

Goal-based theory evaluation

Jaap HAGE

*University of Maastricht, Department of Metajuridica, P.O Box 616,
6200 MD Maastricht. E-mail: jaap.hage@metajur.unimaas.nl*

Abstract: This paper describes how part of Dworkin's Model of Principles, a coherence theory of the law, can be made sufficiently precise to capture it in a logical theory. The result is formalised by means of an addition to Reason-based Logic.

1 Introduction

In his paper 'Some arguments about legal arguments', Thorne McCarty [1] pleads for a theory construction approach in AI and Law research. Legal materials should not be taken at face value, but are amenable to reconstruction. This line of argument is well in line with a teleological tradition in legal theory that has historical roots in the *Interessenjurisprudenz* around 1900 [2, 3] and which is more recently exemplified by the work of, amongst others, Dworkin [4, 5, 6], MacCormick [7], Alexy and Peczenik [8], which emphasises that the law should be constructed as a coherent whole.

In this paper I follow McCarty's suggestion and contribute to a theory construction approach of legal reasoning. In particular I propose a formalisation of one part of Dworkin's version of a coherence theory of the law. In section 2 this version is described as 'The Model of Principles'. In section 3 of this paper I describe part of Dworkin's model informally. Section 4 provides a formal characterisation of this same part, while section 5 puts the formalisation to use in an example.

2 The Model of Principles¹

In a series of papers², culminating in his book *Law's Empire*, Dworkin has developed a coherence theory of the law. This theory, called *The Model of Principles*, recognises three stages in constructing the law [6, pp. 65f]. The first stage, the so-called pre-interpretative stage, consists of a preliminary identification of the rules, standards, and (generalised) decisions that make up the law. In this connection one might think of an inventory of the rules and standards that can be found in statutes, cases, and doctrinal literature. The second, interpretative, stage consists of an identification of the principles (in this connection including values and policies) that underlie (in the sense of explain), or are part of the legal phenomena identified in the first stage. The rules etc. identified in the first stage are to be seen as means to realise the principles identified in the second stage.³

The rules identified in the first stage are often not the best way to realise the principles identified in the second stage. The purpose of the third, reforming, stage is to formulate (relevant parts of) the set of rules, including (generalised) decisions of cases, that best realises the principles identified in the second stage.⁴

The Model of Principles counts two steps. In the first, constructive, step, the raw legal materials identified in the pre-interpretative stage are used to formulate a set of principles that explains them. In the second, reconstructive step the raw materials are replaced by the set of rules that provides the best realisation of the formulated principles. In this paper I will only

pay attention to this second step, and even only to a part of it. I will not concern myself with the formulation of sets of rules that realise sets of principles⁵, but will restrict myself to the formulation of standards that can be used in evaluating competing sets of rules and case decisions in the light of the principles that they aim to realise.

3 An intuitive account of rule evaluation

In his paper *Zum Begriff des Rechtsprinzips* and his book *Theorie der Grundrechten*, Alexy [11, 12] developed the view that legal principles are ‘Optimierungsgebote’. With this he meant that legal principles are a kind of goals which are to be realised as much as possible. There is, however, a tension between goals, because the realisation of the one goal may detract from the realisation of another goal. Take for instance the question how strict ecological rules should be. The stricter the rules, the more the goal to protect the environment is realised. Weaker ecological rules contribute less to the environment, but may contribute more to economic prosperity, which is another goal. The choice which ecological rules should be adopted, depends on the balance between, amongst others, these two goals.

In general it may be said that if (the adoption of) a regulation contributes to a goal that is accepted in the legal system at issue, this fact is a reason that pleads for this regulation. This reason is the stronger, the more important the goal and the contribution of the regulation to the goal are. Analogously, if a regulation detracts from a goal, this is a reason that pleads against the regulation. Again it holds that this reason is the stronger, the more important the goal is and the more the regulation detracts from the goal. In this connection I take a regulation to be a set of one or more related rules.

Given a set of goals and a regulation, the goals can be divided into three subsets. Some goals are promoted by the regulation, others are detracted from, and some are unaffected. The goals that are promoted by the regulation provide reasons that plead for the regulation. The goals that are detracted from by the regulation provide reasons that plead against the regulation. The remaining goals do not lead to reasons for or against the regulation. On the basis of the goals and a regulation one can therefore construct a set of reasons that plead for the regulation, and a set of reasons that plead against it.⁶

Assuming that there are no other reasons pleading for or against the regulation, the regulation can be evaluated by weighing the reasons that plead for and against it. If the reasons that plead for the regulation outweigh the reasons against it, the regulation is said to be *acceptable*. If the reasons against outweigh the pro-reasons the regulation is called *unacceptable*. If neither the pro nor the con reasons outweigh the other ones, the regulation is *indifferent*.

Very often there will be different ways in which a subject can be regulated. The different possible regulations for the same subject can be competitors in the sense that adoption of one of them precludes adoption of the other ones.⁷ As a consequence it is useful to distinguish another way of evaluating regulations, next to the question whether they are acceptable. The other question at issue is which regulation is the best one of the competing regulations for one and the same subject. Normally, a regulation should only be introduced in a legal system if it is both acceptable, and if there is no better one. Such regulations are called *preferable*.⁸

The question whether some set of reasons outweighs another set is not always easy to answer. Reasons usually do not come labelled with a weight, and neither is it the case that a set of reasons with more elements always outweighs sets of reasons with less elements. In general it depends on contingent additional information which set of reasons outweighs which other sets. I call this kind of information *weighing knowledge* [14, pp. 140f.].

As examples in this paper I will use variations on the so-called Lebach-case, which was made familiar by Alexy [11, 12]. The standard case runs as follows: A person, let us call him

E, who was condemned for abduction and subsequent murder of his victim is released from prison after ten years. A tabloid journal jumps on this news and uses the occasion to publish an article on the dangers of abduction in general. The article is illustrated with a photograph of E just after his release. E attempts to prevent circulation of the journal.

