AI & Law, logic and argument schemes

Prakken, Henry

AI & Law, Logic and Argument Schemes

HENRY PRAKKEN
Department of Information and Computing Sciences, Utrecht University, and
Faculty of Law, University of Groningen
The Netherlands
Email: henry@cs.uu.nl

ABSTRACT: This paper reviews the history of AI & Law research from the perspective of argument schemes. It starts with the observation that logic, although very well applicable to legal reasoning when there is uncertainty, vagueness and disagreement, is too abstract to give a fully satisfactory classification of legal argument types. It therefore needs to be supplemented with an argument-scheme approach, which classifies arguments not according to their logical form but according to their content, in particular, according to the roles that the various elements of an argument can play. This approach is then applied to legal reasoning, to identify some of the main legal argument schemes. It is also argued that much AI & Law research in fact employs the argument-scheme approach, although it usually is not presented as such. Finally, it is argued that the argument-scheme approach and the way it has been employed in AI & Law respects some of the main lessons to be learnt from Toulmin's *The Uses of Argument*.

KEYWORDS: argument schemes, legal reasoning, defeasible reasoning, nonmonotonic logic, warrants.

1 INTRODUCTION

In the early days of Artificial Intelligence & Law research there was a sometimes heated debate about the suitability of logic for modelling legal reasoning. Logic, it was said, could not cope with the vagueness, indeterminacy and adversarial nature of the law (e.g. Berman and Hafner, 1987). Now we know that this criticism was not justified and that logical methods can also be applied when arguments for opposite conclusions are possible. For instance, techniques from nonmonotonic logic have been used to model reasoning with hierarchies of possibly conflicting rules (see Prakken and Sartor, 2002 for an overview) and even to model aspects of reasoning with precedents (e.g. Prakken and Sartor, 1998; Prakken 2002). However, even if these modern techniques are used, a logical approach still has some limitations and it needs to be supplemented with other elements. The point is that the notion of a logical inference is very abstract. The logical validity of an inference purely depends on the meaning of some structural words involved in the inferences, such as the connectives (and, or, if, not, ...) and the quantifiers (all, some, most, ...). However, sentences that are indistinguishable from a logical point of view can very well play different roles in an argument. This point was perhaps first stressed by Toulmin (1958), who in his famous scheme for arguments distinguished a *claim* connected to *data* by a *warrant* on account of a *backing*, and subject to exceptions specified by a *rebuttal*. Perhaps more important than Toulmin's particular scheme is his general observation that the various elements of an argument can play different roles, which leads to different standards for evaluating arguments.

Consider by way of illustration the following sentences:

All Dutch men are tall
All email addresses are personal data

From a logical point of view, both sentences are universally quantified implications. However, from an epistemological point of view they are clearly different. The first sentence is an empirical statement about a certain class of animals, while the second sentence is a legal rule interpreting a certain legal concept. Someone who disagrees with the first sentence will use different ways of attacking it than someone who disagrees with the second sentence. Attacks on the first sentence will typically refer to empirical observations ("yesterday I saw a short Dutchman") or to empirical methodology ("your sample is biased"). By contrast, attacks on the second sentence will usually refer to a legal source (relevant statutes), to legal authority ("the Dutch supreme court ruled otherwise") or to principle or policy ("regarding email addresses as personal data allows the use of privacy protection laws against spam"). Sometimes arguments about empirical statements are also source-based ("How do you know that all Dutch men are tall?" "Henry told me, and he is Dutch, so he is in the position to know." "But Henry often lies."). However, even within the class of empirical statements there are clear differences. Compare

Dutch men usually like soccer
Witnesses usually speak the truth

Both sentences are empirical generalisations but the second one is more, since it also expresses a source of knowledge while the first does not. Because of this difference, the second statement can be attacked in ways that do not apply to the first. For instance, it can be attacked on the grounds that a witness is biased, or has malfunctioning senses. Of course, debates about the first sentence could also evolve into a debate about sources, if it is asked what the source is of this generalization. But the sentence itself does not express a source of knowledge.

The general point of these examples is that the use of sentences in arguments does not only depend on their logical form but also on other things, such as their epistemological or pragmatic nature. Logic, with its abstract definition of logical validity (whether deductive or nonmonotonic), is blind to such differences, and should therefore be supplemented with a so-called "argument-scheme" approach. In line with Toulmin (1958), such an approach can identify the different roles that the various elements of an argument can play and thereby paves the way for a field-dependent notion of validity of arguments. The reason is that different types of premises have different ways of being critically examined and, since different fields can have their own typical argumentation schemes, the criteria for evaluating arguments will differ for each field.

Argumentation theory is the only research area where the notion of argument schemes (or "argumentation schemes") is explicitly studied as such. However, in this paper I want to argue (following Gordon 2003) that much AI & Law research in fact also employs the argument-scheme approach and thus takes some of Toulmin's main lessons to heart. In doing so, it is not my aim to give a comprehensive overview of AI & Law research. Rather, I will use a select set of examples from the AI & Law literature to illustrate the argument-scheme approach.

