The three faces of defeasibility in the law

Henry Prakken
Institute of Information and Computing Sciences,
Utrecht University, The Netherlands
Giovanni Sartor
CIRSFID - University of Bologna, Italy

June 5, 2003

1 Introduction

In this paper we will analyse the issue of defeasibility in the law, taking into account research carried out in philosophy, artificial intelligence and legal theory. We will adopt a very general idea of legal defeasibility, in which we will include all different ways in which certain legal conclusions may need to be abandoned, though no mistake was made in deriving them. We will argue that defeasibility in the law involves three different aspects, which we will call inference-based defeasibility, process-based defeasibility, and theory-based defeasibility. Only the integration of these three perspectives allows us to provide a satisfactory account of the role of defeasibility in legal reasoning.

2 Inference-based defeasibility

The first aspect of defeasibility which has attracted the attention of legal thinkers has been what we may call inference-based defeasibility.

Inference-based defeasibility covers the fact that legal conclusions, though correctly supported by certain pieces of information, cannot be inferred when the theory including those information is expanded with further pieces of information (we use the term "theory" to mean in general any set of premises intended to provide an account of a legal domain). This idea is also frequently expressed by saying that common sense reasoning (as opposed to logical deduction) is nonmonotonic: the conclusions of the reasoning process do not grow inevitably as further input information is provided.

This is a feature of human reasoning which has been studied by scholars working in different areas.

In epistemology the idea of defeasible inference has been developed by authors such as [Chi77] and [PC99]. Pollock, in particular, relates defeasible reasoning to a general feature of human cognition. He argues that a reasoning

agent starts with perceptual inputs, and goes on inferring beliefs from the reasons which are available to him (his percepts plus the beliefs which he has previously inferred). Such a belief-formation process must satisfy two apparently incompatible desiderata. Firstly, the agent must be able of forming beliefs on the basis of a partial perceptual input (the agent cannot wait until he has a complete representation of his environment). Secondly, the agent must be able to take into account an unlimited set of perceptual input.

> The only obvious way to achieve these desiderata simultaneously is to enable the agent to adopt beliefs on the basis of small sets of perceptual inputs but then retract them in the face of additional perceptual inputs if those additional inputs conflict in various ways with the original basis for the beliefs. This is a description of *defeasible reasoning*. beliefs are adopted on the basis of arguments that appeal to small sets of previously held information, but the beliefs can later be retracted in the face of new information ([Pol95a], 40).

One usual example for epistemological defeasibility concerns an agent having an image of a red object (see [Pol95a], 40). This percept is usually a good enough reason for that agent to conclude that the object is indeed red. However, if the agent sees that the object is under a red light, then the agent should retract the conclusion that the object is red, since when an object is under a red light, its looking red does not support the conclusion that it is red (also objects having different colours look red under a red light). Similarly, Pollock argues, in statistical reasoning the fact that most F's are G's (e.g. most pieces of information in a newspaper are true), combined with the fact that x is an F (x is a statement in a newspaper) provides a defeasible reason for concluding that x is a G (the statement is true). However, if one has (better) reasons for believing that the x statement is false, one should retract this belief.

In moral reasoning, a similar idea of defeasibility is at the centre of D. Ross's theory of prima-facie moral obligations ([Ros30, Ros39]). In Ross's account, moral intuition plays the same role that perception plays in Pollock's account: both provide starting points for defeasible inferences. As Pollock's account of empirical reasoning is grounded upon the recognition that our reasonings cannot directly take into account all possibly relevant circumstances, so Ross's account of moral reasoning starts with the idea that any single moral principle cannot take into account all morally relevant aspects of our actions:

> Any possible act has many sides to it which are relevant to its rightness or wrongness; it will be pleasure to some people, pain to others; it will be the keeping of a promise to one persons, at the cost of being a breach of confidence to another, and so on. We are quite incapable of pronouncing straight off on its rightness or wrongness on the totality of these aspects; it is only by recognising these different features one by one that we can approach the forming of a judgement on the totality of its nature.

This view implies that our moral principles can lead us to conflicting conclusions. For example D. Ross considers "that sometimes we cannot obey one without disobeying another; that often we cannot obey the principle of telling the truth without disobeying the principle of not causing needless pain, or the principle of keeping promises without disobeying the principle of producing the maximum good" ([Ros39], 83). Correspondingly, Ross argues that the conclusions which are licenced by our moral intuitions may be overridden by other moral intuitions in concrete cases. To establish our real obligations we need to balance our conflicting prima facie conclusions.

> Moral intuitions are not principles by the immediate application of which our duty in particular circumstances can be deduced. They state [...] prima facie obligations. [...] We are not obliged to do that which is only prima facie obligatory. We are only bound to do that act whose prima facie obligatoriness in those respects in which it is prima facie obligatory most outweighs its prima facie disobligatoriness in those aspects in which it is prima facie disobligatory. ([Ros39], 84-85)

In artificial intelligence (AI) defeasible reasoning has also been the focus of a considerable amount of research (for a review, see [PV00], for some major contributions, see [Gin87]). In fact, in the attempt to build intelligent systems which could act in a complex and partially unknown environment, the need emerged to provide such systems with the capacity to "jump" to conclusions which are supported only by a subset of the potentially relevant information [McC87a]. One of the most celebrated examples is Tweety's case, originally put forward by Marvin Minsky. Given the information that Tweety is a bird, and the common sense generalisation that birds can fly, the reasoner should jump to the conclusion that Tweety can fly. However, this conclusion (which is derived only from the fact that Tweety is a bird) should be abandoned if it comes out that Tweety is a penguin, that he has a broken wing, that he is sick, etc. This example well exemplifies the so called qualification problem: an agent cannot be expected to make an action or form a belief only after having checked all possible conditions required for ensuring the successful performance of the action or the truth of the belief ([McC87b], 1). Usually it would be impossible to verify that all such conditions obtain (to verify that Tweety does not belong to any species of non flying bird, that he has no kind of impairment that impedes flying, etc.), or even to spell out all of them in advance. An agent should therefore derive defeasible conclusions from circumstances normally sufficient to support them, that is according to principles which apply under normal circumstances.

Many theories of defeasible reasoning in artificial intelligence have been produced such as default logic [Rei87], circumscription [McC87b], autoepistemic logic [Moo87b, Moo87a], negation by failure [Cla87, Kow79], and various argumentation systems [SL92]. Our preference goes to those models that explicitly ground the "jump" to a certain conclusion upon the articulation of the reasons for and against drawing such conclusion, making the jump dependent upon the interaction of reasons and counter-reasons.

