disagrees with Dworkin's opinion, and argues that rules and principles cannot be strictly distinguished, and do not have a different logical structure.

In this paper, we claim that the differences between rules and principles are merely a matter of degree. We give an integrated view on rules and principles in which rules and principles have the same logical structure, but different behavior in reasoning. In this view, both rules and principles are considered to consist of a condition and a conclusion. The observed differences between rules and principles are, in our view, the result of different types of relations that they have with other rules and principles. In the integrated view, typical rules and typical principles are the extremes of a spectrum.

We support our claim by giving an explicit formalization of our integrated view using the recently developed formal tools provided by Reason-Based Logic. As an application of our view on rules and principles, we give three ways of reconstructing reasoning by analogy.

⋆ This paper integrates two previously published papers, viz. Verheij's (1996b) paper on rules and principles, and Verheij and Hage's (1994) paper on reasoning by analogy. These were also the basis for chapter 3 of Verheij's (1996a, p. 43ff.) dissertation. The dissertation was written with Professor H.J. van den Herik as advisor and Dr J.C. Hage as co-advisor.

1. Reasoning with Rules vs. Reasoning with Principles

Legal reasoning can be warranted by rules and principles.* As a starting point, we consider both rules and principles to consist of a condition and a conclusion. It seems that rules and principles lead to two types of reasoning:

- *Reasoning with rules*
  If the condition of a rule is satisfied, the rule is applied and its conclusion follows directly.
- *Reasoning with principles*
  In contrast with a rule, a principle only gives rise to a reason for its conclusion if it applies. Moreover, there can be other applying principles that give rise to both reasons for and reasons against the same conclusion. A conclusion then only follows by weighing the pros and cons.

Dworkin (1978, p. 22ff. and 71ff.) has made a strict distinction between rules and principles in the field of law.** According to Dworkin, rules have an all-or-nothing character, while principles have a dimension of weight or importance. An example of a typical rule, he says, is 'A will is invalid unless signed by three witnesses'. An example of a typical principle is 'No man may profit from his own wrong'.‡

There are at least three seeming differences between rules and principles. The first is that rules lead directly to their conclusion if they are applied, while principles lead to their conclusion in two steps: first principles give rise to reasons, then these reasons are weighed before a conclusion is drawn.

The second difference between rules and principles appears in the case of a conflict. When rules conflict, i.e., when rules with incompatible conclusions apply to a single case, the rules lead directly to their conclusions, and therefore to a contradiction. When principles conflict, i.e., when principles with incompatible conclusions apply to a single case, no such problems occur. The application of conflicting principles only leads to reasons that plead for incompatible conclusions, so no contradiction is involved. In such cases, a conflict can involve several distinct

* Rules and principles correspond to Toulmin's (1958) warrants of arguments.

** Dworkin (1978, p. 22) identifies not only rules and principles, but also policies, which he characterizes as standards that set out a goal to be reached. As Dworkin remarks, there are close relations between principles and policies, but the distinction frequently collapses. For us, it suffices to note that in a concrete situation it is possible to construct a specific principle corresponding to a policy. For instance, the policy that the inhabitants renting a house should be protected corresponds to the principle that if some act, such as the continuation of the renting contract, protects the inhabitants, the act should be performed. Recently, Hage (1996, 1997) has elaborated on the relation between principles and goals.

‡ As Soeteman (1991, p. 33) notes, the usage of the terms 'rule' and 'principle' is not at all uniform. For instance, 'Ne bis in idem' is called a principle, but has a rule-like nature, while 'A contract must be executed in good faith' is a principle-like rule. Here, we will not deal with the usage of the terms 'rule' and 'principle', but with the nature of typical rules and principles.

Table I. Three seeming differences between rules and reasons.

	Rule	*Principle*
Application	Conclusion	Reason
Conflict	Contradiction	Weighing
Other rules and principles	Independent	Dependent

reasons, some of which plead for a conclusion, others against it. Weighing the pros and cons determines the final conclusion.

The third difference is that rules are independent of other rules and principles and lead to their conclusion in isolation, while principles interact with other principles. For instance, additional reasons arising from other principles can influence the result of the weighing of the reasons.

The three differences are summarized in Table I.

The seeming differences lead to the question whether rules and principles are logically different. There is no agreement. For instance, Dworkin has a strong opinion:

> 'The difference between legal principles and legal rules is a logical distinction' (Dworkin, 1978, p. 24)

Soeteman (1991), in his discussion of rules and principles, takes an apparently opposite stand:

> 'I know of no difference in logical structure between rules and principles.' (Soeteman, 1991, p. 34)*

Apart from differences, there are also clear similarities between rules and principles. We mention two of them. First, rules and principles both are basically a connection of some sort between a *condition* and a *conclusion.*** The difference is only that, in the case of a rule, this connection seems stronger than in the case of a principle.

Second, for a rule or principle *in isolation* the difference disappears. In isolation, the conclusion of both a rule and a principle follows if the condition is satisfied.

These similarities led us to the claim that the seeming differences between rules and principles are merely a matter of degree. There is no clear border between

* Translation from the original in Dutch: 'Ik ken (...) geen verschil in logische structuur tussen regels en beginselen.'

** Often, the wording of a rule or principle in ordinary language does not clearly distinguish a condition and a conclusion. E.g., in the rule 'Someone who commits a tort against someone else has to repair the resulting damages' no condition or conclusion is distinguished. However, the rule can be rephrased as 'If someone commits a tort against someone else, (s)he has to repair the resulting damages', in which the condition and conclusion are clearly distinguished. Rephrasing a rule or principle in this way is not a trivial exercise, but is not at issue in this paper.

reasoning with rules and principles. They are just the two extremes of a spectrum. A similar claim has been made before, e.g., by Soeteman (1991). His argument is based on classical logic. However, in his account, the observed differences between rules and principles have no explicit counterpart. Sartor (1994, p. 189) notes that in a non-classical logic that acknowledges the defeasibility of all norms the distinction between rules and principles tends to be eliminated, or at least that its conceptualization as a structural difference is overcome. However, just as in Soeteman's account, the seeming differences between rules and principles have disappeared.‡

The contribution of this paper is a new integrated view on rules and principles. In the view, rules and principles are the extremes of a spectrum, and the observed differences between rules and principles can be explained by considering the extremes of the spectrum. In order to support the view, it is made explicit by giving an integrated formal representation of rules and principles.

