CrossMark

HYPO'S legacy: introduction to the virtual special issue

T. J. M. Bench-Capon[1]

Abstract This paper is an introduction to a virtual special issue of *AI and Law* exploring the legacy of the influential HYPO system of Rissland and Ashley. The papers included are: *Arguments and cases: An inevitable intertwining*, *BankXX: Supporting legal arguments through heuristic retrieval*, *Modelling reasoning with precedents in a formal dialogue Game*, *A note on dimensions and factors*, *An empirical investigation of reasoning with legal cases through theory construction and application*, *Automatically classifying case texts and predicting outcomes*, *A factor-based definition of precedential constraint* and *An improved factor based approach to precedential constraint*. After describing HYPO, in this introduction to the special issue I look at various aspects of its influence on AI and Law: the developments led by Rissland at Amherst; the developments led by Ashley in Pittsburgh; the expression of these ideas in terms of rule based systems, and their subsequent formalisation; value based theories, which were inspired by a critique of HYPO; and contemporary approaches which revive the idea of dimensions.

1 Introduction

It is now thirty years since HYPO came to the attention of the international AI and Law Community at the very first ICAIL at which Rissland and Ashley (1987) and Ashley and Rissland (1987) were presented. Although Edwina Rissland had published previously on reasoning with hypotheticals in Rissland (1980), Rissland and Soloway (1980) and Rissland (1984) and its application to law in Rissland (1983), it was the PhD project with her student Kevin Ashley which culminated in

✉ T. J. M. Bench-Capon
tbc@csc.liv.ac.uk

[1] Department of Computer Science, University of Liverpool, Liverpool, UK

Ashley (1990) which is generally regarded as the HYPO system (first referred to as such in Rissland et al. (1984)). Over the past 30+ years HYPO has been without doubt the most influential AI and Law project, setting the agenda for reasoning with legal cases, arguing with legal cases, rule based approaches to reasoning with cases and formalisations of precedential reasoning. The papers in this virtual special issue, which are drawn from all three decades, represent a range of significant papers which have built upon and developed ideas from the HYPO system. In this introduction I will use these papers to illustrate HYPO's considerable legacy. There are, of course, many other papers, a number of which were published in other venues, which are needed to tell the whole story, but the selection in this issue is fairly representative of the way thinking in AI and Law has developed. A number of other papers will be referred to in this introduction, so that the complete story can be told.

   Section 2 describes HYPO and its key features, in particular its representation of cases as *dimensions*, and its conception of the form of a *legal* argument. Section 3 describes some direct descendants of HYPO, led by Rissland at Amherst: CABARET (Skalak and Rissland 1992)*,[1] which combined rule-based and case-based reasoning; BankXX (Rissland et al. 1996)* which envisaged constructing legal arguments by heuristic search over a rich network of argument pieces; and SPIRE (Rissland and Daniels 1996), which used HYPO derived techniques to drive a system to retrieve legal cases. Section 4 considers the developments from HYPO led by Ashley following his move to Pittsburgh. The most influential is CATO, carried out with his student (Aleven and Ashley 1995). CATO was never the topic of an article in this journal, and is most completely reported in Aleven (1997, 2003). CATO is best known for its move from dimensions to *factors*, and its organisation of these factors into a *factor hierarchy*. A subsequent project with another of Ashley's students, Stefanie Brüninghaus, led to Issue Based Prediction (IBP) (Ashley and Brüninghaus 2009)*. During the 1980s and the early 1990s, case-based and rule-based approaches to legal reasoning were seen as distinct. Some saw them as potentially complimentary as in CABARET, but others, especially in Europe with its logic programming and civil law traditions, saw them as opposed to one another. An important paper by Prakken and Sartor (1998)*, however, demonstrated how a set of precedents could be represented as a set of rules and priorities between them. This enabled formal accounts of reasoning with cases to be developed Horty and Bench-Capon (2012)* and Rigoni (2015)*. Representation as rules also permitted the use of argumentation schemes to support reasoning with cases: Wyner and Bench-Capon (2007), Bench-Capon (2012) and Prakken et al. (2015). The move to rules and the formailsation of the reasoning is described in Sect. 5. Another direction was initiated by Berman et al. (1993), which led to preferences being explained in terms of purpose and values. This idea was formalised as theory construction in Bench-Capon and Sartor (2003), and explored empirically in Chorley and Bench-Capon (2005b)*. These developments are explored in Sect. 6. Although HYPO is the *fons et origo* of all this work, mostly it has been based on factors, as presented in CATO. There has, however, been interest in reviving

---

[1] The papers in this virtual special issue are indicated with an "*".

dimensions to allow more nuance to be captured. The differences between dimensions and factors are discussed in Rissland and Ashley (2002)*. Although the return to dimensions was argued for in Bench-Capon and Rissland (2001), in is only more recently that work on this has intensified: Prakken et al. (2015), Araszkiewiczet al. (2015) and Al-Abdulkarim et al. (2016b). This work is discussed in Sect. 7, and some concluding remarks are made and open questions identified in Sect. 8.

2 HYPO

Although the ideas relating to HYPO were around earlier, (a panel was held at IJCAI 1985 (Rissland 1985), with panellists including both Rissland and Ashley, together with Michael Dyer, Anne Gardner, Thorne McCarty and Donald Waterman), details of HYPO itself began to emerge at the first ICAIL in 1987 (Rissland and Ashley 1987). By 1989 Ashley had moved to Pittsburgh and so the discussion of HYPO at the second ICAIL was his sole authored (Ashley 1989), and his 1988 PhD thesis appeared in book form as Ashley (1990), providing the most detailed account of the project. HYPO operated in the domain of US Trade Secrets Law. Essentially the law protected trade secrets against misappropriation. To be a trade secret the information needed to have value, and to be regarded as a secret by the plaintiff, shown by making reasonable efforts to maintain secrecy. Misappropriation took the form either of using information knowing it to be confidential, or acquiring the information by dubious means, such as deception of bribery. Defences typically turned on the availability of the information through legitimate sources, or independent development of the information.

   The most important ideas behind HYPO were dimensions, and a particular conception of argumentation, namely three-ply argument. These will be discussed in turn in the following subsections.

2.1 Dimensions

A major difficulty when thinking about how to reason with legal cases is that cases which are held to be similar appear to be very different when we consider their facts. Thus in a famous line of cases considered in Levi (1948), Bench-Capon et al. (2003) and Rissland and Xu (2011) a collection of very different things are seen to play the same role in the various cases[2] . These include a scaffold, an elevator, a bottle of aerated water and a coffee urn which were all held to be relevantly similar to an automobile for the purposes of the particular legal point under consideration. Similarly the series of property cases much studied in AI and Law consider a fox,

---

[2] The need to consider disparate types of thing as if they were the same sort of thing is pervasive in analytic philosophy. Austin wrote in Austin and Warnock (1962) "But then we have to ask, of course, what this class comprises. We are given, as examples, 'familiar objects'—chairs, tables, pictures, books, flowers, pens, cigarettes; the expression 'material thing' is not here (or anywhere else in Ayer's text) further defined. But does the ordinary man believe that what he perceives is (always) something like furniture, or like these other 'familiar objects'—moderate-sized specimens of dry goods?".

wild ducks, a shoal of pilchards and a whale all to be relevantly similar to the baseball disputed in *Popov v Hayashi* (see Atkinson (2012)). Thus deciding whether two legal cases are relevantly similar does not seem to be a matter of considering facts in a simple manner using everyday notions of similarity.

For this reason HYPO introduced the notion of *dimensions* to match and compare cases. These are aspects of the case, ascribed on the basis of facts, which are relevant to the legal issues being considered. Thus what linked the scaffold, elevator, bottle of aerated water and coffee urn was that all were dangerous, but not visibly so. What links the animals and the baseball is that they were being pursued, and the relevant aspects include their value, and their connection to the land they were pursued on Bench-Capon and Bex (2015).

Dimensions are aspects of a case which may or may not be applicable. If applicable, the dimension represents a vector, taking a range of values which entirely favour one of the parties at one end and then increasingly favour the other party, until at the other end of the range the dimension entirely favours that other party. At some point the dimension will cease to favour the plaintiff and at some point (possibly the same point) it will start to favour the defendant. Thus if there are n dimensions, we have a n-dimensional space, with the plaintiff favoured at some locations and defendant favoured at others.

We can illustrate dimensions using the domain of the Automobile Exception to the 4th Amendment, discussed in Rissland (1989), Ashley et al. (2008) and Al-Abdulkarim et al. (2016c). This is a good illustration because it can be seen as having just two dimensions, privacy and exigency, which allows the space to be pictured, as in Fig. 1. Essentially the 4th Amendment provides that

> [t]he right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, ...

The automobile exception permits search of an automobile without a warrant provided that there is probable cause, and that the search is sufficiently exigent. Exigency arises from the possibility that the delay in seeking a warrant might compromise the collection of evidence or public safety. If we consider the dimension of exigency, a completely immovable item would represent one end of the dimension and an automobile in transit on a highway the other. If we arrange our precedents along this dimension we will observe that at one extreme, where there is little exigency, all the cases will require a warrant, and at the other, where the matter is pressing, none of the cases will require a warrant. In between, some cases will need a warrant and some will not, suggesting that other relevant dimensions are in play. The situation is represented pictorially by the vertical grey lines in Fig. 1.[3] Similarly for privacy, at one end (e.g. intimate body search) reasonable expectations of privacy will be such that a warrant will always be required, whereas at the other (in plain view in a public space) there may be little or no expectation of privacy so

---

[3] A similar diagram was actually used in the decision in the Massachussets Trade Secrets case Sissy Harrington-McGill v. Old Mother Hubbard Dog Food, Inc., Middlesex Superior Court C.A. 81-4385.
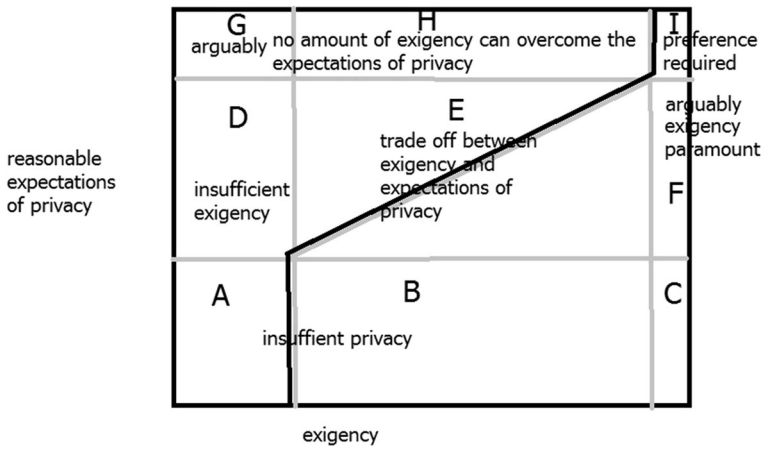
Fig. 1 Automobile exception using two dimensions

that a warrant may never be required. This situation is shown by the horizontal lines of Fig. 1. Putting the two dimensions together, the lines divide the space into 9 two-dimensional areas. In A, D and G, there is insufficient exigency to merit warrantless search, whereas in C and F the exigency is such (perhaps an immediate threat to the life of the President) that no expectations of privacy are sufficient to require a warrant. In A, B and C there are insufficient expectations of privacy to require a warrant given any degree of exigency, while in G and H the expectations of privacy are sufficiently high that no degree of exigency will permit the search. The interesting areas are E and I. In E there is a trade off between the two dimensions, so that some kind of balance must be struck (Lauritsen 2015). In area I, unless there is another dimension to consider, a preference needs to expressed, perhaps based on a preference between values (Bench-Capon and Sartor 2003). The dark line in Fig. 1 thus divides the space into areas where a warrant is required (to the left of the line) and those where it is not (to the right of the line).[4] The line shown in Fig. 1 indicates an even trade off between privacy and exigency (giving the 45^0 line): other trade-offs are possible, which would require a curve rather than a line. It also includes area I in the no-warrant space, indicating a preference for exigency: a different preference would send the line horizontal, excluding I. Note also that trade-off could be possible throughout the range of the dimensions, so that areas A and I are squeezed. The right hand grey line effectively represents the point at which the dimension becomes decisive (subject to preference for a second dimension which is the decisive range for the other side), what (Bruninghaus and Ashley 2003a) call a *knock-out* factor. If both dimensions are such that the area E expands to consume the whole space, we have a situation where the dimensions are never decisive, but require to be balanced across their whole range.

If there is a third dimension the space will become a cube, with 27 partitions, and more dimensions will make visualisation very hard, although mathematicians can

---

[4] Those familiar with machine learning will recognise the similarities with a decision boundary in supervised learning.

happily write equations relating to high dimensional spaces.[5] If we now position the precedents in this dimensional space we can see the problem as being to divide the space into the plaintiff region and the defendant region. In the two dimensional case of Fig. 1, this could be achieved by the black line with the warrant cases on the left and the non-warrant cases on the right. The problem would seem to be the kind of line fitting problem which could be addressed by nearest neighbour or other statistical techniques. But there are two reasons why such techniques are not suitable here: first it is quite clear that this is not the way that people reason with cases, and so the point of the exercise, which was to model lawyer-like reasoning, would not be achieved. But more importantly there are not really sufficient cases to apply the statistical techniques, which depend heavily on having abundant examples. HYPO has 13 dimensions and many of the cases in fact only use some of these, typically no more than 4 or 5. The use of line fitting techniques would require hundreds or even thousands of cases (for a discussion of the required number of cases see Al-Abdulkarim et al. (2015b)). To obtain, let alone perform a dimensional analysis on, that many cases was far beyond what was sensible to ask from a nineteen-eighties PhD project. HYPO has 33 cases in its case base (Ashley 1990). It remains a task few would undertake, even if cases were available in this quantity.[6]

HYPO models argumentation, rather than using a purely statistical approach. Given a set of precedents located in the n-dimensional space, and a new case located in that space, the idea is to find arguments, based on the precedents, for and against a finding for the plaintiff (or defendant). Given a plaintiff precedent P_p and a defendant precedent P_d and a current case C_c, the plaintiff's task is to argue that C_c is on the same side of the line (or its n-dimensional equivalent) as P_p rather than P_d. For example, in Fig. 1, if C_c is South East of P_p, that is a strong argument for the plaintiff: if it is South West (or North East) we have arguments for and against the plaintiff, and if it is North West, it is of no help at all to the plaintiff.

In HYPO the arguments were deployed in a three-ply argumentation structure, as described in Sect. 2.2.

2.2 Three-ply arguments

HYPO deploys its arguments within a three-ply structure. This structure is also found in the presentation of oral arguments in the US Supreme Court (Al-Abdulkarim et al. 2013), but in the Supreme Court the plies take the form of a dialogue between counsel and the Justices, whereas HYPO arguments are presented uninterruptedly (as if by the counsel), rather than in a dialogue. The three plies are:

1. *Cite a Case* One side (say the plaintiff) cites a case found for that side. That case should be as similar (*on-point*) as possible to the current case, and the

---

[5] For example $x^2 + y^2 = r^2$ is a circle of radius r, $x^2 + y^2 + z^2 = r^2$ is a sphere with radius r, $x^2 + y^2 + z^2 + w^2 = r^2$ is a four dimensional "4-sphere" sphere with radius r, and so on.

[6] Where large quantities of cases are available, they tend to be subjected to machine learning techniques rather than manual analysis. A recent example is Aletras et al. (2016), which will be discussed in more detail later (see footnote 12).

suggestion is that the decision in the precedent should be applied to the current case (following the principle of *stare decisis*, which means *let the decision stand*).

2. *Response* The defendant responds to the plaintiff by distinguishing (that is, pointing to significant differences which mean that the precedent should not be followed), and by citing counter examples, precedent cases found for the defendant that are at least as on-point as the case cited by the plaintiff.

