one reason for prosecuting and one reason against. We must assume that there are no other relevant reasons[23]:

```
r+(*should_do(public-prosecutor, prosecute(johnny)) =
   {*thief(Johnny)})

r-(*should_do(public-prosecutor, prosecute(johnny)) =
   {*minor(Johnny)})
```

To draw a conclusion whether the public prosecutor should prosecute, information is needed about the relative weight of these sets of reasons:

```
{*minor(johnny)} > {*thief(johnny)}
```

The (set consisting of the single) fact that Johnny is a minor outweighs the (set consisting of the single) fact that he is a thief (regarding the conclusion that he should be prosecuted).

Given this information, it is possible to derive that the contributive reasons pleading against prosecution outweigh the contributive reasons for prosecution:

```
r-(*should_do(public-prosecutor,prosecute(johnny))) >
    r+(*should_do(public-
prosecutor,prosecute(johnny)))
```

To continue, we also need the information that there are no decisive reasons concerning the issue whether the public prosecutor should prosecute Johnny[24]:

```
~∃*r(Dr(*r,
  *should_do(public-prosecutor,prosecute(johnny))))

~∃*r(Dr(*r,
    *~should_do(public-prosecutor,prosecute(johnny))))
```

Given the absence of decisive reasons, this leads to the conclusion that the public prosecutor should not prosecute Johnny:

```
~Should_do(public-prosecutor, prosecute(johnny))
```

[23] Again, this assumption receives further discussion in section 7

[24] This assumption receives further discussion in section 7 too.

5. RULES

The way RBL deals with rules does not imply a theory about the logical behavior legal rules. It is best considered as a modeling tool that is more or less suitable to deal with legal phenomena. To what extent legal rules should be modeled as RBL-rules depends on how one sees the logical behavior of legal rules. However, the logic of rules in RBL was inspired by a particular view of rules as authoritative decisions about how to deal with types of cases. This view has been exposed by Raz, who stated that mandatory rules are exclusionary reasons and by Schauer, who considers rules to be 'entrenched generalizations'.[25]

5.1 The representation of rules in RBL

Rules are usually assumed to have a conditional structure. They consist of a condition part and a conclusion part and the point of rules is that if their conditions are satisfied, their conclusions obtain. In RBL rules are treated as logical individuals, denoted by a function expression. (Something like: the rule with conditions `a` and conclusion `b`.) RBL has a dedicated function constant that has rules as its values: $\Rightarrow/2$. Both the first parameter and the second parameter are terms denoting abstract states of affairs. The first parameter stands for the rule conditions, the second for the rule conclusion. For instance, the following term denotes the rule that thieves are punishable:

$$*\mathtt{thief}(x) \Rightarrow *\mathtt{punishable}(x)$$

[25] Raz 1975, 73f. and Schauer 1991, 47.

Rules are not considered to be sentences (nor, in general, linguistic entities) and in RBL the expressions that denote rules are used as terms. As a consequence, rules have no truth values. But like ordinary entities, they exist in time, have characteristics and stand in relations to other entities, including other rules. Moreover, it is possible to formulate rules about rules. The following term denotes the rule (of recognition) that rules made by the legislator are valid[26]:

```
*rule(r) & made_by(the_legislator, r) ⇒ *valid(r)
```

5.2 Rule application

Traditionally reasoning with rules is analyzed as arguments of the form Modus Ponens: If the rule conditions are satisfied, the rule conclusion follows from the rule and the description of the facts. Lawyers sometimes say that the facts of a case are subsumed under the rule and for this reason the traditional model may be called the *subsumption model*.[27]

The logic of rule application under RBL is somewhat different from this subsumption model. If a rule applies (is applied) to a case, the conclusion of the rule holds for this case. In terms of reasons, we might say that the application of a rule to a case is a decisive reason for the rule conclusion to hold. For instance the application to John of the rule that thieves are punishable is a decisive reason for the conclusion that John is punishable.

RBL has a dedicated predicate constant to express that a rule applies: `Applies/1`. The parameter of `Applies` is the relevant instantiation of the rule formulation. For instance, the following sentence expresses that the rule that thieves are punishable applies (John is a thief):

```
Applies(*thief(john) ⇒ *punishable(john))
```

[26] One of the major advantages of treating rules as logical individuals is that this facilitates reference to rules and consequently the representation of rules about rules. When rules are treated as full sentences, rules about rules should be formulated in a meta-language, with all complications that result from that. For instance, the following argument is hard to formalise in most logical languages:
- All rules made by the legislator are valid
- The legislator made the rule that thieves are punishable
- John is a thief
- Therefore: John is punishable

[27] This subsumption model of rule application is correct if rules are interpreted as case-legal consequence pairs, as described in chapter 1, section 3.2. The logic of rule application that is described here, is better adapted to rules as analyzed in chapter 6, section 8.