In recent years, several logical 'tools' – an appropriate term used by Prakken (1993) – have been developed that can be used to give a more satisfactory account of rules and principles. Especially dealing with the applicability of rules/principles,⋆ priority relations between rules/principles and reasoning about rules/principles in logic is currently better understood (see, e.g., Prakken, 1993; Hage and Verheij, 1994; Prakken and Sartor, 1995; Yoshino, 1995). The account in this paper uses the tools that are available in Reason-Based Logic (see, e.g., Hage, 1996, 1997; Verheij, 1996a). Using these tools, we are able to give an integrated formal representation of rules and principles. As an application of our view on rules and principles, we give three ways of reconstructing reasoning by analogy.

The paper is organized as follows. We start with an informal presentation of the spectrum of rules and principles (Section 2). Then we discuss the formal tools of Reason-Based Logic (Section 3). After that, the integrated view on rules and principles is formally elaborated (Section 4). The view is applied to reasoning by analogy in Section 5. Section 6 contains the conclusion of the paper.

2. The Spectrum of Rules and Principles

Our integrated view on rules and principles is based on two main assumptions:

- Both rules and principles give rise to reasons if they are applied.
- The differences between reasoning with rules and reasoning with principles result from different types of relations with other rules and principles.

‡ According to Alexy (1979, p. 64 ff.), there are three possible points of view: (1) rules and principles should be strictly distinguished (e.g., Dworkin, 1978); (2) rules and principles coincide, in the sense that any logical characteristic of principles can also occur for rules, and vice versa (e.g., Sartor, 1994); (3) rules and principles should be weakly distinguished, i.e., the differences are a matter of degree (e.g., Soeteman, 1991).

⋆ We use the notation 'rule/principle' (plural: 'rules/principles') for the hybrid of a rule and a principle.

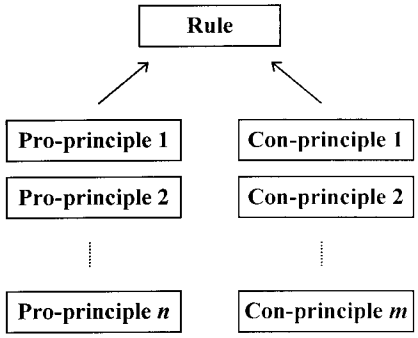

Figure 1. A rule and its underlying principles.

As a basic example of the role of the relations between rules and principles, we discuss a rule and its underlying principles (Section 2.1). Then we discuss our view on a typical rule (Section 2.2), a typical principle (Section 2.3), and a hybrid rule/principle (Section 2.4).

2.1. A RULE AND ITS UNDERLYING PRINCIPLES

If the legislator makes a legal rule, the result is often based on a decision in which several factors are taken into account. These factors, or reasons, are based on principles. If the reasons are in conflict, the legislator decides (either explicitly or implicitly) how they have to be weighed. The principles taken into account by the legislator are said to *underlie* the newly made legal rule. Figure 1 depicts the situation. The principles underlying the rule that can lead to a reason for the conclusion of the rule are indicated as pro-principles, those that can lead to a reason against the conclusion are indicated as con-principles.

As an example, we take the rule from Dutch civil law that the sale of a house should not terminate an existing rent contract (Art. 7A:1612 BW).* This rule has as underlying principles that somebody who lives in a house should be protected against measures that threaten the enjoyment of the house, and that contracts only bind the contracting parties. The first principle pleads against termination of an existing rent contract; the second pleads for termination. As a result, there is (at least) one underlying pro-principle, and one underlying con-principle.

What happens when the legal rule applies? Of course, its underlying principles should normally not also be applicable since the legislator has already considered them. The applying rule is assumed to *replace* its underlying principles. As a result, if the rule of Art. 7A:1612 BW applies, its two underlying principles should not be applicable. The situation is shown in Figure 2.

If the applying rule would not replace its underlying principles, several reasons would arise that already had been taken into account in the rule itself. Due to the

* This example is also discussed by Prakken (1993, pp. 22–23).

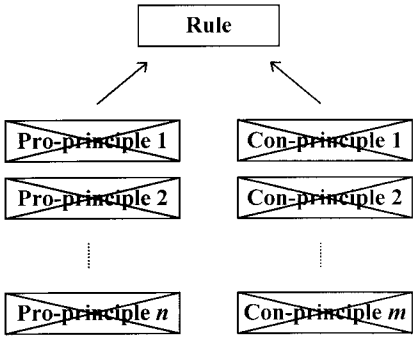

Figure 2. A rule replaces its underlying principles when it applies.

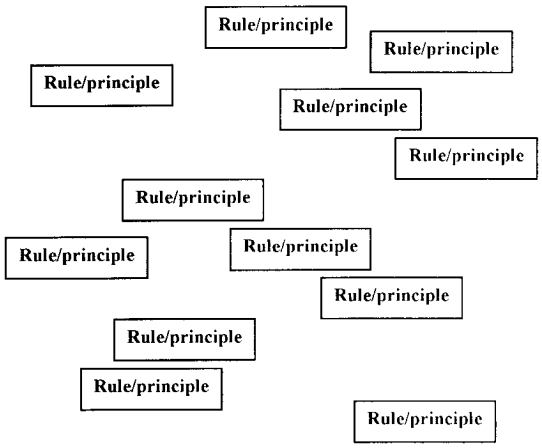

Figure 3. Interfering rules and principles.

special relations of the rule with its underlying principles, the principles should not be applicable.

2.2. A TYPICAL RULE

The relations between rules and principles can be less clear than in the example of a rule and its underlying principles. In the following, we focus on the set of rules and principles that interfere without specifying these relations (Figure 3).

Assume that the rule/principle in the upper left corner is actually a typical rule. Our view on rules and principles implies that, if this typical rule applies, it should block the application of all interfering rules/principles. This situation is shown in Figure 4.

As a result, the conclusion of the rule follows directly.

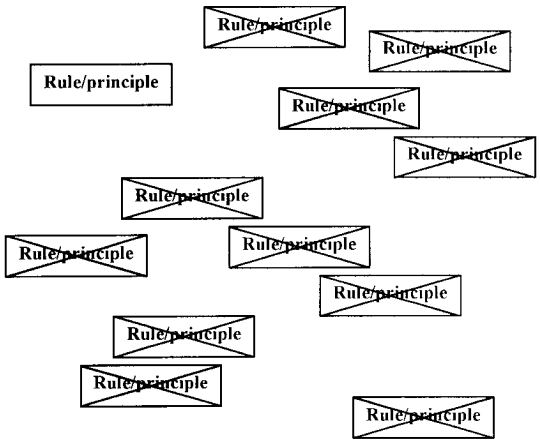

Figure 4. A typical rule applies.