The rest of this paper is organised as follows. In Section 2 I explain the notion of an argument scheme as it is studied in argumentation theory, followed by a brief discussion in Section 3 of how argument schemes can be formalised in AI models of commonsense reasoning. In section 4 I briefly sketch the main phases of legal problem solving, viz. proof of the facts, rule interpretation and rule application. In the subsequent sections (5-7) I then discuss some of the main argument schemes used in each of these phases and some of the AI & Law projects modelling these schemes. In Section 8 I discuss some limitations of the argument-scheme approach to legal reasoning and in Section 9 I conclude by discussing how all this work in AI & Law can be seen as developing the research program suggested by Toulmin (1958).

2 ARGUMENT SCHEMES

In this section I sketch the main ideas behind the argument-scheme approach in argumentation theory. The notion of argument schemes is one of the central topics in current argumentation theory. For a recent overview see Garssen (2001). Important contributions to the study of argument schemes have been made by Douglas Walton (e.g. 1996). As conceived by him, argument schemes technically have the form of an inference rule. Consider, for instance, the following scheme from epistemic reasoning of "arguments from the position to know" (Walton 1996, pp. 61-3):

> *Person W says that p*
> *Person W is in the position to know about p*
> *Therefore (presumably), p*

(Note the resemblance to our example "witnesses usually speak the truth" from the introduction. In fact, the latter sentence is an instance of this scheme.) Or consider the following scheme from practical reasoning, of "arguments from consequences" (Walton 1996, pp. 75-7):

> *If A is brought about, then good (bad) consequences will (may plausibly) occur.*
> *Therefore, A should (not) be brought about.*

Our "all email addresses are personal data" example from the introduction may be transformed into an argument from good consequences:

> *If the term "personal data" of the Dutch Data Protection Act is interpreted to include email addresses, then legal measures against spam become possible, which is good.*
> *Therefore, the term "personal data" of the Dutch Data Protection Act should be interpreted to include email addresses.*

Argument schemes are not classified according to their logical form but according to their content. Many argument schemes in fact express epistemological principles (such as the scheme from the position to know) or principles of practical reasoning (such as the scheme from consequences). Accordingly, different domains may have different sets of such principles. Argument schemes come with a set of critical questions that have to be answered when assessing whether their application in a specific case is warranted. Some of these questions pertain to acceptability of the premises, such as 'is *W* in the position to know about *p*?' or 'is the possibility to use legal means against spam really good?". However, other questions point at exceptional circumstances in which the scheme may not apply, such as 'is *W* sincere?' or "are there better ways to bring about these good consequences?". Clearly, the possibility to ask such critical questions makes argument schemes defeasible, since negative answers to such critical questions are in fact counterarguments, such as "Person *W* is not sincere since he is a relative of the suspect and relatives of suspects tend to protect the suspect". Another reason why argument schemes are defeasible is that they may be contradicted by conflicting applications of the same or another scheme. For instance, a positive instance of the scheme from consequences can be attacked by a negative instance of the same scheme, such as by "interpreting email addresses as personal data also has bad consequences, since the legal system will be flooded with litigation, so the term "personal data" should not be interpreted to include email addresses". Or one person in a position to know (say an eyewitness) may have said that the suspect was at the crime scene while another eyewitness may have said that the suspect was not at the crime scene. Or a witness testimony may be rebutted with an argument from another scheme, such as an argument using camera evidence.

Above I said that argument schemes are classified according to their content. However, from a logical point of view they can be transformed into instances of logical inference rules by adding the connection between premises and conclusion as a conditional premise. Since as just explained most argument schemes are defeasible, this conditional will also be of a defeasible nature. For instance, the scheme from the position to know can be transformed into:

> *Person W says that p*
> *Person W is in the position to know about p*
> <u>*Persons who are in the position to know usually speak the truth*</u>
> *Therefore (presumably), p*

And the scheme from consequences can be transformed into

> *If A is brought about, then good (bad) consequences will (may plausibly) occur.*
> *If bringing about A will (may plausibly) result in good (bad) consequences then,*
> <u>*other things being equal, A should (not) be brought about*</u>
> *Therefore (presumably), A should (not) be brought about.*

Thus both schemes become an instance of the *defeasible modus ponens* rule, which is formalised by many systems of nonmonotonic logic (for an overview see Horty 2001):

> *P*
> <u>*If P then usually Q*</u>
> *Therefore (presumably), Q*

This scheme can be attacked by arguing that there is an exception to the rule that *if P then usually Q* (for instance, *P & R* and *If P & R then usually not-Q*) However, the fact that such a logical reconstruction of argument schemes is possible should not be taken to mean that the notion of argument schemes has no point. The point is that the two argument schemes above are typical ways in which the inference scheme of defeasible modus ponens can be instantiated, each with their own typical ways of critical testing, and that they therefore merit independent study instead of merely as instances of this abstract inference scheme.