The judge who must decide on this case should balance two principles (goals). One is the principle of freedom of the press, the other one is the principle that one should respect other persons' privacy. Let us assume that the judge decides that in cases like this, (the reason based upon) privacy protection outweighs (the reason based upon) freedom of the press. This decision amounts to the adoption of certain weighing knowledge.

There are a few exceptions to the general observation that it depends on contingent information which set of reasons outweighs which other sets. One is that an empty set of reasons is outweighed by any non-empty set. So if there are only reasons that plead for a regulation, and no reasons pleading against it, the regulation is acceptable. Similarly, if there are only reasons that plead against a regulation, and no reasons pleading for it, the regulation is unacceptable.

Other exceptions to the observation that the relative weight of sets of reasons is contingent information depend on a fortiori reasoning with respect to already available weighing knowledge. The best way to explain these exceptions is by means of examples.

Suppose that we have two sets of reasons, PRO_1 and CON_1 pleading respectively for and against a regulation REG_1 , where PRO_1 outweighs CON_1 . Suppose moreover that we have another regulation REG_2 , with as pro- and con-reasons respectively PRO_2 and CON_2 . CON_1 and CON_2 are identical, meaning that the regulations subtract from the same goals in the same degree. PRO_1 and PRO_2 differ from each other, however. PRO_2 contains the same reasons as PRO_1 , with the same weight, but it also contains one or more other reasons that plead for REG_2 . The set of pro-reasons has consequently become stronger in comparison to the reasons for REG_1 , while the set of con-reasons has remained the same. Since the reasons pleading for REG_1 already outweighed the reasons against it, one can conclude a fortiori that the stronger set of reasons pleading for REG_2 outweighs the reasons against REG_2 . So, if REG_1 is acceptable, REG_2 is normally⁹ acceptable too. In fact, REG_2 will even be a better regulation than REG_1 .¹⁰

Let us assume that in our example case, the judge decides that privacy protection outweighs freedom of the press. As a consequence a regulation to the effect that, *ceteris paribus*, it is not allowed to publish photographs of recently released prisoners in a context that connects them with the crime they committed, is an acceptable one.

Now let me change the case a little bit by adding the fact that the news that E was to be released was given to the press on the condition that no photographs would be taken. This forms an additional reason against publication of the article, because the effects of the offence are even enlarged by publishing the photo that was illegally taken. As a consequence the regulation that, *ceteris paribus*, it is not allowed to publish photographs of recently released prisoners in a context that connects them with the crime they committed if the potential publisher undertook the obligation not to take photographs at all, is an even better regulation, in the sense that there is even more support for it. The two regulations are not competitors however, in the sense that adoption of the one precludes adoption of the other one. Therefore the stronger support for the second regulation has no implications for the support for the first one.

It is possible to think of another change in the case which leads to a different conclusion. As yet, the question whether the released prisoner objects against the publication has not been taken into consideration yet. It was tacitly assumed that he did object, but this need not be the case, in particular not if he were to be compensated financially for the publication. A regulation to the effect that publication is only allowed with explicit consent of the person concerned would take a new principle into consideration, namely the principle of autonomy. This regulation would have the pro- and the con-reason of the first case, presumably with the same

weights, but would have an additional pro-reason in that it is supported by the principle of autonomy. Moreover, this last regulation is a competitor for the first proposed regulation which amounts to a blunt prohibition of the publication. As a consequence, the last regulation is better than the first one, and since it is a competitor, the stronger support for the last regulation makes the first regulation less attractive.

A similar argument can be made for the case that PRO_1 and PRO_2 are identical, and that CON_2 is a strict subset of CON_1 . Then the reasons against REG_2 are weaker than those against REG_1 , while the pro-reasons have remained the same. Again one can conclude that the reasons PRO_2 outweigh the reasons CON_2 and that REG_2 is a better regulation than REG_1 .

This would be illustrated by the case in which the tabloid journal has contracted with E that no publication of his photograph would be made. It is arguable that the freedom of the press is not infringed by a prohibition that was voluntarily undertaken by the journal. Since the freedom of the press was a reason against the prohibition, the balance of reasons is moved towards the prohibition if this con-reason is taken away. As a consequence the regulation that publication is prohibited if the potential publisher has voluntarily undertaken the obligation not to publish, has even stronger support than the original prohibition.

It is possible that a set of reasons is strengthened by adding new reasons to it, but also by strengthening the reasons that occur in it. Suppose that the sets of reasons PRO_2 and CON_2 that plead for, respectively against REG_2 , contain the same reasons as the PRO_1 and CON_1 that plead for, respectively against REG_1 . In other words, the regulations REG_1 and REG_2 contribute to and detract from the same goals. However, the contribution of REG_2 to one or more of the goals is bigger than that of REG_1 , while REG_1 does not contribute more to any goal than REG_2 does. Again we can conclude that the set of pro-reasons has become stronger in comparison to the reasons for REG_1 , while the set of con-reasons has remained the same. Since the reasons pleading for REG_1 already outweighed the reasons against it, one can conclude a fortiori that the stronger set of reasons pleading for REG_2 outweighs the reasons against REG_2 . So, if REG_1 is acceptable, REG_2 is acceptable too, and REG_2 is a better regulation than REG_1 .

This is illustrated by the regulation that not only forbids publication of the photograph, but also prescribes that the photograph is destroyed. This regulation provides better protection of privacy, and is therefore better than the simple prohibition.¹¹

A similar argument can be made for the case that PRO_1 and PRO_2 are identical, and that one or more of the reasons in CON_2 are weaker than the corresponding reasons in CON_1 . Then the reasons against REG_2 are weaker than those against REG_1 , while the pro-reasons have remained the same. Again one can conclude that the reasons PRO_2 outweigh the reasons CON_2 and that REG_2 is a better regulation than REG_1 .