In this regard, possibly, the most attracting logical account of defeasible inference is that provided by John Pollock ([Pol95b, Pol95a, PC99]). Pollock provides us with a distinction in two ways in which one inference may defeat another inference. The first is by rebutting. The rebutter of an inference concluding for p is an inference concluding for the negation of p, that is for $\neg p$. For example, the inference that leads me to conclude that Ann is John's killer on the basis of Sylvia's testimony can be rebutted by Mary's testimony that Adrian is (and therefore Ann isn't) John's killer. The second is by undercutting. The undercutter of an inference concluding for p on the basis of a certain ground q is an inference saying that, under the existing circumstances, q is not a proper ground for p. For example, the inference that leads me to conclude that Adrian is John's killer on the basis of Sylvia's testimony, can be undercut by the information that Sylvia was paid for giving such a testimony. This fact is not a reason for concluding that Adrian did not kill John (it does not say anything about it), but it casts doubts on the probational force of Sylvia's testimony.

Besides the idea of defeat, Pollock's account of the process of defeasible justification (for a concise account, cf. [PC99], 197 ff.) includes an analysis of the connections between multiple arguments. This is required since a defeated argument may be reinstated when a broader picture is considered: this happens when the defeater is in turn itself defeated. So, for example, of undercutting above, the conclusion that Ann is John's killer, though attacked by the argument Adrian is John's killer since Sylvia testifies so, is reinstated when Sylvia's testimony is shown to be unreliable (since Sylvia was paid for giving it.). We cannot provide here a formal discussion of Pollock's method, but we will discuss a fairly similar approach by Prakken and Sartor [PS96].

After considering defeasible inference in epistemology, moral philosophy and epistemology, let us go back to legal thinking. The notion of defeasibility, quite usual in legal practice and in doctrinal work, was brought to the attention of legal theorists by Hart:

> When the student has learned that in English law there are positive conditions required for the existence of a valid contract, he has still to learn what can defeat a claim that there is a valid contract, even though all these conditions are satisfied. The student has still to learn what can follow on the words 'unless', which should accompany the statement of these conditions. This characteristic of legal concepts is one for which no word exists in ordinary English. The law has a word which with some hesitation I borrow and extend: this is the word 'defeasible', used of a legal interest in property which is subject to termination or 'defeat' in a number of different contingencies but remains intact if no such contingencies mature ([Har51], 152).

More recently, various ideas related to defeasible inference have been discussed in legal reasoning. Let us first mention the idea of legal reasoning as including non-deductive "jumps", discussed by Peczenik ([Pec83]). Secondly,

we recall the opposition of of *prima-facie* legal obligation and all-considered legal obligations, as proposed by ([Pec89]). This opposition provides a model of legal reasoning similar to the account of morality provided by D. Ross. Single legal norms and principles, when applied to the facts of a case, only provide *prima facie* legal conclusions. All-things-considered legal conclusions can only be obtained by considering all relevant rules and principles, as applied to all relevant facts. This opposition implies a notion of defeasibility, when one considers the limitations of human knowledge. One must be able of forming legal beliefs (or deriving legal conclusions) though one only has a limited legal knowledge (one may not be aware of all legally relevant factors, rules, principles, concepts) and a limited factual knowledge (one may not be aware of all circumstances of the case). In fact, as Peczenik observes "No human being has the resources sufficient to formulate all-things-considered statement *sensu stricto*" ([Pec89] 77). The situation of a legal reasoner, therefore, parallels the situation of Pollock's agent. The legal reasoner, though endowed with limited knowledge, energy, time and resources, must be able of making reasonable decisions. However, he must also be able of taking into account a potentially unlimited amount of further relevant information, as soon as he is aware of it. This is exactly what a method for defeasible legal reasoning should do: it should enable a legal agent to form judgements on the basis of the knowledge he has, and the thinking he is presently able to do, and correct (and possibly withdraw) such conclusions as soon as he is able to take into account further legally relevant information.

A further idea which points to a defeasible model of legal inference is that of weighing and balancing conflicting considerations as the appropriate way of reaching considered legal conclusions. This method of reasoning is not exclusively appropriate to the law. As Stuart Mill said, "any subject on which difference of opinion is possible, the truth depends on a balance to be struck between two sets of conflicting reasons" ([Mil74], 98). However, in practical reasoning, and in particular in legal reasoning, the need to take into account conflicting considerations is particularly important. As A. Peczenick says "In order to justify an all-things-considered practical statement, one must weigh and balance *prima-facie* practical statements which support it against such statements supporting the contrary conclusion ([Pec89], 77).

The weighting and balancing metaphor, that is, the idea that legal conclusions are based upon the comparison of the weight of sets of reasons favouring opposed outcomes, naturally leads to defeasibility. Let us assume that the reasoner believes that outcome p is supported by the set A of reasons, while outcome *not* p is supported by set B, and that he also believes that set A is heavier than set B. When considering only reasons in A and B the reasoner should conclude for p. However, let us assume that reasoner discovers that there is an additional reason r promoting outcome *not* p, and such that $B + r$ is heavier then A. When taking into account also r, the reasoner should retract conclusion p, and accept *not* p. Further reasons, weighing for p may still change the result of the weighing.

Though the idea of defeasibility in legal inference has been discussed in legal theory, as we saw above, it is only within research on AI & law that this idea

has been thoroughly analysed.

This research has pointed to various structures of legal knowledge which support defeasible reasoning. One is the combination of rules and exceptions. Consider for example, the Italian rule saying that if one causes an unjust damage then one is liable, and the exception saying that if one causes an unjust damage in a situation of necessity, one is not liable. Such combinations prompt for defeasible reasoning: one should derive the conclusion of the rule when one knows that its antecedent is satisfied, but one should abandon this conclusion if it appears that one exception was also satisfied. Another structure which prompts for defeasibility is the use of unless clauses. For example, Italian law says that a person is not liable for an action if he is incapable at the moment of that action, unless he put himself in such a state. This is clearly a case of undercutting: having made oneself incapable is not reason for being liable (one is liable for causing an unjust damage, and not because one is drunk), but rather a reason for incapability not to exclude liability.

Moreover, researchers in AI have been investigating defeasibility resulting from clashes of rules and principles. As we have remarked legal inferences can be defeated when competing inferences can be developed. To give more concreteness to the metaphor of weighing and balancing, AI & Law researchers have used various legal criteria for dealing with such conflicts, establishing what inference is defeated. Both well–known general priority rules such as the principles of specificity (Lex Specialis), temporality (Lex Posterior) and hierarchy (Lex Superior) and domain–dependent priority rules have been considered.

A deeper question is whether those specific criteria are to be seen as specific applications of the idea of weighing and balancing reasons (a view taken by [Hag97]), or whether, on the contrary, the idea of combining the weight of reasons is only a possible criterion for solving conflicts, alongside with other criteria (this is the view taken by [PS96]). We cannot here examine all approaches to non-monotonic reasoning in the law. We will focus on an argument-based approach, which has the indiscutibile merit of having been advanced by the authors of this paper. However, it is very close in spirit (though not in the technical solution) to other argument-based models of defeasible legal reasoning [Gor95, Hag97]), and therefore it may be considered representative of the broader research community. Let us briefly present this model of legal inference.

