Formally:

$$\forall s1,s2(\\ (similar\_superset(s2, s1)\ \&\\ (s2 >_w s1) \lor (s2 =_w s1)) \rightarrow\\ Weaker(s1, s2))$$

$$\forall s1,s2(\\ ((s2 >_w s1)\ \&\\ similar\_reasonset(s1, s2) \lor\\ similar\_subset(s1, s2)) \rightarrow\\ Weaker(s1, s2))$$

Equal

A set of reasons is overall equal to another set on logical grounds if (but not necessarily only if):

1. they are similar sets, and
2. they are equal in individual weight.

Formally:

$$\forall s1,s2(((s1 =_w s2)\ \&\ similar\_reasonset(s1, s2)) \rightarrow\\ Equal(s1, s2))$$

9. COMPARATIVE REASONING ABOUT SETS OF CONTRIBUTIVE REASONS

The relations of strength between reason sets as defined in the previous section are quite tight, and will not hold very often. It may therefore seem that they are not very useful for practical reasoning purposes. However, it is thinkable that there are other logical grounds for the existence of one of the mentioned relations between reason sets than the ones discussed in section 8. Moreover, and more importantly, there may also be other than logical grounds on which a set is stronger than, weaker than, or equal to another set. In fact, the determination which of two sets is overall stronger or weaker than the other will most often be just a matter of decision making. But also then the following relations should hold:

$$\forall s1,s2(Stronger(s1,s2) \rightarrow \sim Weaker(s1,s2))$$
$$\forall s1,s2(Stronger(s1,s2) \rightarrow \sim Equal(s1,s2))$$
$$\forall s1,s2(Weaker(s1,s2) \rightarrow \sim Stronger(s1,s2))$$
$$\forall s1,s2(Weaker(s1,s2) \rightarrow \sim Equal(s1,s2))$$
$$\forall s1,s2(Equal(s1,s2) \rightarrow \sim Stronger(s1,s2))$$
$$\forall s1,s2(Equal(s1,s2) \rightarrow \sim Weaker(s1,s2))$$

These relations can be used to reason about the relative strength of reason sets, even when the tight logical relations of the previous section do not exist. And there are even more ways to reason about the relative strength of sets of contributive reasons, based on the transitivity of the stronger than. equal to, and weaker than relations.

If one set of reasons is *on the above mentioned logical grounds* stronger than another set, and this other set is, again *on the above mentioned logical grounds*, stronger than a third set, the first set is stronger than the third set. The same transitivity holds for the equal to and weaker than relations, *to the extent that they exist on the above mentioned logical grounds*.

Because the logical characterizations of these relations as given in the previous section are not definitions, a reason set may be stronger than, equal to, or weaker than another reason set for another reason, for instance just because it was decided to be so. As a consequence, the relations stronger/2, equal/2 and weaker/2 cannot be assumed to be transitive.

Nevertheless it is reasonable to assume that some weaker form of transitivity holds. If set 1 is stronger than set 2 and set 2 is stronger than set 3, then normally set 1 will be stronger than set 3. If set 1 is equal to set 2 and set 2 is equal to set 3, then normally set 1 will be equal to set 3. And, finally if set 1 is weaker than set 2 and set 2 is weaker than set 3, then normally set 1 will be weaker than set 3. This *weak transitivity* can be expressed in terms of abstract reasons as follows:

```
Ar((*stronger(a,b) & stronger(b,c)), *stronger(a,c))
Ar((*equal(a,b) & equal(b,c)), *equal(a,c))
Ar((*weaker(a,b) & weaker(b,c)), *weaker(a,c))
```

Moreover, if a set of contributive reasons *A* is stronger than set *B* and if set *B* is equal to set *C*, then *A* will normally be stronger than *C*:

```
Ar((*stronger(a,b) & equal(b,c)), *stronger(a,c))
```

If a set of contributive reasons *A* is equal to set *B* and if set *B* is stronger than set *C*, then *A* will normally be stronger than *C*:

```
Ar((*equal(a,b) & stronger(b,c)), *stronger(a,c))
```

If a set of contributive reasons *A* is weaker than set *B* and if set *B* is equal to set *C*, then *A* will normally be weaker than *C*:

```
Ar((*weaker(a,b) & equal(b,c)), *weaker(a,c))
```

If a set of contributive reasons *A* is equal to set *B* and if set *B* is weaker than set *C*, then *A* will normally be weaker than *C*:

```
Ar((*equal(a,b) & weaker(b,c)), *weaker(a,c))
```

The practical use of comparative reasoning by means of sets of reasons is greatly increased by:

- the fact that the relations stronger than, weaker than and equal can hold between sets of comparable reasons on other than logical grounds, e.g because of a decision, and
- the existence of (weak) transitivity between these relations, as discussed above.

10. COMPARING ALTERNATIVES

In the qualitative comparison of two alternatives, four sets of reasons are involved:

- the reasons pleading for the first alternative;
- the reasons pleading against the first alternative;
- the reasons pleading for the second alternative;
- the reasons pleading against the second alternative.