3. *Rebuttal* The plaintiff now attempts to rebut the contentions made by the defendant in the second ply: distinguishing the counter examples, emphasising any strengths and attempting to show that any weaknesses are not fatal.

Each of these plies will be explained below.

2.2.1 Citing a case

HYPO measures similarity between cases in terms of four metrics: on pointness (the degree of overlap in terms of dimensions), outcome, the magnitudes of shared dimensions and potential relevance as a near miss. To assess on pointenss, cases are (partially) ordered in a *Claims Lattice*, as shown in Fig. 2. The lattice includes every case which shares at least one dimension with the case being considered. Figure 2 shows the Claims Lattics for *USM Corp v Marson Fastener Corp*.[7] Note that a precedent is considered more on point than another precedent only if the dimensions of the second precedent form a proper subset of the dimensions of the first. Thus *Space Aero v Darling*, which shares three dimensions with *USM*, is not considered more on point than *Automated Systems*, which shares only two, because the dimensions of *Automated Systems* are not a proper subset of *Space Aero v Darling* and so they are on different branches of the lattice. The most on point case, decided for the desired party, and which does not have a more on point case for the other party is the one to cite. Of the three first level cases in Fig. 2, only *Space Aero v Darling* was found for the plaintiff, and so will be cited in the first ply. If there are equally on point cases, several arguments, based on the different cases, could be made in the first ply.

2.2.2 Response

The defendant must now provide a response. First, an attempt to distinguish *Space Aero* should be made. Distinctions can be made by showing that the precedent is stronger for the plaintiff, or by showing that the current case is stronger for the defendant. One way in which this can be done is in terms of different dimensions: if there is a dimension present in one case but not the other, that can provide a reason for saying that the precedent does not govern the current case, although it might still do so if there are also reasons for finding that this distinction is not sufficient. This issue was explored in more depth in CATO, as discussed in Sect. 4.1. To be a good

---

[7] 379 Mass. 90 (1979). This case which concerned the manufacture of blind rivets, turned largely on whether the security measures taken were adequate.
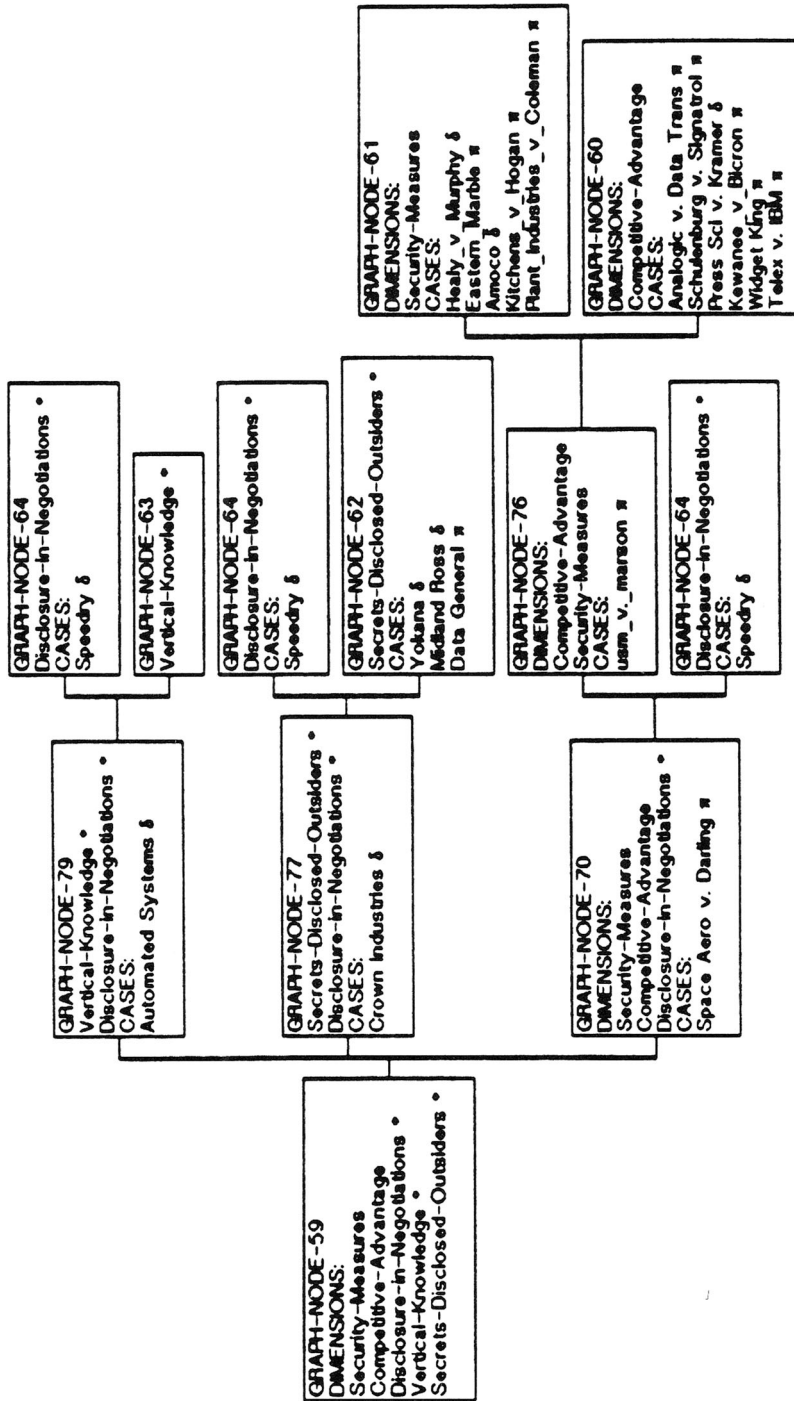
Fig. 2 Claims lattice for *USM*, Figure 8.2 in Ashley ([1990](#))

distinction the different dimension should represent a strength (weakness) for the plaintiff if it is in the precedent (current) case, and a strength (weakness) for the defendant if it is the current (precedent) case. The differences between these two ways of distinguishing is discussed in Bench-Capon (2012). Thus in *USM*, that secrets were disclosed to outsiders in *USM* but not *Space Aero* could be used as a distinction, since it represents a weakness for *USM*. A second way in which a precedent can be distinguished is on a shared dimension. Thus if the information gave a larger competitive advantage to the defendant in *Space Aero* than *USM*, that could be a point of distinction.

Having made the distinctions between the plaintiff's precedent and the current case, the defendant may now cite counter examples of his own. Both *Automated Systems* and *Crown Industries* would be possible counter examples since they are as on point as *Space Aero* and found for the defendant.

2.2.3 Rebuttal

The rebuttal stage gives the plaintiff the opportunity to answer the points made by the defendant in the response. Three types of answer are possible:

- *Distinguish the counter examples* This can be done in the same ways as the defendant distinguished the plaintiff's precedent in the response.
- *Emphasise strengths* If the current case is more favourable to the plaintiff on some dimensions than a precedent which already favours the plaintiff, that is a strength worth emphasising.
- *Show weaknesses not fatal* If the defendant has distinguished a case by showing that the plaintiff is weaker on some shared dimension than a precedent found for the defendant, this can be shown to be not a fatal weakness by finding another precedent which as weak or weaker on that dimension but which was found for the plaintiff. This shows that we are in an area of that dimension which has mixed outcomes, suggesting that it is some other dimension that is decisive.

An important idea in HYPO (the name of the PhD project derived from Rissland's earlier work on hypothetical reasoning) is that if the current case falls between precedents with different outcomes, hypothetical cases can be used, varying the position on the dimension, or by adding or removing a dimension, so that intuitions can suggest exactly where the line should be drawn, and on which side the current case should fall. This aspect of HYPO was perhaps less exploited in the HYPO system in its final form than might have been hoped.

Thus HYPO has a structure within which to deploy the extensive range of arguments made possible by applying its dimensional analysis to the cases. In the next section I will discuss the thirteen dimensions used by HYPO in Ashley (1990).

2.3 Dimensions in HYPO

As mentioned before, HYPO uses thirteen dimensions. But these are not homogeneous. The paradigmatic dimension is continuous, running from the

extreme plaintiff to the extreme defendant location. But some of the dimensions in HYPO are not continuous: rather their range comprises a set of enumerable points. In other examples the dimension degrades to just two points: the extreme plaintiff and the extreme defendant point, and no intermediate points are recognised. Finally the dimension may degrade to a single point: for example if the defendant paid a bribe that is a strong point for the plaintiff, but if no bribe was paid that does not help the defendant's case, rather it means that the dimension does not apply to the case. HYPO's dimensions and their types are listed below:

1. *Competitive Advantage Gained* This dimension combines the development costs in terms of time and in terms of money. It is not entirely clear how these are combined, but the dimension can be regarded as **continuous**. It might have proved convenient to split into two dimensions: time savings and money savings.

2. *Vertical Knowledge* Vertical knowledge is contrasted with technical knowledge, and vertical knowledge favours the defendant. This is a **binary** dimension,

3. *Secrets Voluntarily Disclosed* This is simply the number of people to whom there was voluntary disclosure. It is **continuous**.

4. *Discloses Subject to Restriction* This is the percentage of disclosures which are subject to restriction. As such it is potentially continuous, but in Ashley (1990) only 0% or 100% are recognised and so it is effectively **binary**.

5. *Agreement Supported by Consideration* This refers to a (confidentiality or non-disclosure) agreement between the plaintiff and the defendant, and is a **binary** dimension: the amount of the consideration is not thought important in Ashley (1990), as is usual in law. *Peppercorn rents* are an example where the consideration is purely nominal.

6. *Common Employee Paid to Change Employers* Although this again could in principle be considered continuous, no account is taken of the number of employees concerned, nor of the amount paid, and so it is better regarded as **binary**. In fact, since the dimension does not really apply if there are no common employees, it is effectively **unary**.

7. *Exists Express Noncompetition Agreement* This refers to an agreement between the plaintiff and former employees. It is **binary**.

8. *Common Employee Transferred Product Tools* This is best regarded as **unary**: the dimension does not apply if there are no common employees, or they did not bring tools.

9. *Non-Disclosure Agreement Re Defendant Access* This is an agreement between plaintiff and defendant, and is **binary**.

10. *Common Employee Sole Developer* If the sole developer transferred employer, it may be considered reasonable that he takes his knowledge with him. This is a **binary**, or perhaps even unary, dimension.

11. *Non-disclosure Agreement Specific* This dimension only applies if there is a non-disclosure agreement between the plaintiff and the defendant, and no degrees of specificity are recognised, so that it is **binary**.

12. *Disclosure in Negotiations with Defendant* Such disclosures can weaken the plaintiff's case. It is a **binary** dimension.
13. *Security Measures Adopted* This is an **enumerated** range with eight points, ranging from the extreme pro-defendant point (minimal measures) to the extreme pro-plaintiff point (employee non-disclosure agreements).

After HYPO, Ashley moved to University of Pittsburgh (specifically to the School of Law and Learning, Research and Development Centre). At Pittsburgh, Ashley developed CATO, with Aleven (1997), and SMILE and Issue Based Prediction (IBP) with Ashley and Brüninghaus (2009). Meanwhile Rissland continued to develop the HYPO ideas at Amherst producing CABARET with Skalak and Rissland (1992), BankXX with Skalak and M. Timur Friedman Rissland et al. (1996) and SPIRE with Rissland and Daniels (1996). A major difference between these two strands was their use and understanding of *factors*. As explained in Rissland and Ashley (2002), both groups adopted the more legal sounding term "factors" in preference to the mathematical term "dimensions". But whereas at Amherst, factors were just a different term for HYPO's dimensions, a related but different notion was used in Pittsburgh. The Amherst systems will be discussed in Sect. 3 and the Pittsburgh systems, and the notion of factor as used there, in Sect. 4.

3 Beyond HYPO

Work on reasoning with legal cases was continued at Amherst by Edwina Rissland, working with a succession of PhD students, David Skalak, M. Timur Friedman and Jody Daniels. Two of the resulting systems are described in papers in this special issue, Skalak and Rissland (1992) and Rissland et al. (1996), and the details of these two systems will be left to them. The domains for these systems are the *Home Office Deduction* (an allowance against income tax when part of one's home is used as an office) for CABARET and personal bankruptcy for BankXX and SPIRE.

3.1 CABARET

CABARET was described in conference papers, including Rissland and Skalak (1989b), Rissland and Skalak (1989a) and Skalak and Rissland (1991), before being consolidated in Rissland and Skalak (1991) and Skalak and Rissland (1992) (the very first paper to appear in this journal), which focussed in particular on strategies and argument moves. Because it took as its domain Home Office Deduction, there was a basis for the law in statutes, whereas HYPO had started from case law and its consolidation in the *Restatement of Torts*. This enabled CABARET to take as its focus the interpretation of terms found in the statutes, the idea being that the meaning of such terms is clarified and refined in case law. Thus while the statutes can be represented using rules, much in the manner of Sergot et al. (1986), when "the rules run out" (Gardner 1987), it is necessary to use case based techniques. The process is well explained by Ron Loui in his commentary on Skalak and Rissland (1991) in Bench-Capon et al. (2012). In this way the rules are able to focus the

search and partition the case base so that there are fewer irrelevant distinctions made during the response (and the case base can provide precedents that govern more cases, since differences which relate to terms not at issue are not considered).[8] This use of high level rules to focus and drive the lower level case based reasoning was a significant improvement, also used in IBP (Bruninghaus and Ashley 2003a) and still in use today (Al-Abdulkarim et al. 2016c).

Also CABARET developed a taxonomy of argument moves to represent patterns of actual legal argument, such as straw man and make-weight arguments, and showed how these were supported by the dimensional analysis of precedent cases. This was an important development since it enabled consideration of *strategies* for deploying arguments, an idea later developed through the use of *dialogue games*, such as The Pleadings Game (Gordon 1993), the game of Prakken and Sartor (1996), TDG (Bench-Capon et al. 2000), and PADUA (Wardeh et al. 2009), amongst others.

Both of the aspects introduced in CABARET represent significant developments of HYPO, allowing the dimensional analysis of cases to be employed in the context of rule interpretation, and for arguments with a more natural and familiar form to be used a strategic manner.

3.2 BankXX

BankXX, which operates in the domain of personal bankruptcy (specifically Chapter 13 of the U.S. Bankruptcy Code), was most fully reported in Rissland et al. (1996) (in this issue) and its evaluation described in Rissland et al. (1997), an earlier conference version having appeared as Rissland et al. (1993). BankXX represents a considerable departure from HYPO and CABARET, since it uses the precedents (and other sources) to represent the domain knowledge as a highly interconnected network of building blocks which which is searched to gather *argument pieces*. The nodes in this network encompass a wide variety of ways of representing the domain knowledge including cases as collections of facts, cases as dimensionally-analyzed fact situations, cases as bundles of citations, and cases as prototypical factual scripts, as well as legal theories represented in terms of domain dimensions. Thus cases are represented in several ways. In particular, in its *Domain Factor Space*, cases are represented "by a vector composed of the magnitudes of the case on each dimension that applies to it; non-applicable factors are encoded as NIL. This ... represents a case as a point in an n-dimensional space." Arguments are then formed by performing heuristic search over the network, using evaluation functions at the domain level, the argumentation piece level, and the overall argument level. The result is a highly sophisticated system which can blend the more Boolean nature of CATO's factor-based approach with the more value-oriented nature of HYPO's dimension-based approach. Further, the factors are grouped according to various theories, which, as in CABARET, gives some structure to the set of all factors.

The evaluation in Rissland et al. (1997) is one of the most (if not the most) detailed examples of evaluation in AI and Law. It considers several different forms

---

[8] An alternative approach to reasoning with portions of precedents is offered by Branting (1991).

of the BankXX program, and the evaluation is conducted from several perspectives, and a number of issues relating specifically to the evaluation of programs in the domain of law are noted.