The basic idea (as for other argument-based approaches to legal defeasibility) is that legal rules and principles are to be viewed as reasoning warrants that, when certain antecedent reasons are available (being provided as part of the input, or as the result of previous inference steps), support the derivation of legal conclusions. In this way basic legal inferences are obtained, which consist in chains of warrants leading to legal conclusions.

Such inferences, however, may lead to conflicts, by providing incompatible conclusions. To adjudicate a conflict between two inferences, we need to establish what inference is stronger, and this in general may be done through further inferences (preference inferences), which use further (meta) warrants. By merging these inferences we obtain a complex inference, which can include a basic inference, the comparison inference supporting the first one, and so on.

According to the result of such judgements, we can, as a first approximation, say that any inference which is not stronger then one of its competitors is defeated, and its conclusions need to be retracted, while the conclusions of the stronger inference can be safely asserted. However, the picture is complicated by the idea of reinstatement: defeated inference can be recovered if their defeaters are, in their turn, defeated by stronger inferences. So, we obtain a model of legal reasoning as the dialectical interaction of competing inferences: the outcome of this competition determines what conclusions will be legally justified in the framework of the available legal knowledge.

This account can be precisely stated in a very intuitive way, viz. in the dialectical form of a dispute. The idea is that a statement is warranted (relative to a certain body of information) if its proponent can successfully defend it in a dispute, or 'argument game' with an opponent (obviously, the reasoner who checks on his own whether a statement is warranted should take both positions). One form that such a game can take is the one of [PS96].[1] The proponent starts with an argument for the statement he wants to prove, after which the players take turns, moving one argument at the time. Opponent's task is to attack proponent's last-moved argument with a conflicting argument. Proponent's task then is to prove that this attack is unsuccessful. Proponent can do this in two ways. The first is to move a priority argument which 'neutralises' opponent's attack, i.e., which says that opponent's argument is in fact refuted by its target. For instance, consider the following dispute:

Proponent: 'The newspapers may publish this piece of information about politician B since it is relevant for assessing his political functioning'.

Opponent: 'The newspapers may not publish this piece of information since it is about B's income, information about income is, according to Section x of statute S, private information, so publishing it would violate B's privacy'.

Then proponent might reply:

Proponent: 'For politicians the value of openness about functioning overrides the value of privacy, so my argument is stronger then the opponent's argument'.

(Note that the above rules of the game now allow opponent to attack proponent's priority argument. For instance, opponent could argue that openness about functioning of politicians is not more important than protection of privacy.)

A second option for proponent is to attack some intermediate step in opponent's argument, combined with the reasons why proponent's attack succeeds. For instance, proponent might argue:

[1] Several games are possible, reflecting different notions of defeasible consequence; cf. e.g. [PV00]. For present purposes their differences do not matter.

Proponent: 'Information about income of politicians is not private information, as stated by section y of statute S, and y is a *Lex Specialis* of x'.

(Again, opponent may try to attack proponent's argument, for instance, by arguing that y is not a *Lex Specialis* of x.)

Note that these rules build a form of dialectical asymmetry into the game, capturing the idea that the initial statement must be provably warranted, while objections just need to interfere with the proof for impeding it from succeeding: this is captured by the fact that proponent's counterarguments must refute their target, while opponent's counterarguments are allowed to merely cast sufficient doubt on their target.

In [Pra01a] the game above is refined to cope with shifts of the burden of proof. Suppose, for instance, in a civil dispute that plaintiff claims a contract was created while defendant argues for an exception because of insanity. Then plaintiff will receive the burden of proving offer and acceptance, so plaintiff will be proponent with respect to these statements. However, defendant will be assigned the burden of proving her insanity, so she will be the proponent with respect to that statement. This extension takes us into the second type of defeasibility, process-based defeasibility, which will be examined in the following section.

3 Process-based defeasibility

Theories of inference-based defeasibility (i.e., nonmonotonic logics) essentially take a static view on reasoning, since they just look at a given body of information, and say which beliefs are warranted on the basis of that information. Even the definition of nonmonotonicity is stated in essentially static terms: it just says that the beliefs warranted on the basis of an information set S need not be included in the beliefs warranted on the basis of a larger information set S'. This definition abstracts from whether S precedes S' in the reasoning process or not; and if S did precede S', it is is silent on how the reasoning progressed from S to S'. Most importantly, 'standard' nonmonotonic logics say nothing about in which circumstances S can be taken as a basis for decision, and in which circumstances it is better to search for more information.

Process-based defeasibility addresses such 'dynamic' aspects of defeasible reasoning. As for legal reasoning, a crucial observation here is that it often proceeds according to the rules of legal procedures. In particular, the following features of legal procedures are relevant.

Firstly, the outcome of a legal dispute often depends not only on the statements and arguments exchanged, but also on the 'propositional attitudes' expressed by the parties involved (such as when a claim was disputed, conceded, or not responded to) and on how the burden of proof is allocated. To illustrate this, let us again look at our example, first purely in terms of inference-based

defeat. Then we see that the arguments exchanged make proponent's initial statement provable (since opponent did not find a counterargument against his last argument). However, suppose that the opponent has disputed that the piece of information is about B's income (e.g. since the information was not clearly about B), and that the judge has allocated the burden of proof to proponent, after which the proponent failed to provide any evidence. Procedural laws then often say that proponent has failed to prove his initial statement, since he failed to provide evidence for a premise for which he had the burden of proof; therefore, his final argument should be ignored. In sum, it is not only important which arguments were exchanged, but also which premises of these arguments were disputed, conceded (or withdrawn), and how the burden of proof was allocated.

Even if we only look at arguments and the statements from which they are composed, procedural law has a lot to say that is left out by a purely inference-based notion of defeasibility. A central question here is what it means to have the burden of proof for a certain statement[2]. This involves two aspects, viz. the task to provide an (evidential) argument for a claim and the task to defend this argument in dispute. As for the first task, legal procedures regulate which kinds of evidence are admissible (e.g. illegally obtained evidence and hearsay evidence is often ruled inadmissible). The second task in turn has three aspects: whether the argument is internally acceptable, whether counterarguments are admissible, and whether the argument is dialectically acceptable, i.e., whether it survives the competition with the admissible counterarguments.

Whether an argument is internally acceptable means whether it in itself sufficiently supports its conclusion, apart from possible counterarguments. Sometimes a judge is free to assess the strength of arguments (as is the general rule in Dutch civil law). Of course, this legal-procedural freedom is (at least ideally) bound by rational theories of what counts as a good argument (see e.g. [Wal96]). But some types of evidence are declared *conclusive*, i.e., if no counterevidence is provided, the judge must accept it. For instance, in Dutch law an avidavit is conclusive evidence that its content is true. As for the second aspect, whether counterarguments are admissible, this is purely a matter of procedural law. Some kinds of evidence are declared *incontrovertible*, which means that no counterevidence is allowed. An example in Dutch civil law is an oath of one of the parties. The final aspect of the burden of proof is that one's admissible evidence should survive the competition with all (admissible) counterevidence. Above all, this is, of course, a matter of evaluation by the judge. However, the judge's evaluations are in fact arguments on the strength of other arguments, i.e., undercutting or priority arguments. So, once the evidence has been evaluated, (defeasible) logic comes into play: we end up with a collection of arguments, counterarguments, and priority arguments, and if we add to that the allocation of the burden of proof, we can compute with a logical argument game whether the initial claim is proven.