In principle, a first alternative is better than a second alternative, if either one of the following situations occurs:

1. The first alternative is stronger in pro-reasons than the second, while it is equal or weaker in the con-reasons.
2. The first alternative is weaker in con-reasons than the second, while it is equal in the pro-reasons.
3. The second alternative is weaker in pro-reasons than the first, while it is equal or stronger in the con-reasons.
4. The second alternative is stronger in con-reasons than the first, while it is equal in the pro-reasons.

Formally[25]:

```
Ar((*stronger(r⁺(a1), r⁺(a2)) &
    weaker(r⁻(a1), r⁻(a2)) ∨ equal(r⁻(a1), r⁻(a2))),
  *better(a1, a2))

Ar((*weaker(r⁻(a1), r⁻(a2)) & equal(r⁺(a1), r⁺(a2))),
  *better(a1, a2))
```

[25] That the indicated relation only holds in principle is formalized by specifying the connection in terms of an abstract *contributive* reason.

```
Ar((*weaker(r+(a2), r+(a1)) &
    stronger(r-(a2), r-(a1)) ∨ equal(r-(a2), r-(a1))),
  *better(a1, a2))

Ar((*stronger(r-(a2), r-(a1)) & equal(r+(a1), r+(a2))),
  *better(a1, a2))
```

Similarly, the first alternative is worse than the second alternative, if either one of the following situations occurs:

1. The first alternative is weaker in pro-reasons than the second, while it is equal or stronger in the con-reasons.
2. The first alternative is stronger in con-reasons than the second, while it is equal in the pro-reasons.
3. The second alternative is stronger in pro-reasons than the first, while it is equal or weaker in the con-reasons.
4. The second alternative is weaker in con-reasons than the first, while it is equal in the pro-reasons.

Formally:

```
Ar((*weaker(r+(a1), r+(a2)) &
    stronger(r-(a1), r-(a2)) ∨ equal(r-(a1), r-(a2))),
  *worse(a1, a2))

Ar((*stronger(r-(a1), r-(a2)) & equal(r+(a1), r+(a2))),
  *worse(a1, a2))

Ar((*stronger(r+(a2), r+(a1)) &
    weaker(r-(a2), r-(a1)) ∨ equal(r-(a2), r-(a1))),
  *worse(a1, a2))

Ar((*weaker(r-(a2), r-(a1)) & equal(r+(a1), r+(a2))),
   *worse(a1, a2))
```

Finally, the first alternative and the second alternative are equally good, if they are equal both in pro and con reasons:

```
Ar(*equal(r+(a1), r+(a2)) & equal(r-(a1), r-(a2)),
   *equally_good(a1,a2)
```

The relations `Better/2`, `Worse/2`, and `Equally_good/2` are mutually exclusive:

```
∀a1,a2(
   Better(a1, a2) →
      ~Worse(a1, a2) & ~Equally_good(a1, a2) &
   Worse(a1, a2) →
```

```
    ~Better(a1, a2) & ~Equally_good(a1, a2) &
  Equally_good(a1, a2) →
    ~Worse(a1, a2) & ~Better(a1, a2))
```

The relation `better/2` cannot be taken to be transitive. However, if alternative A is better than alternative B, while B is better than C, it is at least plausible that A is better than C. This weak transitivity can be expressed as follows:

```
Ar((*better(a,b) & better(b,c)), *better(a,c))
```

Moreover, if alternative A is better than alternative B, while B is just as good as C, it is plausible that A is better than C. The same counts if alternative A is just as good as B, while B is better than C :

```
Ar((*better(a,b) & equally_good(b,c)), *better(a,c))
Ar((*equally_good(a,b) & better(b,c)), *better(a,c))
```

The same holds for the relations `worse/2` and `equally_good/2`:

```
Ar((*worse(a,b) & worse(b,c)), *worse(a,c))
Ar((*worse(a,b) & equally_good(b,c)), *worse(a,c))
Ar((*equally_good(a,b) & worse(b,c)), *worse(a,c))

Ar((*equally_good(a,b) & equally_good(b,c)),
  *equally_good(a,c))
```

The practical use of comparative reasoning by means of sets of reasons is greatly increased by:

- the fact that the relations better, worse and equally_good can hold between alternatives on other than logical grounds, e.g because of a decision, and
- the existence of (weak) transitivity between these relations, as discussed above.