When a rule applies to a case, this is a decisive reason for the rule's conclusion:

Consequence applies:

$\forall$*iconds,*iconc(Dr(*applies(*iconds $\Rightarrow$ *iconc), *iconc))

Obviously, only existing rules can apply to a case. In RBL the predicate `Valid/1` is used to represent the validity of rules, which is taken to be the way in which rules exist. The parameter of this predicate is a term denoting an (uninstantiated) rule. For instance:

Valid(*thief(*x*) $\Rightarrow$ *punishable(*x*))

Because application presupposes validity, the following holds:

Application presupposes validity:
Let `ir` denote an instantiation of the rule `r`. Then

$\forall$*r*,*ir*(Applies(*ir*) $\rightarrow$ Valid(*r*))

Whether a rule applies, depends in RBL normally[28] on a balance of reasons for and against application. In particular, a rule can only apply if the reasons for application outweigh the reasons against application. This means that whether the rule that thieves are punishable applies, depends on both the contributive reasons pleading for application and the contributive reasons against application and therefore *not merely on whether the rule conditions are satisfied*. In this respect the RBL-model of rule application differs considerably from the traditional subsumption model. The crucial differences between the RBL-model and the subsumption model are that the RBL-model

1. allows reasons against the application of a rule that collide with reasons for application and
2. does not state which facts count as reasons for application nor as reasons against application.

Although the reason-based model as described above does not specify which facts count as reasons for and against application of a rule, there are very plausible ways to elaborate this model. I will discuss three of such elaborations, which will be formalized as RBL-axioms.

[28] An exception, dealt with in section 5.3 is that there is a decisive reason against application of the rule.

5.3 Applicability as a contributive reason to apply a rule

The first extension is the assumption that if the facts of a case satisfy the conditions of a rule – to be abbreviated as that the rule is *applicable* to the case - this is a *contributive* reason why the rule applies.[29] RBL has a dedicated predicate constant to express that a rule is applicable: `Applicable/1`. The parameter of `Applicable` is the instantiation of the rule formulation that is satisfied by the case facts. For instance, the following sentence expresses that the rule that thieves are punishable, is applicable (John is a thief):

`Applicable(*thief(john) ⇒ *punishable(john))`

A rule is applicable to a case, if the facts of this case satisfy the rule conditions. Moreover, it may be debated whether hypothetical, rather than actually existing rules can be applicable. In the present formalization, it is assumed that applicability presupposes validity.

Applicability of a rule:
Let `*iconds` denote an instantiation of the abstract state of affairs `*conds` and let `*iconcl` denote the instantiation of the abstract state of affairs `*concl` under the same substitution. Then it holds that:

`∀*iconds,*iconcl(Applicable(*iconds ⇒ *iconcl) ≡`
`  Iconds & ∃*conds,*concl(Valid(*conds ⇒ *concl)))`

If a rule is applicable, this is a contributive reason why this rule applies:

Applicable and Applies:
`∀r(Ar(*applicable(r), *applies(r)))`

This relation between the applicability and the application of a rule looks similar to the subsumption model, but there is a crucial difference, because on the subsumption model, the applicability of the rule is a *decisive* reason to apply the rule. What does this difference mean?

First, it means that even if a rule is applicable, there may still be reasons against applying the rule, reasons which may, but need not, outweigh the applicability of the rule as a reason for application. This might, for instance, be the case if application of the rule would be against the purpose of the rule.

Second, it means that there can be a decisive reason against application of the rule and such a decisive reason by definition brings about that the rule does not apply, even if it is applicable. A decisive reason against application

[29] Notice that the applicability of a rule is not the same as its application. The very point of the RBL-model is that applicability is merely a contributive reason for application.

of an applicable rule obtains normally when another rule with an incompatible conclusion is also applicable to the case and this second rule has precedence over the first rule. For example, in Dutch rental law, the rules concerning the rent of business accommodations are sometimes in conflict with the general rules about rent and if such a conflict occurs, the more specific rules concerning the rent of business accommodations have precedence over the general rules about rent. The applicability of a rule that has precedence over another rule is normally a decisive reason against applying the latter rule.