3 A FORMAL FRAMEWORK FOR REASONING WITH ARGUMENT SCHEMES

I now briefly outline a formal framework for modelling reasoning with argument schemes. It especially draws on work of myself and others on argument-based approaches to so-called nonmonotonic reasoning. This outline also illustrates the remarks made in the introduction that (new) logical tools exist that can cope with uncertainty, disagreement and exceptions. The logical account of argument schemes outlined in this section was first proposed by Prakken et al. (2003). Independently, Verheij (2001) suggested a similar account in the context of his "Deflog" logic and developed it in e.g. (Verheij 2003).

The fact that argument schemes leave room for counterarguments naturally points at an argument-based approach to the formalisation of reasoning with such schemes. To this end, so-called logics of defeasible argumentation are in principle very suitable. Such logics were developed in AI to formalise commonsense reasoning (see Prakken and Vreeswijk, 2002 for an overview) and they have been popular in AI & Law as a way to formalise the adversarial nature of legal argument (see e.g. Gordon, 1995; Prakken and Sartor 1996, 1998; Bench-Capon 2003, Verheij 2003). Essentially, such logics define arguments as trees of deductive and/or defeasible inferences (such as deductive or defeasible modus ponens) and they allow for attacks on the defeasible inference steps of an argument (such as attacking defeasible modus ponens by arguing that there is an exception to the rule). For present purposes the work of the philosopher and AI researcher John Pollock (e.g. 1995) is especially relevant, since he classifies defeasible inference rules according to general principles of epistemology and practical reasoning. He calls his defeasible inference rules *prima facie reasons*. One such reason is the perception principle:

Having a percept with content p is a prima facie reason to believe p.

Other *prima facie* reasons studied by Pollock are, for instance, the statistical syllogism (a probabilistic version of defeasible modus ponens) and principles of memory, induction and temporal persistence.

As in all systems of logic, In Pollock's system an argument can be attacked by denying one of its premises (in fact I am ignoring some technical complications here). In addition, Pollock allows arguments to be attacked in two ways. An argument can be *rebut* with an argument for the opposite conclusion, and it can be *undercut* with an argument why a *prima facie* reason does not apply in the given circumstances. Intuitively, undercutting attacks do not argue that the attacked conclusion is false, but only that the conclusion is not sufficiently supported by its premises. Note that thus Toulmin's notion of a rebuttal in fact corresponds with Pollock's notion of an undercutter. An example of an undercutter is that if somebody perceives an object of red colour, then an undercutter of the perception principle is that the object is illuminated by a red light. For an example of an undercutter in a legal context, consider the argument "The suspect was at the murder scene at the time of the murder since witness John saw the suspect there" (applying the *prima facie* reason from perception; note that this reason can be applied only if it is first argued with a position-to-know argument that John saw this because he says that he saw it.). This argument is undercut by "It was too dark, so John could not have made a reliable identification" (applying an undercutter of the perception scheme).

How does Pollock's system relate to the argument-scheme approach? Essentially, the answer is that argument schemes can be formalised as prima facie reasons, that applications of schemes resulting in opposite conclusions can be regarded as rebuttals, while negative answers to critical questions about exceptional circumstances correspond to undercutters. The possibility that arguments can be defeated accounts for the defeasibility of argument schemes (recognised by Toulmin in his notion of a "rebuttal''), while the notion of undercutters allows for field-dependent standards for evaluating arguments, since each scheme has its own undercutters.

This is not all there is to say about argument-based logics. Given a set of conflicting arguments, it must be determined whether some of these arguments prevail. This is done in two steps. Firstly, standards can be used for comparing conflicting arguments to see which one is stronger than the other, if any. For instance, in case of two conflicting arguments from good consequences, one might (other things being equal) prefer the argument about the consequences that are judged the most important (e.g. in our example one might prefer the value of protecting the legal system against too much litigation over the value of fighting spam). Next, when all relations of relative strength between conflicting arguments are determined, the dialectical status of an argument is defined, to identify the defeasibly valid inferences. An important phenomenon here is *reinstatement*: suppose that argument *B* is stronger than argument *A* but that *B* is itself attacked by a stronger argument *C*; in that case *C* reinstates *A*. Consider again our rebutting arguments based on two conflicting witnesses (call them John and Bob). Even if we would prefer Bob's testimony given that, say, he is an adult and John a child, the argument using Bob's testimony may be undercut by a third argument *C* "Bob's testimony is unreliable since he has a strong reason to hate the suspect".

An intuitive way to define the defeasible validity of arguments is in the form of an *argument game* between a proponent and an opponent of an argument. Proponent starts the game with the argument to be tested and then the players take turns, each attacking the preceding argument. Opponent's arguments must be at least as strong as their targets while proponent's arguments must be stronger than their targets. A player has won if the other player has run out of moves. Now an argument A is defeasibly valid if the proponent has a winning strategy (in the game-theoretic sense) in a game beginning with A, i.e., if he can make the opponent run out of moves no matter how she plays.