So far we have only discussed the 'static' aspects of burden of proof, viz. how

[2]The role of the burden of proof in legal defeasibility has been stressed before by e.g. [Mac95] and [Sar95].

it determines the provability of the initial claim at a certain stage of a dispute. However, the burden of proof also has a dynamic role: as soon as one party has fulfilled his burden of proof, the dispute is not over but the other party must be given a fair opportunity to provide counterevidence. That is, once a proof burden is met, the burden should shift to the other party. For instance, (recall Hart) once plaintiff has proven offer and acceptance, defendant must be given the opportunity to prove an exception, such as insanity, or a hidden defect. Of course, this must once come to an end, and this highlights an important dynamic feature of defeasibility, viz. the question when a dispute must be terminated. This relates to the above question under which conditions a certain body of information is sufficient to decide the dispute, and under which conditions it is better to search for more information. In law this is, of course, partly regulated by the relevant procedure; however, such procedures usually leave some freedom to the judge, at which point rational criteria for termination are called for; and such criteria are another process-based aspect of defeasibility.

Consider, for instance, a criminal case where one witness says that the suspect shot the victim to death and another witness says that suspect stabbed the victim to death. This is a case with two conflicting arguments, but whichever argument is preferred, one always ends up with the conclusion that the suspect killed the victim. In the literature on inference-based defeasible reasoning there is a discussion whether such "floating" conclusions should be accepted if this is all the information one has. For instance, [Hor02] argues that it is rational not to accept the conclusion that the suspect killed the victim, since the conflicting witness statements undermine each other's credibility: it may well be that neither speaks the truth. However, whatever the merits are of this analysis, any good lawyer would in such a situation defer judgement and instead ask further questions to the witnesses to investigate their credibility (e.g. "did you hear any sound?"). Again we see that a narrow, inference-based view on the defeasibility of legal reasoning is too limited: it does not take the dynamic aspects of reasoning into account (see [Hin99] for a similar view on standard logic).

Summarising, when one wants to rationally reconstruct a certain piece of legal reasoning, it often does not suffice to apply only the rules of logic, not even if the logic is nonmonotonic. One must also take into account how the relevant legal procedure regulates such things as the burden of proof, the effect of disputing, conceding or ignoring a claim, the admissibility of evidence and counterevidence, and the conclusive force of certain kinds of evidence. Only if these issues have been accounted for, i.e., only if the arguments exchanged in a dispute have gone through a 'procedural filter', the remaining arguments can be used as the input of a nonmonotonic logic. Moreover, this process evolves in time: the arguments that pass the filter at a certain stage in a dispute may favour the plaintiff, but the dispute may proceed (e.g. the burden of proof may switch to defendant) and a new set of arguments passing the procedural filter at a later stage may favour instead the defendant. A final issue that evades a purely inference-based account is at which stage a dispute should come to an end. This issue is governed by issues of fairness, rationality (asking the right question) and resource limitations, especially time and money. All these issues

can be modelled, but clearly not in a purely logical way.

While, as noted in Section 1, formal models of inference-based aspects of defeasibility abound (in the form of nonmonotonic logics), the formal study of its process-based aspects is less developed. Yet a promising research line is evolving, sometimes called 'computational dialectics'. It originated in the seventies, when such logicians as Hamblin, MacKenzie, Woods and Walton embedded standard propositional logic in dialogue game models for argumentation, with speech for such things as claiming, disputing, conceding or retracting a proposition, and providing grounds for a proposition (see [WK95] for an overview). In such games, proponent wins if he succeeds in opponent accepting his initial claim, while opponent wins if she succeeds in making proponent withdraw his claim. In the nineties, AI & Law researchers such as [Gor95, HLL94, Lod99, Lou98, BCGL00, Pra01b] applied such dialogue games to legal disputes, meanwhile adding the very important possibility of counterarguments, thus effectively shifting from deductive to defeasible reasoning. These models regulate when certain speech acts may or must be made, which effect they have on the (current) outcome of a dispute, and (sometimes) when a dispute terminates. Their aim is to provide idealised models of legal dispute, i.e., rules for how a dispute can be conducted and resolved in a fair, rational and effective way. Thus these models are inspired by a procedural view on rationality and justice as expressed by e.g. [Res77] and [Ale78]. Another important feature of these models is that they in fact formalise 'unless shown otherwise' phrases in the law not naively as 'unless logically proven otherwise from the available premises' but as "unless shown otherwise by a process of argumentation and the presenting of evidence to an authorised decision-maker" (cf. [AS89]).

Space limitations prevent us from discussing these models in detail. For two recent overviews see [Hag00] and [PS02]. Rather, we will give an example of the kinds of dialogues regulated by such models. Our main point here is to highlight that such dialogues can now be formalised in ways that capture not only inference-based but also process-based aspects of defeasibility. The example dispute is about contract formation. It contains a claim, concessions and disputations, decisions about the burden of proof, an argument and a counterargument, and a decision that a piece of evidence was illegal.

- *Plaintiff*: I claim that defendant owes me 500 euro.
- *Defendant*: I dispute plaintiff's claim.
- *Judge*: Plaintiff, prove your claim.
- *Plaintiff*: Defendant owes me 500 euro since we concluded a valid sales contract, I delivered but defendant did not pay.
- *Defendant*: I concede that plaintiff delivered and I did not pay, but I dispute that we have valid contract.
- *Judge*: Plaintiff, prove your claim that you have a valid contract.
- *Plaintiff*: This document is an avidavit, signed by us.
- *Defendant*: I dispute that this document is an avidavit.
- *Judge* Defendant, since the document looks like an avidavit, prove that it is not.

- *Defendant*: This lab report shows that the notary's signature was forged.
- *Plaintiff*: That evidence is inadmissible, since I received it too late.
- *Judge*: I agree: the evidence is inadmissible.