Third, the first elaboration of the reason-based model of rule application means that if a rule is applicable and there exists therefore a contributive reason for applying the rule and there is no reason, either contributive or decisive, against application, the rule applies and its consequence is attached to the case.[30] *This is the normal situation and in this situation the reason-based model and the subsumption model of rule application lead to the same results.* It is this kind of situation that has lent some plausibility to the subsumption model, because the shortcomings of that model are not relevant in the normal situation.

5.4 Non-applicability as a contributive reason against application

The second elaboration of the reason-based model of rule application is that if a rule is not applicable to a case, this is a contributive reason against applying the rule to this case:

Non-applicability:
$\forall r$(Ar(*~applicable(r), *~applies(r)))

At first sight this extension seems superfluous, because if a rule is not applicable, there seems to be no reason for applying it, so the issue of application seems not to arise at all. The relevance of the second extension only becomes clear in the light of the third elaboration of the reason-based model of rule application.

This third elaboration is that there can be other reasons for applying a rule than only the applicability of the rule in question. The reason-based model itself does not specify what these other reasons might be; it only leaves the possibility open that there are other reasons for application. The

[30] This is an application of the principle that any non-empty set of reasons for a conclusion outweighs the empty set of reasons pleading against this conclusion. See section 3.6.

obvious role of this possibility to apply a rule when it is not applicable is to allow *analogous application* of a rule.[31]

If a rule applies to a case, although its conditions are not satisfied by that case, the condition part of the rule remains uninstantiated in the parameter of the Applies-predicate. If, for instance, the rule that thieves are punishable is applied analogously to a case of almost-theft (which is not possible in many legal systems), this can be expressed formally as

`Applies(*thief(x) ⇒ *punishable(john))`

If there is a contributive reason to apply a rule analogously, this reason must be weighed against the non-applicability of the rule and possible other contributive reasons against application. Whether the rule in the end applies, depends on the balance of all contributive reasons pleading for and against application.

When a rule is applicable and nevertheless not applied, it is said that there is an exception to the rule. Exceptions have no special role in the present version of RBL[32], but nevertheless a definition comes in helpful:

Definition rule exception:

`∀r(Exception(r) ≡ Applicable(r) & ~Applies(r))`

If there is an exception to a rule, then either the reasons against application outweigh the reasons for application, or there is a decisive reason against application.

6. REASONING WITH RULES

The RBL model of rule application is somewhat more complicated than the simple deductive model according to which rule application is represented as an argument of the form Modus Ponens. To illustrate the differences, both with the deductive model and with reasoning with principles, I will use more or less the same example as in section 4. The case deals with a thief, Johnny, who is a minor. First I disregard the fact that Johnny is a minor and discuss the simple situation in which the rule that thieves are punishable is applied to Johnny's case. Then I take another rule into consideration, namely the rule

[31] A more extensive discussion of analogous rule application can be found in Verheij and Hage 1994 and in Hage 1997 (RwR), 118f.

[32] In Hage 1996 and 1997 (RwR), I used the Excluded-predicate more or less for what I now call exceptions.

that minors are not punishable and the example becomes an example of rule conflict.

There are two other interesting cases, namely when there are contributive reasons not to apply the rule that thieves are punishable and when this rule might be applied analogously. Because the treatment of these two cases under the present version of RBL is not very different from the ways in which they are treated under RBL as presented in *Reasoning with Rules*, I refer the interested reader to that work.[33]

6.1 Simple rule application

The case at issue can be represented by the following premises:

```
Thief(johnny)
```

and

```
Minor(johnny)}
```

The validity of the rule that thieves are punishable is represented by:

```
Valid(*thief(x) ⇒ *punishable(x))
```

The facts of the case instantiate the rule conditions and as a consequence it is possible to derive that:

```
Applicable(*thief(johnny) ⇒ *punishable(johnny))
```

and

```
Cr(*applicable(*thief(johnny) ⇒ *punishable(johnny)),
    *applies(*thief(johnny) ⇒ *punishable(johnny)))
```

We have derived one contributive reason why the rule that thieves are punishable applies in Johnny's case. What we need, however, are the sets of all reasons pleading for and against application of this rule and weighing knowledge that tells us which one of these sets outweighs the other set. The case facts do not provide us with this information and there are three ways to deal with this 'problem'. One is to be content with the outcome that nothing relevant can be derived. This is obviously the wrong 'solution', because we should be able to derive that Johnny is punishable – at least if we disregard that he is a minor. The second way is to add information to the case, to the effect that there are no reasons why the rule that thieves are punishable

[33] Hage 1997 (RwR), 187f. and 191f.