In sum, argument-based logics conceive of argumentation as a tree of trees: individual arguments are trees in which statements are linked with each other by inference rules, and the dialectical status of arguments is determined by forming a dialectical tree of all possible ways to play an argument game for this argument. An argument is defeasibly valid if the proponent can choose his arguments in the dialectical tree in such a way that he always ends in a leaf with one of his own arguments.

4 A BRIEF SKETCH OF LEGAL PROBLEM SOLVING

Let us now take a closer look at legal reasoning, to identify some of the main argument schemes used in it. (The analysis in this section is not meant to be original; it merely serves as a basis for the further discussions.) I will take my starting point in the phenomenon of a statutory rule that has to be applied to the facts of a case. A typical legal rule is of the form

If Conditions then Legal consequence

A legal rule connects the factual world with the normative world. Consider the following paraphrase of a legal rule from the Dutch Data Protection Act.

If personal data is reused without the subject's permission for means irreconcilable with the aims for which the data were collected, then the reuse is not allowed.

Consider now the following (entirely real) case. Somebody claims, firstly, that Utrecht University has given the addresses of all their students without their permission to the local police in order to enable the police to start a campaign against bicycle theft (a very common criminal offence in Dutch university towns) by sending all students a letter to warn them that buying a stolen bicycle is itself a criminal offence. The person also claims that this was not allowed since this goal is irreconcilable with the aim for which these addresses were collected, viz. to manage the university administration with respect to their students. Now if such a case is taken to court, at least four questions have to be answered.

The first question is whether all these events indeed happened. This is a matter of evidence. It will be decided on the basis of the available 'sense data', such as, for instance, the letter of the police and further documentary evidence (e.g. a letter of request from the police to the university board) and/or witness testimonies (e.g. a statement by a police officer that they obtained the addresses from the University).

Suppose that on the basis of this evidence the court is convinced that the events indeed occurred. Then a second step has to be taken to classify the events under the conditions of the rule, viz. interpreting the rule's conditions to decide whether it subsumes the events as proven by the 'sense data'. A well-known problem here is that often there are no clear criteria for this decision, because of the vagueness or open-texturedness of the rule's conditions. This is the question to which most of the AI & Law research on legal argument is devoted.

Suppose now that the court has decided that the events as proven indeed classify as an instance of the rule's conditions, for example, on the grounds that preventing theft of bicycles has nothing to do with running the university administration. Then two further questions have to be answered. The first is whether the rule is legally valid, i.e., whether it is from a legally recognised source of law. This question must be answered independently of the facts of the case. This is different for the final question to be answered, viz., whether the rule must be applied to the case at hand, or whether there are circumstances that prevent the rule's application (e.g. a conflicting rule also applies, or applying the rule would be manifestly unjust or unreasonable). Perhaps the university could argue that the negative consequences for its students and employees are so small and that the problem of bicycle theft in Utrecht is so serious while no other measure has worked that applying the rule in this case is unreasonable. Perhaps the university can even cite a precedent where the rule was set aside in a similar case.

In sum, statutory rule application involves (at least) four steps: proving that the facts to which the rule is claimed to apply have indeed occurred (evidence), deciding that the facts as proven are subsumed under the rule's conditions (classification/interpretation), deciding that the rule is valid law (rule validity) and deciding that the rule ought to be applied (rule application).

It should be noted that this four-step process is in general not sequential. For instance, the choice of facts to be proven is determined not only by the available evidence but also by the possible rules that may fit the facts once proven. If a university student in our example thinks that a claim on the basis of breach of contract is more promising than a claim on the basis of privacy violation, she might select different facts-to-be-proven, such as that a contract exists, and she might accordingly search for different items of evidence. Modelling this process of reinterpreting the facts to fit a certain rule has so far proven too hard for AI & Law research (cf. e.g. Branting, 2003).

Let us now have a closer look at each of the four steps, discussing some of the main argument schemes involved and some of the AI & Law research on modelling these schemes.

5 SCHEMES FOR REASONING ABOUT EVIDENCE

Until the 2003 Conference on AI & Law in Edinburgh, reasoning about evidence was a largely neglected area of AI & Law research. One exception was Lutomski (1989), who presented an early application of the argument-scheme approach to evidential reasoning at ICAIL-1989 in Vancouver. His system, which was implemented but, to my knowledge, never used in practice, was meant to assist an attorney in dealing with statistical evidence in the domain of employment discrimination. The system stored typical arguments based on statistics in a Toulmin structure, together with typical critical questions (for instance, "have all and only relevant data be collected?").

At ICAIL-2003 a considerable number of papers addressed the topic of evidential reasoning. One of them was my paper with Chris Reed and Douglas Walton (Prakken et al. 2003), in which some first steps were taken to develop an explicit argument-scheme analysis of evidential reasoning. This work was further developed in Bex et al. (2003) and Prakken (2004). In this section I briefly summarize this work.