For instance, a regulation that allows photographs, as long as the persons on the photographs are not recognisable, makes a smaller infringement on the freedom of the press, while the protection of the privacy remains the same. Such a regulation would therefore be better than a mere prohibition of publishing photographs.

The two kinds of a fortiori argument, based on changes in the sets of reasons and on the weights of the reasons, can be combined. If a set of reasons R_{1a} outweighs R_2 , and if R_{1a} is strengthened both by adding new reasons and by increasing the weight of the existing reasons, the resulting set R_{1b} will still outweigh the set R_2 . If a set of reasons R_1 outweighs R_{2a} , and if R_{2a} is weakened both by removing one or more reasons from it and by decreasing the weight of the existing reasons, the resulting set R_{2b} will still be outweighed by the set R_1 .

The combination of strengthening the stronger set (in one or both of the two mentioned ways) and weakening the weaker set (also in one or both of the two mentioned ways) should also lead to the conclusion that the strengthened set outweighs the weakened set.

4 Formalisation¹²

4.1 The language

The language in which the theory of the above section is formalised is that of predicate logic, with a number of additional conventions. The most important convention is an ontological one. The language recognises states of affairs, rules, and regulations (sets of rules) as logical individuals. A state of affairs is what is expressed by a sentence. All sentences are assumed to express states of affairs, while true sentences express states of affairs that actually obtain.

All predicates (zero- or more placed) start with an uppercase letter. All function expression (zero- or more placed) and variables start with a lowercase letter. If *S* is a sentence, then **s* is a term that denotes the state of affairs expressed by *S*. For instance, the sentence *Loves(john, mary)* expresses the state of affairs that John loves Mary, which is denoted by the term **loves(john, mary)*.

4.2 Dedicated predicates

In formalising the theory about regulation evaluation I will use a number of predicates that fulfil a special role:

Goal/1 is used to express that a particular state of affairs is a goal within the legal system at issue. For instance *Goal(*freedom_of_expression)* expresses that freedom of expression is a goal (in a particular legal system).

Acceptable/1 is used to express that a regulation is acceptable.

Preferable/1 is used to express that a regulation is preferable.

Competitors/2 is a symmetric a-reflexive relation between regulations. In this paper it remains undefined, but see note 7.

Better/2 is a weakly transitive¹³, a-symmetric, a-reflexive relation on regulations. *Better(regulation1, regulation2)* expresses that regulation 1 is better than regulation 2.

>_g/2 is a weakly transitive, a-symmetric, a-reflexive relation on the relative importance of goals. For instance **goal1 >_g *goal2* expresses that goal 1 is more important than goal 2.

Ct/3 is used to express that a regulation contributes to a certain degree to a goal. For instance, *Ct(regulation1, *goal1, d1)* expresses that regulation 1 contributes to goal 1, to degree *d1*.

Df/3 is analogously used to express that a regulation detracts to a certain degree from a goal.

>_d/2 is a transitive, a-symmetric, a-reflexive relation on the degrees to which regulations contribute to, or detract from goals. For instance, *d1 >_d d2* expresses, in the light of the examples above, that the contribution of regulation 1 to goal 1 is bigger than the contribution of regulation 2 to goal 2.

Reason/3 is used to express that a certain fact is a reason for or against a particular conclusion, expressed as a state of affairs. The third parameter is either *pro*, or *con*. For instance, *Reason(*ct(regulation1, *goal1, d1), *acceptable(regulation1), pro)* expresses that the fact that regulation 1 contributes to goal 1, to a degree *d1* is a reason for the acceptability of regulation 1.¹⁴

$\text{Reason}(*df(\text{regulation1}, *goal1, d1), *acceptable(\text{regulation1}), \text{con})$
expresses that the fact that regulation 1 detracts from goal 1, to a degree $d1$ is a reason against the acceptability of regulation 1.

$r^+(*s)$ denotes the set of reasons that plead for the presence of state of affairs $*s$. Formally:

$r^+(*s) = \{*r: \text{Reason}(*r, *s, \text{pro}) \text{ is true}\}$

$r^-(*s)$ denotes the set of reasons that plead against the presence of state of affairs $*s$. For-

formally: $r^-(*s) = \{*r: \text{Reason}(*r, *s, \text{con}) \text{ is true}\}$

$w/2$ is a function the value of which is the weight of a fact as a reason for a particular conclusion. For instance,

$w(*ct(\text{regulation1}, *goal1, d1), *acceptable(\text{regulation1})) = w1$

denotes that the fact that regulation 1 contributes to goal 1, to a degree $d1$ is a reason (for the acceptability of regulation 1) with weight $w1$.

Analogously

$w(*ct(\text{regulation2}, *goal2, d2), *acceptable(\text{regulation1})) = w2$

expresses that the fact that regulation 2 contributes to goal 2, to a degree $d2$ is a reason (for the acceptability of regulation 2) with weight $w2$.

$=_w/2$ is an equivalence relation on the weight of *individual* reasons. For instance, $w1 =_w w2$ expresses, in the light of the examples above, that the reason consisting of the contribution of regulation 1 to goal 1 weighs precisely as much as the reason consisting of the contribution of regulation 1 to goal 2.

$>_w/2$ is a weakly transitive, a-symmetric, a-reflexive relation on the weight of *individual* reasons. For instance, $w1 >_w w2$ expresses, in the light of the examples above, that the reason consisting of the contribution of regulation 1 to goal 1 weighs more than the reason consisting of the contribution of regulation 2 to goal 2.