BankXX has been rather unduly neglected[9], and has no obvious descendants in current AI and Law research. This does scant justice to the importance of the work. Construction of cases by performing heuristic search was also carried out by AGATHA (Chorley and Bench-Capon 2005a), but the search tree was over only a collection of cases represented as bundles of CATO-style factors, rather than the highly sophisticated network of knowledge used in BankXX. Perhaps AI and Law should make more use of traditional AI techniques such as heuristic search. It is to be hoped that the appearance of BankXX in this special issue may revive interest in this work, and perhaps lead to more thorough evaluations of work in AI and Law. Evaluation remains an issue in AI and Law (Conrad and Zeleznikow 2015).

3.3 SPIRE

SPIRE Rissland and Daniels (1995), Rissland and Daniels (1996) and Daniels and Rissland (1997), which was applied to both the domain of home office deduction and the domain of personal bankruptcy, used case based reasoning to kick-start (*seed*) an information retrieval system. The idea was that in this way a large case base could be searched, but only a small number of these cases need be given the detailed analysis required for dimensional case based reasoning. From a case based reasoning view point this gives access to very much larger *corpora* of cases, while from an information retrieval point of view it enables the formulation of better queries.

The system operates in two stages (originally developed separately: the name SPIRE was not used in the earlier papers).

– First (Rissland and Daniels 1995, 1996), the analysed cases are used to retrieve documents (case decisions) that are relevant to the presented problem case, and
– Second, within those retrieved documents, passages that contain information relevant to specific case features are highlighted. Here the "case base" (more like an examples base) was a collection of excerpts; one set for each term of interest (such as the *sincerity* of the debtor). Differences with other systems in the HYPO family are that: they are not structured objects like cases, (but only snippets of text; the same set of excerpts is used for all attempts to locate passages and thus, there is no attempt to make the selection of "see" excerpts "relevant" to the case at hand. The idea was then to do something analogous for case retrieval, but this time on an individual document and to locate relevant passages within it.

These two stages were put together as SPIRE (Daniels and Rissland 1997). For SPIRE, the input was a fact situation and the output was a set of relevant passages in

---

[9] As of March 2017, (Rissland et al. 1996) had received only 83 citations on Google Scholar, and the evaluation in Rissland et al. (1997) a mere 19.

each of the most relevant cases. It would then be possible to use these passages to query the main case base.

The engine used to query the main case base was INQUERY (Callan et al. 1992), developed by the University of Massachusetts Information Retrieval group in the Center for Intelligent Information Retrieval (CIIR). The use of HYPO is in selecting the seed cases: the top two layers of the case lattice proved to be a good choice for determining the cases to use.

Several different queries were tried , including bag of words, set of words and what was termed a "sum" query which required terms from within an excerpt to be found co-occurring in the passage, and some weighting methods. Among the results was that bag of words performed better than set of words, and bag of words and sum performed better than the other query types tried.

Although an interesting use of legal cased based reasoning, and with some encouraging results, SPIRE was not widely influential. This may be because Daniels went on to have a successful career in the US Army after finishing at Amherst, but in any case information retrieval is a hard field for research to make a lasting impact in because of the rapid development of technology. The tools now available for information retrieval are such as could only have been dreamed of in the mid-90s when the SPIRE research was carried out (which was before Google transformed the way we access information).

4 Doing things with factors

Following HYPO, Kevin Ashley moved to Pittsburgh in 1989, and continued to work on reasoning with legal cases, using many of the ideas developed in HYPO. Significantly, Ashley moved specifically to the Learning and Development Research Center. This connection with the Learning Center had a significant influence on Ashley's projects, in that the task focus became legal education. Also the position gave access to real classes of legal students for evaluation of these systems. Ashley took full advantage of this opportunity and his work at Pittsburgh is marked by a focus on empirical evaluation and the use of carefully designed experiments to produce empirical results. His student, Vincent Aleven, with whom he developed CATO (first mentioned as such in Aleven and Ashley (1993)), entitled his PhD *Teaching case-based argumentation through a model and examples* (Aleven 1997), and it made extensive use of law students in its evaluation. The key systems we will look at in this introduction, because of their influence on AI and Law, are CATO and IBP (Bruninghaus and Ashley (2003a) and Ashley and Brüninghaus (2009)). Other projects carried out by Ashley at this time included Sirocco (McLaren and Ashley 1999) with Bruce McLaren.

4.1 CATO

The task CATO was intended to address was to support teaching law students to distinguish cases. Not every difference between a case and a precedent represents a usable distinction. If we think think in two dimensions, as in Fig. 1, the plaintiff

"owns" the area to the south-east of the precedent. Therefore a distinction has to place the current case to the north or west of the precedent since otherwise the difference actually strengthens the case, and the current case can be decided in the same way as the precedent using an *a fortiori* argument. Even if the distinction is usable, however, it may be that there are other differences between the current case and the precedent which can be used to compensate for or cancel out the identified distinction. Such a difference enables the distinction to be *downplayed*. If a distinction cannot be downplayed it may be considered significant, and so can be emphasised. Thus the task of CATO was to teach students:

– which differences between a case and a precedent can be used as distinctions,
– which distinctions can be downplayed,
– how to downplay distinctions.

4.1.1 Factors in CATO

A major difference between HYPO and CATO was the use of *factors* instead of dimensions. In HYPO cases had been represented in terms of their facts, and the facts used to determine whether the dimension was active in a particular case and, if active, where on the dimension the case lay. Factors are also related to facts, but are attributed by the analyst on the basis of the case facts, so that factors are simply present in or absent from a case and cases can be represented as bundles of factors. Thus in CATO facts are seen only by the analyst. Another important feature of factors, is that the presence of a factor always favours the same side: a factor is *either* pro-plaintiff *or* pro-defendant. In sum factors can be considered as *stereotypical patterns of fact that favour one party or the other*. While this is something of a simplification when compared to dimensions, it does not affect the task being taught: the difference between cases can be considered in terms of differences between the factors present in the cases: the process of attributing factors is not important to this particular task.

Because HYPO and CATO both use the same domain, US Trade Secrets, direct comparison can be made between CATO's factors and HYPO's dimensions. CATO used 26 factors (F1–F27—there was no F9), as shown in Table 1. The Table also shows any relation of the factors to HYPO dimensions: Dimensions are identified using the numbers in Sect. 2.3.

All but one of the dimensions (which relates to contract law rather than trade secrets) in HYPO correspond to one or more factors. In addition there are a number of factors which are unrelated to HYPO dimensions: these mostly relate to legitimate and questionable means of discovering the information. CATO analysed more cases (148 as against 33) and so it is to be expected that new factors or dimensions will emerge from the analysis. In fact if we replace D2 with a dimension *Questionable Means*, and introduce a new dimension *Legitimate Means* we can locate most of these new factors on these dimensions. Or we could see these as a single dimension with F22 as the extreme pro-plaintiff point and F24 as the extreme pro-plaintiff point. F15 and F18 are harder to locate on this dimension, and so a

Table 1 Base level factors in CATO

ID	Factor	HYPO dimension
F1	DisclosureInNegotiations (d)	Defendant end of D12
F2	BribeEmployee (p)	D6
F3	EmployeeSoleDeveloper (d)	D10
F4	AgreedNotToDisclose (p)	Plaintiff end of D9
F5	AgreementNotSpecific (d)	Defendant end of D11
F6	SecurityMeasures (p)	Any point of D13 except extreme defendant end
F7	BroughtTools (p)	D8
F8	CompetitiveAdvantage (p)	Plaintiff end of D1
F10	SecretsDisclosedOutsiders (d)	Any point of D3 except extreme plaintiff end
F11	VerticalKnowledge (d)	Defendant end of D2
F12	OutsiderDisclosuresRestricted (p)	Plaintiff end of D4
F13	NoncompetitionAgreement (p)	Plaintiff end of D7
F14	RestrictedMaterialsUsed (p)	Not in HYPO
F15	UniqueProduct (p)	Not in HYPO
F16	InfoReverseEngineerable (d)	Not in HYPO
F17	InfoIndependentlyGenerated (d)	Not in HYPO
F18	IdenticalProducts (p)	Not in HYPO
F19	NoSecurityMeasures (d)	Defendant end of D13
F20	InfoKnownToCompetitors (d)	Not in HYPO
F21	KnewInfoConfidential (p)	Not in HYPO
F22	InvasiveTechniques (p)	Not in HYPO
F23	WaiverOfConfidentiality (d)	Not in HYPO
F24	InfoObtainableElsewhere (d)	Not in HYPO
F25	InfoReverseEngineered (d)	Not in HYPO
F26	Deception (p)	Not in HYPO
F27	DisclosureInPublicForum (d)	Point on D3 towards defendant end

second new dimension may be required for them. The important thing is that there is a mapping from dimension points to factors, which would allow the HYPO cases to be expressed as factors (the missing dimension D5 is used only to invalidate agreements). In general it is possible to see factors as points (or ranges) on HYPO dimensions. But this does mean that the *analyst* decides at what point the dimension ceases to favour the plaintiff and begins to favour the defendant, rather than this question forming part of the argumentation.

Representing cases as bundles of factors make it very easy to discover differences between cases. As noted in Wyner and Bench-Capon (2007) this means that when comparing cases the factors can be placed in one of seven partitions:

A Plaintiff factors in both current case and precedent.
B Defendant factors in both current case and precedent.
C Plaintiff factors in current case not in precedent.
D Defendant factors in current not in precedent case.

E   Defendant factors in precedent case not in current case.
F   Plaintiff factors in precedent not in current case.
G   Factors (both plaintiff and defendant) not in either current case or precedent

This partitioning was first used in Allen et al. (2000) from which Fig. 3 is taken.

(A) and (B) represent what is similar in the cases. For the precedent case to serve as a precedent for current case there must be at least one factor in at least one of these partitions. (C) and (D) represent aspects in which current case is stronger for the plaintiff than the precedent, and so cannot be used to distinguish the precedent. (E) and (F) represent aspects in which current case is weaker for the plaintiff than the precedent, and so can be used to distinguish the precedent. (G) contains the factors which are not relevant in the comparison. Thus the law students will be advised to consider only (E) and (F) when looking for distinguishing factors that the defendant can use.

4.1.2 Downplaying and emphasising distinctions

Having identified which factors could serve as a distinction, the question arises as to which will be the basis of good distinctions. The answer is provided by CATO's second innovation: the introduction of abstract factors, and the organisation of factors into a factor hierarchy. The factor hierarchies of CATO are shown in Fig. 4. Note that the plural is used because the root of a factor hierarchy is a legal issue, and
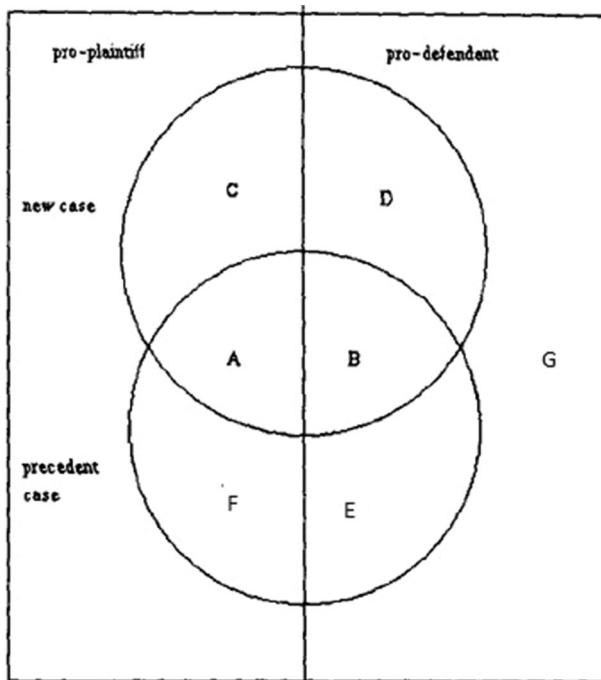

Fig. 3 Partioning CATO factors from Allen et al. (2000)

CATO recognises five issues giving rise to five hierarchies. The hierarchies may have factors in common, but the issues are kept separate. It is important to recognise that these hierarchies are not is-a hierarchies. Like base level factors, abstract factors are also either present of absent in a case, and always favour the same side. Their children are *reasons* for their presence (if they favour the same side), or absence (if they favour different sides). Where there are reasons for and against the conflict must be resolved: sometimes the resolution is obvious: for example a waiver of confidentiality (F26) will cancel the factor F4 (Agreed not to disclose) so that the abstract factor F121 (Express confidentiality agreement) is not present. In other cases, such as F111 (Questionable Means), if both F26 (Deception) and F25 (Information Reverse Engineered) are present, it may be difficult to determine whether the abstract factor is present or not. Indeed there may be no general answer, and the particular case facts may need to be considered.

Thus a consideration of the factor hierarchy can tell us the significance of a distinction. If the "missing" factor has a sibling present in the case this can be used as an alternative way of establishing the presence of the abstract factor which represents a strength (when they favour the same side) or, when they favour different sides, as the reason for the absence of an abstract factor which would otherwise represent a weakness. This is how we can downplay a distinction. On the other hand if any additional factors are no better than cousins, or even pertain to an entirely different issue, the distinction is significant and can be emphasised. These different ways of making and challenging distinctions were expressed as argumentation schemes in Wyner and Bench-Capon (2007).

4.1.3 Argument moves

A third influential feature of CATO was its identification of a set of *argument moves*, several based on various operations carried out in HYPO. Although other sets of argument moves have been proposed, such as those used in CABARET (Skalak and Rissland 1992), the CATO set has proved enduringly useful, especially when attempting to identify a procedure, such as Bench-Capon (1997) or when devising a dialogue game to undertake the reasoning (Wardeh et al. 2009). The argument moves used in CATO (as listed in Aleven (1997), Figure 2.2) are:

- Analogising a case to a past case with a favourable outcome;
- Distinguishing a case with an unfavourable outcome;
- Downplaying the significance of a decision;
- Emphasising the significance of a distinction;
- Citing a favourable case to emphasise strengths;
- Citing a favourable case to argue that weaknesses are not fatal;
- Citing a more on point counterexample to a case cited by an opponent;
- Citing an as on point counter example to a case cited by an opponent.