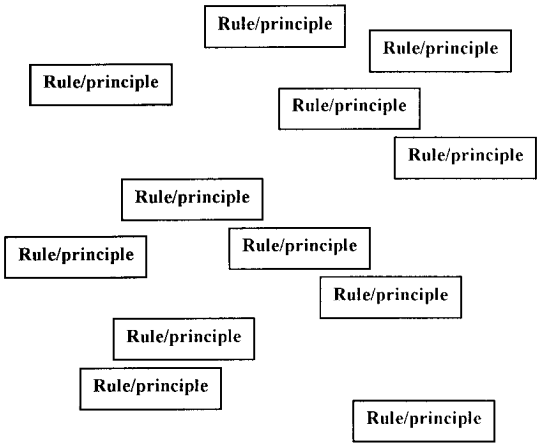

Figure 5. A typical principle applies.

2.3. A TYPICAL PRINCIPLE

If the rule/principle in the upper left corner is a typical principle, it should not block any of the interfering rules/principles in case it applies. The situation is shown in Figure 5.

As a result, the conclusion of the principle does not follow directly, but only after weighing the reasons arising from the other rules/principles.

2.4. A HYBRID RULE/PRINCIPLE

Typical rules and typical principles are the extreme cases. Most rules/principles are hybrid: they are neither a typical rule, nor a typical principle. A hybrid rule/principle blocks some, but not all interfering rules/principles. The situation

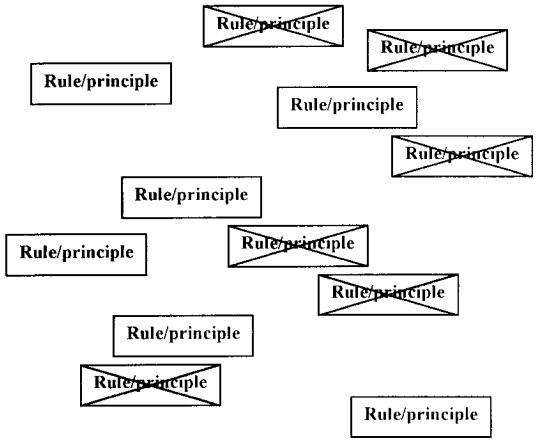

Figure 6. A hybrid rule/principle applies.

that the rule/principle in the upper left corner is a hybrid rule/principle which applies is shown in Figure 6.

As a result, the conclusion of the hybrid rule/principle does not follow directly, but only after weighing the reasons arising from the other rules/principles that are not blocked.

In Section 4, this informal sketch of an integrated view on rules and principles will be formalized using the formal tools provided by Reason-Based Logic. These are introduced in the next section.

3. Reason-Based Logic: A Brief Overview

Below, we give a brief overview of Reason-Based Logic (RBL), based on the description by Verheij (1996a, Chapter 2). This logic contains several formal tools that are useful to deal with exceptions to rules and principles, the weighing of reasons and reasons for and against the application of rules and principles.[⋆] In the language of Reason-Based Logic, several specific types of facts concerning rules and principles can be expressed (Section 3.1). The relations between these types of facts are discussed in Section 3.2.[⋆⋆] We do not discuss the monotonic and nonmonotonic consequence relations of Reason-Based Logic (see Verheij, 1996a, p. 38ff.).[‡]

[⋆] It is well-known that classical logic does not suffice to deal with such issues. For additional information, the reader is referred to Hage (1997) and Verheij (1996a).

[⋆⋆] Over the years, there have been many versions of Reason-Based Logic. Hage (1991) started the development of Reason-Based Logic; later it was continued in cooperation with Verheij. Hage and Verheij (1994) describe the first version that is formally satisfactory. Recent full descriptions of Reason-Based Logic are given by Hage (1996, 1997) and Verheij (1996a).

[‡] The relevance of nonmonotonic reasoning with defeasible arguments for law is, e.g., discussed by Prakken (1993, 1995), Hage (1996, 1997), Peczenik (1996), and Prakken and Sartor (1996).

3.1. TYPES OF FACTS

The language of Reason-Based Logic is basically a classical first-order language, with some adaptations to express specific types of sentences.

Since we do not only need to *express* states of affairs by means of sentences, but also to *refer* to them in other sentences, we assume a translation from logical sentences to logical terms.[⋆] Any sentence begins with an uppercase character, and any term with a lowercase. Each sentence translates to a term by changing its initial uppercase to a lowercase. The translation extends to the metavariables, written in italics.

For example, the sentence

Is_thief(mary)

translates to the term

is_thief(mary).

The logical connectives are treated as if they also are function symbols. In this way, the translation can be kept as simple as it is now. For example, the sentence

Is_guilty(mary) $\wedge$ $\neg$Punish(mary)

translates to the term

is_guilty(mary) $\wedge$ $\neg$punish(mary).

To stay as close as possible to the usual notation of sentences, the logical connectives are *infix* function symbols. For instance, instead of writing terms of the form $\wedge(term_1, term_2)$, we write $term_1 \wedge term_2$.

In Reason-Based Logic, a number of function and predicate symbols are used to express specific types of facts concerning rules and reasons. Below we provide an overview of these function and predicate symbols and their use.

- rule(*condition*, *conclusion*)
  As stated before, we consider both rules and principles basically as a relation between a condition and a conclusion,[⋆⋆] corresponding to Toulmin's (1958) warrants of arguments (see note 5).[‡] Since we treat rules and principles as objects, they are denoted in Reason-Based Logic as terms. In this way it is possible to express facts about rules. Since in our view on rules and principles, there is no structural difference between a rule and a principle, both are denoted as a term of the form rule(*condition*, *conclusion*). Here *condition* and *conclusion*

[⋆] This is related to an often-encountered technique, known as *reification*. We refer to the overview of meta-languages, reflection principles and self-reference by Perlis and Subrahmanian (1994).

[⋆⋆] Another approach is taken by Hage (1997, p. 134ff.), who uses different logical structures for rules and principles.

[‡] Toulmin (1958) also discusses backings of warrants. Loui and Norman (1995) give a partial taxonomy of types of rationales for the adoption of rules. Their rationales correspond to Toulmin's backings of warrants. It would be interesting to investigate how the different types of rationales influence the relations between rules and principles, as described in this paper.

are terms with free variables. As we will see, the distinction between rules and principles stems from their relations with other rules/principles. We call a formal rule/principle an RBL rule when we need to distinguish it from its natural language counterpart.

- {*fact*$_1$, *fact*$_2$, . . . , *fact*$_n$} (for $n = 1, 2, \ldots$)
  Symbols of this form are used to refer to the sets of facts that are reasons for some conclusion. We use an unusual syntax of terms to stay as close as possible to the normal notation of sets. The term {thief(mary), minor(mary)} refers to the set of the two reasons expressed by the sentences Thief(mary) and Minor(mary). The term { } is used to denote an empty set of reasons.

  There is a problem here with different terms that denote identical sets, such as {thief(mary), minor(mary)} and {minor(mary), thief(mary)}. Axioms should be included in Reason-Based Logic to the effect that formulas that are equal up to different terms denoting identical sets, are logically equivalent. We don't do this explicitly.