At this point, both plaintiff and defendant have constructed one argument, displayed in Figures 1 and 2 (the 'warrants' connecting premises and conclusion are left implicit, as is so often done in legal reasoning). Plaintiff's argument passes the procedural filter, since its premise (8) was not replied to by defendant, its premises (3) and (5) were conceded by defendant, while with respect to the remaining premise (7), the burden of proving the opposite was assigned to defendant. Defendant's argument, on the other hand, does not pass the procedural filter, since one of its premises, (11) is declared inadmissible evidence. Observe, finally, that the judge has not attacked plaintiff's argument with an undercutter, so that the judge has implicitly ruled this argument internally acceptable. The argument game then trivially computes that plaintiff wins (at least at this stage of the dispute!).

$$\dfrac{\dfrac{\text{(3) } I \text{ delivered} \qquad \dfrac{\dfrac{\text{(7) } avidavit \quad \text{(8) } signed}{\text{(6) } offer \text{ \& } acceptance}}{\text{(4) } valid\ contract} \qquad \text{(5) } you\ didn't\ pay}{\text{(2) } breach\ of\ contract}}{\text{(1) } you\ owe\ me\ 500}$$

Figure 1: Plaintiff's argument

$$\dfrac{\dfrac{\text{(11) } lab\ report}{\text{(10) } notary's\ signature\ forged}}{\text{(9) } no\ avidavit}$$

Figure 2: Defendant's argument

Of course, current process-based models of legal reasoning are far from complete or perfect. To start with, they should be complemented with theories on what counts as a good legal argument, and on what are good ways to question or attack such arguments. Some important investigations of this kind already exists, such as [Ale78, Hag97, Wal02], to name but a few. However, more work remains to be done, especially on formalising these accounts. Other important research issues are what are good strategies for playing a dialogue game, e.g. what are the right questions to ask, and what are good criteria for termination of disputes? Finally, it should be studied how process-based models of legal reasoning can be combined with an important kind of legal reasoning, viz.

coherence-based theory construction. Let us now turn to this kind of reasoning.

4 Theory-based defeasibility

The third form of defeasibility we will consider is theory-based defeasibility. This results from the evaluation and the choice of theories which explain and systematise the available input information: when a better theory becomes available, inferior theories are to be abandoned.

The basic inspiration is given by theories of the development of science centered upon the idea of competition between alternative theories. As is well known, in the first half of the past century the idea of an incremental growth of knowledge was countered by the Popperian idea of theories as refutable conjectures [Pop59], to be abandoned when contradicted by observations. However, this idea also appears inadequate in relation to the fact that humans need theories, to interpret, systematise, and anticipate their experience. We only abandon a theory when a better theory is available to us ([Lak78]). Substituting a better theory for the old one can be viewed as a move that is appropriate when a scientific revolution ([Tha92]) is taking place, i.e. in those crucial moments in the development of science when a new paradigm is emerging and displacing the once dominating view, now irremediably degenerating for its incapacity to address new issues (cf. [Kuh62], and on paradigm change in the law, cf. [Aar84])

However, the idea of theory change has a broader significance: it also covers cases where only a portion of a theory is substituted with a new theory, that is, where from theory A_1 a new theory A_2 is produced, by cancelling from A_1 subtheory S_1 and substituting S_1 with the new subtheory S_2. As is well known, the issue of theory change has been considered to a large extent in philosophy, law and computing. However, the focus in most approaches (such as [AGM85, Gar87]) has been on how to update a consistent theory into a new consistent theory which contains some new information.

In the study of change in legal theories one has to adopt a different focus, since legal theories are defeasible theories, as we said in section 2, so that they can accommodate new knowledge without the need for retracting previous information, even when there is a contradiction. Moreover, if one agrees that defeasible inference and in particular weighing and balancing of conflicting rules and principles is the appropriate way of approaching legal issues, then one should reject all attempts to provide a consistent representation of the law. Such a representation would indeed not respect the intuitive understanding of legal knowledge, and would not support its normal utilisation. A set of decisional criteria, each one supported by its rationale, susceptible of being compromised through defeasible inferences according to further criteria and rationales, would be translated into a set of much more complex rules, having no recognisable grounding. Such a representation would not be able to support the dynamic adjustment which takes place whenever new information concerning the conflicting principles and the criteria for adjudicating their contest is taken into consideration.

The rejection of consistency does not imply that legal knowledge should be a chaotic bag where anything can be thrown into. We should build our representations of the law (our legal theories) by using (only) the materials that we believe to be useful (or needed) in such a construction. Each one of those representations may (and should indeed) include conflicting (and therefore logically inconsistent) knowledge, when this is appropriate, but it should nevertheless provide an adequate framework for legal thinking and problem solving.

Since different theories of the same legal domain are possible, we need ways of comparing those theories and of selecting the most appropriate one. Note that the need of a choice is not excluded by the fact that the largest part of such theories may overlap. What matters is that we have to compare the competing theories as wholes: if we have to decide whether to substitute, within theory A_1, subtheory S_1 with the new subtheory S_2, we do not have to compare S_1 against S_2, but rather we need compare the whole theory A_2 with the whole theory A_2 that would be obtained by subtracting S_1 from A and adding S_2. The most general way of synthesising the idea of the adequateness of a body of knowledge is by appealing to the idea of coherence. We cannot get here into a discussion of the general notion of coherence (cf., for example [Leh90, Tha01]). This would also require discussing the connections between this idea and other notions such as the notions of consilience [Wil99] and integrity [Dwo86]. As an example of an analysis of this notion in legal theory, we could mention the contribution of Alexy and Peczenik ([AP90]) who have focused their attention on the frequency and the intensity of connections between statements in a theory. In a recent work, Peczenik defines coherence as follows:

> The more the statements belonging to a given theory approximate a perfect supportive structure, the more coherent the theory is.
>
> *Ceteribus paribus*, the degree of coherence of a theory depends on such circumstances as how great a number of supported statements belong to it, how long chains of reasons support on and the same conclusion, how great number of general concepts belong to it, how high the degree of generality of these concepts, how great number of cases it covers, and how great number of fields it covers. The degree of coherence is determined by weighing and balancing of the criteria ([Pec97], 196)

In legal theory, coherence has been usually approached as an ideal which should inspire legal theory and legal decision-making, rather than as a criterion for choosing between alternative theories. The latter idea has been especially discussed in research on artificial intelligence and law. So, McCarty observes that:

> A judge rendering a decision in a case is constructing a theory of that case. It follows, then, that a lawyer's job (in an appellate argument) is to construct a theory of the case, too, and one that

> just happens to coincide with his client's interest. Since the opposing lawyer is also constructing a theory of the case, which coincides with her client's interest, the argument boils down to this: which theory should the judge accept? There are several constraints here, but they are weak ones. Legal theories should be consistent with past cases, and they should give acceptable results on future cases. They should not be too hard to understand, or too hard to administer, which means that they should be free of major anomalies. One term that is sometimes used to describe these requirements is "coherence". Although difficult to specify precisely, the link between coherent legal theories and persuasive legal arguments seems to be quite strong [McC97].

The notion of coherence so sketched by McCarty has a pragmatic flavour. It points to the features of a legal theory that make it useful or appropriate as an instrument for legal thinking and problem solving. A similar notion of coherence has been used a criterion for theory choice in some recent contributions by Bench-Capon and Sartor ([Sar02, BCS01b, BCS01a, BCS03]).