Not all of these moves are new in CATO: some can be found in HYPO. But the third and the fourth are new - made available by the factor hierarchy - and it is this list
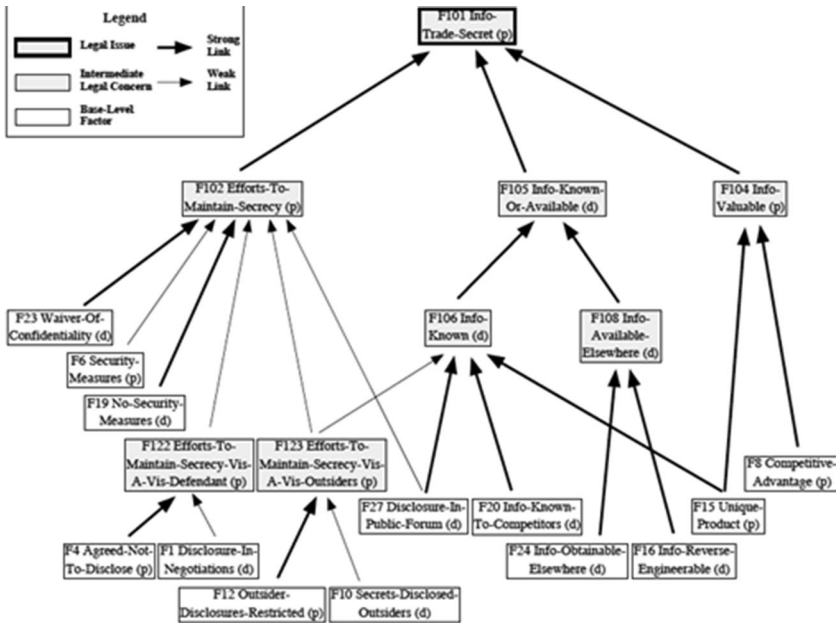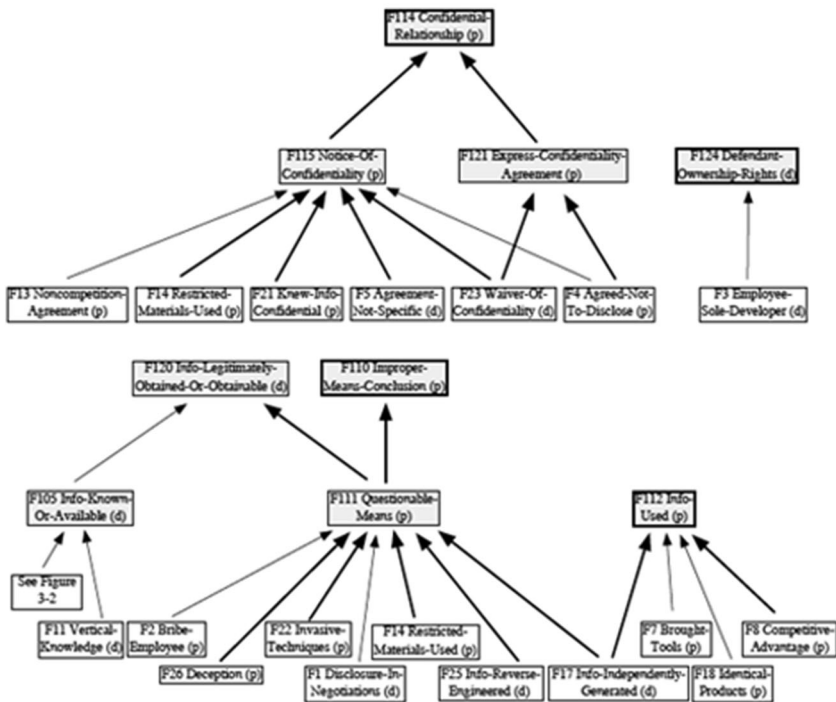
Figure 3-2: CATO's Factor Hierarchy, part 1

Figure 3-3: CATO's Factor Hierarchy, part 2

Fig. 4 CATO abstract factor hierarchy from Aleven (1997)

that is usually cited in the later work mentioned above. The moves are also important in identifying the critical questions (Walton 1996) posed when using the argumentation schemes in Wyner and Bench-Capon (2007).

4.1.4 CATO summary

CATO has been extremely influential on subsequent developments in AI and Law. Even when HYPO is cited as the key influence, the actual approach taken often owes as much to CATO as to HYPO. All three aspects introduced in CATO have had influence:

– CATO style factors, and the representation of cases as bundles of binary factors have been the predominant form of representation of cases in work such as dialogue (Prakken and Sartor 1998), theory construction (Bench-Capon and Sartor 2003) and argumentation (Wyner and Bench-Capon 2007) and are the starting point for formalisations (Horty and Bench-Capon 2012).
– The factor hierarchy and its introduction of intermediate predicates have been found useful in works such as Bruninghaus and Ashley (2003b), Atkinson and Bench-Capon (2005), Lindahl and Odelstad (2008) and Grabmair and Ashley (2011). The idea has also been used in work on methodologies for representing cases such as Al-Abdulkarim et al. (2016c).
– Argument moves, in particular the set identified by CATO, have also influenced AI and Law work on dialogues for reasoning about cases (Wardeh et al. 2009).

Also heavily influenced by CATO, was the work of Ashley with another student, Stefanie Brüninghaus, which continued the exploration of reasoning with Trade Secrets law cases. This will be discussed in the next section.

4.2 IBP

Brüninghaus worked with Ashley on two projects, SMILE (*SMart Index Learner*) and IBP (*Issue-Based Prediction*) both of which are described in a paper included in this special issue (Ashley and Brüninghaus 2009). The two programs were supposed to act at different ends of the case based reasoning process: SMILE would take cases input as natural language and identify which factors were present, and then IBP would deploy CATO style reasoning, adapted for prediction rather than teaching, to predict the outcome of the case. Empirical evaluations are also reported in Ashley and Brüninghaus (2009), showing a strong performance from IBP, but a rather weaker performance from SMILE. Here we will concentrate on IBP, since SMILE draws mainly on classification techniques rather than the case based reasoning techniques we are focusing on here, albeit that SMILE was attempting to ascribe CATO's factors to cases.

4.2.1 Logical model

The key modification necessary for converting CATO into a program capable of predicting case outcomes is to combine the five issue hierarchies into a single structure. As noted in Sect. 4.1.2, and as is evident from Fig. 4, CATO uses five distinct hierarchies. IBP organises these using what it terms a logical model, taken from the *Restatement of Torts*. This logical model is shown in Fig. 5. The logical model serves the same purpose as the top layers of rules in CABARET (Skalak and Rissland 1992). Essentially the idea is that the overall question of whether a Trade Secret was misappropriated requires both that the information was a trade secret and that it was misappropriated. To be considered a secret, it needs to be shown both that the information was valuable and that the owner had made efforts to maintain its secrecy. Misappropriation could either be shown by the use of improper means to obtain the information, or by the breach of a confidential relationship in using the information. Thus the model has five leaf nodes, corresponding to the five hierarchies in CATO.

The 26 base level factors of CATO are thus split over the five issues, as indicated in Table 2. Notice that several factors relate to two issues, and one, F3, is not used in IBP.

IBP will now use the factors to decide which party if favoured on each of the issues, and then put these into the logical model to predict the outcome. The precise algorithm is given in Fig. 2 of Ashley and Brüninghaus (2009) in this issue, and is fully described there, so no details are required in this introduction. One additional feature introduced by IBP is that of *strength* of factors. *Knock Out Factors* are those "representing behavior paradigmatically proscribed or encouraged under trade secret law and for which the probability that a side wins when the Factor applies is at least 80% greater than the baseline probability of the side's winning." Conversely *Weak Factors* are those "for which the probability of the favored side's winning, given that one knows the Factor applies, is less than 20% over the baseline probability of the side's winning." Weak factors appear to act only as supports for other factors, and IBP will not consider an issue raised if only weak factors are present. Given that factors can be related to points on HYPO style dimensions, it is not surprising that differences in strength need to be considered if we are resolving issues: the use of Knock Out and Weak factors allows this while maintaining much
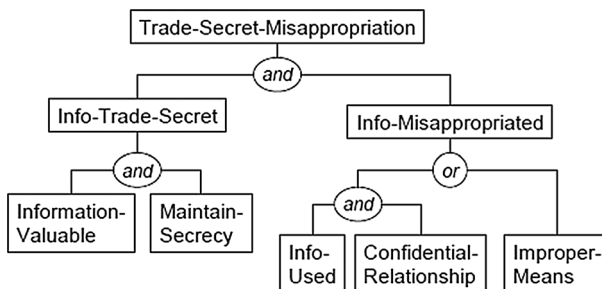

Fig. 5 IBP logical model from Bruninghaus and Ashley (2003a)

Table 2 Mapping of base level factors to issues in IBP

ID	Factor	CATO issues
F1	DisclosureInNegotiations (d)	Confidential relationship
F2	BribeEmployee (p)	Improper means
F3	EmployeeSoleDeveloper (d)	
F4	AgreedNotToDisclose (p)	Maintain secrecy, confidential relationship
F5	AgreementNotSpecific (d)	Defendant end of D11
F6	SecurityMeasures (p)	Maintain secrecy
F7	BroughtTools (p)	Information used, improper means
F8	CompetitiveAdvantage (p)	Information valuable, information used
F10	SecretsDisclosedOutsiders (d)	Maintain secrecy
F11	VerticalKnowledge (d)	Defendant end of D2
F12	OutsiderDisclosuresRestricted (p)	Maintain secrecy
F13	NoncompetitionAgreement (p)	Confidential relationship
F14	RestrictedMaterialsUsed (p)	Information used, improper means
F15	UniqueProduct (p)	Information valuable
F16	InfoReverseEngineerable (d)	Information valuable
F17	InfoIndependentlyGenerated (d)	Information used, improper means
F18	IdenticalProducts (p)	Information used
F19	NoSecurityMeasures (d)	Maintain secrecy
F20	InfoKnownToCompetitors (d)	Information valuable
F21	KnewInfoConfidential (p)	Confidential relationship
F22	InvasiveTechniques (p)	Improper means
F23	WaiverOfConfidentiality (d)	Confidential relationship
F24	InfoObtainableElsewhere (d)	Information valuable
F25	InfoReverseEngineered (d)	Information used, improper means
F26	Deception (p)	Improper means
F27	DisclosureInPublicForum (d)	Information valuable

of the attractive simplicity of factors as compared with dimensions. The need to consider different strengths of factors remains a current concern (Al-Abdulkarim et al. 2016b).

4.3 Empirical results for IBP

Another feature of IBP was that it was given a thorough empirical evaluation based on 184 cases (the 148 used in CATO together with another 36 analysed specifically for IBP). This evaluation was undertaken using versions of HYPO and CATO, a version of IBP with just the logical model, and several different machine learning approaches in addition to IBP itself. IBP performed best with a 91.8% success rate, and only a single abstention. The next best performer was a Rule Learning program which achieved 88%.[10] The version of CATO achieved 77% (with 22 abstentions),

---

[10] This program was not part of the evaluation in Al-Abdulkarimet al. (2015a) in which Naive Bayes with 86.5% was the second best performer.

and HYPO 67.9% with 50 abstentions. This is, of course, a little unfair on CATO and HYPO, since they were not designed to predict outcomes, and the abstentions count heavily against them. In fact, HYPO gets 93.3% of the cases for which it offers an opinion correct.

These results have been used as a basis for comparison when evaluating later programs, including AGATHA (Chorley and Bench-Capon 2005a), a program based on the approach of Chorley and Bench-Capon (2005b) in this issue, and ANGELIC (Al-Abdulkarim et al. 2015a), a project which produced a methodology for encapsulating case based domains, which was applied to several domains, including Trade Secrets. These programs were forced to use only a subset of the cases, since the majority of the case analyses have unfortunately not been made available to researchers, but evaluations showed them able to achieve 90+% on these cases. The results from Al-Abdulkarim et al. (2015a) are shown in Table 3. An earlier program based on neural networks (Bench-Capon 1993) was able to reach 98+%, on a different set of cases.[11] The full set of results for the IBP project are given as Table 3 of Ashley and Brüninghaus (2009). These results provide a useful benchmark for other programs attempting the task, and a success rate of 90% would seem a sensible aspiration.[12]

5 Integration with rule based reasoning

In the 1980s and early 1990s Case Based and Rule Based systems were sometimes seen as competing approaches (although others, including both Gardner and Rissland, recognised that statutory interpretation (at least in the common law traditions) would require both rules and cases). Discussion of the two approaches was held at two of the first three ICAIL conferences, in 1987, and 1991, which was chaired by Don Berman (Berman 1991). Berman's view seemed to be that the two approaches had different aspirations: rule based systems were more likely to

---

[11] This dataset was artificially generated, allowing for the use of a very large number (1500) cases. The dataset has been subsequently reused in Johnston and Governatori (2003), Možina et al. (2005), Wardeh et al. (2009) and elsewhere.

[12] Although just 79% was considered enough for the work reported in Aletras et al. (2016) to make a stir in the national press and radio in the UK in October 2016: e.g The Guardian at www.theguardian.com/technology/2016/oct/24/artificial-intelligence-judge-university-college-london-computer-scientists and the BBC, www.bbc.co.uk/news/technology-37727387 This perhaps a reflection of the upsurge in interest in applications of AI and Law in the UK. See for example the 2016 Law Society of England and Wales panel at www.youtube.com/watch?v=8jPB-4Y3jLg and a special edition of the BBC Radio 4 programme Law in Action at www.bbc.co.uk/programmes/b07dlxmj. The program in Aletras et al. (2016) took the text of published judgements from the European Court of Human Rights (i.e material produced ex post) acting as proxies for the unavailable applications and briefs. As the authors observe, ex ante documents would be the input to an actual prediction system. The system produced a binary classification into violation and no violation. The task in Aletras et al. (2016) is thus more like that of of SMILE+IBP (although SMILE was looking for individual factors rather than the outcome) which, because of the difficulty of the SMILE task compared with the IBP task, performs much closer to 70%. SMILE. moreover, worked on squibs, summaries of the cases, and although these were taken from, and used the language of, the decision, did not contain phrases related to to the outcome of the case. All websites accessed March 2017.

Table 3 Results, including some from Bruninghaus and Ashley (2003a), Chorley and Bench-Capon (2005a) and Al-Abdulkarim et al. (2016c)

	Correct	Error	Abstain	Accuracy	No-abst
ANGELIC secrets II	31	1	0	96.8	96.8
ANGELIC refined	31	1	0	96.8	96.8
AGATHA brute force	30	2	0	93.7	93.7
IBP	170	15	1	91.4	91.9
AGATHA A*	29	3	0	90.6	90.6
ANGELIC secrets	29	3	0	90.6	90.6
CATO-coin	163	30	0	89.0	89.0
Naive Bayes	161	25	0	86.5	86.5
HYPO-coin	152	34	0	81.7	81.7
ANGELIC CATO	25	7	0	78.1	78.1
CATO	152	19	22	77.8	88.8
HYPO	127	9	50	68.3	93.4
IBP-model	99	15	38	72.6	86.8

produce practical applications, with rules derived either from statutes as in Sergot et al. (1986) or from an expert as in Smith and Deedman (1987) in which J.C. Smith expressed his understanding as a set of rules. Berman, however, saw case based reasoning as more capable of capturing legal reasoning: "A major goal of pure AI is to represent accurately human intelligence and, therefore, to represent legal thought. CBR research in the legal domain must continue" (Berman 1991). In general Rule-based systems were associated with Europeans, because of the widespread use of Prolog and logical representations of legislation, inspired by the success of Sergot et al. (1986) and the prevalence of Civil Law in continental Europe, while Case-based systems were seen as the approach of choice of US researchers, with their preference for LISP and the presence of a highly adversarial common law tradition. In the late eighties the two were seen (perhaps especially in Europe) as separate approaches and research tended to be conducted by different groups. However, exposure to each others approaches through ICAIL and elsewhere led to a greater interest in and understanding of the position of the "other side". Increasingly in the 1990s practitioners of the different approaches were now recommending integration, with both approaches required to produce a complete system. Thus in CABARET (Skalak and Rissland 1992) we can see rules used to provide a top level structure in which CBR can be deployed, with an agenda mechanism to control the processing of two co-equal CBR and RBR reasoners. CABARET used observations and control rules to post and order tasks on the agenda: control in CABARET is best discussed in Rissland and Skalak (1991). Within the rule-based approaches, a requirement for cases to provide sufficient conditions to enable interpretation of the terms of legislation was recognised in the work of proponents of rule based systems such as Bench-Capon (1991).[13] A very real advance in allowing the approaches to

---

[13] In many ways this is just a return to the very early approach of Gardner (1987) which began with rules, but which turned to cases when "the rules ran out".

be integrated was, however, made by Prakken and Sartor (1998) which provided an elegant way of moving from cases to a set of rules. This paper will be discussed in the next section.

5.1 Central idea of Prakken and Sartor (1998)

A number of different issues are addressed in Prakken and Sartor (1998), but here I will focus on its most influential idea, the movement from precedents represented as bundles of factors to sets of rules. The main aim of the translation was to enable the precedent to be used in a formal dialogue game, but the rules can equally be used as the basis of an executable logic program. Once rewritten as rules, the set of precedents are available in a form convenient for use in argumentation generally, as knowledge bases, and as the basis of formalisations of the reasoning involved.