$\text{Stronger}_w/2$ (stronger in individual weight) is an a-symmetric, a-reflexive relation on *sets* of reasons. This relation holds between two sets of reasons iff from the reasons which the two sets *have in common* at least one reason of the first set weighs more than the corresponding reason from the second set, while the opposite is not the case. Notice that this relation does *not* imply that the related sets consist of the same reasons. (This relation is defined formally in Axiom 11 in section 4.3.)

$\text{Weaker}_w/2$ (weaker in individual weight) is an a-symmetric, a-reflexive relation on *sets* of reasons. This relation holds between two sets of reasons iff from the reasons which the two sets *have in common* at least one reason of the first set weighs less than the corresponding reason from the second set, while the opposite is not the case. Notice that this relation does *not* imply that the related sets consist of the same reasons. (This relation is defined formally in Axiom 12 in section 4.3.)

$\text{Equal}_w/2$ (equal in individual weight) is a symmetric and reflexive relation on *sets* of reasons. This relation holds between two sets of reasons if the reasons which the two sets have in common, pair wise have the same weights. Notice that this relation does *not* imply that the related sets consist of the same reasons. (This relation is defined formally in Axiom 13 in section 4.3.)

$\text{Stronger}/2$ is a weakly transitive, a-symmetric, a-reflexive relation on sets of reasons pleading for or against the same conclusion. It is used to express that the first set of reasons provides stronger support for or against the conclusion than the second set. For instance,

$\text{Stronger}(s1, s2)$ expresses that $s1$ as a set of reasons pleading for conclusion $*c$ provides more support for $*c$ than $s2$, which is also a set of reasons pleading for $*c$. (This relation is defined formally in Axiom 14 in section 4.3.)

$\text{Weaker}/2$ is a weakly transitive, a-symmetric, a-reflexive relation on sets of reasons pleading for or against the same conclusion. It is used to express that the first set of reasons provides less support for or against the conclusion than the second set. (This relation is defined formally in Axiom 15 in section 4.3.)

$\text{Equal}/2$ is an equivalence relation on sets of reasons pleading for or against the same conclusion. It is used to express that the first set of reasons provides precisely as much support for or against the conclusion as the second set. (This relation is defined formally in Axiom 16 in section 4.3.)

$\text{Outweighs}/2$ is an a-symmetric, a-reflexive, a-transitive relation between two sets of reasons the elements of one of which pleads for a particular conclusion, while the elements of the other plead against this conclusion. So $\text{Outweighs}(s1, s2)$ presupposes that there is some conclusion $*c$, such that either both $s1 \subseteq r^+(*s)$ and $s2 \subseteq r^-(*s)$ are true, or both $s1 \subseteq r^-(*s)$ and $s2 \subseteq r^+(*s)$ are true.

4.3 Axioms

The following axioms can be added to Reason-based Logic in order to characterise goal-based theory evaluation. They are not intended to characterise a full logic by themselves.

A1. $(\text{Goal}(*g) \ \& \ \text{Ct}(r, *g, d)) \rightarrow$
 $\text{Reason}(*\text{ct}(x, *g, d), *\text{acceptable}(r), \text{pro})$

Informally: If g is a goal, and if regulation r contributes to g to some degree d , then this last fact is a reason for the acceptability of r .¹⁵

A2. $(\text{Goal}(*g) \ \& \ \text{Df}(r, *g, d)) \rightarrow$
 $\text{Reason}(*\text{df}(x, *g, d), *\text{acceptable}(r), \text{con})$

Informally: If g is a goal, and if regulation r detracts from g to some degree d , then this last fact is a reason against the acceptability of r .¹⁶

Comments:

The axioms A1 and A2 indicate how the contribution to, or detraction from a regulation towards a goal leads to reasons pleading for, respectively against the regulation.

A3. $\text{Ct}(r, *g, d1) \rightarrow \forall d2(\text{Ct}(r, *g, d2) \rightarrow (d2 = d1))$

Informally: A regulation cannot contribute in different degrees to one and the same goal

A4. $\exists d1(\text{Ct}(r, *g, d1)) \rightarrow \sim \exists d2(\text{Df}(r, *g, d2))$

Informally: A regulation cannot both contribute to, and detract from one and the same goal.

A5. $\forall \text{reg1}, \text{reg2}, *g, d(\text{Ct}(\text{reg1}, *g, d) \ \& \ \text{Ct}(\text{reg2}, *g, d) \rightarrow$
 $w(*\text{ct}(x, *g, d), *\text{acceptable}(\text{reg1})) =$
 $w(*\text{ct}(x, *g, d), *\text{acceptable}(\text{reg2})))$

Informally: If two regulations contribute to the same degree to the same goal, these facts are reasons for the acceptability of these two regulations with the same weight.

$$\text{A6. } \forall \text{reg1, reg2, *g, d} (\text{Df}(\text{reg1}, *g, d) \ \& \ \text{Df}(\text{reg2}, *g, d) \rightarrow \\ \text{w}(*\text{df}(x, *g, d), *\text{acceptable}(\text{reg1})) = \\ \text{w}(*\text{df}(x, *g, d), *\text{acceptable}(\text{reg2})))$$

Informally: If two regulations detract to the same degree from the same goal, these facts are reasons against the acceptability of these two regulations with the same weight.

$$\text{A7. } (\text{Goal}(*g) \ \& \ \text{Ct}(r1, *g, d1) \ \& \ \text{Ct}(r2, *g, d2) \ \& \ d1 >_d d2) \rightarrow \\ \text{w}(*\text{ct}(x, *g, d1), *\text{acceptable}(r1)) >_w \\ \text{w}(*\text{ct}(x, *g, d2), *\text{acceptable}(r2))$$

Informally: If g is a goal, and if regulation r1 contributes to a larger degree to g than r2, then the contribution by r1 is a stronger reason for the acceptability of r1, than the contribution by r2 is a reason for the acceptability of r2.

$$\text{A8. } (\text{Goal}(*g) \ \& \ \text{Df}(r1, *g, d1) \ \& \ \text{Df}(r2, *g, d2) \ \& \ d1 >_d d2) \rightarrow \\ \text{w}(*\text{df}(x, *g, d1), *\text{acceptable}(r1)) >_w \\ \text{w}(*\text{df}(x, *g, d2), *\text{acceptable}(r2))$$

Informally: If g is a goal, and if regulation r1 detracts to a larger degree from g than r2, then the detraction by r1 is a stronger reason against the acceptability of r1, than the detraction by r2 is a reason against the acceptability of r2.