should not be applied. The third way is to assume this information by default. In both cases, we can derive that:

r^{+}(*applies(*thief(johnny) ⇒ *punishable(johnny))) =
{*applicable(*thief(johnny) ⇒ *punishable(johnny))}

and

r^{-}(*applies(*thief(johnny) ⇒ *punishable(johnny))) = ∅

Then we can apply the weighing knowledge that, by default, any non-empty set outweighs the empty set of reasons and draw the conclusions that

Applies(*thief(johnny) ⇒ *punishable(johnny))

and

Punishable(johnny)

6.2 Rule conflicts

Let us now consider the rule that minors are not punishable, the validity of which is represented by:

Valid(*minor(*x*) ⇒ *~punishable(*x*))

If the rule that thieves are punishable would be disregarded, treatment of this rule analogous to that of the rule that thieves are punishable would lead to the conclusion that Johnny is not punishable. Apparently the joint application to Johnny's case of the rules that thieves are punishable and that minors are not punishable leads to inconsistency.

Rather than accepting this inconsistency, the law deals with such cases by means of conflict rules, that specify which of two conflicting rules precedes. One such conflict rule says that the more specific rule precedes the more general one. Arguably the rule that minors are not punishable is more specific than the rule that thieves are punishable.[34] Instead of representing the argument from specificity to precedence, I will directly represent the precedence of the rule about minors to the rule about thieves in Johnny's case as follows:

[34] Arguably, but not from a logical point of view. One needs legal knowledge to see that the rule about minors is meant to make an exception to general rules about punishability and is *in that sense* more specific. Apparently specificity is a conclusion, rather than a premise of precedence.

```
Precedes(*minor(johnny) ⇒ *~punishable(johnny),
 *thief(johnny) ⇒ *punishable(johnny))
```

Because precedence makes only sense in case of applicable rules (if the rules would not both be applicable, the precedence issue would not arise), the applicability of the precedence relation presupposes the applicability of the rules for which this relation holds. Therefore the following might be added as an axiom to RBL:

Precedence implies applicability:

```
∀*cond1,*conc1,*cond2,*conc2 (
  Precedes(*cond1 ⇒ *conc1, *cond2 ⇒ *conc2) →
   Applicable(*cond1⇒*conc1) &
     Applicable(*cond2 ⇒*conc2))
```

If some rule has precedence over another rule, this is in general a decisive reason[35] not to apply the last rule:

Effect of precedence:

```
∀r1(Valid(*∃r2(precedes(r2, r1)) ⇒ *~applies(r1)))
```

Given this axiom, it follows from the precedence of the rule about minors over the rule about thieves that the rule that thieves are punishable does not apply. As a consequence, only the rule that minors are not punishable applies and the conclusion that follows is that Johnny is not punishable (because he is a minor).

7. REASON-BASED LOGIC AS A NON-MONOTONIC LOGIC

The examples in the sections 4 and 6 illustrated amongst others that many reasoning tasks presuppose information that is often not explicitly available. This includes information about all the reasons for or against a particular conclusion and about possible exceptions to rules. For practical purposes this lack of explicit information is seldom problematic, because we are often willing to draw conclusions in the absence of relevant information and

[35] I represent this decisive reason by means of the validity of a rule, rather than by means of a material conditional. The difference is that there cannot be exceptions to material conditionals, while rules can have exceptions. Notice that this possibility of exceptions does not conflict with the fact that a decisive reason determines its conclusion. If there is an exception to the rule underlying the decisive reason, this means that *there is no decisive reason*, not that the decisive reason does not determine its conclusion.

remain prepared to withdraw these conclusions again if what we assumed to be the case (e.g. that there are no other relevant reasons than the ones already taken into account) turns out to be incorrect. In other words, for practical purposes we work with default assumptions and recognize that the conclusions based on these assumptions are justified only to the extent that these assumptions are correct.

The question how this kind of reasoning 'by default' can be implemented in a system of logic is quite challenging, because there is no obvious way in which we tend to deal with defeasible reasoning. For instance, we can take a default assumption to be true until:

- the contrary was actually proven (possibly according to some specific procedure),
- the contrary is provable (according to some logic),
- the contrary was actually accepted (for whatever reasons),
- or until it is more reasonable to accept the contrary (given as yet unspecified standards for rationality).

