



## Try to see it my way: Modelling persuasion in legal discourse<sup>1</sup>

TREVOR J. M. BENCH-CAPON

*Department of Computer Science, University of Liverpool, Liverpool, UK*  
*E-mail: T.J.M. Bench-Capon@csc.liv.ac.uk*

**Abstract.** In this paper I argue that to explain and resolve some kinds of disagreement we need to go beyond what logic alone can provide. In particular, following Perelman, I argue that we need to consider how arguments are ascribed different strengths by different audiences, according to how accepting these arguments promotes values favoured by the audience to which they are addressed. I show how we can extend the standard framework for modelling argumentation systems to allow different audiences to be represented. I also show how this formalism can explain how some disputes can be resolved while in others the parties can only agree to differ. I illustrate this by consideration of a legal example. Finally, I make some suggestions as to where these values come from, and how they can be used to explain differences across jurisdictions, and changes in views over time.

**Key words:** argumentation, audiences, legal reasoning, persuasion

### 1. Introduction

The heart of a legal dispute can be seen as disagreement. In essence the parties disagree about a decision that is to be made, and attempt to *persuade* the judge that the case should be decided according to their wishes. Once the judge has made the decision, he must then justify his decision, so as to *persuade* the world, or at least any superior court which can be appealed to, that the decision was correct. So we must first consider why people disagree.

Some disputes are capable of relatively simple resolution. These have their roots in ignorance, weakness, or deliberate fault. It may be that one of the parties to the disagreement is ignorant of some crucial piece of information. Simply supplying the missing piece of information will suffice to settle the dispute. Similar is the case where the parties disagree as to a matter of fact. Resolving such a dispute requires only that the truth of the matter be established. Weakness can lead to disagreement when one of the parties is in possession of all the information needed to draw a conclusion, but is unable to derive the conclusion from the premises. Here all that is necessary is to present a step by step proof of the conclusion, spelling out the reasoning required in sufficient detail to be followed. Finally the disagreement may be simply

perverse: although in possession of all facts and capable of drawing the required conclusion, the party flatly refuses to do so. Here the dispute can be resolved by inducing the objector to follow normal procedures, or by ruling the objection illegitimate. All of these disputes *should*, in some sense, end in agreement.

There are, however, disagreements which can remain even when both parties are in agreement as to the facts, the logic that should be applied to them, the conclusions that can be drawn from them and the rules appropriate to the debate. Such disagreement is hard to resolve because reasons can be given for it, and no demonstration of the correctness of one of the parties is possible. In such situations it often appears rational to agree to differ.

We can illustrate this situation with an example. One decision that any Government must take is to decide on an appropriate rate of income tax. On the one hand, there will be an argument in favour of increasing the rate of taxation, since this progressive form of taxation will reduce income inequalities. On the other hand, it can be argued that a decrease in taxation will promote more enterprise, increasing Gross National Product, and so raising the absolute incomes of everyone. It is possible to see both these arguments as valid, since both supply a reason to act: and yet a choice must be made, since the actions are incompatible. Which choice is made will depend on whether the chooser prefers equality or enterprise. Thus two parties may be in agreement as to the consequences of a movement in the tax rate, and yet disagree as to the choice to be made because they differ in their fundamental aspirations. Different people will prize social values differently, and one may prefer equality to enterprise, while another prefers enterprise to equality. Thus while both arguments are agreed to be valid, one *audience* will ascribe more force to one than the other, whilst a different *audience* will make a different choice. In such cases these different audiences will rationally disagree, and agreement can only be reached by coming up with additional arguments which convince both audiences *in terms of their own preferences*, or by converting one to a different appraisal of social values. It is this situation, where disagreement resides in differing value preferences, that I will attempt to model in this paper.

In Section 2 I will cite the work of Perelman, who drew attention to the need to address this kind of disagreement. In Section 3 I will present a model which can capture this kind of disagreement, and point to some interesting properties of this model. In Section 4 I will apply this model to a much discussed legal example, and in Section 5 I will discuss the sources of different value rankings. Finally I will offer some concluding remarks in Section 6.

## 2. The importance of the audience

We may begin by considering some quotations from Perelman, who in *The New Rhetoric* (Perelman and Olbrechts-Tytca 1969) and a number of essays

(Perelman 1980) persuasively argued that logic is insufficient to capture all the nuances required for practical reasoning, and that the audience to whom the reasoning is addressed is a crucial determinate of its acceptability. Perelman writes:

“If men oppose each other *concerning a decision to be taken*, it is not because they commit some error of logic or calculation. They discuss apropos the applicable rule, the ends to be considered, the meaning to be given to values, the interpretation and characterisation of facts.” (Perelman, 1980, *italics mine*).

Again,

“Arguments which justify our opinions, choices and decisions are never as compelling as demonstrative proofs: they are more or less strong, relevant or convincing. A demonstration is correct or incorrect, it is imposed absolutely or lacks value; but in argumentation it is always possible to plead for or against, because arguments which support one thesis do not entirely exclude the opposite one; *this in no way means that all arguments are of the same value*” (Perelman, 1980, *italics mine*).