  We do not consider infinite sets of reasons.
- Reason(*fact*, *state_of_affairs*)
  A sentence of this form expresses that the fact referred to by the term *fact* is a reason for the state of affairs referred to by the term *state_of_affairs*. The sentence *Fact* translates to the term *fact*, the sentence *state_of_affairs* translates to the term *state_of_affairs*. If *State_of_affairs* is an atom *Atom*, *Fact* is a reason for *Atom* and a reason against ¬*Atom*; similarly, if *State_of_affairs* is a negated Atom ¬*Atom*, *Fact* is a reason for ¬*Atom* and a reason against *Atom*.
- Valid(rule(*condition*, *conclusion*))
  A sentence of this form expresses that the rule/principle with condition *condition* and conclusion *conclusion* is valid.
- Excluded(rule(*condition*, *conclusion*), *fact*, *state_of_affairs*)
  A sentence of this form expresses that the rule/principle with condition *condition* and conclusion *conclusion* is excluded, for the instance *Fact* of the rule's condition *Condition*. *Fact* must be an instance of *Condition*, and *State_of_affairs* the corresponding instance of *Conclusion*.
- Applicable(rule(*condition*, *conclusion*), *fact*, *state_of_affairs*)
  A sentence of this form expresses that the rule/principle with condition *condition* and conclusion *conclusion* is made applicable by the fact expressed by the term fact. If a rule/principle is applicable, it may give rise to a reason for the state of affairs expressed by the term *state_of_affairs*. *Fact* must be an instance of *Condition*, and *State_of_affairs* the corresponding instance of *Conclusion*.
- Applies(rule(*condition*, *conclusion*), *fact*, *state_of_affairs*)
  A sentence of this form expresses that the rule/principle with condition *condition* and conclusion *conclusion* applies on the basis of the fact expressed by *fact* and therefore generates a reason for the state of affairs expressed by *state_of_affairs*. *Fact* must be an instance of *Condition*, and *State_of_affairs* the corresponding instance of *Conclusion*. The predicate Applies should not

be confused with the predicate Applicable. The difference in meaning is made precise in Section 3.2.

- Outweighs(*reasons_pro*, *reasons_con*, *state_of_affairs*)
  A sentence of this form expresses that the reasons in the set referred to by the term *reasons_pro* outweigh the reasons in the set referred to by the term *reasons_con* (as reasons concerning *state_of_affairs*). The terms *reasons_pro* and *reasons_con* must both have the form $\{fact_1, fact_2, \ldots, fact_n\}$, where $n \geq 0$. The reasons in *reasons_pro* are reasons for *State_of_affairs*, and the reasons in *reasons_con* are reasons against *State_of_affairs*. Equivalently, if *Not_state_of_affairs* is the literal that is the opposite of *State_of_affairs*, the reasons in *reasons_pro* are reasons against *Not_state_of_affairs*, and the reasons in *reasons_con* are reasons for *Not_state_of_affairs*.

3.2. RELATIONS BETWEEN FACTS

In this section, we describe the relations that hold between the described facts concerning rules and reasons. We do it in terms of the truth values of the corresponding sentences. The basis is again the semantics of first-order logic. For instance, the following relations hold:

NOT
For all sentences *State_of_affairs*,
Either *State_of_affairs* is true or $\neg$*State_of_affairs* is true.

AND
For all sentences $State\_of\_affairs_1$ and $State\_of\_affairs_2$,
$State\_of\_affairs_1$ is true and $State\_of\_affairs_2$ is true if and only if
$State\_of\_affairs_1 \wedge State\_of\_affairs_2$ is true.

OR
For all sentences $State\_of\_affairs_1$ and $State\_of\_affairs_2$,
$State\_of\_affairs_1$ is true or $State\_of\_affairs_2$ is true if and only if
$State\_of\_affairs_1 \vee State\_of\_affairs_2$ is true.

The relations that hold between sentences that are typical for Reason-Based Logic are defined in a similar way. They are called VALIDITY, EXCLUSION, APPLICABILITY, APPLICATION, WEIGHING, and WEIGHING_AXIOMS.* We assume in the following that all mentioned sentences are well-formed, i.e., are sentences of the language of Reason-Based Logic.

VALIDITY
For all sentences *Condition*, *Conclusion*, *Fact* and *State_of_affairs*,
If Excluded(rule(*condition*, *conclusion*), *fact*, *state_of_affairs*),

* These relations could also be given as a set of axioms. We have chosen the present form in order to stress that in Reason-Based Logic the standard logical connectives, such as $\neg$ and $\wedge$, are not treated differently from the non-standard logical constants, such as Valid and Applicable.

Applicable(rule(*condition*, *conclusion*), *fact*, *state_of_affairs*) or Applies(rule(*condition*, *conclusion*), *fact*, *state_of_affairs*) is true, then Valid(rule(*condition*, *conclusion*)) is true.

Informally, VALIDITY says that a rule/principle can only be excluded, be applicable, or apply if it is valid.

EXCLUSION

For all sentences *Fact* and *State_of_affairs*,

If *Fact* and Valid(rule(*condition*, *conclusion*)) are true, then either Excluded(rule(*condition*, *conclusion*), *fact*, *state_of_affairs*) or Applicable(rule(*condition*, *conclusion*), *fact*, *state_of_affairs*) is true.

Informally, EXCLUSION says that a rule/principle is either excluded or applicable if its condition is satisfied. Here *Fact* stands for the fact that satisfies the condition of the rule/principle.

APPLICABILITY

For all sentences *Fact* and *State_of_affairs*,

a. Applicable(rule(*condition*, *conclusion*), *fact*, *state_of_affairs*) is true if and only if Reason(*fact*, applies(rule(*condition*, *conclusion*), *fact*, *state_of_affairs*)) is true.
b. If Applicable(rule(*condition*, *conclusion*), *fact*, *state_of_affairs*) is true, then *Fact* is true.

Informally the first part of APPLICABILITY says that if and only if a rule/principle is applicable, the fact that makes the rule/principle applicable is a reason to apply the rule/principle. The second part says that a rule/principle can only be applicable if its condition is satisfied. Again, *Fact* stands for the fact that satisfies the condition of the rule/principle.

APPLICATION

For all sentences *Fact* and *state_of_affairs*,

There are terms *condition* and *conclusion*, such that Applies(rule(*condition*, *conclusion*), *fact*, *state_of_affairs*) is true if and only if Reason(*fact*, *state_of_affairs*) is true.

Informally this relation says that if and only if a rule/principle applies, the fact that makes the rule/principle applicable is a reason for the rule/principle's (instantiated) conclusion, or, equivalently, a reason against the opposite of the rule/principle's conclusion.