The starting point is a set of cases represented as sets of factors together with their outcome. The factors can be divided into pro-plaintiff factors and pro-defendant factors, with each factor representing a reason to decide the case for the side they favour. It is assumed that an additional factor for a party will always strengthen the reasons to decide for that party. This may not be true in general (see Prakken (2005)), but is true of CATO, and can be imposed as a constraint on what is acceptable as a factor. Now the strongest pro-plaintiff reason will be the set of all the pro-plaintiff factors present in the case (F_p) and the strongest pro-defendant reason is the set of all the pro-defendant factors present in the case (F_d). We can express this as two rules:

r1: $F_p \rightarrow p$
r2: $F_d \rightarrow d$

where p and d represent decisions for the plaintiff and defendant respectively. Now the outcome of the case will indicate which of these reasons was preferred, and so we add a rule expressing a priority between these two rules, so that

r3: $r1 \succ r2$

represents a preference for F_p over F_d, and holds if the plaintiff won the case. Each of the precedent cases can be represent as three rules in this way, so that the whole case base can be rewritten as a set of rules to find for the plaintiff, a set of rules to find for the defendant, and an incomplete set of preferences between them. The result is a knowledge base with many similarities to the Reason Based Logic of Hage Hage (1993) and Hage (1996), further developed with his student Bart Verheij (Verheij 1995) and Verheij et al. (1998), but the timing and simplicity of the technique of Prakken and Sartor (1998), together with its strong connection to case based reasoning systems, meant that it made a more lasting impression on the AI and Law community.

The major limitation with the resulting knowledge base is that many possible rule conflicts will not be resolved by the priorities derived from the precedents. As noted in Al-Abdulkarim et al. (2015b), the 26 factors of CATO, split evenly as they are between plaintiff and defendant, give rise to more than 8000 sets of factors

favouring each of the sides, and in excess of 67 million potential comparisons. Even allowing for subsumption of proper subsets, the number of required priorities to ensure that a conflict can be resolved with precedential authority will exceed the cases available (and most probably all the cases that have ever been decided). There is, however, a desire to go beyond what is contained in the set of precedent cases, so that we can say something about cases which are not decidable *a fortiori* with respect to the precedents. In Prakken and Sartor (1998) this is achieved through *rule broadening*, allowing one or more factors to be removed from the antecedent, but then maintaining the priority of the original rule as established in the case from which it derived. The problem is that this lacks justification, unless or until endorsed by the judges in an actual case. Essentially, if the precedent does not provide a definitive answer, it can be distinguished by pointing to a factor missing from the plaintiff rule, or an additional factor in the rule in the precedent, and we need to rely on the user to say whether the distinction blocks the argument or not. Broadening, and several kinds of distinction, are fully discussed in Prakken and Sartor (1998). Although this is a difficult problem, addressed differently in Horty and Bench-Capon (2012), discussed in the following subsubsection, and in the value based theory construction approaches described in Sect. 6, the techniques of Prakken and Sartor (1998) at least enable a clean expression of the problem and potential solutions to it.

5.2 Formalisation of precedential reasoning

Although (Prakken and Sartor 1998) provides a formal dialogue game which can be regarded as a formalization of the aspects of legal theories on judicial reasoning discussed in that paper, it is very much tied to the dialogical context, and so to a process of argumentation. Horty's main aim was to answer the question *How is it, exactly, that precedents constrain future decisions?*, and attempts to give a logic of precedent. Horty pursued this through a series of papers beginning with Horty (1999), moving through Horty (2004), Horty (2011a) and Horty (2011b) to its culmination in the paper included in this special issue, Horty and Bench-Capon (2012). Horty is motivated by a desire to counter the arguments of Alexander (1989). As Horty explains in Horty (2004), Alexander (who is working in Law rather than AI and Law) had identified three different models of precedent:

- the *natural model* by which "a precedent decision might figure into the reasoning of a court in its attempt to reach the correct decision in a current case; but on the natural model, this is the extent of precedential constraint" (p. 19);
- the *rule model* by which precedents include a "rule that carries the precedential constraint. Constraint by precedent is just constraint by rules; a constrained court must apply the rules of precedent cases in reaching current decisions ... There is no room for narrowing the rule, or distinguishing the current case from the precedent;" (p. 20);
- the *result model* by which "a precedent controls all and only *a fortiori* cases - that is, all and only those cases that are as least as strong for the winning side of the precedent as the precedent case itself." (p. 20).

In Horty ([2004](#)) Horty wishes to defend the result model against the rule model advocated by Alexander. Both Horty and Alexander reject the natural model. To defend the result model, Horty draws extensively on AI and Law research, citing amongst others, Ashley ([1989](#)) and Ashley ([1990](#)), Aleven ([1997](#)) and Prakken and Sartor ([1998](#)).

The main idea of Horty is to represent precedents in rules, in much the same way as Prakken and Sartor ([1998](#)), but with one very important difference. Suppose the case was decided for the plaintiff. Now the defendant's reason, in both Prakken and Sartor ([1998](#)) and Horty and Bench-Capon ([2012](#)) will be the strongest possible, that is the set of all pro-defendant factors present in the case, but the plaintiff's reason is treated differently. Whereas in Prakken and Sartor ([1998](#)) it is the strongest pro-plaintiff reason, that is all the pro-plaintiff factors present in the case, Horty reasons that a weaker reason may have been enough to defeat the defendant. Thus the plaintiff rule in Horty and Bench-Capon ([2012](#)) can be a subset of the factors present in the case favouring the *winner*. The makes the resulting rule stronger, because it constrains future cases to a greater extent. The effect is the same as rule broadening, but is is done at the theory level rather than the case level, and is supposed to apply to all the cases in the case base. The situation is illustrated using the pictorial notation of Bench-Capon ([1999](#)) (which followed Prakken and Sartor ([1998](#)) in using narrow preferences) in Fig. [6](#).

The figure shows a Factor Lattice for a domain with six factors: three pro-plaintiff and three pro-defendant. These are grouped into two lattices one pro-plaintiff and one pro-defendant. These contain every possible combination of factors favouring the appropriate party, arranged in a partial order embodying the assumption that more factors beat fewer factors. A precedent provides a link between one node from each lattice, and so enables comparison between sets of factors from the two different lattices. Thus a narrow interpretation of the precedent with all six factors decided for the plaintiff, (shown in Fig. [6](#)) will indicate a preference for {P1, P2, P3} over {D1 D2 D3}. This will not constrain any of the subsets. There are, however, several broad interpretations possible in the manner of Horty and Bench-Capon ([2012](#)): any of the two member or single member sets could be said to be preferred to {D1 D2 D3}, one of which, {P3}, is shown in Fig. [6](#). This means that {P3} is preferred to every combination of pro-defendant factors (since {D1 D2 D3} is preferred to every subset). Now only three pro-plaintiff subsets ({P1, P2}, {P1} and {P2}) are not constrained, and so await future decisions. For example a case with factors {P1, P2, D1} would, if decided for the defendant, now fully determine all subsequent cases. This would incidentally mean that {P3} was preferred to {P1, P2}, but this is not significant, since these subsets are never competing.

With these mechanisms Horty can formally characterise what it is for a case base with the given (broad or less broad) interpretations to be consistent. Now cases must be decided so as to maintain consistency of the case base: this may constrain the outcome, or just how broad the rule taken for the winner of the case can be. The open question is how the subset of pro-winner factors is to be determined. Horty does not specify this, although one suggestion might be that it should be the *ratio*
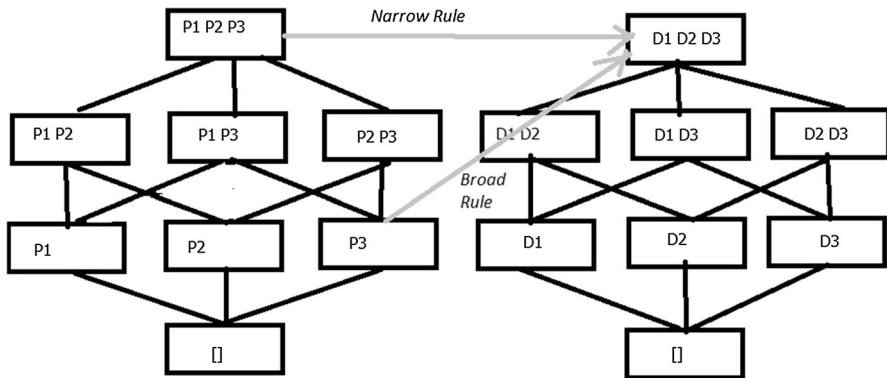
Fig. 6 Factor lattice for domain with 3 pro-P and 3 pro-D factors. Given a precedent with all six factors present decided for the plaintiff, the plaintiff rule can be given the narrow interpretation of Prakken and Sartor (1998), or several broad interpretations as in Horty and Bench-Capon (2012), one of which is shown

deciendi of the case (see Branting (1993)), which can often be discovered by an examination of the text of the decision. Horty also envisages the subset being determined when the precedent is decided. A possible alternative might be that the theory is built anew for each new case, the constraint being that a consistent theory explaining the previous precedents would need to be produced. This might well, especially when there are relatively few precedents available, lead to some considerable revision of the rules to be taken from the precedent cases, but as the number of precedents increases the scope to revise rules while maintaining consistency with cases already decided will decrease. This view of reasoning with cases has some strong similarities with the theory construction approaches found in McCarty (1995) and Bench-Capon and Sartor (2003), and with the notion of a life cycle of case law found in Levi (1948) in which a period of turbulence is followed by a period of stability.

Horty is the starting point for Rigoni (2015), who likes the rule based representation of cases, especially because it is able to maintain the venerable distinction between *ratio deciendi* and *obiter dicta*. Rigoni associates Horty's broadened rules with *ratio deciendi*. Rigoni, however, offers a number of improvements to Horty's account. First he allows for multiple rules to be associated with a single precedent in order to handle *over-determined cases*. This would also allow the incorporation of *obiter dicta*. Second he recognises that not all precedents serve the same purpose: identifying a class of precedents he calls *framework cases*. His example of a framework precedent is *Lemon v. Kurtzman*.[14] Rigoni summarises the case:

> In that case the US Supreme Court addressed the question of whether Pennsylvania's and Rhode Island's statutes that provided money to religious primary schools subject to state oversight violated the Establishment Clause of

---

[14] Lemon v. Kurtzman 403 U.S. 602 (1971).

the First Amendment. The court introduced a three-pronged test and ultimately ruled that both programs did violate the Establishment Clause.

Rigoni's insight is that not all precedents express preferences: some rather supply tests which provide a framework for deciding the cases, in a manner similar to the use of the statute in CABARET and the Restatement of Torts in CATO. The three pronged test can thus be seen as three issues to be considered in *Lemon*'s domain. Rigoni provides a way of accommodating framework cases in a Horty-style formalisation. For Rigoni's approach, the reader is referred to Rigoni's paper in this special issue.

It is a feature of HYPO and CATO that all the precedent cases play the same role: to express a preference or, in HYPO, to constrain the n-dimensional space in some way. Thus cases are analysed against a framework of dimensions and factors developed with respect to the whole set of cases. In practice, however, the dimensions, factors and issues have developed over time, and some of the cases that introduced a factor, issue or dimension may well be in the case base. For example, if we consider the cases related to the Automobile Exception to the Fourth Amendment described in Rissland ([1989](#)), Bench-Capon ([2011](#)) and Al-Abdulkarim et al. ([2016c](#)), we see that no mention was made of privacy issues in the case which is seen as introducing the exception, *Carroll v US*.[15] The need to consider privacy is stressed in *South Dakota v. Opperman*.[16] Thus the issue of privacy must have been introduced in *Opperman*, or some earlier framework precedent. The much discussed *Califormia v Carney*,[17] which, especially in the oral argument (Rissland [1989](#); Ashley et al. [2008](#); Al-Abdulkarim et al. [2013](#)) was important for its identification of the factors needed to distinguish between an automobile being used as a vehicle and an automobile being used as a residence. This was extensively discussed in the Oral Argument for that case (Rissland [1989](#); Al-Abdulkarim et al. [2013](#)) and a number of new factors made their way into the decision. Such factors are rarely considered as present in previous cases, and so it may be a mistake to analyse earlier cases in terms of dimensions, factors and issues introduced subsequently. Understanding a body of cases as a *sequence*, although the topic of Henderson et al. ([2001](#)) and Rissland and Xu ([2011](#)), is perhaps a topic which merits further exploration. Keeping a representation current in the face of changes in case law is also explored in Al-Abdulkarim et al. ([2016a](#)).

5.3 Argumentation schemes

The motivation for expressing a set of precedents as rules in Prakken and Sartor ([1998](#)) was so that the precedents could be used to deploy arguments in a dialogue game. At the time, following Gordon ([1993](#)), dialogue games were a popular way of expressing legal procedures, such as particular forms of legal argumentation. Since their introduction to explain fallacies in propositional logic Hamblin ([1970](#)) and

---

[15] Carroll v. United States, 267 U.S. 132 (1925).

[16] South Dakota v. Opperman, 428 U.S. 364 (1976).

[17] California v. Carney, 471 U.S. 386 (1985).

Mackenzie (1979), such games handed tended to be built on an underlying rule base. An excellent overview of such systems is given in Prakken (2006). As this century has progressed, however, the use of dialogue games has tended to make way for argumentation schemes Walton (1996) and Walton et al. (2008). Although argument schemes had been used before in AI and Law: arguably the three ply argumentation of HYPO is a scheme, CABARET certainly uses schemes and Toulmin's argumentation scheme (Toulmin 1958) was widely used (e.g Marshall (1989)), often to drive dialogues as in Bench-Capon et al. (2000), explicit use of argumentation schemes moved to the forefront when Walton's understanding of them in terms of premises, conclusions and critical questions, as expressed in Walton (1996), became prevalent.

A particular argumentation scheme (originally designed for value based practical reasoning) was applied to reasoning with legal cases in Greenwood et al. (2003), and argumentation schemes were used to model a particular case in Gordon and Walton (2006). The explicit use of CATO's argumentation moves was carried out in Wyner and Bench-Capon (2007). These schemes were used to model particular cases in Bench-Capon (2012), and formalised, further developed and extended to include dimensions in Wyner et al. (2011), Atkinson et al. (2013) and Prakken et al. (2015). The last of these, in particular, offers a set of schemes designed to model HYPO/CATO style reasoning in a formal framework, ASPIC+ Modgil and Prakken (2014).

6 Purposes and values

HYPO and CATO, as mentioned above treat all their cases as homogeneous. Whether considered as bundles of facts or bundles of factors, they are abstracted from all aspects of context. There is no consideration of the level of court, or the jurisdiction. Thus the HYPO cases ranged across a number of different states. Nor was there any consideration of date or sequence. HYPO cases ranged from 1845 to 1980, although the majority dated from the nineteen sixties and seventies (Ashley 1990). This should not be seen as a defect: there is no reason why an AI model has to cover every aspect, and classroom discussions often use cases in the context independent fashion.[18] The intention of HYPO was to model one *necessary* part of actual appellate reasoning, and so it is not fair to complain that it is not *sufficient* in the real world. This aspect of HYPO was the subject of a series of critiques by Don Berman and Carole Hafner, who noted the absence of procedural context (Berman and Hafner 1991), consideration of purpose (Berman et al. 1993), and temporal aspects (Berman and Hafner 1995). These papers were consolidated in a paper in a special issue of this journal (10:1–3) in memory of Don Berman (Hafner and Berman 2002). Even though HYPO was not intended to address these aspects, it is

---

[18] In fact the influential critique in Berman et al. (1993) was focused on three cases: Pierson v Post (New York, 1805), Keeble v Hickergill (an English Queen's Bench case from 1707) and Young v Hitchens (an English Queen's Bench case of 1844). The critique is itself thus made using three cases from different jurisdictions, with very different dates, presented out of sequence. None the less these cases are often used when teaching property law in the US.

important to tell (some) computer scientists that HYPO is not the whole of the story, and to explore how these aspects can be represented to augment the basic HYPO approach. These critiques were a little slow in being taken up, but meanwhile within the computational argumentation community interest was growing in Perelman's notion of an audience (Perelman and Olbrechts-Tyteca 1971). This had been the subject of Grasso et al. (2000) and was integrated with Dung's argumentation frameworks (Dung 1995) in Bench-Capon (2003a). This led to a revisiting of the ideas of Berman et al. (1993), which had argued that the preferences between factors reflected the purposes attributed to the law. These could change across times and jurisdictions (Christie 2000), and so explain why similar cases might be decided differently at different times and in different places. As such the resolution of the conflicts could be seen as relative to the audiences to which the laws and arguments were addressed, and could be made computational using the formal characterisation of audiences in Bench-Capon (2003a), that is as an ordering of the social values to which the arguments related, by equating the notion of purpose with the promotion of social values as in Bench-Capon (2003b) and Bench-Capon et al. (2005).