All of these different variants would lead to different logics and all of them have at least something to speak for them. Probably all of these variants on defeasibility play a role in actual reasoning practices. In this book I will not attempt to develop one or more of such logical systems. Instead I will confine myself to pointing out that the phenomenon to be captured by a logic for defeasible reasoning is that an acceptance that is justified relative to a particular acceptance set, need not be justified relative to another acceptance set. This means that the notion of validity that is at stake is that of justified acceptance, not that of necessary truth of the conclusion given the truth of the premises. Because the phenomenon to be captured is justification relative to the premises of the argument, the only information that has to be taken into account to judge the acceptability of the conclusion is the information given in the premises. All other information is irrelevant. This means that the only reasons to be taken into account are the reasons that 'follow' from the premises of the argument and that there are no 'unexpected' decisive reasons or contributive reasons. Because the exclusion of abstract reasons is reason-based, the absence of unexpected reasons means that there are no unexpected exclusions either.

Even the insight that there can be no relevant 'unexpected' reasons does not suffice to overcome the problem that reasoning with contributive reasons is global. With this I mean that the conclusions based on the balancing of contributive reasons must be based on *all* the reasons that 'follow' from the premises and not merely some of them. One or more arguments that establish the presence of one or more reasons can by themselves not establish that these reasons are all the relevant reasons. Somehow one must

have disposal over the sets of all reasons that 'follow from' the premises and that plead for and against a particular conclusion. Reasoning *within* the formalism of a logical system cannot lead to this kind of information. To obtain all relevant reasons, one must resort to reasoning *about* the logical system.

Instead of trying to develop a metalogical theory about RBL by means of which can be proven which reasons follow according to RBL from a particular set of premises, I propose to deal with the defeasibility of conclusions in RBL by means of a dialectical setting. If somebody shows that there is a reason for a conclusion, the conclusion is provisionally justified. His opponent can take this justification away by producing a reason against this conclusion. The proponent can then either argue for weighing knowledge according to which his reason is stronger than the reason adduced by the opponent, or he can produce additional reasons, etc.[36]

[36] These ideas have been elaborated in Hage e.a. 1994 and in Lodder 1999. See also chapter 9.

APPENDIX

The present version and presentation of RBL differs in several main aspects from the version described in *Reasoning with Rules* (RwR). I will briefly mention them in turn.

RBL AS A FLEXIBLE EXTENSION OF PREDICATE LOGIC

In RwR I presented RBL as a non-monotonic logic that was especially made to deal with rules, principles, goals and reasons. Here I presented RBL as an extension of predicate logic with only one special characteristic, namely that it deals with (reasoning with) reasons. The parts that deal with rules, principles, goals and with the comparison of alternatives (see chapter 4), can be added in the form of additional axioms. This style of presentation reflects the underlying philosophy that there is no clear boundary between pre-formal logic and domain knowledge and that it is a matter of choice which parts of a domain are considered as sufficiently fixed to treat them as logically necessary and incorporate them in a system of formal logic.[37]

RBL AS A MONOTONIC LOGIC

RBL is presented here as a monotonic, even a deductive, logic. The special needs of the legal domain which ask for a non-monotonic logic (see chapter 1, section 4) are delegated to an unspecified dialectical setting within which the present version of RBL can function.[38]

REPLACEMENT OF PRINCIPLES BY ABSTRACT REASONS

In RwR, principles took a central place. If a principle is valid, its instantiated conditions would normally be a (contributive) reason for its instantiated conclusion. In the present version of RBL, principles are replaced by abstract reasons. My main motivation for doing so, is that abstract reasons play an important role in arguments in which alternatives are compared (see chapter 4), and that these abstract reasons could not easily be modeled by means of principles. On the other hand, the operation of principles can well be described in terms of abstract reasons.

[37] Cf. Hage 2001 (LL).
[38] See also chapter 8.

Relation between applicability, application and exceptions to rules

In RwR, a rule would be applicable if its conditions are satisfied and if it is not excluded. Applicability would then be a (contributive) reason why the rule ought to be applied. In the present version of RBL, the technical notion of applicability is used as shorthand for the satisfaction of the rule conditions. The notion of exclusion does not play a role anymore in connection with rules (but it does in connection with abstract reasons). Moreover, in the present version I use the notion of an exception to a rule. Exceptions do not play a real role in reasoning with rules, however. They are merely a name for the phenomenon that a rule is applicable, but nevertheless not applied.

Absence of deontic notions

In RwR the logic of rules was described in deontic terms. Applicability would be a reason why the rule *ought to be* applied. For the sake of logical simplicity, I have dropped this deontic element. I still believe, however, that the peculiar phenomena connected to the deontic element of reasoning with rules that I described in RwR (deontic collapse and deontic inflation; see RwR 205) are interesting and in need of explication.