Comments:

The axioms A7 and A8 play in particular a role if two competing regulations are to be compared which contribute to and detract from the same goals, but where one of the regulations contributes or detracts more than the other regulation. According to the axioms A7 and A8, this difference in contribution leads to a difference in weight of the corresponding reasons. This difference in weight can *ceteris paribus* lead to the conclusion that the one regulation is better than the other.

$$\text{A9. } (\text{Goal}(*g1) \ \& \ \text{Goal}(*g2) \ \& \ \text{Ct}(r1, *g1, d) \ \& \ \text{Ct}(r2, *g2, d) \ \& \\ *g1 >_g *g2) \rightarrow \\ \text{w}(*\text{ct}(x, *g1, d), *\text{acceptable}(r1)) >_w \\ \text{w}(*\text{ct}(x, *g2, d), *\text{acceptable}(r2))$$

Informally: If g1 and g2 are goals, while g1 is more important than g2, and if regulation r1 contributes to g1 to the same degree as r2 contributes to g2, then the contribution by r1 to g1 is a stronger reason for the acceptability of r1, than the contribution by r2 to g2 is a reason for the acceptability of r2.

$$\text{A10. } (\text{Goal}(*g1) \ \& \ \text{Goal}(*g2) \ \& \ \text{Df}(r1, *g1, d) \ \& \ \text{Df}(r2, *g2, d) \ \& \\ *g1 >_g *g2) \rightarrow \\ \text{w}(*\text{df}(x, *g1, d), *\text{acceptable}(r1)) >_w \\ \text{w}(*\text{df}(x, *g2, d), *\text{acceptable}(r2))$$

Informally: If g1 and g2 are goals, while g1 is more important than g2, and if regulation r1 detracts from g1 to the same degree as r2 detracts from g2, then the detraction by r1 from g1 is a stronger reason against the acceptability of r1, than the detraction by r2 from g2 is a reason against the acceptability of r2.

Comments:

The axioms A9 and A10 play in particular a role if two competing regulations are to be compared which contribute to or detract from different goals, in the same degree. According to the axioms A9 and A10, the difference in importance of the goals leads to a difference in weight of the corresponding reasons. This difference in weight can *ceteris paribus* lead to the conclusion that the one regulation is better than the other.

The practical importance of A9 and A10 is very limited, however, since it is (almost) impossible to establish whether regulation 1 contributes more to goal A than regulation 2 contributes to goal B. Arguably, the axioms A9 and A10 may therefore be left out, and if this is done, the ordering on goals on the basis of their relative importance loses its practical relevance too.

Let s_1 and s_2 be sets of reasons for different conclusions. Then:

A11. $\text{Stronger}_w(s_1, s_2) \equiv$

$$\begin{aligned} & \exists *r_1 \exists *r_2 (*r_1 \in s_1 \ \& \ *r_2 \in s_2 \ \& \ *r_1 = *r_2 \ \& \ r_1 >_w r_2) \ \& \\ & \sim (\exists *r_1 \exists *r_2 (*r_1 \in s_1 \ \& \ *r_2 \in s_2 \ \& \ *r_1 = *r_2 \ \& \ r_2 >_w r_1)) \end{aligned}$$

A12. $\text{Weaker}_w(s_1, s_2) \equiv$

$$\begin{aligned} & \exists *r_1 \exists *r_2 (*r_1 \in s_1 \ \& \ *r_2 \in s_2 \ \& \ *r_1 = *r_2 \ \& \ r_1 <_w r_2) \ \& \\ & \sim (\exists *r_1 \exists *r_2 (*r_1 \in s_1 \ \& \ *r_2 \in s_2 \ \& \ *r_1 = *r_2 \ \& \ r_2 <_w r_1)) \end{aligned}$$

A13. $\text{Equal}_w(s_1, s_2) \equiv$

$$\begin{aligned} & \forall *r_1 \forall *r_2 ((*r_1 \in s_1 \ \& \ *r_2 \in s_2 \ \& \ *r_1 = *r_2) \rightarrow \\ & w(*r_1, *c_1) = w(*r_2, *c_2)) \end{aligned}$$

Comments:

The axioms A11 to A13 are explained informally in section 4.2.

A14. $((s_1 \supset s_2) \ \& \ (\sim \text{Weaker}_w(s_1, s_2)) \vee \text{Equal}(s_1, s_2)) \vee ((s_1 = s_2) \ \& \ \text{Stronger}_w(s_1, s_2)) \equiv \text{Stronger}(s_1, s_2)$

Informally: The relation $\text{Stronger}/2$ holds between two sets of reasons if and only if either:

- the first set is a proper superset of the second, and the first set is equal to or not weaker in individual weight than the second, or
- the first set is identical to the second set, and the first set is stronger in individual weight than the second set.

A15. $((s_1 \subset s_2) \ \& \ (\sim \text{Stronger}_w(s_1, s_2) \vee \text{Equal}(s_1, s_2))) \vee ((s_1 = s_2) \ \& \ \text{Weaker}_w(s_1, s_2)) \equiv \text{Weaker}(s_1, s_2)$

Informally: The relation $\text{Weaker}/2$ holds between two sets of reasons if and only if either:

- the first set is a proper subset of the second, and the first set is equal to or not stronger in individual weight than the second, or
- the first set is identical to the second set, and the first set is weaker in individual weight than the second set.