These papers provide a model of legal reasoning which is centered upon theory construction and theory comparison. The parties in a case, given a shared legal background, develop alternative legal theories, and victory goes to the party who develops the better theory. This leads to the idea of theory-based argumentation, i.e., the idea that legal debates consist in the dialectical exchange of competing theories, supporting opposed legal conclusions in the issue at stake. A central aspect of theory-based argumentation is, as we shall see, theory-based defeasibility, in the sense that the weaker (less coherent) theory is defeated by the stronger one.

Before introducing this model, let us briefly present the original example of [BH93] which is used as a test example in all those contributions. The example consists of three American cases, where, the plaintiff (π) was chasing wild animals, and the defendant (δ) interrupted the chase, preventing π from capturing those animals. The issue to be decided is whether π has a legal remedy (a right to be compensated for the loss of the game) against δ or not. In the first case, Pierson v Post, π was hunting a fox on open land in the traditional manner using horse and hound, and had not yet gained possession of the fox, when δ killed and carried off the fox. In this case π was held to have no right to the fox. In the second case, Keeble v Hickeringill, π owned a pond and made his living by luring wild ducks there with decoys and shooting them. Out of malice δ used guns to scare the ducks away from the pond. Here π won. In the third case, Young v Hitchens, both parties were commercial fisherman. While π was closing his nets, δ sped into the gap, spread his own net and caught the fish. The issue at stake is how this third case should be decided, that is, whether δ should compensate π for the loss of the game, or whether no such liability exists.

The shared legal background of the parties includes the recognition that certain possible features of cases in this domain are factors pointing to certain

conclusions. For example, π's owning the land where he is hunting favours the conclusion that he ought to be compensated. On the other hand, π's not having yet gained possession of the game favours the conclusion that he should not be compensated. Those (defeasible) connections between factors and conclusions are called factor-links. The shared background also contains an understanding of why factor-links should be adopted: this is because the adoption of a certain factor-link (its being used by legal agents as a standard for their reasoning and practice) advances the achievement of certain values. This connection between a factor-link and a value is called a teleological link. Here is an example of a possible teleological link: compensating hunters who were disturbed while chasing on their grounds is a way of promoting the value of security (in enjoying one's own possessions). The assumption that factor, values, and their connections are shared only is a simplifying assumption in this model, but it also corresponds to the idea that usually, within a certain culture, people tend to agree both on the evaluative recognition on certain values, and in the factual recognition that these values are going to be promoted by adopting certain decisional patterns (factor-links), while they are more likely to disagree on the relative importance of the values ([Pec89], 58 ff).

Given this background, the parties try to build theories. We will not consider the mechanism for building theories, which consists, according to the elementary model of [Sar02, BCS01b], in combining factors into rules and in establishing priorities between such rules Let us just remark that in [Sar02] each theory is viewed as an explanatory hypothesis, i.e. as an *explanans*, in the terminology of [Hem66], 51). The theory aims both at accounting for the precedents, which are its *explanandum*, and to suggest an outcome for the new case. This suggestion may be interpreted either as a normative justification or as a forecast, according to the objective of the construction, and corresponds to the anticipation of future experimental data in scientific theories.

For example, let us consider a pro-δ theory according to which a disappointed hunter is to be compensated whenever he owns the land where he is chasing, in order to promote the value of security of possession. This theory allows the defendant to explain why in Pearson π lost (he was hunting in open ground), why in Keeble he won (he was hunting in his property), and why in Young π should lose (he was hunting in open sea).

However, this is not the only possible theory. Another possible theory, which also explains all precedents, but gives a different result in the new case (Young), is the theory that the disappointed hunter π should be compensated whenever he is pursuing his livelihood, in order to promote the values of work and productivity. This would still explain why π lost in Pearson (where he was hunting for his pleasure), why π won in Keeble (where he was a professional hunter), but would imply that π should win also in Young (where he was a professional fisher).

The possibility that different theories can be developed on the basis of the same legal background leads to the issue of what theory is to be preferred. As a standard for comparing theories [Sar02, BCS01b, BCS01a] proposes the idea of coherence, though intended in the broad (and pragmatic) sense indicated above,

that is, as the set of the requirements that make a legal theory a good (and useful) one. This idea is more specifically articulated, in a precedent-domain, into the following requirement: case-coverage (a better theory explains more precedents): factor-coverage (a better theory takes into account more features of those precedents), value-coverage (a better theory takes into account a larger set of values), analogical connectivity (a better theory includes more analogical connections between its components), non-arbitrariness (a better theory contains fewer ad-hoc statements, required neither for explaining the past evidence nor for implementing the shared assessment of value-priorities).

For example, this would be a good theory (among those generated with elementary theory constructors) according to the account of [Sar02, BCS01b].

> When π is hunting in his grounds he should be compensated (in order to promote the value of security of possession). This rule prevails over the rule that if π has not got hold of the game he should not be compensated (in order to limit litigation). When δ is pursuing its livelihood, in the open ground, and π has not yet caught the game, then π should not be compensated (to promote competition and prevent litigation). The latter rule prevails over the rule that when π is a professional hunter he should be compensated (since under those circumstances the need to prevent litigation favours δ who also is a professional hunter).

The implications of theories such as the one above are to be evaluated according to the defeasible logic of [PS96] (though other defeasible logics could equally be used). In particular, a theory explains a case when the decision of that case is a justified conclusion derivable from the theory plus the facts of the case.

To assess the goodness of a theory we need to combine all coherence requirements into a comprehensive evaluation of the theory. This issue is discussed in [BCS01a] where a connectionist approach is proposed, following the model of [Tha92]. Thagard's essential idea is to represent the evidence to be accounted for by a theory and the tenets of a theory as nodes connected by links representing support and conflict. A set of initial scores (between 1 and -1) is assigned to these nodes (this captures the idea that coherence in science concerns the ability of accounting for external inputs). Then these scores are propagated, support links increasing the scores of nodes, and conflict links decreasing them. Moreover, links are subject to a rate of decay so that the score of isolated nodes decreases. This propagation is continued through a number of cycles, until the values of the nodes stabilise. In Thagard's interpretation of this process, nodes which end with a high activation can be considered part of a coherent, and hence acceptable theory, while those with a low activation should be rejected. Thagard's theory is attractive for two main reasons: first of all it provides an effective way of "computing coherence", and secondly it offers a reconciliation between foundationalism (at least of the defeasible type [PC99]) and coherentism.