The basic idea is to formalise evidential argument schemes in John Pollock's framework (see Section 3) as *prima facie* reasons and to regard the critical questions attached to the schemes as pointers to undercutting defeaters. Some of Pollock's own reasons directly apply to evidential reasoning, such as the perception principle discussed in Section 3 and principles based on memory, induction and the statistical syllogism. The latter principle is Pollock's probabilistic version of defeasible modus ponens. I paraphrase it in a version without numbers:

> *'c is an F' and 'F's are usually G's' is a prima facie reason for 'c is a G'*

This principle drives reasoning with empirical generalisations. The main undercutter is subproperty defeat, which captures exceptions to a generalisation:

> *'c is an F&H' and 'it is not the case that F&H's are usually G's' is an undercutter of the statistical syllogism.*

For example, an argument using "Dutch men usually like soccer" may be undercut by an argument using "It is not the case that Dutch men with a PhD degree usually like soccer". Applying both generalisations to the author of this paper results in a defeated argument that Henry likes soccer, although no argument for the opposite conclusion can be built.

In addition, Bex et al. discuss argument schemes for applying witness testimonies (essentially a variant of the scheme of arguments from the position to know), expert testimonies (another variant of this scheme) and temporal persistence. The latter reason can be used to argue from the fact that a fact *F* is true at a time *T1* that *F* is still true at a later time *T2,* if there is no evidence that *F* became false between *T1* and *T2*. Temporal persistence is an important aspect of evidential reasoning. For instance, in civil cases the usual way to prove that one has a legal right (e.g. ownership) is to prove that the right was created (e.g. by sale plus delivery). The other party must then usually prove later events that terminated the right. Prakken (2004) also discusses several ways to attack empirical generalisations that do not employ the subproperty defeater but that attack the sources of the generalisations (such as 'common sense' or 'science'). Bex et al. (2003) applies the approach of Prakken et al. (2003) to a small part of Kadane and Schum's (1996) reconstruction of the famous Sacco and Vanzetti case (viz. their Chart no. 4). The arguments that resulted from this reconstruction turned out to be based on Pollock's reasons from memory, perception, temporal persistence and the statistical syllogism and on the "position to know" scheme as reconstructed as an instance of the statistical syllogism (cf. the end of Section 2 above). The counterarguments that were not rebuttals could all be analysed as undercutters of these schemes.

This work is still preliminary. One direction of future research is attempting to formalise the schemes that practicing lawyers use in their cases. For common-law jurisdictions interesting sources of such schemes exist, viz. manuals for trial advocacy, such as Bergman (1997). By and large, this manual (implicitly) follows the argument-scheme approach, listing typical evidential arguments and typical ways to attack them. Another valuable research direction is capturing the available knowledge about the reliability of eyewitness testimonies in a knowledge-based system (cf. e.g. Bromby and Hall, 2002).

6 SCHEMES FOR RULE APPLICATION

Applying a legal rule to the facts is perhaps the central element of legal problem solving. One of the most elaborate AI & Law accounts of what it takes to apply a legal rule is given by Hage and Verheij in an application of their "reason-based logic" (see e.g. Hage 1997, Verheij 1996). Their central claim is that applying a legal rule involves much more than just applying the logical inference rule of *modus ponens* (whether defeasible or not). Their account of rule application can be briefly summarised as follows. First it must, of course, be determined whether the rule's conditions are satisfied (the interpretation question). If this hurdle is cleared (see also Section 7), it has to be determined whether the rule is legally valid (for instance, by arguing that it is from a certain legally recognised source). Then it has to be determined whether the rule's applicability is not excluded in the given case (for instance, the Dutch Data Protection Act does not apply to the police.). If this is the case, it must finally be determined that the rule can be applied (i.e., that no conflicting rules or principles apply). Interestingly, while Hage and Verheij mainly discuss how rule application can be blocked by legal principles, the CABARET system of Skalak and Rissland (1992), only allows rule application to be blocked by citing a precedent where the rule was not applied (see also Section 8). CABARET's approach is based on Gardner's (1987) point of view that if a legal principle or value justifies setting aside a rule, this will usually have been decided in a precedent.

Both Hage and Verheij and others (such as Gordon, 1995 and Prakken and Sartor 1996) have shown how arguments about these issues can be formalised in nonmonotonic logics. Among other things, these techniques can model the fact that in legal practice the validity and applicability of legal rules is usually presumed, a presumption which can be overturned only by an argument that it does not hold. The details of the techniques used are beyond the scope of the present paper, except for the remark that they in fact formalise and further develop Toulmin's notion of a rebuttal; see Prakken and Sartor (2002) for an overview of the various techniques. For present purposes the main conclusion is that the argument scheme for rule application involves various steps and that each of these steps can be attacked in stereotypical ways.

7 SCHEMES FOR PRECEDENT-BASED REASONING

In the previous section I briefly mentioned that most AI & Law research on the modelling of legal argument concerns the interpretation of legal concepts. This is a very hard research problem, since often a large gap exists between the concrete nature of the facts of a case and the abstract nature of legal concepts. This induces legal uncertainty in (at least) two ways.