Notice the difference between a rule/principle's being applicable and its being applied. If a rule/principle is applicable, this only indicates that there is a reason for applying the rule/principle (see APPLICABILITY, part a). In general, there can also be reasons against applying a rule/principle.

WEIGHING

For all sentences Pro_1, Pro_2, ..., Pro_n (for some natural number n), Con_1, Con_2, ..., Con_m (for some natural number m), *State_of_affairs*, and its opposite *Not_state_of_affairs*,

If Reason(pro_1, *state_of_affairs*), Reason(pro_2, *state_of_affairs*), ..., Reason(pro_n, *state_of_affairs*), Reason(con_1, *Not_state_of_affairs*), Reason(con_2, *Not_state_of_affairs*), ..., Reason(con_m, *Not_state_of_affairs*), and also Outweighs({pro_1, pro_2, ..., pro_n}, {con_1, con_2, ..., con_m}, *state_of_affairs*) is true, then *State_of_affairs* is true, or there is a term *con*, different from con_1, con_2, ..., and con_m, such that Reason(*con*, *not_state_of_affairs*) is true.

Informally the first part of this relation says that reasons make a conclusion true if the pros outweigh the cons, provided that no con is overlooked. It is allowed that one or more of the pros is overlooked: if a subset of the pros already suffices to outweigh all cons, the conclusion certainly follows if there are even more pros. It may seem that a similar relation between facts is required for the case that the cons outweigh the pros. However, since in Reason-Based Logic a reason against a state of affairs is just a reason for the opposite state of affairs, the relation above suffices.

WEIGHING_AXIOMS

For all sentences $Fact_1$, $Fact_2$, ..., $Fact_n$ (for some positive natural number n), *State_of_affairs*, and its opposite *Not_state_of_affairs*, and all terms *pros* and *cons*,

a. Outweighs(*pros*, *cons*, *state_of_affairs*) and Outweighs(*cons*, *pros*, *not_state_of_affairs*) are not both true.
b. If Reason($fact_1$, *state_of_affairs*), Reason($fact_2$, *state_of_affairs*), ..., Reason($fact_n$, *state_of_affairs*) are true, then Outweighs({$fact_1$, $fact_2$, ..., $fact_n$}, { }, *state_of_affairs*) is true.

The first part of this relation says that the pros as reasons for *State_of_affairs* cannot outweigh the cons and the other way around at the same time. However, the first weighing axiom does not make it impossible that ¬Outweighs(*pros*, *cons*, *state_of_affairs*) and ¬Outweighs(*cons*, *pros*, *state_of_affairs*) are both true.

Reason-Based Logic does in general not determine which set of reasons outweighs another set. However, for the case that all reasons point in the same direction, i.e., all reasons are either pros or cons, the second part of the relation gives the result: any non-empty set of reasons outweighs the empty one.

4. A Logical Reconstruction of Rules and Principles

We now return to our integrated view on rules and principles, as introduced in Section 2. Recall that our view was based on two assumptions:

- Both rules and principles give rise to reasons if they are applied.
- The differences between reasoning with rules and principles result from different types of relations with other rules and principles, which may interfere.

In Section 4.1, we discuss our basic example, namely a rule with underlying principles. In Section 4.2, we return to the differences between rules and principles as discussed in Section 1.

4.1. A RULE AND ITS UNDERLYING PRINCIPLES

In Section 2.1, we discussed the Dutch legal rule of Art. 7A:1612 BW that sale of a house should not terminate an existing rent contract. This rule can be represented in Reason-Based Logic as follows:

```
Valid(rule(sale_house,
   ought_to_be_done(continuation_contract)))
```

We considered two principles underlying this rule, namely a pro-principle that somebody who lives in a house should be protected against measures that threaten the enjoyment of the house, and a con-principle that contracts only bind the contracting parties. These principles can be represented as RBL rules as follows:

```
Valid(rule(protects_inhabitants(act),
   ought_to_be_done(act)))
Valid(rule(¬party_bound_by_contract,
   ¬ought_to_be_done(continuation_contract)))
```

The fact that these principles underlie the rule of Art. 7A:1612 BW is represented as:

```
Underlies(rule(protects_inhabitants(act),
      ought_to_be_done(act)),
   rule(sale_house,
      ought_to_be_done(continuation_contract)))
Underlies(rule(¬party_bound_by_contract,
      ¬ought_to_be_done(continuation_contract)),
   rule(sale_house,
      ought_to_be_done(continuation_contract)))
```

The rule and its underlying principles are schematically shown in Figure 7.

If a house with renting inhabitants is sold, the two principles lead to conflicting reasons, since continuation of an existing rent contract protects the inhabitants of a house, while the new owner is not bound by the contract with the inhabitants. We have

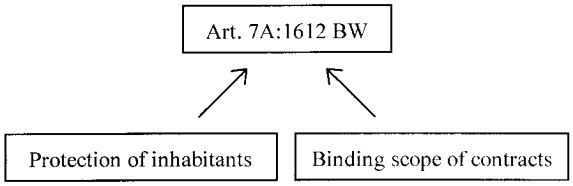

Figure 7. The rule of Art. 7A:1612 BW and its underlying principles.

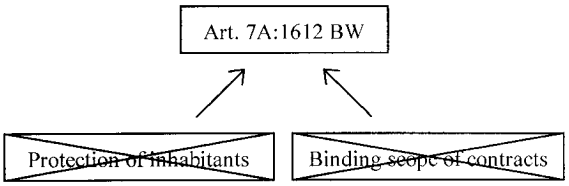

Figure 8. The rule of Art. 7A:1612 BW replaces its underlying principles if it applies.

Protects_inhabitants(continuation_contract)
¬Party_bound_by_contract

and therefore the two RBL rules about the protection of inhabitants and about the binding scope of contracts lead to the conflicting reasons:

Reason(protects_inhabitants(continuation_contract),
  ought_to_be_done(continuation_contract))
Reason(¬party_bound_by_contract,
  ¬ought_to_be_done(continuation_contract))

However, by making the legal rule of Art. 7A:1612 BW, the legislator has balanced the conflicting principles, and decided how the reasons generated by them should be weighed against each other. Therefore, if we have the fact

Sale_house

the rule of Art. 7A:1612 BW should lead to the conclusion

Ought_to_be_done(continuation_contract)

without the interference of the two underlying principles: the rule of Art. 7A:1612 BW replaces its underlying principles if it applies (see Section 2.1), and the two principles should not be applicable. The required situation is shown in Figure 8.