11. APPLICATION OF THE FORMALIZATION

To illustrate the formalization described above, I will formalize the example about the caustic soda case and the yew case where QCR is applied to case based reasoning. I will use the follow abbreviations:

`Ds`: defendant created a dangerous situation to which plaintiff fell victim
`Ec`: it was easy and cheap to avoid the danger
`Dh`: potential damages were high
`Na`: defendant was not aware that he created a danger

`N`: defendant was negligent

`Csds`: defendant created a dangerous situation to which plaintiff fell victim in the caustic soda case
`Csec`: it was easy and cheap to avoid the danger in the caustic soda case
`Csdh`: potential damages were high in the caustic soda case
`Csna`: defendant was not aware that he created a danger in the caustic soda case
`Csn`: defendant was negligent in the caustic soda case

`Yds`: defendant created a dangerous situation to which plaintiff fell victim in the yew case
`Yec`: it was easy and cheap to avoid the danger in the yew case
`Ydh`: the potential damages were high in the yew case
`Yna`: defendant was not aware that he created a danger in the yew case
`Yn`: defendant was negligent in the yew case

The following premises are assumed:

```
Csds & Csec & Csdh & Csna & Csn

Yds & Yec & Ydh & Yna

Ar(*ds, *n) & Ar(*ec, *n) & Ar(*dh, *n) & Ar(*na, *~n)

∃i(*csds = instantiation(*ds, i))
∃i(*yds= instantiation(*ds, i))
∃i(*csec = instantiation(*ec, i))
∃i(*yec= instantiation(*ec, i))
∃i(*csdh = instantiation(*dh, i))
∃i(*ydh= instantiation(*dh, i))
∃i(*csna = instantiation(*na, i))
∃i(*yna = instantiation(*na, i))
∃i(*csn= instantiation(*n, i))
∃i(*yn = instantiation(*n, i))
```

Given these premises and the assumption that the derivable reasons are the only ones, it is possible to derive that:

$r^{+}(*csn) = \{*csds, *csec, *csdh\}$
$r^{-}(*csn) = \{*cna\}$

$r^{+}(*yn) = \{*yds, *yec, *ydh\}$
$r^{-}(*yn) = \{*yna\}$

$\text{similar\_reasonsets}(r^{+}(*csn), r^{+}(*yn))$

```
similar_reasonsets(r⁻(*csn), r⁻(*yn))

r⁺(*csn)=w r⁺(*yn)
r⁻(*csn)=w r⁻(*yn)
```

From these last four sentences it follows that:

```
equal(r⁺(*csn), r⁺(*yn))
equal(r⁻(*csn), r⁻(*yn))
```

and

```
equally_good(*csn, *yn)
```

To draw the additional conclusion that the defendant in the yew case acted negligently, an additional premise is necessary. If in one case a particular decision was taken, and a similar decision in another case would be at least as good (equally good, or even better), this is a reason why this similar decision should be taken in the other case[26]:

```
Ar(
   *decision(c1,*d1) & similar(*d1,*d2) &
    (equally_good(*d1,*d2) ∨ better(*d2,*d1)),
   *sb(decision(c2,*d2))
```

In the caustic soda case the decision was that defendant acted negligently:

```
Decision(caustic_soda_case, *csn)
```

Moreover, the (possible) decisions `*csn` and `*yn` both instantiate `*n` and are therefore similar.

As a consequence it follows that:

```
Cr(*decision(caustic_soda_case, *csn)&
    similar(*csn,*yn) &
      (equally_good(*csn,*yn) ∨ better(*csn,*yn)),
  *sb(decision(yew_case,*yn))
```

There are no contributive reasons why the decision in the yew case should not be `*yn`[27] :

[26] In the formula below, the function expression `sb/1` is used to express that the decision in the second case *should be* `*d2`.

[27] The reason that defendant was not aware that he created a danger was already taken into account in the case comparison and is therefore not taken into consideration anymore (excluded; see chapter 3, section 3.5).

```
r⁻(*sb(decision(yew_case,*yn))= ∅
```

Therefore the reasons why the decision in the yew case should be to assume negligence outweigh the reasons against this conclusion:

```
r⁺(*sb(decision(yew_case,*yn))>
  r⁻(*sb(decision(yew_case,*yn))
```

It therefore follows that in the yew-case negligence should be assumed:

```
Sb(decision(yew_case,*n))
```

12. RELATED RESEARCH

The topics discussed in the previous sections have been dealt with before in the literature on AI and/or law. QCR is the topic of Brewka and Gordon 1994. Presently I am not aware of similar work. Quantitative comparative reasoning is dealt with in Keeney and Raiffa 1993.

Legal theory construction, although not always under that name, is the topic of an enormous amount of literature. Most relevant in connection with the present research is the work of MacCormick 1978, Dworkin 1978 and 1986, McCarty 1997, Bench-Capon and Sartor 2001 and 2003, Hage 2000 (GTE) and 2001 (FLC), Peczenik and Hage 2000 and Hage and Sartor 2003. Concerning the logic of goals (legal principles, human rights), the work of Alexy 1979, 1996, 2000 and 2003 is particularly relevant.

The core texts with regard to case based reasoning in connection with AI and Law are still Ashley 1990, 1991 and 1992. Relevant other work is Rissland and Skalak 1991, Aleven 1997 and 2003, Hage 1997 (RwR), section V,9, and Roth 2003. Prakken and Sartor 1998 do not deal so much with case-based reasoning in the sense used here, but rather with the justification of rules on the basis of cases

Legal proof is also the subject of a tremendous amount of literature, e.g. Twining 1985 and 1991; Edwards 1988; Prakken e.a. 2003. To my knowledge (supported by the expert advice of H. Crombag), comparative reasoning as a tool for reasoning about proof has not been the subject of much discussion yet.