The first way is the existence of conflicting interpretation rules (based, for instance, on opinions of legal experts, on commonsense interpretations of natural language or on the rationale of a precedent). For instance, one judge (or legal scholar) may say that email addresses are always personal data since when combined with an IP address of a computer they enable the identification of the user, while another judge or legal scholar may argue that an email address is not personal data if the left part of the address does not resemble the user's name. These are simply conflicting if-then rules, and any suitable technique from nonmonotonic logic can be used to formalise reasoning with such rules (see again the survey in Prakken and Sartor 2002). This is essentially the approach taken by Gardner (1987). Her (implemented) system, which investigated whether a contract was created by offer and acceptance, stored possibly conflicting interpretation rules derived from legal experts, commonsense and case law, and applied these rules in modus-ponens style, using a priority mechanism to give precedence to case law rules over conflicting expert or commonsense rules.

However, sometimes interpreting a legal concept is not a matter of simply formulating or selecting a suitable interpretation rule. Sometimes all there is, is a set of different factors, possibly with different magnitudes, that somehow have to be weighed in each particular case to determine its outcome. A well-known AI & Law example is HYPO's modelling of the American precedent-based domain of trade secret law (Ashley 1990). However, this phenomenon is not confined to common law jurisdictions. For instance, the Dutch Data Protection Act, when defining the concept of irreconcilable reuse of personal data, states five factors that "at least" have to be taken into account, without stating how they should be combined in a given case:

- the similarity between the aim of the reuse and the original aim for which the data were collected;
- the nature of the data involved;
- the consequences of the reuse for the person to which the data pertain;
- the manner in which the data were obtained;
- the extent to which suitable measures are taken to protect the privacy of the person to which the data pertain.

In such 'factor-based' domains, a decision in a new case is often made by referring to past decisions, i.e., to precedents. However, as shown by e.g. Ashley, the rationales of precedents often do not directly apply to a new case since different cases often have different constellations of the relevant factors and their values. Therefore, the rationales must often be adapted to fit the new case. A typical way to do so is to point at the similarities to a precedent with the desired outcome, to argue that because of these similarities the same decision should be made in the new case. And two typical ways to attack such an argument are, firstly, distinguishing the precedent by pointing at the differences and, secondly, pointing at a counterexample, i.e., at another precedent that is at least as similar and that has the opposite outcome. All this (and more) is modelled in the HYPO system. It is interesting to note that HYPO uses such precedent-based arguments in the context of an argument game (see Section 3 above) between a plaintiff and defendant in a certain case. In fact, the disputes thus generated by HYPO are at most three moves long (plaintiff-defendant-plaintiff) but nothing prevents a generalisation to disputes of arbitrary length. Thus HYPO illustrates that reasoning with precedent-based argument schemes can be modelled as a logical argument game.

The HYPO system is now more than 15 years old and much research has followed it. Most of the subsequent research consists of enriching HYPO's scheme for representing precedents and exploiting the added expressiveness for generating new kinds of arguments and counterarguments. While HYPO just distinguishes sets of pro-plaintiff and pro-defendant factors and a simple decision (plaintiff won or defendant won), in the CATO system (Aleven 1997) hierarchies of more and less abstract factors can be defined so that, for instance, a distinction can be downplayed by arguing that at a more abstract level the cases are still similar. Consider the following example in the context of Dutch privacy law, where one case is about a single sending of a warning letter by the police to students while another case (also entirely real) is about a single sending of a fund raising letter by the university to their students and employees. The cases could be distinguished at this factual level but the distinction could be downplayed by arguing that both cases are about one-time letters about matters of public interest. Others, e.g. Prakken (2002), Bench-Capon and Sartor (2003) and Atkinson et al. (2005) have tried to represent the values that are advanced or endangered by deciding a case in one way or another, resulting in the modelling of new, teleological argument schemes, related to Walton's argument scheme from consequences. For instance, in Prakken's (2002) approach a distinction can be emphasised by saying that because of the differences between the precedent and the current case, following the precedent in the current case will not advance the same values as were advanced by the precedent's outcome.

8 LIMITATIONS OF THE ARGUMENT SCHEME APPROACH

Concluding this brief overview of AI & Law research, we have seen that an argument-scheme approach to the modelling of legal argument is a useful supplement to a purely logic-based approach. In particular, an argument-scheme approach can model the different roles that the various statements in an argument can have and thus allow for different standards for evaluating arguments. We have also seen that much AI & Law research in fact employs the argument-scheme approach, although it usually is not presented as such.

Perhaps at this point the reader has the impression that all that modelling legal reasoning is about is modelling the relevant argument schemes and associated critical questions, and using them in a logical argument game as explained in Section 3. However, this would be a severe simplification of legal reasoning, and much interesting work in AI & Law goes beyond this simple approach.