In Reason-Based Logic, replacement can be modeled using exclusionary reasons. We need the following rule:

Valid(rule(underlies($rule_1$, $rule_2$) ∧ applies($rule_2$),
  excluded($rule_1$)))*

* Henry Prakken has correctly noted that $rule_2$ also excludes $rule_1$ in case there is another rule or principle that does not underlie $rule_2$, but nevertheless interferes. As a result, there can be no

Since we can conclude

Applies(rule(sale_house,
ought_to_be_done(continuation_contract)),
sale_house,
ought_to_be_done(continuation_contract))

we find:

Excluded(rule(protects_inhabitants(act),
ought_to_be_done(act)),
protects_inhabitants(continuation_contract),
ought_to_be_done(continuation_contract))
Excluded(rule(¬party_bound_by_contract,
¬ought_to_be_done(continuation_contract)),
¬party_bound_by_contract,
¬ought_to_be_done(continuation_contract))

The principles about the protection of inhabitants and about the binding scope of contracts do no longer lead to reasons. As a result, the rule of Art. 7A:1612 BW leads without interference to the conclusion

Ought_to_be_done(continuation_contract),

just as required.

4.2. WHAT REMAINS OF THE DIFFERENCES BETWEEN RULES AND PRINCIPLES?

We can now finish our integrated view on rules and principles as represented in Reason-Based Logic. As in the case of a rule that replaces its underlying principle, a typical rule is an RBL rule that leads to exclusionary reasons against the applicability of any interfering rule or principle. A typical principle is an RBL rule that does not exclude any interfering rule/principle. Interfering rules and principles are typically rules and principles with equal or opposite conclusion.

This is in line with our two main assumptions:

- Both rules and principles give rise to reasons if they are applied. The difference between the two is that an applying rule not only generates a reason for its conclusion, but also exclusionary reasons for the principles it replaces.
- The differences between reasoning with rules and reasoning with principles result from different types of relations with other rules and principles: rules lead to exclusionary reasons to interfering rules and principles, while principles lead to reasons that are weighed in case of a conflict.

interaction of the other rule or principle with $rule_1$ if $rule_2$ applies. This does not always seem desirable, and deserves further study. Interestingly, in this situation $rule_2$ is not a typical rule.

It is clear that in this view there is no clear border between rules and principles. For instance, an isolated rule cannot be distinguished from an isolated principle. Only if there are interfering rules and principles, gradual differences can be seen. On the one extreme there is the typical principle that, if it applies, does not generate exclusionary reasons for any of the rules and principles that interfere with it. On the other extreme there is the typical rule that, if it applies, excludes all interfering rules and principles. In between the two extremes there are many degrees of hybrid rules/principles, some more principle-like, others more rule-like.

In Section 1, we discussed three differences between rules and principles. First, it seemed that rules lead directly to their conclusion if they apply, while principles lead to reasons that have to be weighed. This difference has disappeared since in our view both rules and principles generate reasons. Therefore both rules and principles first lead to reasons that are then weighed. Nevertheless, also in our view, rules seem to lead directly to their conclusion. This is the result of the fact that in the case of an applying rule no weighing of reasons is necessary since all interfering rules and principles are excluded. Therefore, the step from reason to conclusion seems immediate.

Second, it seemed that conflicting rules lead to a contradiction if they apply, while conflicting principles merely lead to conflicting reasons. In our representation, no real contradiction can arise by the application of rules with opposite conclusions, since rules just as principles only generate reasons. Moreover if an apparent rule gives rise to a reason that conflicts with another reason, this is a sign that it is not a typical rule, but has a somewhat more principle-like character.

Third, it seemed that rules lead to their conclusion in isolation, while principles interact with other principles: additional relevant reasons arising from other principles can influence the result of weighing. In our view, this seeming difference is beside the point since rules in isolation do not differ from principles in isolation. The rule-like character of a rule can only be appreciated if there are interfering rules or principles.

5. An Application to Reasoning by Analogy

The last topic that we discuss is reasoning by analogy. As an application of our integrated view on rules and principles, we describe three different ways of reconstructing* reasoning by analogy.** To avoid misunderstanding, we stress that

* Chalmers et al. (1995, p. 181ff.) are disappointed about much of the current work in the (computational) modeling of analogy since it often bypasses the process of perception. By focusing on reconstructing reasoning by analogy, we do the same, although we agree with Chalmers et al. that reasoning by analogy is highly influenced by the way of perceiving a situation. However, this does not imply that different ways of reconstructing reasoning by analogy cannot be fruitfully studied in their own right.

** Kaptein (1994) claims that in the context of justification reasoning by analogy presupposes no more logical structure than simple modus ponens, and defers the interesting part of reasoning by

our approach to reasoning by analogy is not based on cases,[‡] but on rules and principles. Instead of using the similarity and dissimilarity of cases as criteria to justify reasoning by analogy, we use the relations between rules and principles.

We assume that in reasoning by analogy there is a rule that does not apply because its condition is not satisfied, but that nevertheless its conclusion holds on the basis of additional information about the relations between the rule and other rules and principles. We distinguish three forms of reasoning to analyze reasoning by analogy:

- Application of principles that underlie the original rule that does not apply itself.
- Application of an analogous rule/principle that has the same underlying principles as the original rule that does not apply.
- Analogous application of the original rule, i.e., application of the rule with a 'non-standard' justification, based on, for instance, a principle.

We do not claim that these three forms of reasoning are always cases of reasoning by analogy, but that they are useful means to analyze a given case of reasoning by analogy. Below we use one example, and analyze it by the three mentioned forms of reasoning.

5.1. APPLICATION OF UNDERLYING PRINCIPLES

In the first form of reasoning by analogy, the principles apply that underlie the original rule that does not apply itself.

The example we use is based on Art. 7A:1612 BW. It was also used in the sections 2.1 and 4.1 to explain the replacement of the principles underlying a rule. Again, we have one rule and two underlying principles:

```
Valid(rule(sale_house,
   ought_to_be_done(continuation_contract)))
Valid(rule(protects_inhabitants(act),
   ought_to_be_done(act)))
Valid(rule(¬party_bound_by_contract,
   ¬ought_to_be_done(continuation_contract)))
Underlies(rule(protects_inhabitants(act),
      ought_to_be_done(act)),
   rule(sale_house,
      ought_to_be_done(continuation_contract)))
```

analogy to the context of discovery. However, we reconstruct reasoning by analogy strictly in the context of justification, retaining a deeper logical structure than merely modus ponens.

[‡] See, for instance, Ashley (1990), Yoshino et al. (1993) and Tiscornia (1994).

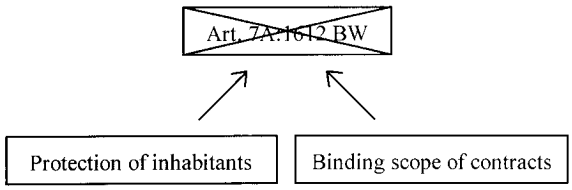

Figure 9. The principles underlying the rule of Art. 7A:1612 BW apply.