A16. $(s_1 = s_2) \ \& \ \text{Equal}_w(s_1, s_2) \equiv \text{Equal}(s_1, s_2)$

Informally: The relation $\text{Equal}/2$ holds between two sets of reasons if and only if the two sets are identical, and equal in individual weight.

A17. $\forall s1, s2, s3, s4((\text{Outweighs}(s1, s3) \ \& \ \text{Equal}(s3, s4) \ \& \ (\text{Stronger}(s2, s1) \vee \text{Equal}(s2, s1))) \rightarrow \text{Outweighs}(s2, s4))$

Informally: If a set of reasons is outweighed by another, its equals are outweighed by all sets of reasons that are stronger than, or equal to this other set.

A18. $\forall s1, s2, s3, s4((\text{Outweighs}(s1, s3) \ \& \ \text{Equal}(s2, s1) \ \& \ (\text{Weaker}(s4, s3) \vee \text{Equal}(s4, s3))) \rightarrow \text{Outweighs}(s2, s4))$

Informally: If a set of reasons outweighs another set, its equals outweigh all sets of reasons that are weaker than, or equal to this other set.

Comments:

The axioms A14 to A16 indicate how characteristics of sets of reasons which plead for or against the same conclusion, such as the individual weights of their elements and a super-set/subset relation can be used to draw conclusions about the relative weight of these sets of reasons.

Notice that two sets of reasons need not always stand in one of the relations stronger, weaker, or equal to each other. In particular sets can be incomparable either because neither one is a subset of the other, or because each set has a reason that is also in the other set, which has a greater individual weight than the corresponding reason in the other set.

A19. $\forall \text{reg1} \forall \text{reg2} (\text{competitors}(\text{reg1}, \text{reg2}) \ \& \ (\text{Stronger}(\text{r}^+(\text{*acceptable}(\text{reg1})), \text{r}^+(\text{*acceptable}(\text{reg2}))) \ \& \ (\text{Equal}(\text{r}^-(\text{*acceptable}(\text{reg1})), \text{r}^-(\text{*acceptable}(\text{reg2}))) \vee \ (\text{Weaker}(\text{r}^-(\text{*acceptable}(\text{reg1})), \text{r}^-(\text{*acceptable}(\text{reg2})))))) \vee \ \text{Weaker}(\text{r}^-(\text{*acceptable}(\text{reg1})), \text{r}^-(\text{*acceptable}(\text{reg2}))) \ \& \ (\text{Equal}(\text{r}^-(\text{*acceptable}(\text{reg1})), \text{r}^-(\text{*acceptable}(\text{reg2}))) \vee \ (\text{Stronger}(\text{r}^-(\text{*acceptable}(\text{reg1})), \text{r}^-(\text{*acceptable}(\text{reg2})))))) \vee \ \text{Weaker}(\text{r}^+(\text{*acceptable}(\text{reg2})), \text{r}^+(\text{*acceptable}(\text{reg1}))) \ \& \ (\text{Equal}(\text{r}^-(\text{*acceptable}(\text{reg1})), \text{r}^-(\text{*acceptable}(\text{reg2}))) \vee \ (\text{Stronger}(\text{r}^-(\text{*acceptable}(\text{reg1})), \text{r}^-(\text{*acceptable}(\text{reg2})))))) \vee \ \text{Stronger}(\text{r}^-(\text{*acceptable}(\text{reg2})), \text{r}^-(\text{*acceptable}(\text{reg1}))) \ \& \ (\text{Equal}(\text{r}^-(\text{*acceptable}(\text{reg1})), \text{r}^-(\text{*acceptable}(\text{reg2}))) \vee \ (\text{Weaker}(\text{r}^-(\text{*acceptable}(\text{reg1})), \text{r}^-(\text{*acceptable}(\text{reg2})))))) \equiv \text{Better}(\text{reg1}, \text{reg2}))$

Informally: One regulation is better than a second regulation if and only if they are competitors and either:

- the reasons pleading for the first are stronger than the reasons pleading for the second, while the reasons pleading against the first are equal to or weaker than those pleading against the second, or
- the reasons pleading against the first are weaker than the reasons pleading against the second, while the reasons pleading for the first are equal to or stronger than those pleading for the second, or
- the reasons pleading for the second are weaker than the reasons pleading for the first, while the reasons pleading against the second are equal to or stronger than those pleading against the first, or

- the reasons pleading against the second are stronger than the reasons pleading against the first, while the reasons pleading for the second are equal to or weaker than those pleading for the second.

A20. $\forall reg(\text{Acceptable}(reg) \equiv \text{Outweighs}(r^+(\text{*acceptable}(reg), r^-(\text{*acceptable}(reg))))))$

Informally: That a regulation is acceptable is equivalent to that the reasons for the acceptability of this regulation outweigh the reasons against the acceptability.

A21. $\text{Preferable}(reg1) \equiv \text{Acceptable}(reg1) \ \& \ \exists \sim reg2(\text{Competitors}(reg1, reg2) \ \& \ \text{Better}(reg2, reg1))$

Informally: A regulation is preferable iff it is acceptable, and there is no better competing regulation.

5 Application of the formalisation

To illustrate the operation of the formalism presented above, I will formalise one example of section 3.¹⁷ It is the example that a regulation to the effect that publication is only allowed with explicit consent of the person concerned, is better than a blunt prohibition of publication. In formalising this example, I assume that there are no other reasons involved than those which can be derived from the facts of the case. This assumption is abbreviated as Ass.