Finally, this chapter, and in particular the second part of it, elaborates the theory of defeasible reasoning in the law, and in particular reason-based logic as a logical tool to deal with it. Relevant related research has been published in, amongst others, Verheij 1996, Prakken and Sartor (eds.) 1997, Hage 1997 (RwR), Prakken 1997, Brozek 2003 and Sartor 2005.

Chapter 5

RULE CONSISTENCY

1. INTRODUCTION

The purpose of this chapter is to introduce and develop a theory about the consistency of rules. There are at least three reasons why the consistency of rules differs from the consistency of descriptive sentences. First, many rules have a conditional structure, but their consistency cannot be treated as the consistency of conditional sentences. Second, consistency of both sentences and rules is relative to a set of constraints that determine which states of affairs can go together. Part of the complications in connection with rule consistency is that rules themselves can function as constraints relative to which consistency has to be judged. And, finally, there can be exceptions to rules that block the application of applicable rules.[1] Such exceptions can prevent threatening rule conflicts, thereby making seemingly inconsistent rules consistent. I will try to develop a theory of rule consistency that takes all these three aspects into account.

To my knowledge, the consistency of rules has not received much attention yet in the literature about legal logic. A topic that may seem related and that has received attention is that of *deontic consistency*, also called normative consistency.[2] This concerns questions as whether there *can* be logical relations between deontic sentences, prescriptions, or norms such as

[1] A rule is applicable if its conditions are satisfied. See chapter 3, section 5.3.

[2] An overview can be found in Den Haan 1996. See also Hamner Hill 1987 and Lindahl 1992.

that it is forbidden to steal and that it is permitted to steal.[3] Deontic consistency is only to a limited extent related to rule consistency, however, because the rules that can be (in)consistent need not be deontic at all. The question whether the conceptual rules (legal definitions) that surf boards count as vehicles and that nothing without wheels counts as a vehicle, are consistent falls, for instance, under the topic of this paper, but has nothing to do with deontic consistency.

Before continuing the discussion of rule consistency, I want to point out that consistency in connection with rules is not exactly the same phenomenon as consistency in connection with descriptive sentences. The use of the term ‘consistency’ in connection with rules may therefore give rise to some confusion. However, the phenomenon with which I will deal in this chapter has enough in common with ‘ordinary’ consistency, to justify the use of the same term. Moreover, an alternative term, such as ‘compatibility’, may give rise to other misunderstandings. Therefore I will continue to write about rule consistency, under the acknowledgment that the phenomenon for which the term stands is in some respects different than the consistency of descriptive sentences.

2. RULES AS CONDITIONALS

At first sight, it seems attractive to treat rule consistency as a special form of the consistency of conditional sentences. At second sight, this approach turns out to be less attractive, because if the consistency of rules were the same as that of conditional sentences, the following two rules would be consistent:

Thieves are punishable.
Thieves are not punishable.

[3] Deontic consistency is the issue at stake in the discussion of norm conflicts in the sense of Kelsen 1979, of norm contradictions and norm collisions in the sense of Hamner Hill 1987, of norm consistency in the sense of Von Wright 1991, of disaffirmation conflicts and compliance conflicts in the sense of Lindahl 1992, and of norm conflicts in the sense of Ruiter 1997.

Their consistency would follow from the fact that the following sentences are *not* inconsistent[4]:

$\forall x(\texttt{Thief}(x) \rightarrow \texttt{Punishable}(x))$
$\forall x(\texttt{Thief}(x) \rightarrow {\sim}\texttt{Punishable}(x))$

Instead of being inconsistent, these sentences allow the derivation of

${\sim}\exists x(\texttt{Thief}(x))$

Intuitively, however, a legislator should not be able to remove thieves from the world, merely by issuing both the rules that thieves are punishable and that they are not punishable.

The first conclusion to draw, therefore, is that a theory in which rules are treated like conditional sentences and that considers rule consistency as similar to the consistency of conditional sentences, is unsatisfactory. There is reason to search for a notion of consistency that is especially suited to rules.[5] A relevant intuition in this connection is that the consistency of rules should not depend on whether certain states of affairs obtain. We want, for instance, the rules that thieves are punishable and that they are not punishable to be inconsistent, independent of whether there are thieves. If there are thieves, the two rules can, barring exceptions, be used to derive an inconsistency in the traditional sense, because then it can be derived that these thieves are both punishable and not. However, we want the inconsistency of the rules to be independent of whether there are facts that satisfy their conditions.

Yet, it is important for the consistency of rules whether the rules *can* be applied to the same case. For instance, the rules that thieves are punishable and that non-thieves are not punishable are intuitively consistent. In this connection, three kinds of situations can be distinguished. The first situation, exemplified by the rules that thieves are both punishable and not punishable, is that two rules attach incompatible consequences to the same category of cases. If the one rule is applicable, the other rule is applicable too. The

[4] Arguably this formalization in the form of conditional sentences is incorrect. Saying that thieves are punishable is not the same as saying that if somebody is a thief, he is punishable. The rules that thieves are punishable and that they are not punishable are inconsistent because they attach incompatible consequences to the same category of cases. It is well defendable that they are therefore more similar to the statements that John is punishable and that John is not punishable, than to conditional sentences. In the rest of this paper, I will ignore this line of thought, if only because the following treatment of rule consistency is compatible with it.