Underlies(rule(¬party_bound_by_contract,
¬ought_to_be_done(continuation_contract)),
rule(sale_house,
ought_to_be_done(continuation_contract)))

Here we assume that a house with renting inhabitants is not sold, but donated. So, we have the facts:

¬Sale_house
Donation_house

As a result, the condition of the rule of Art. 7A:1612 BW is not satisfied, and the rule does not apply. But just as in the case of sale, continuation of the existing rent contract is a way to protect the inhabitants, while the new owner is not bound by the existing contract:

Protects_inhabitants(continuation_contract)
¬Party_bound_by_contract

Therefore, the conditions of the principles about the protection of inhabitants and about the binding scope of contracts are satisfied. Since the rule of Art. 7A:1612 BW does not apply, the replacement rule

Valid(rule(underlies($rule_1$, $rule_2$) ∧ applies($rule_2$),
excluded($rule_1$)))

does not give exclusionary reasons for the two underlying principles. They apply and give rise to the reasons:

Reason(protects_inhabitants(continuation_contract),
ought_to_be_done(continuation_contract))
Reason(¬party_bound_by_contract,
¬ought_to_be_done(continuation_contract))

The situation is shown in Figure 9.

So, in the case of donation two reasons arise that are based on the same principles as those taken into account by the legislator, when the original rule was made.

There are good reasons to assume that the weighing of these reasons has the same outcome as in the reasoning of the legislator:

```
Outweighs({protects_inhabitants(continuation_contract)},
    {¬party_bound_by_contract},
    ought_to_be_done(continuation_contract))
```

and leads to the same conclusion that the contract should be continued:

```
Ought_to_be_done(continuation_contract)
```

In this analysis, two principles applied in the case of donation. They are precisely the two principles that were replaced in the case of sale. The case of donation is therefore in a sense of *the same kind* as the case of sale. Therefore we speak of a form of reasoning by analogy. If only some of the underlying principles apply, or more goals and principles are relevant, we cannot always speak of a case of reasoning by analogy. The case might even be solved differently, since the reasons might be weighed differently.

5.2. APPLICATION OF AN ANALOGOUS RULE/PRINCIPLE

In the second form of reasoning by analogy, a analogous rule/principle applies that has the same underlying principles as the original rule. This leads to another analysis of the same example.

In our example the analogous rule/principle might be:

```
Valid(rule(donation_house,
    ought_to_be_done(continuation_contract)))
```

The legal decision maker that wants to base his reasoning on this rule has to justify its validity. This justification can be based on the same reasons as the ones used by the legislator when he made Art. 7A:1612 BW:

```
Reason(protects_inhabitants(continuation_contract),
    valid(rule(donation_house,
        ought_to_be_done(continuation_contract))))
Reason(¬party_bound_by_contract,
    ¬valid(rule(donation_house,
        ought_to_be_done(continuation_contract))))
```

In this line of reasoning, the two reasons are not relevant for the conclusion that the contract should be continued, but for the validity of the new RBL rule about donation. In their new role, the reasons might be weighed the same way as before:

```
Outweighs({protects_inhabitants(continuation_contract)},
    {¬party_bound_by_contract},
    valid(rule(donation_house,
        ought_to_be_done(continuation_contract)))
```

The conclusion is that the RBL rule about donation is valid.

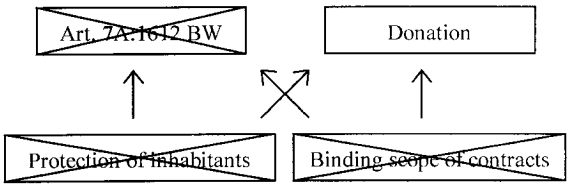

Figure 10. The rule about donation applies having the same underlying principles as the original rule of Art. 7A:1612 BW.

It may seem that there is a problem here with the separation of powers: while the legislator can make rules, the legal decision maker cannot. However, this problem is only seeming, and due to the different meanings of rule validity in law and in reasoning. We use the term 'rule validity' in the latter sense. For rule validity in that sense the separation of powers is irrelevant.*

If the rule about donation applies, the principles about the protection of inhabitants and about the binding scope of contracts are again replaced by the rule about donation and do not apply. An overview of the relations of the rules and principles involved in this reasoning is shown in Figure 10.

Since the rule about donation has the same underlying principles as the rule of Art. 7A:1612 BW we say that a rule is applied analogous to the original rule.

5.3. ANALOGOUS APPLICATION OF THE ORIGINAL RULE

The third form of reasoning by analogy is typical for Reason-Based Logic, since it involves reasons for and against applying a rule.

In this third analysis of the example, the rule of Art. 7A:1612 BW is not applicable, since its condition is not satisfied, just as in the previous two analyses. As a result, the standard reason for applying the rule, based on the relation between facts APPLICABILITY (Section 3.2), does not arise. However, a rule that is not applicable can apply, since there can be other reasons that lead to its application.

In our case, the reasons are again those for and against the continuation of the contract having a new role. They now are represented as follows:

```
Reason(protects_inhabitants(continuation_contract),
   applies(rule(sale_house,
         ought_to_be_done(continuation_contract)),
      sale_house,
      ought_to_be_done(continuation_contract)))
```

* Verheij and Hage (1994) put it differently, and wrote that the legal decision maker can only validate legal principles (and not legal rules) because of the separation of powers. However, in the line of reasoning described in the text the two underlying principles are replaced if the RBL rule about donation applies. Otherwise the reasons arising from these principles would be taken into account twice. As a result, the RBL rule about donation has a rule-like character.

```
Reason(¬party_bound_by_contract,
   ¬applies(rule(sale_house,
         ought_to_be_done(continuation_contract)),
      sale_house,
      ought_to_be_done(continuation_contract)))
```

Here the reasons protects_inhabitants(continuation_contract) and ¬party_bound_by_contract are reasons for and against applying the rule of Art. 7A:1612 BW, respectively. Again the result of weighing these reasons might be the same in this new role, as in Section 5.2:

```
Outweighs(protects_inhabitants(continuation_contract),
   {¬party_bound_by_contract},
   applies(rule(sale_house,
         ought_to_be_done(continuation_contract)),
      sale_house,
      ought_to_be_done(continuation_contract)))
```

As a result, we can conclude that the rule of Art. 7A:1612 BW applies, even though its condition is not satisfied and it is not applicable:

```
Applies(rule(sale_house,
         ought_to_be_done(continuation_contract)),
   sale_house,
   ought_to_be_done(continuation_contract))
```

Since the rule of Art. 7A:1612 BW applies, it replaces its underlying principles by the replacement rule, just as any applying rule: the principles about the protection of inhabitants and about the binding scope of contracts are excluded and do not apply. Figure 8 shows the relations of the rules and principles involved (but does not show the reasons in their new role). These relations are the same as in the case of normal rule application. Since in this example the rule does apply, but not for the standard reason that its condition is satisfied, we call this *analogous* rule application.