In the latter regard, we must observe that Thagard's approach assumes that the input nodes (the experimental evidence, in scientific theories) have an initial positive value which is independent of their role in the theory. It is up to the input nodes to provide value to the theory, though the theory, once it has received value from the input nodes can impact negatively on some of them, possibly conducing to their rejection. In [BCS01b] Thagard's method is adapted as follows to the legal domain. First, more in line with an adversarial view of legal argument, the purpose is not to extract the best theory out of the body of existing evidence and hypotheses, but to measure the comparative coherence of competing theories. Therefore, nodes having a low activation are not rejected, but they lower the overall coherence of the theory containing them. Secondly, two types of inputs are given a positive initial score: on the one hand precedents (which play the role of empirical data), on the other hand values. Therefore, cases and values are viewed as the sources from which a legal theory draws its merit. This corresponds to the idea that a coherent legal theory should be more than an internally connected web of concepts. It should account for certain experiences, i.e., in this model, the decisions of past cases and the appreciation of values. Values and cases have a prima-facie positive significance in their own (independently of their connection to the theory), which is reverberated on the theory (starting from the statements explaining cases or promoting values), though the theory may then retroact on cases and values, diminishing or increasing their individual score. Thirdly, in [BCS01b]'s approach, theories are assumed to be defeasible: therefore the fact that a theory includes conflicting information detracts very little from its merit, when those conflicts are solved by further information in the theory (according to which one of the conflicting rules is overruled). On the contrary, unsolved conflicts are very damaging for a theory, because they compromise its ability to explain the cases.

5 An overall framework for legal defeasibility

We believe that all three forms of defeasibility need to be integrated into a comprehensive account of legal reasoning. Inference-based defeasibility is an essential feature of rational reasoning within a legal theory, that is, reasoning that takes place when one is applying just one legal theory to a specific issue. Theory-based defeasibility is not reducible to inference-based or process-based defeasibility, since it concerns the holistic choice between theories, rather than the use of a theory. One may argue that such a reduction is possible, since the moves through which we add information to a theory are defeasible inferences, and the moves through which one deletes information from a theory are the defeat of this information. However, this does not seem to be the way in which human reasoners proceed. We rather tend to distinguish the construction and modification of our theories (the premises of our reasoning) from the use of such theories in making inferences. And we distinguish occasional conclusions we may derive from our premises from the premises constituting knowledge. The first only possess the strength they inherit, via inferences, from other premises.

The latter derive their value from the function they play in the knowledge base in which they take part, and from the global performance of that knowledge base, as compared to alternative ones.

Finally, process-based defeasibility, when aimed at rational decision-making must make use of both other types of defeasibility. That is, it provides the dynamic context in which inference-based and theory based defeasible reasoning take place, through the interaction of multiple agents.

References

[Aar84] A. Aarnio. Paradigms in legal science. In A. Peczenik, editor, *Theory of Legal Science*. Reidel, 1984.

[AGM85] C.E. Alchourrón, P. Gardenfors, and D. Makinson. On the logic of theory change: Partial meet functions for contractions and revisions. *Journal of Symbolic Logic*, 50:510–530, 1985.

[Ale78] R. Alexy. *Theorie der juristischen Argumentation. Die Theorie des rationalen Diskurses als eine Theorie der juristischen Begründung.* Suhrkamp Verlag, Frankfurt am Main, 1978.

[AP90] R. Alexy and A. Peczenik. The concept of coherence and its significance for discursive rationality. *Ratio Juris*, 3:130–147, 1990.

[AS89] L.E. Allen and C.S. Saxon. Relationship of expert systems to the operation of a legal system. In *Preproceedings of the III International Conference on "Logica, Informatica, Diritto" (Appendix)*, pages 1–15, Florence, 1989.

[BCGL00] T.J.M. Bench-Capon, T. Geldard, and P.H. Leng. A method for the computational modelling of dialectical argument with dialogue games. *Artificial Intelligence and Law*, 8:233–254, 2000.

[BCS01a] T.J.M. Bench-Capon and G. Sartor. Theory based explanation of case law domains. In *Proceedings of the Eighth International Conference on Artificial Intelligence and Law*, pages 12–21, New York, 2001. ACM.

[BCS01b] T.J.M. Bench-Capon and G. Sartor. Using values and theories to resolve disagreement in law. In J. Breuker, Leenes R., and R. Winkels, editors, *Proceedings of the The Thirteenth Annual Conference on Legal Knowledge and Information Systems JURIX 2000*, pages 73–84, Amsterdam, 2001. IOS Press.

[BCS03] T.J.M. Bench-Capon and G. Sartor. A model of legal reasoning with cases incorporating theories and values. To appear in *Artificial Intelligence*, 2003.

[BH93] D.H. Berman and C.D. Hafner. Representing teleological structure in case-based reasoning: The missing link. In *Proceedings of the Fourth International Conference on Artificial Intelligence and Law*, pages 50–59, New York, 1993. ACM.

[Chi77] R. Chisholm. *Theory of Knowledge*. Prentice Hall, Englewood Cliffs (NJ), [1966] 1977.

[Cla87] K.L. Clark. Negation as failure. In M.L. Ginsberg, editor, *Readings in Nonmonotonic Reasoning*, pages 311–325. Morgan Kaufmann, Los Altos (California), [1978] 1987. First published in Logic and Data Bases. Ed. H. Gallaire and J. Minker, 293-322. New York: Plenum Press.

[Dwo86] R.M. Dworkin. *Law's Empire*. Kermode, London, 1986.

[Gar87] P. Gardenfors. *Knowledge in Flux*. MIT, Cambridge (MS), 1987.

[Gin87] M.L. Ginsberg. *Readings in Nonmonotonic Reasoning*. Morgan Kaufmann, Los Altos (California), 1987.

[Gor95] T.F. Gordon. *The Pleadings Game. An Artificial Intelligence Model of Procedural Justice*. Kluwer, Dordrecht, 1995.

[Hag97] J.C. Hage. *Reasoning With Rules. An Essay on Legal Reasoning and Its Underlying Logic*. Kluwer, Dordrecht, 1997.

[Hag00] J.C. Hage. Dialectical models in artificial intelligence and law. *Artificial Intelligence and Law*, 8:137–172, 2000.

[Har51] H.L.A. Hart. The ascription of responsibility and rights. In A. Flew, editor, *Logic and Language*, pages 145–166, Oxford, [1948-1949] 1951. Blackwell. First published in 1948-1949. Proceedings of the Aristotelian Society 49: 171-194.

[Hem66] C.G. Hempel. *Philosophy of Natural Sciences.* Prentice-Hall, Englewood Cliffs (NJ), 1966.

[Hin99] J. Hintikka. Is logic the key to all good reasoning? In *Inquiry as Inquiry: a Logic of Scientific Discovery*, volume 5 of *Jaakko Hintikka Selected Papers*, pages 1–24. Kluwer Academic Publishers, Dordrecht/Boston/London, 1999.

[HLL94] J.C. Hage, R.E. Leenes, and A.R. Lodder. Hard cases: a procedural approach. *Artificial Intelligence and Law*, 2:113–166, 1994.

[Hor02] J. Horty. Skepticism and floating conclusions. *Artificial Intelligence*, 135:55–72, 2002.

[Kow79] R.A. Kowalski. *Logic for Problem Solving*. North Holland, Elsevier, New York, 1979.

[Kuh62] T. Kuhn. *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago (Illinois), 1962.