[5] Obviously, the reasons given here why rule consistency is different from the consistency of descriptive sentences are not decisive. It is well possible to treat rule consistency in the same way as consistency of descriptive sentences and take the phenomenon that rules are only inconsistent if certain facts are present, into the bargain.

second situation is when one rule deals with a subset of the cases with which the second rule deals. An example would be the rules:

Thieves are punishable.
Thieves below the age of twelve are not punishable.

The third situation is when two rules deal with sets of cases that are logically unrelated, but which may have members in common. An example would be the rules:

Thieves are punishable.
Minors are not punishable.

In general, the inconsistency of rules depends both on the incompatibility of the conclusions of the rules and on the compatibility of the rule conditions. The basic idea is that a set of rules is inconsistent if it is possible that there is a case in which the conditions of all the rules are satisfied, while the consequences that are attached to this case by these rules are incompatible. This basic idea needs to be refined, however. For instance, the rules that thieves are punishable and that minors are not, are inconsistent, because there is a possible case (a minor thief) to which the rules attach incompatible consequences. If the rule that non-thieves are not punishable is added to these two rules, there cannot be a case anymore that satisfies the conditions of all the three rules, because it is not possible that somebody is both a thief and a non-thief. According to the basic idea about rule consistency, the resulting set of three rules would be consistent, because there cannot be a case anymore in which all three rules are applicable. It should not be possible, however, to make an inconsistent set of rules consistent by adding a rule with conditions that are incompatible with those of one of the rules in the inconsistent set. To avoid this complication, we can require that a set of consistent rules does not contain an inconsistent subset. In other words, a set of rules would be inconsistent if it contains an inconsistent subset.[6] This leads me to the following provisional theory about rule consistency[7]:

The rules in a set s are consistent if and only if it is not so that there are a subset s' of s and a possible case f such that

a. *the facts in f satisfy the conditions of all the rules in s', and*
b. *the rules in s' attach incompatible consequences to f.*

6 Obviously, the subset need not be a proper one.

7 For the purpose of this theory, and the improvements upon it that will be proposed, a case is taken to be a set of facts, and a possible case is therefore a set of states of affairs. More about facts, states of affairs and their mutual relations in chapter 3, section 2.

This provisional theory will be developed in the rest of this paper.

3. CONSISTENCY, COMPATIBILITY AND CONSTRAINTS

Descriptive sentences are called *consistent* if it is possible that they are all true. For instance, the sentences 'John is a thief' and 'John is a minor' are consistent, because it is possible that John is both a thief and a minor. In other words, because the *states of affairs* that John is a thief and that he is a minor are *compatible*, the *sentences* that express these states of affairs are *consistent*. The sentences 'John is a thief' and 'John is not a thief' are inconsistent, because it is not possible that John both is and is not a thief. It is the *incompatibility* of the *states of affairs* that John is a thief and that he is not a thief that makes the corresponding *sentences inconsistent*.

Compatibility of states of affairs is always relative to some background of *constraints*.[8] The states of affairs that John is a thief and that he is not a thief are incompatible because of the constraint that a state of affairs cannot both obtain and not obtain. A similar constraint is that the compound state of affairs that John is both a thief *and* a minor can only obtain if *both* the states of affairs that John is a thief and that he is a minor obtain. Such constraints are usually called *logical* constraints. Besides logical constraints, there are also other constraints. There are *physical* constraints that prevent somebody from being in two non-adjacent countries at the same time. It is, for instance, physically impossible that John is both in France and in Austria. *Conceptual* constraints make it impossible that something is both a square and a circle.

This is the occasion to introduce a terminological convention. The expressions 'compatible' and 'incompatible' will be used for *states of affairs* which can, or cannot, go together relative to a set of constraints. The expressions 'consistent' and 'inconsistent' are used for both *descriptive sentences* and for *rules*, with different criteria for sentence and rule consistency.

What is possible depends on the constraints that are taken into account. I will develop this idea by means of the notion of a possible world. A world is a set of states of affairs that is possible relative to some set of constraints c, in the sense that the facts of that world satisfy the constraints in c. This set should be maximal in the sense that it is not possible to add a state of affairs

[8] This point has, in a different context, also been made by Prakken and Sartor 1996, 184/5.

to it without violating a constraint. A state of affairs is possible (can obtain), if there is at least one possible world in which this state of affairs obtains.

Next to the familiar logical and physical constraints, there can also be legal constraints on possible worlds. For instance, it might be the case that in a legally possible world somebody cannot both be a thief and not punishable.[9] As this example shows, the constraints on possible worlds can be the result of human culture. By adopting rules, humans can impose additional constraints on the world in which they live. Rule-based constraints are contingent in the sense that they are absent in a world in which these rules do not exist. But when they exist, they rule out certain combinations of states of affairs as impossible, and necessitate other states of affairs.