Chapter 4

COMPARING ALTERNATIVES

1. RIGHT AND BETTER

When your old car has broken down and you must decide which brand your new one should be, your main problem will probably not be of a logical nature. Nevertheless, if you have to make a choice between for instance a Mercedes, a Volvo, and a Porsche, the logic underlying the decision is interesting. Each brand of car has advantages and disadvantages, and rational decision making requires a form of balancing these (dis)advantages. The easiest case would be if there were a common scale against which all brands could be measured, because then the only 'logic' involved would be to pick the brand with the highest score. However, when this method is not available, other ways to rationalize the decision making process must be looked for.

Another way to deal with this kind of question is to transform it into the issue whether one should buy a particular brand of car, for instance a Volvo. Logically this would boil down to balancing the reasons for and against buying a Volvo. It is well imaginable that if the question is framed this way, the reasons for buying a Volvo outweigh the reasons against buying one. It is, however, equally well imaginable that in a similar way the reasons for buying a Mercedes outweigh the reasons against buying one, and that the reasons for buying a Porsche outweigh the reasons against buying a Porsche. If the decisions are taken as independent from each other, one might well end up with buying three cars! What we need to know is not merely whether it is *right* to buy a particular brand of car, but whether it is *better* to buy a Mercedes, a Volvo, or a Porsche.

Obviously the decisions are not independent from each other, and one way to deal with this is to treat every reason for buying a Porsche as a reason against buying a Mercedes and against buying a Volvo. On this approach, the mutual dependence between the decisions is taken into account and a decision to buy one particular brand of car is implicitly a decision not to buy a car of one of the other brands.[1]

This is a viable way to deal with the issue, as long as the number of alternatives is limited. If the number of alternatives is large, the situation becomes problematic, because every reason to buy a particular brand of car becomes a reason against buying a car of any other brand. Apart from the complexity this leads to, it is unrealistic, because some reasons to buy a particular brand of car will also be reasons pleading for other brands. For instance, one reason to buy a Mercedes is that it is a safe car. This would also be a reason to buy a Volvo. Another reason to buy a Volvo might be that it fits with the image that one wants to create. This same reason might also plead for buying a Porsche. A reason against buying a Volvo would be that it is less suitable for car racing, and this reason pleads also against buying a Mercedes. And so on …

A more attractive way would be to collect for each brand of car the reasons pleading for it and the reasons pleading against it. Every brand that has stronger reasons pleading against it than for it, can be disregarded. The remaining brands should be compared. If brand A is in some respects better than brand B, and in no respect worse, brand A is preferable to brand B and brand B can be disregarded. It is well possible that this process of elimination leaves only alternative over, and then the decision can be taken purely on basis of qualitative reasoning. If more than one alternative remains, additional decision making is necessary.

2. QUALITATIVE COMPARATIVE REASONING

Suppose that one must choose between buying a Volvo and a Mercedes. A Volvo has two reasons pleading for it, namely that it is a safe car, and that there is a Volvo dealer next door. It has the disadvantage that it is an expansive car. A Mercedes is also expensive, but has (in the example) only one advantage, namely that it is a safe car. There happens to be no Mercedes dealer in the neighborhood. Under these circumstances, everything that pleads for a Mercedes also pleads for a Volvo, but a Volvo has an additional

[1] This approach is taken in Brewka and Gordon 1994 and in Gordon and Karacapilidis 1997.

reason pleading for it, namely the availability of a dealer nearby. Moreover, a Volvo and a Mercedes have the same reason pleading against it, namely that they are expensive. It seems, therefore, that a Volvo is preferable to a Mercedes. This is a reasonable conclusion, even in the absence of any information concerning the (relative) weights of the reasons that the cars are safe, that there is a Volvo dealer nearby, and that the cars are expensive. Analogously it is reasonable to conclude that a Mercedes is preferable to a Porsche, if a Mercedes and a Porsche have the pro-reason in common that they are German cars (for those who like German cars), and they also share the con-reason that they are expensive, while a Porsche has the additional disadvantage that it liable to be stolen. In general alternative A is preferable to alternative B if either:

1. the set of reasons pleading for A is 'stronger' than the set of reasons pleading for B, while the set of reasons pleading against A is not 'stronger' than the set of reasons pleading against B; or
2. the set of reasons pleading against B is 'stronger' than the set of reasons pleading against A, while the set of reasons pleading for B is not 'stronger' than the set of reasons pleading for A; or
3. both 1 and 2 hold.