[Lak78] I. Lakatos. Falsification and methodology of scientific research programmes. In *The Methodology of Scientific Research Programmes*, pages 8–101. Cambridge University Press, Cambridge, [1970] 1978.

[Leh90] K. Lehrer. *Knowledge*. Routledge, London, 1990.

[Lod99] A.R. Lodder. *DiaLaw. On Legal Justification and Dialogical Models of Argumentation*. Law and Philosophy Library. Kluwer Academic Publishers, Dordrecht/Boston/London, 1999.

[Lou98] R.P. Loui. Process and policy: resource-bounded non-demonstrative reasoning. *Computational Intelligence*, 14:1–38, 1998.

[Mac95] N. MacCormick. Defeasibility in law and logic. In Z. Bankowski, I. White, and U. Hahn, editors, *Informatics and the Foundations of Legal Reasoning*, Law and Philosophy Library, pages 99–117. Kluwer Academic Publishers, Dordrecht/Boston/London, 1995.

[McC87a] J. McCarthy. Epistemological problems of artificial intelligence. In M.L. Ginsberg, editor, *Readings in Nonmonotonic Reasoning*, pages 46–52. Morgan Kaufmann, Los Altos (California), [1977] 1987. First published in Proceedings of the Fifth International Joint Conference in Artificial Intelligence, 1038-1044. Cambridge (Massachusetts).

[McC87b] J. McCarthy. Circumscription – a form of non-monotonic reasoning. In M.L. Ginsberg, editor, *Readings in Nonmonotonic Reasoning*, pages 145–151. Morgan Kaufmann, Los Altos (California), [1980] 1987. First published in First published in Artificial intelligence 13: 27-39.

[McC97] L.T. McCarty. Some arguments about legal arguments. In *Proceedings of the Sixth International Conference on Artificial Intelligence and Law*, pages 215–224, New York, 1997. ACM.

[Mil74] J. Stuart Mill. *On Liberty*. Harmondsworth, Penguin, [1859] 1974.

[Moo87a] R.C. Moore. Semantical considerations on nonmonotonic logic. In M.L. Ginsberg, editor, *Readings in Nonmonotonic Reasoning*, pages 127–136. Morgan Kaufmann, Los Altos (California), [1984] 1987. First published in Artificial Intelligence 25: 79-94.

[Moo87b] R.C. Moore. Possible-world semantic for autoepistemic logic. In M.L. Ginsberg, editor, *Readings in Nonmonotonic Reasoning*, pages 137–142. Morgan Kaufmann, Los Altos (California), [1985] 1987. First published in Proceedings 1984 Non-monotonic Reasoning Workshop, New Palz. New York: American Association for Artificial Intelligence.

[PC99] J.L. Pollock and J. Cruz. *Contemporary Theories of Knowledge*. Rowman and Littlefield, Totowa (New York), [1986] 1999.

[Pec83] A. Peczenik. *The Basis of Legal Justification*. Peczenick, Lund, 1983.

[Pec89] A. Peczenik. *On Law and Reason*. Kluwer, Dordrecht, 1989.

[Pec97] A. Peczenik. The passion for reason. In Wintgens L.J, editor, *The law in philosophical perspective*, pages 173–223. Kluwer, Dordrecht, 1997.

[Pol95a] J.L. Pollock. *Cognitive Carpentry: A Blueprint for How to Build a Person*. MIT, New York, 1995.

[Pol95b] J.L. Pollock. *Knowledge and Justification*. Princeton University Press, Princeton (New Jersey), 1995.

[Pop59] K.R. Popper. *The Logic of Scientific Discovery*. Hutchinson, London, [1934] 1959.

[Pra01a] H. Prakken. Modelling defeasibility in law: logic or procedure? *Fundamenta Informaticae*, 48:253–271, 2001.

[Pra01b] H. Prakken. Modelling reasoning about evidence in legal procedure. In *Proceedings of the Eighth International Conference on Artificial Intelligence and Law*, pages 119–128, New York, 2001. ACM Press.

[PS96] H. Prakken and G. Sartor. A dialectical model of assessing conflicting arguments in legal reasoning. *Artificial Intelligence and Law*, 4:331–368, 1996.

[PS02] H. Prakken and G. Sartor. The role of logic in computational models of legal argument: a criticial survey. In A. Kakas and F. Sadri, editors, *Computational Logic: Logic Programming and Beyond. Essays in honour of Robert A. Kowalski – Part II*. Springer Lecture Notes in Computer Science 2048, pp. 342–380. Springer Verlag, Berlin, 2002.

[PV00] H. Prakken and G.A.W. Vreeswijk. Logical systems for defeasible argumentation. In D. Gabbay & F. Guenthner, editors, *Handbook of Philosophical Logic*, second edition, Vol 4, pp. 219–318. Kluwer, Dordrecht/Boston/London, 2002.

[Rei87] R. Reiter. A logic for default reasoning. In M.L. Ginsberg, editor, *Readings in Nonmonotonic Reasoning*, pages 145–151. Morgan Kaufmann, Los Altos (California), [1980] 1987. First published in Artificial intelligence 13: 81–132, 1980.

[Res77] N. Rescher. *Dialectics: a Controversy-oriented Approach to the Theory of Knowledge*. State University of New York Press, Albany, N.Y., 1977.

[Ros30] W.D. Ross. *The Right and the Good*. Clarendon Press, Oxford, 1930.

[Ros39] W.D. Ross. *Foundations of Ethics*. Clarendon Press, Oxford, 1939.

[Sar95] G. Sartor. Defeasibility in legal reasoning. In Z. Bankowski, I. White, and U. Hahn, editors, *Informatics and the Foundations of Legal Reasoning*, Law and Philosophy Library, pages 119–157. Kluwer Academic Publishers, Dordrecht/Boston/London, 1995.

[Sar02] G. Sartor. Teleological arguments and theory-based dialectics. *Artificial Intelligence and Law* 10:95–112, 2002.

[SL92] G.R. Simari and R.P. Loui. A mathematical treatment of defeasible reasoning and its implementation. *Artificial Intelligence*, 53:125–157, 1992.

[Tha92] Paul Thagard. *Conceptual Revolutions*. Princeton University Press, Princeton (NJ), 1992.

[Tha01] Paul Thagard. *Coherence in Thought and Action*. MIT, Cambridge (MA), 2001.

[Wal96] D.N. Walton. *Argumentation Schemes for Presumptive Reasoning*. Lawrence Erlbaum Associates, Mahwah, NJ, 1996.

[Wal02] D.N. Walton. *Legal Argumentation and Evidence*. Pennsylvania State University Press, University Park, PA, 2002.

[Wil99] Edward O Wilson. *Consilience. The Unity of Knowledge*. Abacus, London, [1998]1999.

[WK95] D.N. Walton and E.C.W. Krabbe. *Commitment in Dialogue. Basic Concepts of Interpersonal Reasoning*. State University of New York Press, Albany, NY, 1995.