It might be objected that rules should not be treated as constraints on possible worlds, but rather as entities that obtain in some possible worlds and are absent from other possible worlds. Only the 'logical' consequences of the existence of rules should be treated as constraints on possible worlds. For instance, it would be a constraint on possible worlds that if the rule that thieves are punishable exists in some of them, in those worlds thieves are punishable. But it would not be a constraint on possible worlds in general that thieves are punishable. This objection presupposes that there is a sharp demarcation between facts that happen to obtain in a world (such as the existence of a rule), and constraints on possible world that hold non-contingently, such as the logical consequences of the existence (validity) of rules. In my opinion there is no such sharp demarcation, however. It is a matter of choice, or at least something that is mind-dependent, what counts as a constraint and what as merely contingent.

Moreover, sometimes the legislator explicitly wants some rules to count as background in order to prevent other rules from being inconsistent. Take for instance the Dutch rules about theft and embezzlement.[10] The Dutch legislator threatens embezzlement with a lesser penalty than theft. If some concrete act would count both as embezzlement and as theft, this would lead to an inconsistency. However, by defining theft and embezzlement such that an act cannot fall under both classifications, the potential inconsistency is avoided. The purpose of these definitions is precisely that they function as constraints on legally possible worlds that make it impossible that an act is both a case of theft and of embezzlement.

Apparently it is possible to treat the existence of a rule as a merely contingent fact in a possible world, but also to consider the effects of a rule

9 Exceptions to rules are disregarded at this stage of the presentation.
10 For the purpose of this example, I ignore some complications in the Dutch law.

as constraints on worlds that one counts as possible. This double role of rules, both as part of a contingent set of rules that is judged on its consistency, and as a constraint on possible worlds that determines what counts as consistent, is explored in the rest of this paper.

If the compatibility of states of affairs is relative to a set of constraints, this has implications for our provisional definition of rule consistency:

The rules in a set s are consistent relative to a set of constraints c if and only if it is not so that there are a subset s' of s and a possible case f such that

- *the states of affairs in f are compatible relative to c,*
- *the states of affairs in f satisfy the conditions of all the rules in s', and*
- *the rules in s' attach consequences to f that are incompatible relative to c.*

Let me illustrate the implications of the above theory of rule consistency by means of some examples.[11]

EXAMPLE 1

```
1: *thief(x) ⇒ *punishable(x)
2: *thief(x) ⇒ *~punishable(x)
```

The rules 1 and 2 are logically inconsistent, because if John is a thief, this fact satisfies the conditions of both rules, while the conclusions of the two are logically incompatible.[12] Notice that the inconsistency of the rules does not depend on the presence of the fact that John is a thief. This fact merely illustrates the inconsistency.

EXAMPLE 2

```
1: *thief(x) ⇒ *punishable(x)
3: *minor(x) ⇒ *~punishable(x)
```

The rules 1 and 3 are logically inconsistent, for the same reasons as in example 1. The inconsistency is illustrated by the case of John who is both a thief and a minor.

[11] In these examples I use the formalism of RBL as exposed in chapter 3.

[12] I assume that it is clear what is logically incompatible. In section 7 the notion of logical compatibility is made more precise. Moreover, in this, and some of the following examples I do not make the logical constraints relative to which the rules are inconsistent explicit. In general, logical constraints are left implicit, while other constraints are explicitly indicated by specifying the set of constraints that, together with the unspecified logical constraints, determine what counts as possible.

EXAMPLE 3

```
1: *thief(x) ⇒ *punishable(x)
4: *minor(x) ⇒ *protected(x)
5: *protected(x) ⇒ *~punishable(x)
```

The rules 1, 4 and 5 are logically inconsistent, because the rules 1 and 5 are logically inconsistent. Rule 4 plays no role in this connection. In example 4 in the next section, we will encounter a related situation in which rule 4 does have a role to play.

4. RULES AS CONSTRAINTS

Rules function as constraints on the worlds in which they exist. In the Netherlands the rule exists that thieves are punishable. As a consequence the states of affairs that somebody is a thief and that he is not punishable are, barring exceptions to the rule, legally incompatible. In a legal system where this rule does not exist, these states of affairs might be compatible. The phenomenon that rules can function as constraints on possible worlds has implications for the above theory of rule consistency. To illustrate this, I will adapt example 3:

EXAMPLE 4

```
1: *thief(x) ⇒ *punishable(x)
4: *minor(x) ⇒ *protected(x)
-----------------------------------------------
c = {L; 5: *protected(x) ⇒ *~punishable(x)}
```

The third rule of example 3 is removed from the set of rules that is evaluated with regard to its consistency, and added to the set c of constraints that govern the world in which the rules 1 and 4 are evaluated.[13] The first thing to notice is that the remaining rules 1 and 4 are *logically* consistent. This is not surprising, because the inconsistency of the rules 1, 4 and 5 in example 3 depended on the presence of rule 5. If this rule is removed from the set, the logical consistency is restored.