For instance, CABARET (Skalak and Rissland 1992) defines strategies and tactics for using and combining rule- and precedent-based schemes for certain dialectical purposes, such as confirming or discrediting a rule. Thus Cabaret in fact defines rational strategies for playing an argument game. And part of the HYPO and CATO systems are mechanisms for interpreting existing material before using it in an argument. For instance, CATO still uses HYPO's simple case representation scheme but CATO's factor hierarchy can be used to generate different arguments about why a case was decided the way it was, by suggesting different 'paths' from the factors to the decision through the hierarchy. The idea of reinterpreting precedents was further developed by Loui and Norman (1995), who model five ways to reinterpret the (precedent-based) arguments of one's opponent in order to reveal new attacking points so that they can be better attacked. One way is to argue that in the precedent the outcome was based on choosing between two conflicting arguments, and that in the new case the winning argument does not apply since one of its premises is missing in the new case, so that the argument that was overruled in the precedent should now prevail. Finally, Bench-Capon and Sartor (2003) have addressed the problem of theory formation, by modelling constructors for theories that explain a certain set of precedent decisions. All this is very important work but it goes beyond the argument-scheme approach. Either it provides the material from which arguments can be built (CATO, Loui and Norman, Bench-Capon and Sartor) or it defines tactics and strategies for how an argument game can be played (CABARET, Loui and Norman).

9 AI & LAW AND TOULMIN

Finally, it is interesting to discuss how all this AI & Law work respects the observations of Toulmin (1958). To start with, the reader might wonder why AI & Law has not made more direct use of Toulmin's argument scheme. Here Toulmin's own reflections on his 1958 work are relevant. In the preface of his second edition (Toulmin, 2003) he expresses his surprise that many regarded his 1958 scheme as a proposal for a theory of argumentation while his aim had been different, viz. to criticise the view that all arguments can be put in deductive form. We should therefore not be surprised that subsequent research has replaced Toulmin's original single argument scheme with revised and more refined classifications of schemes. In line with this, it seems more worthwhile to investigate how AI & Law has taken the general lessons of Toulmin (1958) to heart than to discuss how AI & Law has used Toulmin's particular scheme.

In my opinion, AI & Law has taken to heart the following three of Toulmin's (1958) lessons. Firstly, Toulmin stressed that premises of an argument can have different roles. We have seen that AI & Law has identified many argument schemes with stereotypical roles for premises. Secondly, Toulmin stressed that everyday arguments are defeasible, which he captured in his notion of a rebuttal. It has been argued many times before that the field of nonmonotonic logic has formalised and further developed this aspect of Toulmin's scheme, and the logical account of reasoning with argument schemes outlined in Section 3 illustrates this point. Finally, Toulmin stressed that the standards for evaluating arguments are field-dependent. His original argument scheme captured this by allowing for different backings of warrants. The account of Section 3 illustrates that it is possible to capture similar distinctions in a system of formal logic (be it a nonmonotonic logic), if the logic allows for the formalisation of different argument schemes, each with their own set of typical premises. Then the field-dependency of the standards can be captured by identifying the schemes employed in a certain field and formulating the undercutters that correspond to the critical questions of these schemes. Since different fields can have their own typical argumentation schemes, the criteria for evaluating arguments will thus differ for each field.

Concluding, while AI & Law has replaced Toulmin's original argument scheme with more refined analyses, it has done so fully within the spirit of Toulmin's challenge to develop an account of the validity of reasoning that applies to everyday argument. What is especially exciting is that AI & Law has shown that such an account can still be formal and computational.

ACKNOWLEDGEMENTS

An earlier version of this paper was presented at a session on Artificial Intelligence and Law of the International Congress of Comparative Cultures and Legal Systems, Mexico City, 9-14 February 2004. I am indebted to Trevor Bench-Capon and Thomas Gordon for illuminating discussions about the topic of this paper.

REFERENCES

Aleven, V.: 1997, *Teaching Case Based Argumentation Through an Example and Models*, PhD Thesis, The University of Pittsburgh.

Ashley, K.D.: 1990, *Modeling Legal Argument: Reasoning with Cases and Hypotheticals*, MIT Press, Cambridge, MA.

Atkinson, K., Bench-Capon, T.J.M. and McBurney, P.: 2005, `Arguing about cases as practical reasoning', in *Proceedings of the Tenth International Conference of Artificial Intelligence and Law*, ACM Press, New York, pp. 35-44.

Bench-Capon, T.J.M.: 2003, `Persuasion in Practical Argument Using Value-based Argumentation Frameworks', *Journal of Logic and Computation* **13**, 429-448.

Bench-Capon, T.J.M. and Sartor, G.: 2003, `A model of legal reasoning with cases incorporating theories and values'. *Artificial Intelligence* **150**, 97-143.

Bergman, P.: 1997, *Trial Advocacy in a Nutshell*, 3rd edition, West Publishing Co, St. Paul, MN.

Berman, D.H., and Hafner, C.D.: 1987, `Indeterminacy: a challenge to logic-based models of legal reasoning', in *Yearbook of Law, Computers and Technology*, Butterworths, London, Vol. 3, pp. 1-35.