6. Conclusion

In this paper, we claimed that the differences between reasoning with rules and principles, as for instance put forward by Dworkin (1978), do not require a strict logical distinction between rules and principles. The distinction is merely gradual, and rules and principles are the extremes of a spectrum. The observed differences between rules and principles can be explained by considering the extremes of the spectrum. We have supported our claim by giving a formal elaboration of an integrated view on rules and principles using the logical tools provided by Reason-Based Logic. As an application of our integrated view on rules and principles, we discussed three different ways of reconstructing reasoning by analogy.

Acknowledgments

The research reported here was partly financed by the Dutch Foundation for Knowledge-based Systems (SKBS) as part of the B3.A project. The first author thanks in particular Ron Loui and the Computer Science Department at the Washington University in Saint Louis (Missouri) for their hospitality in the summer of 1996. His stay there was sponsored by the National Science Foundation (NSF) under number 9503476.

References

Alexy, R. 1979. Zum Begriff des Rechtsprinzips. *Rechtstheorie. Beiheft 1. Argumentation und Hermeneutik in der Jurizprudenz.* In W. Krawietz, K. Opalek, A. Peczenik, and A. Schramm (eds.), Duncker & Humblot, Berlin, pp. 59–87.

Ashley, K. 1990. *Modeling Legal Argument. Reasoning with Cases and Hypotheticals.* The MIT Press, Cambridge (Massachusetts).

Chalmers, D., French, R., and Hofstadter, D. 1995. High-level perception, representation, and analogy: A critique of artificial-intelligence methodology. In D. Hofstadter & the Fluid Analogies Research Group (eds.), *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thoughts.* Basic Books, New York (New York), pp. 169–193.

Dworkin, R. 1978. *Taking Rights Seriously. New Impression with a Reply to Critics.* Duckworth, London.

Hage, J. C. 1991. Monological reason based reasoning. in J. A. Breuker, R.V. de Mulder and J. C. Hage (eds.), *Legal Knowledge based Systems. Model-Based Legal Reasoning.* Vermande, Lelystad, pp. 77–91.

Hage, J. C. 1996. A theory of legal reasoning and a logic to match. *Artificial Intelligence and Law* 4, 199–273.

Hage, J. C. 1997. *Reasoning with Rules. An Essay on Legal Reasoning and Its Underlying Logic.* Kluwer Academic Publishers, Dordrecht.

Hage, J. C. and Verheij, B. 1994. Reason-based logic: a logic for reasoning with rules and reasons. *Law, Computers and Artificial Intelligence* 3(2/3), 171–209. Also published as report SKBS/B3.A/94–10.

Kaptein, H. 1994. The Redundancy of Precedent and Analogy. Or: Eat S∗∗t, Five Billion Flies Can't Be Wrong. *1994 ISSA Congress.*

Loui, R. P. and Norman, J. 1995. Rationales and argument moves. *Artificial Intelligence and Law* 3, 159–189.

Peczenik, A. 1996. Jumps and logic in the law. What can one expect from logical models of legal argumentation? *Artificial Intelligence and Law* 4, 297–329.

Perlis, D. and Subrahmanian, V. S. 1994. Meta-languages, reflection principles and self-reference. In D. M. Gabbay, C. J. Hogger and J. A. Robinson (eds.), *Handbook of Logic in Artificial Intelligence and Logic Programming. Volume 2. Deduction Methodologies.* Clarendon Press, Oxford, pp. 323–358.

Prakken, H. 1993. *Logical Tools for Modelling Legal Argument.* Doctoral thesis, Free University, Amsterdam.

Prakken, Henry 1995. From logic to dialectics in legal argument. *The Fifth International Conference on Artificial Intelligence and Law. Proceedings of the Conference,* ACM, New York, pp. 165–174.

Prakken, H. and Sartor, G. 1995. On the relation between legal language and legal argument: assumptions, applicability and dynamic properties. *The Fifth International Conference on Artificial Intelligence and Law. Proceedings of the Conference,* ACM, New York, pp. 1–10.

Prakken, H. and Sartor, G. 1996. A dialectical model of assessing conflicting arguments in legal reasoning. *Artificial Intelligence and Law* 4, 331–368.

Sartor, G. 1994. A formal model of legal argumentation. *Ratio Juris* 7(2), 177–211.

Soeteman, A. 1991. Hercules aan het werk. Over de rol van rechtsbeginselen in het recht. *Rechtsbeginselen*, Ars Aequi, Nijmegen, pp. 41–56.

Tiscornia, D. 1994. Three meanings of analogical reasoning in law. In I. Carr, and A. Narayanan (eds.), *Proceedings of the Fourth National Conference on Law, Computers and Artificial Intelligence* University of Exeter, pp. 137–153.

Toulmin, S. E. 1958. *The Uses of Argument.* University Press, Cambridge.

Verheij, B. 1996a. *Rules, Reasons, Arguments. Formal Studies of Argumentation and Defeat*. Dissertation Universiteit Maastricht. A summary and table of contents are available on the World-Wide Web at http://www.metajur.unimaas.nl/∼bart/proefschrift/.

Verheij, B. 1996b. An integrated view on rules and principles. In R. W. van Kralingen, H. J. van den Herik, J. E. J. Prins, M. Sergot, and J. Zeleznikow (eds.), *Legal Knowledge Based Systems. Foundations of Legal Knowledge Systems.* Tilburg University Press. Also published as report SKBS/B3.A/96–05, pp. 25–38.

Verheij, B. and Hage, J. C. 1994. Reasoning by analogy: a formal reconstruction. In H. Prakken, A. J. Muntjewerff and A. Soeteman (eds.), *Legal Knowledge Based Systems. The Relation with Legal Theory*. Koninklijke Vermande, Lelystad, pp. 65–78. Also published as report SKBS/B3.A/94-14.

Yoshino, H. 1995. The systematization of legal meta-inference. *The Fifth International Conference on Artificial Intelligence and Law. Proceedings of the Conference*, ACM, New York, pp. 266–275.

Yoshino, H., Haraguchi, M., Sakurai, S. and Kagayama, S. 1993. Towards a legal analogical reasoning system: knowledge representation and reasoning methods. *The Fourth International Conference on Artificial Intelligence and Law. Proceedings of the Conference*, ACM, New York, pp. 110–116.