However, if the removed rule is added to the constraints relative to which consistency is evaluated, the rules 1 and 4 become inconsistent relative to the constraints in c, since these constraints make the states of affairs that

[13] The L in the set of constraints is shorthand for the set of logical constraints that is not mentioned explicitly if the only constraints are those of logic.

somebody is punishable and that he is protected incompatible. So the rules 1 and 4 are logically consistent, but they are inconsistent relative to the set c of constraints which includes rule 5. It makes no difference whether the rules 1, 4 and 5 are evaluated on logical consistency, or that the consistency of the rules 1 and 4 is evaluated relative to constraints including rule 5. Nevertheless there is a difference if only the consistency of the rules 1 and 4 is considered. They are logically consistent, but relative to rule 5 they are inconsistent.
The following two examples aim to illustrate that it makes a difference whether a rule is part of a set that is evaluated on logical consistency, or whether this rule is taken as part of the constraints:

EXAMPLE 5

```
1: *thief(x) ⇒ *punishable(x)
3: *minor(x) ⇒ *~punishable(x)
6: *minor(x) ⇒ *~thief(x)
```

EXAMPLE 6

```
1: *thief(x) ⇒ *punishable(x)
3: *minor(x) ⇒ *~punishable(x)
-----------------------------------
c = {L; 6: *minor(x) ⇒ *~thief(x)}
```

We have seen in example 3 that the rules 1 and 3 are *logically* inconsistent. This logical inconsistency is maintained if rule 6 is added to the rules 1 and 3, because the resulting set still has an inconsistent subset and is therefore inconsistent. However, the situation changes if rule 6 is taken *as one of the constraints* relative to which the consistency of the rules 1 and 3 is evaluated, as in example 6. The conditions of the rules 1 and 3 are not compatible relative to a background that contains rule 6. As a consequence the rules 1 and 3 are consistent *relative to this set of constraints*, even thought the rules 1, 3 and 6 are logically inconsistent. Apparently it makes a difference whether a rule is considered as part of the set that is evaluated with regard to its logical consistency, or as part of the constraints for the consistency of the other rules.

In the examples 3 and 4 it did not matter for the consistency of the set whether rule 5 was part of the set, or part of the background, while in the examples 5 and 6 it makes a difference whether rule 6 is part of the set of evaluated rules, or part of the constraints. This difference can be explained by pointing out that in the examples 5 and 6, rule 6 made the *conditions* of the rules 1 and 2 incompatible, while in the examples 3 and 4, rule 5 made the *conclusions* of the rules 1 and 2 incompatible.

This is an important observation regarding the influence of the constraints on the consistency of a set of rules. The more demanding the background, the more strict are the constraints on the states of affairs that are compatible. If a set of states of affairs is incompatible relative to a certain background c, it will be incompatible relative to any background c' which imposes more constraints than c.[14] The consistency of a set of rules varies positively with the compatibility of the conclusions of these rules and negatively with the compatibility of the rule conditions. As a consequence, the addition of constraints to the background contributes to the consistency of rules by making the rule conditions incompatible. Addition of constraints detracts from the consistency of rules by making the rule conclusions incompatible. The overall effect of adding to the background of constraints on the consistency of a set of rules depends on the conditions and conclusions of the rules that are evaluated with regard to their consistency and the contents of the constraints that are added to the background.

5. CONDITIONLESS RULES

Until now we have only considered conditional rules. There are also rules without conditions, such as the rule that it is forbidden to steal. Such rules share some characteristics with rules that have conditions, in particular that they can have exceptions. For the evaluation of their consistency they seem a little different, however. The first part of the definition of rule consistency, that there is a set of compatible states of affairs that satisfies the conditions of all the rules, seems not to apply to conditionless rules.

This seeming complication can be remedied by treating conditionless rules as rules with conditions that are always satisfied. If conditions that are always satisfied are denoted by the term `*true`, conditionless rules are represented as rules with `*true` as their condition part. For instance[15]:

`*true ⇒ *od(x, ¬steal)`

The rules that it is forbidden to steal (all actors ought to refrain from stealing) and that it is permitted to steal are then easily shown to be

[14] This observation does not hold without restrictions if exception-introducing constraints are added to the background.

[15] In this example, the two place predicates `Od` and `Pd` stand for ought-to-do and permitted-to-do respectively. Their first parameter is a set of actors (all actors if it is a variable), and the second parameter stands for an action type. The operator ¬ maps action types onto action types. The intended interpretation is that `¬action` stands for refraining from `action`.

inconsistent with regard to the constraint that an action is not both forbidden and permitted for the same actors[16]:

EXAMPLE 7

```
7: *true ⇒ *od(x, ¬steal)
8: *true ⇒ *pd(x, steal)
―――――――――――――――――――――――――――――――――――――――――――――――――――――――――――――
c = {L; ∀x,action(Od(x, ¬action) ≡ ~Pd(x, action))}
```

The inconsistency of the rules 7 and 8 against the background c is illustrated by any case, since any case satisfies the conditions of these two rules.