6.1 Value based theory construction

Interest in Berman et al. (1993) and examination of the purpose of laws, interpreted as the values they were intended to promote and protect, was revived by a group of three articles that appeared in *Artificial Intelligence and Law*, volume 10 numbers 1–3: Bench-Capon (2002), Prakken (2002) and Sartor (2002). Following this Sartor and Bench-Capon developed the ideas in a series of papers, including Bench-Capon and Sartor (2000) and Bench-Capon and Sartor (2001), culminating in Bench-Capon and Sartor (2003). These papers gave an account of theory construction with disagreements resolved according to the value preferences of the audience (here society, as represented by the judges).

The overall idea is shown in Fig. 7. The starting point is at the bottom left where we have a set of cases (described as bundles of factors) and their outcomes. These outcomes reveal preferences between sets of factors, which will explain the case outcomes. The preferences are essentially the priority rules of Prakken and Sartor (1998). Each factor is associated with a value: a value that will be promoted by deciding the case for the side favoured by that factor. This enables us to rewrite sets of factors as sets of values, and transfer the preferences to these sets of values. The heart of the theory is this set of value preferences. This takes us to the top of the diagram, and we have constructed a theory and so we can start to apply the theory and work our way down the right hand side. The value preferences determine and explain preferences between sets of factors which are not to be found in the existing precedents. These preferences can then be used to determine the outcomes of as yet undecided cases which have these sets of factors.

The theories are constructed using a set of operators, defined in Bench-Capon and Sartor (2003). The theory starts from nothing and construction begins by including a case from the background. This will bring with it a set of factors and their associated values. Each factor is associated with a simple rule stating that the factor is a reason

Fig. 7 Value based theory construction from Bench-Capon and Sartor (2001)
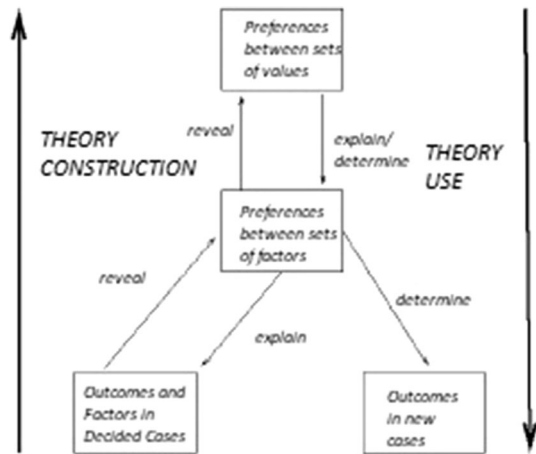
to decide for the side favoured by the factor. The theory can then be extended by including more cases, combining simple rules into rules with multiple antecedents and establishing preferences between rules and between values. This continues until the theory can be applied to give an outcome for the case under consideration. At this point the onus moves to the other party who must attempt to extend the theory to produce a better theory with an outcome for its favoured side, whereupon, it is again the turn of the original side. This process of extending and refining the theory continues until there is no possible extension of the theory which changes the outcome. The paper illustrates the process with the wild animals cases introduced in Berman et al. (1993). The paper also includes a discussion of how the theory operators relate to the argument moves of CATO.

This approach was tested empirically in Chorley and Bench-Capon (2005b) (included in this issue) and Chorley and Bench-Capon (2005a). The paper included in this special issue explored the use of CATE (CAse Theory Editor) in a series of experiments intended to explore a number of issues relating to the theories constructed using the operators of Bench-Capon and Sartor (2003), including how the theories should be constructed, how sets of values should be compared, and the representation of cases using structured values (which are akin to dimensions) as opposed to factors. In CATE, the construction of theories is done by the user, supported by the CATE toolset. The second paper described AGATHA (Argument Agent for Theory Automation) which is designed to automate the theory construction process, by constructing the theory first as a search over the space of possible theories, and then as a two player dialogue game (which could be played with the AGATHA program playing both sides). A set of search operators and argument moves are defined in terms of the theory constructors and the resulting theories are evaluated according to their explanatory power and their simplicity. The search or game continues until it is not possible to produce a better theory. Several search methods were investigated: brute force and heuristic search using A* and adversarial search using α/β pruning. The results proved to be good: they were reported in Chorley and Bench-Capon (2005a):

AGATHA produces better theories that the hand constructed theories reported in Chorley and Bench-Capon (2005b), and theories comparable in explanatory power to the best performing reported technique, IBP (Ashley and Brüninghaus 2009). Note also that AGATHA can be used even when there is no accepted structural model of the domain, whereas IBP relies on using the structure provided by the Restatement of Torts.

Table 3 of this introduction includes the results of some versions of AGATHA.

6.2 Other uses of values

Apart from their pivotal role in the approach to theory construction advocated in Bench-Capon and Sartor (2000), values have been applied to case based reasoning in several other ways. Their original use, to tailor an argumentation framework to the preferences of an audience using Value based Argumentation Frameworks (VAF) (Bench-Capon 2003a) was applied to reasoning with cases in Atkinson and Bench-Capon (2005), Wyner and Bench-Capon (2007) and Atkinson and Bench-Capon (2007), but recent work by Atkinson and her colleagues has placed moved away from this use of values in Al-Abdulkarim et al. (2015b) and Al-Abdulkarim et al. (2016d). They continue to use the VAF approach for legislative law making and e-democracy Atkinson et al. (2006), Atkinson et al. (2011) and Bench-Capon et al. (2015), and in the justification of norms (Bench-Capon and Modgil 2017).

Values have also been used for reasoning with cases in Araszkiewicz (2011), which recognised that values could be used to justify rules (their traditional role in VAF based systems) and also the inclusion of particular antecedents in rules, and in Grabmair and Ashley (2011), from which developed a different formalism intended to capture dynamic aspects of reasoning with legal cases (Grabmair and Ashley 2013). The most recent work in this strand is Grabmair (2017). Finally, another approach to using values in the representation of cases can be found in Verheij (2016).

One issue relating to values that has yet to be fully explored is that it seems that they can be used in a variety of ways. Sometimes they appear to give rise to conflicting positions which can be resolved by preferring one value to another, which is essentially the approach of VAF based systems, At others they appear to be better regarded a *sets* of reasons, for which issues such as accrual are pertinent (Prakken 2005; Modgil and Bench-Capon 2010; Bench-Capon et al. 2011, which seems applicable in domains such as US Trade Secrets, when a number of different aspects of the case must be considered. Also, however, there seem to be situations where a balance (Lauritsen 2015; Gordon and Walton 2016) must be struck between two values: it cannot be said that one is preferred to another, but rather both must be respected to an appropriate degree. This is perhaps the case with the Automobile Exception cases, where a balance must be struck between exigency and privacy, Bench-Capon and Prakken (2010). Unpicking these different aspects is something that AI and Law may need to address in future work.

7 Back to dimensions

The work we have been discussing has tended to adopt CATO's version of factors, features either present or absent and always favouring the same side, rather than HYPO's dimensions, which represent a range of values, and may favour either side according to where on the dimension the case lies. This is understandable and the use of such factors is a very sensible simplification for the purpose which CATO was intended to address, the teaching of distinguishing cases. Although there are other sensible simplifications (see, for example, Rissland et al. (1996)), CATO's notion of factors has been widely adopted and this has enabled understanding of reasoning with legal cases to make the significant progress represented by the papers included in this special issue. It remains the case, however, that factors do represent a simplification and so, if we are to make further progress, it will be necessary to improve our understanding of dimensions. The differences between dimensions and factors are discussed in Rissland and Ashley (2002) in this special issue: in the next section I will look at some more recent attempts to make use of dimensions.

7.1 Factors and dimensions

In the Restatement of Torts (quoted in Atkinson and Bench-Capon (2005)) we find (italics mine):

> Some factors to be considered in determining whether given information is one's trade secret are: 1. the *extent* to which the information is known outside of his business; 2. the *extent* to which it is known by employees and others involved in his business; 3. the *extent* of measures taken by him to guard the secrecy of the information; 4. the *value* of the information to him and to his competitors; 5. the *amount of effort* or money expended by him in developing the information; 6. the *ease or difficulty* with which the information could be properly acquired or duplicated by others.

Although we can identify all of these "factors to consider" with factors in CATO, the language here is, as the italicised phrases show, not really consistent with the all or nothing, present or absent, nature of CATO-style factors. Each of them seems to require some kind of quantitative estimate of *how much?*. Thus F16, *ReverseEngineerable*, a CATO factor which has presented a number of difficulties for subsequent systems attempting to use CATO's analysis, such as that reported in Al-Abdulkarimet al. (2015a), cries out for a more nuanced representation. *Reverse Engineerable* could mean that a person could build a version after a cursory inspection, or could require many person years of expert effort. Arguably, anything is reverse engineerable with sufficient effort, expertise and ingenuity. There are three ways in which issues requiring dimensions can arise:

1. Whether the facts favour the plaintiff or defendant: if a hunter is chasing a wild animal, how close does he have to get to count as possessing the animal? The famous case of *Pierson v Post* turned on precisely this point, as discussed in Bench-Capon and Rissland (2001).

2. Whether the factor should or should not be ascribed to the case: at what level of effort does a product cease to be considered reverse engineerable?

3. In comparisons: for example, thinking in terms of values, is the weak promotion of a preferred value to be preferred to the strong promotion of a lesser value (Sartor 2010)? This issue is also explored empirically in Chorley and Bench-Capon (2005b), and more recently in Al-Abdulkarim et al. (2016b).

The first two of these concern what can be argued about. When using factors, the decisions are made by the analyst and once made cannot be debated when attempting to resolve the case. (1) is effectively decided when defining the factors: thus whether Justinian, who demands actual physical possession for the animal to count as *caught* should be chosen over other authorities who make less stringent demands, such as mortal wounding or certain capture, is something that the analyst needs to be told along with the existence of *caught* as a factor. Once the definition has been chosen, ascription to cases is relatively easy. (2) relates to the analysis of particular cases. Given that the information was not in fact reverse engineered (in which case F25 would apply instead), the analyst must decide whether "the *ease or difficulty* with which the information could be properly acquired or duplicated by others" is such that F16 can be deemed to be present. (3) rather concerns computation, and allows mapping onto numbers to facilitate accrual and comparison. We need to be able to argue about whether a factor applies to a case and which side is favoured, but using CATO-style factors these issues are resolved by the analyst and so outside the scope of systems which begin with cases represented as bundles of factors.

7.2 Current use of dimension for computation

The computation aspect is key in Chorley and Bench-Capon (2005b). The main concern in that paper is values, but there is a section on what it terms *structured values*. The 26 CATO factors are associated with one of five values: *CA* for respect for confidentiality agreements, *QM* for questionable methods used by the defendant, *LM* for legitimate methods by the defendant, *RE* for reasonable efforts to maintain secrecy of the part of the plaintiff, *and* MW for material worth of the information.. A structured value runs from an extreme pro-plaintiff point to an extreme pro-defendant point, and the various CATO factors relating to the value are assigned positions along this line according to how strongly they promote or demote the value. They are thus dimensions in all but name. The values can themselves be ordered, and to the contribution of the factor becomes a function of the importance of the value and the extent to which the particular factor promotes it. The empirical consequences of the additional degrees of freedom this enables are given as part of the results in Al-Abdulkarim et al. (2016b)[19] and Table 3 of this paper.

---

[19] Several papers cited in this section report work carried out by Latifa Al-Abdulkarim as part of her PhD project (Al-Abdulkarim 2016). This project developed the ANGELIC (ADF for kNowledGe Encapsulation of Legal Information from Cases) methodology, which was applied to several domains including US Trade Secrets (Angelic Secrets) and the wild animals cases (Angelic Animals).

This approach is also adopted in Bench-Capon and Bex (2015) which provides dimensions for the wild animal cases of Berman et al. (1993) and Al-Abdulkarim et al. (2016b) which applies the same process to CATO's factors. In Al-Abdulkarim et al. (2016b) seven dimensions are used rather than the five structured values of Chorley and Bench-Capon (2005b), and the dimension runs from 10 (extreme pro-plaintiff) to 0 (extreme pro-defendant) rather than from $+10$ to -10, as used by Chorley and Bench-Capon (2005b).

The seven dimensions in Al-Abdulkarim et al. (2016b) are:

– *Agreement* This relates to the existence of explicit confidentiality and non-disclosure agreements.
– *Dubious* This considers whether the means used by the defendant were illegal or otherwise dubious or questionable
– *Legitimate* This covers the independent discovery of the information.
– *Measures* This relates to the security measures used by the plaintiff.
– *Worth* This relates to the value of the information, in terms of the time and effort that it might save, and the value it would add to the defendant's product.
– *Disclosure* This relates to the number of people to whom the information and been disclosed and the circumstances of such disclosures.
– *Availability* This considers how readily available the information was to the defendant.

The distribution of the 26 CATO factors is shown in Table 4.

These dimensions allowed computation of predictions of case outcomes using propositions with truth values between 0 and 1 to represent the dimension points, and interpreting the logical operators in the manner of fuzzy logic (Zadeh 1965), as proposed in Bench-Capon and Gordon (2015). The results of the dimensional version of Angelic Secrets reported in Al-Abdulkarim et al. (2016b) showed an improvement on previously reported versions which had used boolean factors (Al-Abdulkarimet al. 2015a). This seemed to centre in particular on a better handling of

Table 4 Dimensions in Al-Abdulkarim et al. (2016b), their values and their factors

	Agreement CA	Dubious QM	Legitimate LM	Measures RE	Worth MW	Disclosure RE	Availability LM
10	F4	F22		F6	F8		
9		F2					
8	F13	F26	F7		F7	F12	
7	F21	F14					
6	F14				F18		F15
5							
4						F1	
3	F5		F25		F16		F20
2							
1	F23		F17			F10	
0		F3		F19	F11	F27	F24

the vague factor F16, reverse engineerability. Two other cases with factors which it had been suggested in Al-Abdulkarim et al. (2016c) were wrongly ascribed, were wrongly decided by the program, but when the factors in these cases were corrected in the ways suggested in that earlier paper a success rate of 96.8% was achieved. This suggests that it is possible to achieve better results if we allow for degrees of support, rather than insisting on the all or nothing choices required by the use of factors. Note, however, that even with the dimensional approach of Al-Abdulkarim et al. (2016b), the description of the cases in terms of factors/dimensions is still a matter for the analyst: it cannot be disputed within the program. This relates to issues (1) and (2) above and will be discussed in Sect. 7.3. Note, however, that that there is no obligation for the computer system to cover the whole process: as discussed in Al-Abdulkarim et al. (2016d) a system may be scoped so as to start with factors or dimensions. It may be preferable to allow an analyst to represent the case in terms of dimensions and factors rather than attempt to enable their derivation from the brute facts. There is a degree of trade off here: see also Ashley (2009) and the commentary on it by Thorne McCarty in Bench-Capon et al. (2012).