2.1 Comparing reason sets

Until now, the examples dealt with the qualitative comparison of *alternatives* in terms of reasons pleading for and against them. It is also possible to apply qualitative comparative reasoning to *sets of reasons*. These sets can be compared qualitatively with regard to their relative 'strength'.

In the above characterization of when one alternative is preferable to (better than) another alternative, I placed the word 'stronger' between quotes, because the notion of strength involved needs to be elaborated. In the examples, I implicitly assumed that supersets were 'stronger' than their subsets, but intuitive as this may be at first sight, it ignores that individual reasons have a dimension of weight and that this dimension may interfere with the sheer number of reasons. For instance, if a Volvo is much more expensive than a Mercedes, its additional expensiveness might be more important than the presence of a dealer nearby, with as consequence that a Mercedes might be preferable to a Volvo, even though a Volvo has more reasons pleading for it.

Moreover, the suggestion that the same reason can plead for or against several alternatives is somewhat misleading. It may seem that their safety is a reason that pleads both for buying a Volvo and a Mercedes, but on closer examination the concrete reason for buying a Mercedes is that *a Mercedes* is

a safe car, while the reason for buying a Volvo is that *a Volvo* is a safe car. Buying a Mercedes and buying a Volvo share the abstract pro-reason 'being a safe car', but they do not share concrete reasons.[2] However, the actual reasons for buying these cars are the concrete reasons, not the abstract ones. I will deal with this issue in terms of 'similar reasons', concrete reasons that instantiate the same abstract reason. For instance, that a Volvo is a safe car is as reason similar to the reason that a Mercedes is a safe car.

The issue of weights has also to do with this distinction between abstract and concrete reasons, because the weights of reasons are attached to concrete reasons. This means that the concrete reason that a Mercedes is a safe car may have a different weight than the concrete reason that a Volvo is a safe car. One might argue that abstract reasons have a dimension of weight too and that concrete reasons inherit this weight 'by default'. In this case, the reason that a Mercedes is a safe car would by default have the same weight as the reason that a Volvo is safe car. Let us assume that this is correct, but nevertheless it may occur that similar reasons in a concrete case have different weights and that this interferes with the number of reasons pleading for and against alternatives. Only if the weights of the similar reasons are identical, the strengths of two sets of reasons can be compared qualitatively by means of the numbers of their elements.

The last point can also be turned around: if two sets of reasons have similar elements, their relative strengths can be compared on the basis of the weights of their elements. For instance, if both a Volvo and a Mercedes have one reason pleading for them, namely that they are safe cars, the relative strength of these unitary sets is determined by the weights of these reasons. For instance, if a Mercedes is safer than a Volvo, the weight of the reason that a Mercedes is a safe car is by default bigger than the weight of the reason that a Volvo is a safe car. Then the set of reasons consisting of the reason that a Mercedes is a safe car is 'stronger' than the set of reasons consisting of the reason that a Volvo is a safe car.

2.2 Degrees and probabilities

The same example also illustrates a different point, concerning the relation between the 'degree' in which a fact obtains, and the weight of the reason that this fact constitutes. Let me be more concrete. A Mercedes is not just safe or not safe, but it is safe to a certain degree. In a similar way it is expensive to a certain degree. Some kind of facts – one might call them

[2] Abstract and concrete reasons are discussed in chapter 3, section 3.5.

'dimensions'[3] – do not just obtain or not obtain, but they obtain in degrees. If such facts are concrete reasons, the weights of these reasons will normally depend on the degree in which these facts obtain. If car A is more expensive than car B, which is also expensive, the fact that car A is expensive is a stronger reason against buying car A than the fact that car B is expensive is a reason against buying car B.

A similar phenomenon occurs in connection with probabilities. Reasons pleading for and against alternatives, especially when these alternatives are lines of action, often will concern the consequences of adopting the alternatives. These consequences have a certain degree of probability and an attractive consequence will lead to a stronger pro-reason if the probability of this consequence is higher. Similarly, an unattractive consequence will lead to a stronger con-reason if the probability of this bad consequence is higher.[4]

As these examples illustrate, the dimension of weight of reasons can be used to reflect two other dimensions of reasons, namely the degree in which the reason-giving facts obtain and the probability of their consequences.