6. EXCEPTIONS TO RULES

It is not uncommon that two rules in a legal system attach incompatible consequences to a case. For instance, the rule that an owner is allowed to do anything he likes with his property collides with many rules that limit his property right. In such cases, the law contains a *prima facie* rule conflict. Many prima facie rule conflicts turn out not to be *actual* conflicts, because one of the prima facie conflicting rules is left out of application by making an exception to it. Two or more rules are in actual conflict when they *actually* apply to one and the same case, and attach incompatible consequences to this case.

I will again use an example to sharpen our intuitions concerning the effect of exceptions to the consistency of rules. Take the following three rules:

1: Thieves are punishable.
3: Minors are not punishable.
9: In case of minors the rule that thieves are punishable does not apply (there is an exception to it).

These rules interact in case of a minor who is a thief. If rule 9 is left out of consideration, the rules 1 and 3 are inconsistent, because they lead to incompatible results in case of a minor thief. Rule 9 prevents that rule 1 is applied, however, so that the prima facie rule conflict is not actualized.[17]

[16] Some would want to include this constraint into the set of logical constraints. In general the example leaves a lot to be said concerning deontic logic. This is beyond the scope of this paper, however.

[17] A logical account of the operation of exceptions to rules can be found in chapter 3.

Therefore the rules 1 and 3 are in my view consistent with regard to a set of constraints that includes rule 9, although they are logically inconsistent.

The observation that rules can have exceptions which prevent them to come into an actual conflict leads to the following adapted version of the above theory of rule consistency:

The rules in a set s are consistent relative to a set of constraints c if and only if it is not so that there are a subset s' of s and a possible case f such that

a. *the states of affairs in f are compatible relative to c,*
b. *the states of affairs in f satisfy the conditions of all the rules in s',*
c. *there is no exception to either one of the rules in s', and*
d. *the rules in s' attach consequences to f that are incompatible relative to c.*

Exceptions ought to be exceptional. I take this to mean that there are no exceptions to rules unless there are special reasons to make them. Such reasons exist if there are rules that attach the presence of an exception to a rule to the presence of some facts. For instance, if somebody is a minor, the rule that thieves are punishable should not be applied to him. Since this exception holds, in principle, for all minors, there is a rule[18] to the effect that in case of minors there is an exception to (amongst others) the rule that thieves are punishable. In general I propose the theory that there can only be an exception to a rule if there is another rule, the conditions of which are satisfied, and which is not subject to an exception itself, that has as its conclusion that there is an exception to the first mentioned rule.[19] I will call exceptions that do not satisfy the above mentioned constraint *free-floating exceptions*. A good theory about rule exceptions should, in my opinion, make such free-floating exceptions impossible.

The implications of the amendment to the theory of rule consistency which takes exceptions into account are illustrated by the following examples:

[18] Logically the presence of an exception to a rule can be based on any reason against applying this rule that outweighs the rule's applicability and possible other reasons for applying it. (See chapter 3, section 5.4.) From the perspective of a legislator, however, the obvious way to create exceptions to rules is to make rules to this effect.

[19] This theory will be formalized in section 11.

EXAMPLE 8

```
1: *thief(x) ⇒ *punishable(x)
3: *minor(x) ⇒ *~punishable(x)
-----------------------------------------
c = {L; 9: *minor(x) ⇒
    *exception(*thief(x) ⇒ punishable(x))}
```

Rule 9 holds that if somebody is a minor, the rule that thieves are punishable does not apply to him.

The rules 1 and 3 by themselves are logically inconsistent. Inclusion of rule 9 in the background makes that if the conditions of rule 3 are satisfied, there is an exception to rule 1, which takes the rule conflict away. As a consequence, the rules 1 and 3 are consistent relative to a background that contains rule 9.

Exceptions can also make a consistent set of rules inconsistent[20]:

EXAMPLE 9

```
1: *thief(x) ⇒ *punishable(x)
3: *minor(x) ⇒ *~punishable(x)
c = {9: *minor(x) ⇒
        *exception(thief(x) ⇒ *punishable(x));
10: *second_offender(x) ⇒ *exception(rule-9)}
```

We have seen in example 8 that the rules 1 and 3 are consistent with respect to a set of constraints that includes rule 9. The addition of rule 10 to the background makes that there is no guarantee anymore that, in case of a minor, there is an exception to rule 1. This is illustrated by the case that John is not only a thief and a minor, but also a second offender. In that case there is an exception to rule 9, and presumably no exception to rule 1. This illustrates that if rule 10 is added to the background, there are possible cases in which the conditions of both the rules 1 and 3 are satisfied, and in which these two rules are in actual conflict.

7. MODEL THEORY FOR RULES

Model-theoretic semantics for logic specifies the meanings of logical operators by means of the truth conditions of sentences in which these

[20] In the formalization of rule 10, rule 9 is referred to by 'rule-9'.