The case is described by the following premises:

- (1:) $r1 = \{ \text{true} \Rightarrow \text{*prohibited}(\text{publication}) \}$
- (2:) $r2 = \{ \text{*}\sim\text{consent} \Rightarrow \text{*prohibited}(\text{publication}) \}$
- (3:) $\text{Competitors}(r1, r2)$
- (4:) $\text{Goal}(\text{*privacy})$
- (5:) $\text{Goal}(\text{*free_press})$
- (6:) $\text{Goal}(\text{*autonomy})$,
- (7:) $\text{Ct}(r1, \text{*privacy}, d1)$
- (8:) $\text{Df}(r1, \text{*free_press}, d2)$
- (9:) $\text{Ct}(r2, \text{*privacy}, d1)$
- (10:) $\text{Ct}(r2, \text{*autonomy}, d2)$
- (11:) $\text{Df}(r2, \text{*free_press}, d3)$
- (12:) $\text{Outweighs}(\{\text{*ct}(x, \text{*privacy}, d1)\}, \{\text{*df}(x, \text{*free_press}, d2)\})$

By means of the axioms A1 and A2 it follows that:

- (13:) $\text{Reason}(\text{*ct}(x, \text{*privacy}, d1), \text{*acceptable}(r1), \text{pro}) \quad (4, 7, A1)$
- (14:) $\text{Reason}(\text{*df}(x, \text{*free_press}, d2), \text{*acceptable}(r1), \text{con}) \quad (5, 8, A2)$
- (15:) $\text{Reason}(\text{*ct}(x, \text{*privacy}, d1), \text{*acceptable}(r2), \text{pro}) \quad (4, 9, A1)$
- (16:) $\text{Reason}(\text{*ct}(x, \text{*autonomy}, d3), \text{*acceptable}(r2), \text{pro}) \quad (6, 10, A1)$
- (17:) $\text{Reason}(\text{*df}(x, \text{*free_press}, d2), \text{*acceptable}(r2), \text{con}) \quad (5, 11, A2)$
- (18:) $r^+(\text{*acceptable}(r2)) \supset r^+(\text{*acceptable}(r1)) \quad (13, 15, 16, \text{Ass})$
- (19:) $r^-(\text{*acceptable}(r2)) = r^-(\text{*acceptable}(r1)) \quad (14, 17, \text{Ass})$

By means of axiom A5 it follows that

- (20:) $w(\text{*ct}(x, \text{*privacy}, d1), \text{*acceptable}(r1)) = w(\text{*ct}(x, \text{*privacy}, d1), \text{*acceptable}(r2)) \quad (7, 9, A5)$
- (21:) $w(\text{*df}(x, \text{*free_press}, d2), \text{*acceptable}(r1)) = w(\text{*df}(x, \text{*free_press}, d2), \text{*acceptable}(r2)) \quad (8, 11, A5)$

And then:

- (22:) $\text{Equal}_w(r^+(*\text{acceptable}(r2)), r^+(*\text{acceptable}(r1)))$ (20, A13)
- (23:) $\text{Stronger}(r^+(*\text{acceptable}(r2)), r^+(*\text{acceptable}(r1)))$ (18, 22, A14)
- (24:) $\text{Equal}_w(r^-(*\text{acceptable}(r1)), r^-(*\text{acceptable}(r2)))$ (21, A13)
- (25:) $\text{Equal}(r^-(*\text{acceptable}(r2)), r^-(*\text{acceptable}(r1)))$ (19, 24, A16)
- (26:) $\text{Outweighs}(r^+(*\text{acceptable}(r2)), r^-(*\text{acceptable}(r2)))$ (12, 23, 25, A17)
- (27:) $\text{Acceptable}(r2)$ (26, A20)
- (28:) $\text{Better}(r2, r1)$ (23, 25, A19)
- (29:) $\text{Preferable}(r2, r1)$ (27, 30, A21)¹⁸

6 Conclusion, future research and acknowledgments

In this paper I hope to have shown how one part of Dworkin's model of principles can be cast in the conceptual framework of Reason-based Logic, and how this makes it possible to apply logical techniques for the comparison of regulations in the light of the goals they aim to realise. It seems to me that this is also a step towards putting the relatively abstract logical framework that was developed during the nineties to practical use.

Although the present work is strongly based upon Reason-based Logic, there are in my opinion no major objections to basing a similar theory on other logical foundations, such as the work by Prakken and Sartor, e.g. [18 and 20]. Moreover, I see a close connection with the formal work on theory formation, as exemplified by [10 and 19]. I hope that these different, but related, research lines can be brought together in the future.

Another topic for future research is to find a technique that allows reference to regulations in reasons that plead for or against these regulations, and which nevertheless allows comparison of sets of reasons related to different regulations, on purely logical grounds. Cf. note 15.

The author wants to thank Bob Brouwer for fruitful discussions on the present subject and Giovanni Sartor, and the anonymous referees for the Jurix conference for useful comments on a draft of this paper that have led to a number of improvements.