The use of dimensions can also be used to generate case narratives, as described in Bench-Capon and Bex (2015), using methods akin to the scripts of Schank and Abelson (1977). These narratives can then be used to recast the story of the case in terms of legally significant facts, and to support improved explanations.[20]

7.2.1 Comparison with HYPO dimensions

The seven dimensions used in Al-Abdulkarim et al. (2016b) are somewhat different from those used in the original HYPO program as reported in Ashley (1990). This should come as no surprise: Al-Abdulkarim et al. (2016b) was able to draw on three decades of work on representations of US Trade Secrets, above all CATO, but also the work of Ashley and Brüninghaus (2009) and Chorley and Bench-Capon (2005b). Table 5 shows the relation between the two sets of dimensions. In some cases we see several of HYPO's dimensions arranged on same Angelic dimension: confidentiality, for example embraces four of HYPO's (binary) dimensions. In contrast, the dimensions relating to the acquisition of the information seem underrepresented in HYPO, and the dimension Availability, corresponds to none of HYPO's dimensions, and so is omitted from Table 5. These differences are probably because the cases introduced in CATO contained several cases which related to this issue. HYPO, in contrast, contained a number of cases relating to the validity of agreements: the older cases related to validity of contracts not specifically associated with Trade Secrets. This is reflected in HYPO's dimensions, in particular include the notion of *consideration*, which relates specifically to the validity of any agreements and does not feature in either CATO or Angelic Secrets (where an agreement without consideration simply does not appear as a factor).

---

[20] This also revives an idea from the past: for example (Rissland et al. 1993) discussed recurring prototypical fact patterns or legal stories in terms of legal scripts. Schank himself was an invited speaker at the second ICAIL (1989).

Table 5 Dimensions in Al-Abdulkarim et al. (2016b) and Ashley (1990)

	Confidentiality agreement CA	Dubious means QM	OK means LM	Security measures RE	Material worth MW	Disclosure RE
10	Agreed Not To Disclose			Security Measures (P)	Competitive Advantage (P)	Secrets Disclosed Outsiders (P)
9		Bribe Employee			Vertical Knowledge (P)	
8	Non competition Agreement				Brought Tools	Outsider Disclosures Restricted (P)
7						
6						
5						
4						Disclosure In Negotiations
3	Agreement Not Specific					
2						
1					Vertical Knowledge (D)	Outsider Disclosures Restricted (D)
0	No Consideration		Employee Sole Developer	Security Measures (D)	Competitive Advantage (D)	Secrets Disclosed Outsiders (D)

7.3 Dimensions as a bridge to facts

The transition from the facts of a case as reported using "world facts" to a set of legally significant "legal facts" has long been as issue in AI and Law, at least since Breuker and Den Haan (1991). The difficulty is that the world facts are diverse and peculiar to particular cases, whereas the legal facts exhibit more uniformity and have a more definite relation to the issues to decided. It is often said that once the world facts have been qualified into legal facts, the decision itself is often obvious. This clearly relates to issues (1) and (2) from Sect. 7.1. Both these questions need to be able to be made the subject of argument.

As can be seen from the use of structured values in Chorley and Bench-Capon (2005b) and the Tables in this paper, especially Table 4, factors can be regarded as points (or ranges) on dimensions. So the first move is to see the legal facts of a case as a bundle of dimension points. Then a case description can be seen as a frame with

slots corresponding to dimensions and fillers as points/ranges of those dimensions. This is precisely the structure advocated in Bench-Capon and Bex (2015) and Al-Abdulkarim et al. (2016b). It is also the method of representing cases in Prakken et al. (2015) to enable argumentation about the how the factors present in a case relate to facts and dimensions.

The matter is made more explicit in Al-Abdulkarim et al. (2016d), which distinguishes various types of statements in the journey from evidence to verdict, with a particular emphasis of how we can move from the world facts to the legal facts. It is important not to confuse issues of fact and issues of law. Table 6 shows the different types of statement used in that paper.

Note that the points/ranges on dimensions play three roles. First, they are the conclusion of the reasoning about the world, and the input to the legal reasoning. Thus we start with a mass of evidence from which we need to arrive at a set of well structured legal facts which will relate to the law governing the case. As such they are typically not certain but established with a certain degree of belief. The methods used here will be the standard ways in which people reason about the world: evidence takes many forms: witness testimony, documents, forensics, video footage etc. Each of these are associated with their own kind of reasoning, coherence (Bex 2011), probability (Timmer et al. 2015), etc, and the courts should make use of these established ways of reasoning. There is nothing distinctively *legal* in this phase, and the decisions are often made by lay juries, and often facts cannot be revisited at the Appeal stage. At his point the appropriate standard of proof (Farley and Freeman 1995; Gordon and Walton 2009) is applied. The appropriate proof standard will depend on the jurisdiction and the nature of the case: typically criminal cases impose a higher standard of proof than civil cases.

Those factual conclusions that meet the required standards become legal facts. At this point they become either true or false: all factual conclusions that meet the proof standard are "equally true": the question is not the extent to which they are believed, but whether they are believed sufficiently to be taken as true in the legal reasoning. This is the bridge between the world and realm of law: the proof standard is the toll which gives entry to the bridge. On leaving the bridge the legal facts

Table 6 Summary of Statement Types. Base Level Factors, Legal Facts and Factual Conclusions all Correspond to Ranges on Dimensions

Statement type		"Truth" value	Justification
Verdict	V	0, 1	Entailed by II/LI
Intermediate issue	II	0, 1	Entailed by LI
Leaf issue	LI	0, 1	$af \in AF$ Sufficiently favours plaintiff
Abstract factor	AF	−1 1	Deduced from BF
Base level factor	BF	−1 1	Mapped from LI
Legal fact	LF	0, 1	$if \in FC$ meets proof standard
Factual conclusion	FC	0 ...1	Entailed by IF
Intermediate fact	IF	0 ... 1	Derived from E
Evidence	E	1	Given

assume their third role: they become base level factors. As such they again become associated with a number: 1 for extreme pro-plaintiff factors, −1 for extreme pro-defendant factors and somewhere in between otherwise. Note that the number is entirely independent of the degree of belief with the factual conclusion was established: the base level factor number is determined by the dimension and the position of the factor on that dimension. Conflation of these roles can lead to problems: perhaps applying the standard of proof during legal reasoning, or allowing the degree of belief to contaminate the contribution of the factor.

All this is quite recent work and nothing like the consensus about, let alone the understanding of, reasoning with CATO-style factors has emerged with respect to dimensions. None the less, this is an active area of work, and points to a direction which may serve to enhance our understanding of reasoning with legal cases.

8 Conclusion

Reasoning with legal cases has been a central concern of AI and Law since its earliest days. As can be seen from the papers in this special issue and the many other papers referred to in this introduction, HYPO and it descendants—especially CATO—have had an enormous influence, setting the agenda for a variety of investigations carried out by different groups over three decades. As a result, the role of legal cases in moving from legal facts through intermediate concepts to issues and a final verdict is relatively well understood and has been formalised in Horty and Bench-Capon (2012) and Rigoni (2015). Also it has given rise to implementations which have performed well in empirical evaluations, capable of predicting case outcomes in more than 90% of cases, as shown in Ashley and Brüninghaus (2009) and Chorley and Bench-Capon (2005b).

All of this, however, has used CATO-style factors, and so has dealt with factors as present or absent, all or nothing, and so does not capture nuances associated with degrees of presence and degrees of support: the kind of factors suggested by the language used in the Restatement of Torts. This has led to a refocus of HYPO's original mode of representation, using dimensions, which does allow for these notions.

There are currently a number of open questions which we would expect to be pursued in the near and middle future. These include:

– Can we modify formalisations of the sort found in Horty and Bench-Capon (2012) to accommodate factors associated with magnitudes, rather than just binary true and false?
– Horty and Bench-Capon (2012) is concerned with comparing sets of factors. What about cases where factors need to be balanced?
– Can we use dimensions to be more precise about the transition from world facts to legal facts?
– Can we exploit the structure of cases into dimensions and points/ranges on these dimensions when reasoning about evidence?

- What roles can precedents have? We have seen that there are precedents indicating preferences, precedents providing a framework of issues as in Rigoni (2015), and tentatively identified precedents which introduced additional factors. Are there others?
- What is the role of purposes and values? If values can play several roles, as suggested in Sect. 6, how can these be distinguished, and how do they relate?

These are some of the questions that will need to be answered if reasoning with cases in the manner associated with HYPO and its successors is to be further developed. Thus HYPO will continue to inspire work in AI and Law for years to come.

References

Al-Abdulkarim L (2016) Representation of case law for argumentative reasoning. PhD thesis, University of Liverpool

Al-Abdulkarim L, Atkinson K, Bench-Capon T (2013) From oral hearing to opinion in the US Supreme Court. In: Legal knowledge and information systems—JURIX 2013: the twenty-sixth annual conference, pp 1–10

Al-Abdulkarim L, Atkinson K, Bench-Capon T (2015a) Evaluating the use of abstract dialectical frameworks to represent case law. In: Proceedings of the 15th international conference on artificial intelligence and law, ACM, pp 156–160

Al-Abdulkarim L, Atkinson K, Bench-Capon T (2015b) Factors, issues and values: revisiting reasoning with cases. In: Proceedings of the 15th international conference on artificial intelligence and law, ACM, pp 3–12

Al-Abdulkarim L, Atkinson K, Bench-Capon T (2016a) Accommodating change. Artif Intell Law 24(4):409–427

Al-Abdulkarim L, Atkinson K, Bench-Capon T (2016b) Angelic secrets: bridging from factors to facts in US Trade Secrets. In: Legal knowledge and information systems—JURIX 2016: the twenty-ninth annual conference, pp 113–118

Al-Abdulkarim L, Atkinson K, Bench-Capon T (2016c) A methodology for designing systems to reason with legal cases using abstract dialectical frameworks. Artif Intell Law 24(1):1–49

Al-Abdulkarim L, Atkinson K, Bench-Capon T (2016d) Statement types in legal argument. In: Legal knowledge and information systems—JURIX 2016: the twenty-ninth annual conference, pp 3–12

Aletras N, Tsarapatsanis D, Preoţiuc-Pietro D, Lampos V (2016) Predicting judicial decisions of the European Court of Human Rights: a natural language processing perspective. PeerJ Comput Sci 2. doi:10.7717/peerj-cs.93

Aleven V (1997) Teaching case-based argumentation through a model and examples. PhD thesis, University of Pittsburgh

Aleven V (2003) Using background knowledge in case-based legal reasoning: a computational model and an intelligent learning environment. Artif Intell 150(1–2):183–237

Aleven V, Ashley KD (1993) What law students need to know to win. In: Proceedings of the 4th international conference on artificial intelligence and law, ACM, pp 152–161

Aleven V, Ashley KD (1995) Doing things with factors. In: Proceedings of the 5th international conference on artificial intelligence and law, ACM, pp 31–41

Alexander L (1989) Constrained by precedent. South Calif Law Rev 63:1–64

Allen M, Bench-Capon T, Staniford G (2000) A multi-agent legal argument generator. In: Proceedings of the 11th international workshop on database and expert systems applications, IEEE, pp 1080–1084

Araszkiewicz M (2011) Analogy, similarity and factors. In: Proceedings of the 13th international conference on artificial intelligence and law, ACM, pp 101–105

Araszkiewicz M, Łopatkiewicz A, Zienkiewicz A, Zurek T (2015) Representation of an actual divorce dispute in the parenting plan support system. In: Proceedings of the 15th international conference on artificial intelligence and law, ACM, pp 166–170

Ashley KD (1989) Toward a computational theory of arguing with precedents. In: Proceedings of the 2nd international conference on artificial intelligence and law, ACM, pp 93–102

Ashley KD (1990) Modeling legal arguments: reasoning with cases and hypotheticals. MIT press, Cambridge

Ashley KD (2009) Ontological requirements for analogical, teleological, and hypothetical legal reasoning. In: Proceedings of the 12th international conference on artificial intelligence and law, ACM, pp 1–10

Ashley KD, Brüninghaus S (2009) Automatically classifying case texts and predicting outcomes. Artif Intell Law 17(2):125–165

Ashley KD, Rissland EL (1987) But, see, accord: generating blue book citations in hypo. In: Proceedings of the 1st international conference on artificial intelligence and law, ACM, pp 67–74

Ashley KD, Lynch C, Pinkwart N, Aleven V (2008) A process model of legal argument with hypotheticals. In: Legal knowledge and information systems: JURIX 2015: the twenty-first annual conference, pp 1–10

Atkinson K (2012) Introduction to special issue on modelling Popov v. Hayashi. Artif Intell Law 20(1):1–14

Atkinson K, Bench-Capon T (2005) Legal case-based reasoning as practical reasoning. Artif Intell Law 13(1):93–131

Atkinson K, Bench-Capon T (2007) Argumentation and standards of proof. In: Proceedings of the 11th international conference on artificial intelligence and law, ACM, pp 107–116

Atkinson K, Bench-Capon T, Cartwright D, Wyner A (2011) Semantic models for policy deliberation. In: Proceedings of the 13th international conference on artificial intelligence and law, ACM, pp 81–90

Atkinson K, Bench-Capon T, McBurney P (2006) PARMENIDES: facilitating deliberation in democracies. Artif Intell Law 14(4):261–275

Atkinson K, Bench-Capon T, Prakken H, Wyner A (2013) Argumentation schemes for reasoning about factors with dimensions. In: Legal knowledge and information systems—JURIX 2013: the twenty-sixth annual conference, pp 39–48

Austin JL, Warnock GJ (1962) Sense and sensibilia. Oxford University Press, Oxford

Bench-Capon T (1991) Practical legal expert systems: the relation between a formalisation of law and expert knowledge. Computers, Law and AI, Ablex, New York

Bench-Capon T (1993) Neural networks and open texture. In: Proceedings of the 4th international conference on artificial intelligence and law, ACM, pp 292–297

Bench-Capon T (1997) Arguing with cases. In: Legal knowledge and information systems—JURIX 1997: the tenth annual conference, pp 85–100

Bench-Capon T (1999) Some observations on modelling case based reasoning with formal argument models. In: Proceedings of the 7th international conference on artificial intelligence and law, ACM, pp 36–42

Bench-Capon T (2002) The missing link revisited: the role of teleology in representing legal argument. Artif Intell Law 10(1–3):79–94

Bench-Capon T (2003a) Persuasion in practical argument using value-based argumentation frameworks. J Log Comput 13(3):429–448

Bench-Capon T (2003b) Try to see it my way: modelling persuasion in legal discourse. Artif Intell Law 11(4):271–287

Bench-Capon T (2011) Relating values in a series of supreme court decisions. In: Legal knowledge and information systems—JURIX 2011: the twenty-fourth annual conference, pp 13–22

Bench-Capon T (2012) Representing Popov v Hayashi with dimensions and factors. Artif Intell Law 20(1):15–35

Bench-Capon T, Bex F (2015) Cases and stories, dimensions and scripts. In: Legal knowledge and information systems: JURIX 2015: the twenty-eighth annual conference, vol 279, IOS Press, pp 11–20

Bench-Capon T, Gordon T (2015) Two tools for prototyping legal CBR. In: Legal knowledge and information systems—JURIX 2015: the twenty-eighth annual conference, pp 177–178

Bench-Capon T, Modgil S (2017) Norms and value based reasoning: justifying compliance and violation. Artif Intell Law 25(1):1–36

Bench-Capon T, Prakken H (2010) Using argument schemes for hypothetical reasoning in law. Artif Intell Law 18(2):153–174

Bench-Capon T, Rissland E (2001) Back to the future: Dimensions revisited. In: Legal knowledge and information systems—JURIX, vol 2001, pp 41–52