2.3 The 'logic' of comparison

If two 'similar' sets (sets that contain similar reasons) have more than one element, they can only be compared qualitatively on basis of the weights of their reasons if all the differences in weight work in the same direction. Suppose, for instance, that a Volvo and a Porsche have the same pro-reasons, namely their social status and their suitability for holiday purposes. If a Volvo is both better for status and for holiday purposes, the set of pro-reasons for a Volvo is by default stronger than the set of pro-reasons for a Porsche. But if a Volvo is better for holiday purposes, but a Porsche better for social status, the sets of reasons cannot be compared qualitatively, at least not without additional information.[5]

The above can be generalized as follows. Each alternative in a set of alternatives has one (possibly empty) set of reasons pleading for it (the pro-reasons) and one (possibly empty) set of con-reasons. Two alternatives can be compared by pair wise comparing the sets of pro- and con-reasons. For this purpose the relations stronger, weaker and equal are used. A set of reasons can be stronger than, weaker than, or equal to another set. These

[3] This is the term used by Ashley 1990 and 1991. See also Bench-Capon and Rissland 2001, about the relevance of these dimensions.

[4] This is a familiar theme from decision theory. See for instance Keeney and Raiffa 1993, 5f.

[5] However, see section 9.

three relations are mutually exclusive. They are not exhaustive, however, because in some cases sets of reasons cannot be compared qualitatively.

Given these relations between sets of reasons, it is sometimes possible to establish on logical grounds, without additional decision making, which of two alternatives, if any, is preferable to the other. If the two alternatives are called A and B, and the relevant sets of reasons pleading for and against A and B are called Pro-A, pro-B, con-A and con-B, it holds that:

Alternative A should be preferred to (is better than) alternative B (and then B is worse than A) if (but not necessarily only if):

- Pro-A is stronger than pro-B, and con-B is either equal to or stronger than con-A; or
- Pro-A is equal to pro-B, and con-B is stronger than con-A.

Alternative A is equal to alternative B if (but not necessarily only if) both:

- Pro-A is equal to pro-B, and
- Con-A is equal to con-B.

If either

- Pro-A is stronger than pro-B, while con-A is stronger than con-B, or
- Pro-A is weaker than pro-B, while con-A is weaker than con-B, or
- Pro-A and pro-B, or con-A and con-B cannot be compared qualitatively

then it is not possible to establish *on the above mentioned grounds* which alternative is better than the other, or whether the alternatives are equal to each other.[6]

Sometimes it is possible to determine on logical grounds whether a set of reasons is stronger than, weaker than, or equal to another set. In this connection two aspects of these sets are taken into account, namely:

1. whether one set is a proper (similar-)superset of the other, or - in other words - whether one set contains all similar elements of the other and then some more, and
2. whether one or more of the reasons in one of the sets weigh more than the similar reasons in the other set.[7]

[6] It may nevertheless be possible to establish a ranking between alternatives by means of additional decision making.

[7] As described above, the degree in which a reason (a dimension) obtains, is taken into account via the weight of the reason.

A set is stronger than another set (and the other set is weaker) if (but not necessarily only if):

- it is a proper (similar-) superset of the other and none of its reasons weighs less than the similar reason in the other set (if there is such a similar reason); or
- all its elements are similar to elements of the other set and the other way round, none of its reasons weighs less than the similar reason in the other set, and at least one of its reasons weighs more than the similar reason in the other set.

A set is equal to another set if (but not necessarily only if):

- all its elements are similar to elements of the other set and the other way round; and
- all of its reasons have the same weight as the similar reasons in the other set.

2.4 Weak Transitivity

Often two sets of alternatives will not be comparable on logical grounds alone. Then additional decision making is necessary to establish which one is better. For instance, if a Volvo is a safer car then a Porsche, but a Porsche is better for one's social status, and these are the only relevant reasons, it is not possible to establish on logical grounds which brand is better. Suppose that a decision is made that a Volvo is better than a Porsche. Suppose, moreover, that a Mercedes is just as safe as a Volvo and is even better for one's social status, and there are no other relevant reasons, then it is possible to determine on logical grounds that a Mercedes is better than a Volvo. Since it has been established by decision making that a Volvo is preferable to a Porsche, it seems rational to assume that a Mercedes must, in the absence of special circumstances, be better than a Porsche too.

This can be generalized as follows: If alternative A is better than alternative B and if C is better than A, then, *by default*, C will be better than B too. Another way to say this is that the better than relation is *weakly transitive*. The weakness of the transitivity consists in the defeasibility of the application of transitivity. Analogously, the equal to-relation between alternatives is also weakly transitive.

Weak transitivity does not only hold for the better and worse than relation as applied to alternatives, but also for the stronger than, weaker than, and equal to relations as they hold between sets of reasons. For instance, if a Mercedes and a Volvo are both reliable and save cars, while the Volvo is safer, but the Mercedes is more reliable, the sets of pro-reasons for a