Notes

- ¹ Much of the analysis in this section was inspired by Edwige Burg's lucid thesis 'The Model of Principles' [9].
- ² The papers in question are in particular the paper 'Hard Cases', included in [4], and the papers in part two (Law as interpretation) of [5].
- ³ In [4, p. 23/4], Dworkin also assigns another role to legal principles, namely that of a factor that influences the right outcome of a case directly. Here, I do not address this logical role of legal principles.
- ⁴ From a logical point of view, there need not be one single set of best rules. Dworkin, however, has gone to some length in arguing that every case has one right answer, thereby implicitly assuming that there is one best set of rules.
- ⁵ This subject is addressed by Sartor and Bench-Capon in [10 and 19].
- ⁶ I assume here that the goals are independent. It is for instance not the case that one goal is (merely) a subgoal of another goal.
- ⁷ In this paper I will leave the notion of competitors undefined. One way to elaborate the idea that two regulations are competitors of each other is to say that two regulations are competitors if a. they deal with approximately the same cases, and b. they lead to different outcomes for at least one possible case. See in this connection [16] or [17] on rule consistency.
- ⁸ If it is assumed that it is always possible to 'regulate' a subject by doing nothing, the two tests for a regulation can be combined into one. Doing nothing does not make any changes in the realisation of goals and therefore neither contributes, nor detracts from any goal. As a consequence, there are neither goal-based reasons for or against doing nothing. If a regulation is better than this 'empty regulation', it is automatically also an acceptable one. The best regulation is then either the empty one, or an acceptable one.
- ⁹ This inclusion of the qualifier 'normally' is to indicate that the a fortiori argument involved here is defeasible. I will leave this defeasibility and all other forms of defeasibility out of consideration in this paper. For details, see [14, pp. 203f.] on the accrual of reasons.
- ¹⁰ There is a close relation between the present theory and the economic notion of a Pareto-improvement, if the goals that are involved in the comparison between regulations are compared to the customers involved in an allocation of assets. See [9, p. 141]. See also the logical characterisation of the mechanism employed in the Hypo-system as described in [13, pp. 254f.] or [14, pp. 185f].
- ¹¹ Arguable this regulation would infringe the property right of the journal, but for the sake of the example, this complication is ignored.
- ¹² This formalisation builds on the theory of Reason-based Logic as developed in [13] and [14].
- ¹³ I call a relation R *weakly transitive* if the following is a valid principle in the sense of Reason-based Logic [14]: $((a R b) \ \& \ (b R c)) \supset (a R c)$.
- ¹⁴ But see note 15.
- ¹⁵ Arguably, the resulting reason should be not that some unidentified entity (x) contributes to the goal, but rather that the rule in question (r) contributes to the goal. That would have an unattractive consequence however. If two different rules contribute to the same degree to the same goal, they should have the same reason pleading for them. If the rule is mentioned in the reason, this is no longer the case anymore, because the fact that rule 1 contributes to a goal is not the same reason as that rule 2 contributes to that same goal.
To avoid this complication, the reference to the rule in the reason is replaced by an unbound variable. Technically this does the trick, but the chosen solution is not fully satisfactory. Therefore a major topic for future research is to look for a solution that maintains the reference to the regulation in the reason, while allowing comparison of the sets of reasons that deal with different regulations.
- ¹⁶ See the previous note.
- ¹⁷ This formalisation presupposes Reason-based Logic as exposed in [14]. There are slight differences in formalisation, which should be understandable from the present text.
- ¹⁸ This step assumes that the only regulations involved are r_1 and r_2 .

References

- [1] L.T. McCarty, Some Arguments about Legal Arguments. *Proceedings of the Sixth International Conference on Artificial Intelligence and Law*, ACM, New York, 1997, pp. 215-224.
- [2] M. Marx, Interessenjurisprudenz und Freirechtslehre, in A. Kaufmann und W. Hassemer (hrsg.), *Einführung in Rechtsphilosophie und Rechtstheorie der Gegenwart*, Müller Juristischer Verlag, Heidelberg, 1977, pp. 97-102.
- [3] K. Larenz, *Methodenlehre der Rechtswissenschaft*, 5th ed., Springer Verlag, Berlin, 1983.
- [4] R. Dworkin, *Taking Rights Seriously*, 2nd ed., Duckworth, London, 1978.
- [5] R. Dworkin, *A Matter of Principle*. Clarendon Press, Oxford, 1985.
- [6] R. Dworkin, *Law's Empire*, Fontana, London, 1986.
- [7] D.N. MacCormick, *Legal Reasoning and Legal Theory*, Clarendon Press, Oxford, 1978.
- [8] R. Alexy, and A. Peczenik, The concept of coherence and its significance for discursive rationality, *Ratio Juris* **3**, (1990), no. 1bis.
- [9] E. Burg, *The Model of Principles. The quest for rationality in the implementation of conflicting principles*. PhD thesis, Amsterdam, 2000.
- [10] G. Sartor, Teleological Arguments and Theory Based Dialectics. Submitted to *Artificial Intelligence and Law*.
- [11] R. Alexy, Zum Begriff des Rechtprinzips, *Rechtstheorie*, Beiheft 1, (1979), pp. 59-87.
- [12] R. Alexy, *Theorie der Grundrechten*, Nomos Verlagsgesellschaft, Baden-Baden, 1985.
- [13] J.C. Hage, A Model of Legal Reasoning and a Logic to Match. *Artificial Intelligence and Law* **4**, nos. 3-4 (1996), pp. 199-273. Also in H. Prakken and G. Sartor (eds.), *Logical Models of Legal Argumentation*, Kluwer Academic Publishers, Dordrecht, 1997, pp. 43-117.
- [14] J.C. Hage, *Reasoning with rules*, Kluwer Academic Publishers, Dordrecht, 1997.
- [15] Bart Verheij, Rules, Reasons, Arguments. Formal studies of argumentation and defeat. PhD thesis Maastricht, 1996.
- [16] J.C. Hage, (1999). Rule consistency. H.J. van den Herik e.a. (eds.), *Legal Knowledge Based Systems. Jurix 1999: The Twelfth Conference*, GNI, Nijmegen, 1999, pp. 47-60.
- [17] J.C. Hage, Rule consistency. *Law and Philosophy* **19** (2000): p. 369-390.
- [18] H. Prakken and G. Sartor, A Dialectical Model of Assessing Conflicting Arguments in Legal Reasoning. *Artificial Intelligence and Law* **4** (1996) no. 3/4. pp. 331-368. Also in H. Prakken and G. Sartor (eds.), *Logical Models of Legal Argumentation*. Kluwer Academic Publishers, Dordrecht 1997.
- [19] T. Bench-Capon and G. Sartor, *Using Values and Theories To Resolve Disagreement in Law*. In this volume.
- [20] H. Prakken, *An exercise in formalising teleological case-based reasoning (extended abstract)*. In this volume.