Bench-Capon T, Sartor G (2000) Using values and theories to resolve disagreement in law. In: Legal knowledge and information systems—JURIX 1997: the thirteenth annual conference, pp 73–84

Bench-Capon T, Sartor G (2001) Theory based explanation of case law domains. In: Proceedings of the 8th international conference on artificial intelligence and law, ACM, pp 12–21

Bench-Capon T, Sartor G (2003) A model of legal reasoning with cases incorporating theories and values. Artif Intell 150(1):97–143

Bench-Capon T, Araszkiewicz M, Ashley K, Atkinson K, Bex F, Borges F, Bourcier D, Bourgine P, Conrad JG, Francesconi E et al (2012) A history of AI and Law in 50 papers: 25 years of the international conference on AI and Law. Artif Intell Law 20(3):215–319

Bench-Capon T, Atkinson K, Chorley A (2005) Persuasion and value in legal argument. J Log Comput 15(6):1075–1097

Bench-Capon T, Atkinson K, Wyner A (2015) Using argumentation to structure e-participation in policy making. In: Hameurlain A, Küng J, Wagner R, Decker H, Lhotska L, Link S (eds) Transactions on large-scale data-and knowledge-centered systems XVIII. Springer, Berlin, pp 1–29

Bench-Capon T, Freeman JB, Hohmann H, Prakken H (2003) Computational models, argumentation theories and legal practice. In: Reed C, Norman T (eds) Argumentation machines. Springer, Dordrecht, pp 85–120

Bench-Capon T, Geldard T, Leng P (2000) A method for the computational modelling of dialectical argument with dialogue games. Artif Intell Law 8(2–3):233–254

Bench-Capon T, Prakken H, Visser W (2011) Argument schemes for two-phase democratic deliberation. In: Proceedings of the 13th international conference on artificial intelligence and law, ACM, pp 21–30

Berman DH (1991) Developer's choice in the legal domain: the Sisyphean journey with CBR or down hill with rules (a working paper for the case-rules panel at the third international conference on artificial intelligence and law). In: Proceedings of the 3rd international conference on artificial intelligence and law, ACM, pp 307–309

Berman DH, Hafner CD (1991) Incorporating procedural context into a model of case-based legal reasoning. In: Proceedings of the 3rd international conference on artificial intelligence and law, ACM, pp 12–20

Berman DH, Hafner CD (1993) Representing teleological structure in case-based legal reasoning: the missing link. In: Proceedings of the fourth international conference on artificial intelligence and law, pp 50–59

Berman DH, Hafner CD (1995) Understanding precedents in a temporal context of evolving legal doctrine. In: Proceedings of the 5th international conference on artificial intelligence and law, ACM, pp 42–51

Bex F (2011) Arguments, stories and criminal evidence: a formal hybrid theory, vol 92. Springer, Dordrecht

Branting LK (1991) Reasoning with portions of precedents. In: Proceedings of the 3rd international conference on artificial intelligence and law, ACM, pp 145–154

Branting LK (1993) A computational model of ratio decidendi. Artif Intell Law 2(1):1–31

Breuker J, Den Haan N (1991) Separating world and regulation knowledge: where is the logic. In: Proceedings of the 3rd international conference on artificial intelligence and law, ACM, pp 92–97

Bruninghaus S, Ashley KD (2003a) Predicting outcomes of case based legal arguments. In: Proceedings of the 9th international conference on artificial intelligence and law, ACM, pp 233–242

Bruninghaus S, Ashley KD (2003b) A predictive role for intermediate legal concepts. In: Legal knowledge and information systems: JURIX 2003: the sixteenth annual conference, ISO Press, pp 153–162

Callan JP, Croft WB, Harding SM (1992) The INQUERY retrieval system. In: Min Tjoa A, Ramas I (eds) Database and expert systems applications. Springer, Berlin, pp 78–83

Chorley A, Bench-Capon T (2005a) Agatha: using heuristic search to automate the construction of case law theories. Artif Intell Law 13(1):9–51

Chorley A, Bench-Capon T (2005b) An empirical investigation of reasoning with legal cases through theory construction and application. Artif Intell Law 13(3–4):323–371

Christie G (2000) The notion of an ideal audience in legal argument, vol 45. Springer, Law and Philsophy Library, Dordrecht

Conrad JG, Zeleznikow J (2015) The role of evaluation in AI and Law: an examination of its different forms in the AI and Law journal. In: Proceedings of the 15th international conference on artificial intelligence and law, ACM, pp 181–186

Daniels JJ, Rissland EL (1997) Finding legally relevant passages in case opinions. In: Proceedings of the 6th international conference on artificial intelligence and law, ACM, pp 39–46

Dung PM (1995) On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artif Intell 77(2):321–357

Farley AM, Freeman K (1995) Burden of proof in legal argumentation. In Proceedings of the 5th international conference on artificial intelligence and law, ACM, pp 156–164

Gardner A (1987) An artificial intelligence approach to legal reasoning. MIT press, Cambridge

Gordon TF (1993) The pleadings game. Artif Intell Law 2(4):239–292

Gordon TF, Walton D (2009) Proof burdens and standards. In: Rahwan I, Simari G (eds) Argumentation and artificial intelligence. Springer, Dordrecht

Gordon TF, Walton D (2006) Pierson vs. Post revisited—a reconstruction using the Carneades argumentation framework. In: Computational models of argument: proceedings of COMMA 2006, pp 208–219

Gordon TF, Walton D (2016) Formalizing balancing arguments. In: Computational models of argument—proceedings of COMMA 2016, pp 327–338

Grabmair M (2017) Predicting trade secret case outcomes using argument schemes and learned quantitative value effect tradeffs. In: Proceedings of the 16th international conference on artificial intelligence and law, ACM (**in Press**)

Grabmair M, Ashley KD (2011) Facilitating case comparison using value judgments and intermediate legal concepts. In: Proceedings of the 13th international conference on artificial intelligence and law, ACM, pp 161–170

Grabmair M, Ashley KD (2013) Using event progression to enhance purposive argumentation in the value judgment formalism. In: Proceedings of the fourteenth international conference on artificial intelligence and law, ACM, pp 73–82

Grasso F, Cawsey A, Jones R (2000) Dialectical argumentation to solve conflicts in advice giving: a case study in the promotion of healthy nutrition. Int J Hum Comput Stud 53(6):1077–1115

Greenwood K, Bench-Capon T, McBurney P (2003) Towards a computational account of persuasion in law. In: Proceedings of the 9th international conference on artificial intelligence and law, ACM, pp 22–31

Hafner CD, Berman DH (2002) The role of context in case-based legal reasoning: teleological, temporal, and procedural. Artif Intell Law 10(1–3):19–64

Hage J (1993) Monological reason-based logic: a low level integration of rule-based reasoning and case-based reasoning. In: Proceedings of the 4th international conference on artificial intelligence and law, ACM, pp 30–39

Hage J (1996) A theory of legal reasoning and a logic to match. Artif Intell Law 4(3–4):199–273

Hamblin CL (1970) Fallacies. Methuen, London

Henderson J, Bench-Capon T (2001) Dynamic arguments in a case law domain. In: Proceedings of the 8th international conference on artificial intelligence and law, ACM, pp 60–69

Horty JF (1999) Precedent, deontic logic, and inheritance. In: Proceedings of the 7th international conference on artificial intelligence and law, ACM, pp 63–72

Horty JF (2004) The result model of precedent. Leg Theory 10(1):19–31

Horty JF (2011a) Reasons and precedent. In: Proceedings of the 13th international conference on artificial intelligence and law, ACM, pp 41–50

Horty JF (2011b) Rules and reasons in the theory of precedent. Leg Theory 17(1):1–33

Horty JF, Bench-Capon TJ (2012) A factor-based definition of precedential constraint. Artif Intell Law 20(2):181–214

Johnston B, Governatori G (2003) Induction of defeasible logic theories in the legal domain. In: Proceedings of the 9th international conference on artificial intelligence and law, ACM, pp 204–213

Lauritsen M (2015) On balance. Artif Intell Law 23(1):23–42

Levi EH (1948) An introduction to legal reasoning. Univ Chic Law Rev 15(3):501–574

Lindahl L, Odelstad J (2008) Intermediaries and intervenients in normative systems. J Appl Log 6(2):229–250

Mackenzie JD (1979) Question-begging in non-cumulative systems. J Philos Log 8(1):117–133

Marshall CC (1989) Representing the structure of a legal argument. In: Proceedings of the 2nd international conference on artificial intelligence and law, ACM, pp 121–127

McCarty LT (1995) An implementation of Eisner v. Macomber. In: Proceedings of the 5th international conference on artificial intelligence and law, ACM, pp 276–286

McLaren B, Ashley K (1999) Case representation, acquisition, and retrieval in sirocco. In: International conference on case-based reasoning, Springer, pp 248–262

Modgil S, Bench-Capon T (2010) Integrating dialectical and accrual modes of argumentation. In: Computational models of argument: proceedings of COMMA 2010, pp 335–346

Modgil S, Prakken H (2014) The ASPIC+ framework for structured argumentation: a tutorial. Argum Comput 5(1):31–62

Možina M, Žabkar J, Bench-Capon T, Bratko I (2005) Argument based machine learning applied to law. Artif Intell Law 13(1):53–73

Perelman C, Olbrechts-Tyteca L (1971) The new rhetoric: a treatise on argumentation. University of Notre Dame Press, Notre Dame

Prakken H (2002) An exercise in formalising teleological case-based reasoning. Artif Intell Law 10(1–3):113–133

Prakken H (2005) A study of accrual of arguments, with applications to evidential reasoning. In: Proceedings of the 10th international conference on artificial intelligence and Law, ACM, pp 85–94

Prakken H (2006) Formal systems for persuasion dialogue. Knowl Eng Rev 21(02):163–188

Prakken H, Sartor G (1996) A dialectical model of assessing conflicting arguments in legal reasoning. Artif Intell Law 4(3–4):331–368

Prakken H, Sartor G (1998) Modelling reasoning with precedents in a formal dialogue game. Artif Intell Law 6(3–4):231–287

Prakken H, Wyner A, Bench-Capon T, Atkinson K (2015) A formalization of argumentation schemes for legal case-based reasoning in ASPIC+. J Log Comput 25(5):1141–1166

Rigoni A (2015) An improved factor based approach to precedential constraint. Artif Intell Law 23(2):133–160

Rissland EL (1980) Example generation. In: Third biennial conference of the canadian society for computational studies of intelligence, pp 280–288

Rissland EL (1983) Examples in legal reasoning: legal hypotheticals. In: Proceedings of the 8th international joint conference on artificial intelligence, pp 90–93

Rissland EL (1984) The ubiquitous dialectic. In: Proceedings of the sixth European conference on artificial intelligence, pp 367–372

Rissland EL (1985) AI and legal reasoning. In: Proceedings of the 9th international joint conference on artificial intelligence, vol 2, Morgan Kaufmann Publishers Inc., pp 1254–1260

Rissland EL (1989) Dimension-based analysis of hypotheticals from supreme court oral argument. In: Proceedings of the 2nd international conference on artificial intelligence and law, ACM, pp 111–120

Rissland EL, Ashley KD (1987) A case-based system for trade secrets law. In: Proceedings of the 1st international conference on artificial intelligence and law, ACM, pp 60–66

Rissland EL, Ashley KD (2002) A note on dimensions and factors. Artif Intell Law 10(1–3):65–77

Rissland EL, Daniels JJ (1995) A hybrid CBR-IR approach to legal information retrieval. In: Proceedings of the 5th international conference on artificial intelligence and law, ACM, pp 52–61

Rissland EL, Daniels JJ (1996) The synergistic application of CBR to IR. Artif Intell Rev 10(5–6):441–475

Rissland EL, Skalak DB (1989a) Combining case-based and rule-based reasoning: a heuristic approach. In: Proceedings of the 2nd international joint conference on artificial intelligence, pp 524–530

Rissland EL, Skalak DB (1989b) Interpreting statutory predicates. In: Proceedings of the 2nd international conference on artificial intelligence and law, ACM, pp 46–53

Rissland EL, Skalak DB (1991) Cabaret: rule interpretation in a hybrid architecture. Int J Man Mach Stud 34(6):839–887

Rissland EL, Soloway E (1980) Overview of an example generation system. In: Proceedings of the 1st annual national conference on artificial intelligence, pp 256–258

Rissland E. L, Xu X (2011) Catching gray cygnets: an initial exploration. In: Proceedings of the 13th international conference on artificial intelligence and law, ACM, pp 151–160

Rissland EL, Skalak DB, Friedman MT (1993) Bankxx: a program to generate argument through case-base research. In: Proceedings of the 4th international conference on artificial intelligence and law, ACM, pp 117–124

Rissland EL, Skalak DB, Friedman MT (1996) Bankxx: supporting legal arguments through heuristic retrieval. Artif Intell Law 4(1):1–71

Rissland EL, Skalak DB, Friedman MT (1997) Evaluating a legal argument program: the bankxx experiments. Artif Intell Law 5(1–2):1–74

Rissland EL, Valcarce EM, Ashley KD (1984) Explaining and arguing with examples. In: Proceedings of the fourth AAAI conference on artificial intelligence, AAAI Press, pp 288–294

Sartor G (2002) Teleological arguments and theory-based dialectics. Artif Intell Law 10(1–3):95–112

Sartor G (2010) Doing justice to rights and values: teleological reasoning and proportionality. Artif Intell Law 18(2):175–215

Schank R, Abelson R (1977) Scripts, plans, goals and understanding, an inquiry into human knowledge structures. Lawrence Erlbaum, Hillsdale

Sergot MJ, Sadri F, Kowalski RA, Kriwaczek F, Hammond P, Cory HT (1986) The British Nationality Act as a logic program. Commun ACM 29(5):370–386

Skalak DB, Rissland EL (1991) Argument moves in a rule-guided domain. In: Proceedings of the 3rd international conference on artificial intelligence and law, ACM, pp 1–11

Skalak DB, Rissland EL (1992) Arguments and cases: an inevitable intertwining. Artif Intell Law 1(1):3–44

Smith J, Deedman C (1987) The application of expert systems technology to case-based law. In: Proceedings of the 1st international conference on artificial intelligence and law. ACM Press, New York, pp 84–93

Timmer ST, Meyer J.-J.C., Prakken H, Renooij S, Verheij B (2015) A structure-guided approach to capturing bayesian reasoning about legal evidence in argumentation. In: Proceedings of the 15th international conference on artificial intelligence and law, ACM, pp 109–118

Toulmin S (1958) The uses of argument. Cambridge University Press, Cambridge

Verheij B (1995) Arguments and defeat in argument-based nonmonotonic reasoning. In: Portuguese conference on artificial intelligence, Springer, pp 213–224

Verheij B (2016) Formalizing value-guided argumentation for ethical systems design. Artif Intell Law 24(4):387–407

Verheij B, Hage JC, Van Den Herik HJ (1998) An integrated view on rules and principles. Artif Intell Law 6(1):3–26

Walton D (1996) Argumentation schemes for presumptive reasoning. Lawrence Erlbaum Associates, Mahwah

Walton D, Reed C, Macagno F (2008) Argumentation schemes. Cambridge University Press, Cambridge

Wardeh M, Bench-Capon T, Coenen F (2009) Padua: a protocol for argumentation dialogue using association rules. Artif Intell Law 17(3):183–215

Wyner A, Bench-Capon T (2007) Argument schemes for legal case-based reasoning. In: Legal knowledge and information systems—JURIX 2013: the twentieth annual conference, pp 139–149,

Wyner A, Bench-Capon T, Atkinson K (2011) Towards formalising argumentation about legal cases. In: Proceedings of the 13th international conference on artificial intelligence and law, ACM, pp 1–10

Zadeh L (1965) Fuzzy sets. Inf Control 8(3):338–353