Bex, F.J., Prakken, H., Reed, C. and Walton, D.N.: 2003, `Towards a formal account of reasoning about evidence: argumentation schemes and generalisations', *Artificial Intelligence and Law* **11**, 125-165.

Branting, L.K.: 2003, `An agenda for empirical research in AI and Law', in *Workshop Notes of the ICAIL-2003 Workshop on Evaluation of Legal Reasoning and Problem-Solving Systems*, Edinburgh, UK, pp. 28-35.

Bromby, M.C. and Hall, M.J.J.: 2002, `The development and rapid evaluation of the knowledge model of ADVOKATE: an advisory system to assess the credibility of eyewitness testimony', in T. J. M. Bench-Capon, A. Daskalopulu and R. G. F. Winkels (eds), *Legal Knowledge and Information Systems, JURIX 2002: The Fifteenth Annual Conference*, Amsterdam, IOS Press, pp. 143-152.

Gardner, A. von der Lieth.: 1987, *An Artificial Intelligence Approach to Legal Reasoning*, MIT Press, Cambridge, MA.

Garssen, B.: 2001, `Argument schemes', in F.H. van Eemeren (ed.), *Crucial Concepts in Argumentation Theory*, Amsterdam University Press, Amsterdam, 81-99.

Gordon, T.F.: 1995, *The Pleadings Game. An Artificial Intelligence Model of Procedural Justice*, Kluwer Academic Publishers, Dordrecht/Boston/London.

Gordon, T.F.: 2003, Artificial Intelligence Models of Legal Argumentation, *Tutorial notes Ninth International Conference on Artificial Intelligence and Law*, Edinburgh, UK.

Hage, J.C.: 1997, *Reasoning with Rules. An Essay on Legal Reasoning and its Underlying Logic*, Kluwer Academic Publishers, Dordrecht/Boston/London.

Horty, J.F.: 2001, `Nonmonotonic logic', in L. Goble (ed.), *The Backwell Guide to Philosophical Logic*, Blackwell Publishers, Oxford, pp. 336-361.

Kadane, J.B. and Schum, D.A.: 1996, *A Probabilistic Analysis of the Sacco and Vanzetti Evidence*, John Wiley & Sons, New York.

Loui, R.P. and Norman, J.: 1995, `Rationales and argument moves', *Artificial Intelligence and Law* **3**, 159-189.

Lutomski, L.S.: 1989, `The design of an attorney's statistical consultant', in *Proceedings of the Second International Conference of Artificial Intelligence and Law*, ACM Press, New York, pp. 224-233.

Pollock, J.L.: 1995, *Cognitive Carpentry. A Blueprint for How to Build a Person*, MIT Press, Cambridge, MA.

Prakken, H.: 2002, `An exercise in formalising teleological case-based reasoning', *Artificial Intelligence and Law* **10**, 113-133.

Prakken, H.: 2004, `Analysing reasoning about evidence with formal models of argumentation', *Law, Probability & Risk* **3**, 33-50.

Prakken, H., Reed, C. and Walton, D.N.: 2003, `Argumentation schemes and generalisations in reasoning about evidence', in *Proceedings of the Ninth International Conference of Artificial Intelligence and Law*, ACM Press, New York, pp. 32-41.

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

Prakken, H. and Sartor, G.: 1998, `Modelling reasoning with precedents in a formal dialogue game', *Artificial Intelligence and Law* **6**, 231-287.

Prakken, H. and Sartor, G.: 2002, `The role of logic in computational models of legal argument: a critical survey', in A. Kakas and F. Sadri (eds.), *Computational Logic: Logic Programming and Beyond. Essays In Honour of Robert A. Kowalski*, Part II. Springer Lecture Notes in Computer Science 2048, Berlin, pp. 342-380.

Prakken, H. and Vreeswijk, G.A.W.: 2002, `Logics for defeasible argumentation', in D. Gabbay and F. Guenthner (eds.), *Handbook of Philosophical Logic*, Kluwer Academic Publishers, Dordrecht/Boston/London, second edition, Vol 4, pp. 219-318.

Skalak, D.B. and Rissland, E.L.: 1992, `Arguments and cases: an inevitable intertwining', *Artificial Intelligence and Law* **1**, 3-44.

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

Toulmin, S.E.: 2003, *The Uses of Argument*, Cambridge University Press, Cambridge, 2nd edition.

Verheij, B.: 1996, *Rules, Reasons, Arguments. Formal Studies of Argumentation and Defeat*, Doctoral Dissertation University of Maastricht, the Netherlands.

Verheij, B.: 2001, `Legal decision making as dialectical theory construction with argumentation schemes' (research abstract), in *Proceedings of the Eighth International Conference of Artificial Intelligence and Law*, ACM Press, New York, pp. 225-226.

Verheij, B.: 2003, `Dialectical argumentation with argumentation schemes: an approach to legal logic', *Artificial Intelligence and Law* **11**, 167-195.

Walton, D.N.: 1996, *Argumentation Schemes for Presumptive Reasoning*, Erlbaum, Mahwah, N.J.