Finally he concludes the introduction to Perelman and Olbrechts-Tytca (1969) with

“Logic underwent a brilliant development during the last century when, abandoning the old formulas, it set out to analyze the methods of proof used effectively by mathematicians ... One result of this development is to limit its domain, since everything ignored by mathematicians is foreign to it. Logicians owe it to themselves to complete the theory of demonstration obtained in this way by a theory of argumentation”.

The point is that whether an argument is acceptable can only be asked in the context of a particular audience, since different audiences will accept different arguments. Moreover, the difference in acceptability derives from the way audiences assess the strength of arguments in terms of the values that accepting them promotes. Thus we need to be able to represent the differences between audiences which allow them to assess the strength of arguments, namely the *values* which will be supported by accepting the arguments, and the way in which an audience orders these values.

We can see the form of a practical argument (one about a decision to be taken or a choice to be made) as follows:

1. in these circumstances, C,

2. you should choose A,
3. because choosing A in C will promote some good G.

This is different from the standard reasoning that we should choose A in C to realise D, where D is some state of affairs. Rather G is a good promoted by the bringing about the state of affairs D by choosing A in C. The role of G is to explain why A *should* be chosen in C: it takes us beyond the factual into the realm of value, beyond the realm of truth into the realm of opinion. Of course such an argument must be *sound* on standard logical grounds: it must really be the case that choosing A will indeed promote G in C, and G must be acceptable as a social value which *prima facie* is worth promoting. The disagreement we wish to model comes when an alternative choice, B, promotes some other value, H, and H is preferred by the audience in question to G.<sup>2</sup>

In Section 3 I will introduce a formalism which allows this style of reasoning to be represented. I shall build on the Argumentation Framework formalism of Dung (1995), extending it in the manner of Bench-Capon (2003a).

### 3. Argumentation frameworks and values

Let us begin by reviewing Dung's original definition of Argumentation Frameworks<sup>3</sup>. For Dung the notion of an argument is as abstract as it can be: arguments are characterised only by the arguments they attack and are attacked by. This is especially suitable for modelling informal, natural language arguments, since the arguments are unconstrained in form, and there are no restrictions on what we can choose to count as an attack of one argument on another.

A formal definition of an Argumentation Framework, and the central notions concerning Argumentation Frameworks, is given as Definition 1.

**DEFINITION 1.** An Argumentation Framework (*AF*) is a pair  $AF = \langle AR, A \rangle$ , where  $AR$  is a set of arguments and  $A \subset AR \times AR$  is the attack relationship for *AF*. *A* comprises a set of ordered pairs of distinct arguments in  $AR$ . A pair  $\langle x, y \rangle$  is referred to as “*x* attacks *y*”.

For  $R, S$ , subsets of  $AR$ , we say that

- (a)  $s \in S$  is attacked by  $R$  if there is some  $r \in R$  such that  $\langle r, s \rangle \in A$ .
- (b)  $x \in AR$  is *acceptable* with respect to  $S$  if for every  $y \in AR$  that attacks  $x$  there is some  $z \in S$  that attacks  $y$  (i.e.,  $z$ , and hence  $S$ , defends  $x$  against  $y$ ).
- (c)  $S$  is *conflict free* if no argument in  $S$  is attacked by any other argument in  $S$ .

- (d) A conflict free set is *admissible* if every argument in  $S$  is acceptable with respect to  $S$ .
- (e)  $S$  is a *preferred extension* if it is a maximal (with respect to set inclusion) admissible subset of  $AR$ .

A useful way to picture an  $AF$ , to which we will appeal on occasion, is as a directed graph with arguments as vertices and directed arcs representing the attacks relation.

The key notion here is the *preferred extension* which represents a position which is

- internally consistent,
- can defend itself against all attacks,
- cannot be further extended without becoming inconsistent or open to attack.

From Dung (1995) we know that every  $AF$  has a preferred extension (possibly the empty set if a cycle of odd length exists in  $AF$ ), and that it is not generally true that an  $AF$  has a unique preferred extension. In fact any  $AF$  that contains a cycle of even length may have multiple preferred extensions (see Bench-Capon 2003a for a proof). In the special case where there is a unique preferred extension we say the dispute is *resolvable*, since there is only one set of arguments capable of rational acceptance. Where there are multiple preferred extensions, we can view a *credulous* reasoner as one who accepts an argument if it is in *at least one* preferred extension, and a *sceptical* reasoner as one who accepts an argument only if it is in *all* preferred extensions.

Once we allow that arguments may have different strengths, we open the possibility that an attack can fail, since the attacked argument may be stronger than its attacker. Thus, if an argument attacks an argument whose value is preferred it can be accepted, and yet not defeat the argument it attacks. To represent this possibility of unsuccessful attacks we must extend the standard argumentation framework so as to include the notion of value.

To record the values associated with arguments we need to add to the standard argumentation framework a set of values, and a function to map arguments on to these values.

**DEFINITION 2.** A value-based argumentation framework ( $VAF$ ) is a 5-tuple:

$$VAF = \langle AR, attacks, V, val, P \rangle,$$

where  $AR$ , and  $attacks$  are as for a standard argumentation framework,  $V$  is a non-empty set of values,  $val$  is a function which maps from elements of  $AR$  to elements of  $V$  and  $P$  is the set of possible audiences. We say that an argument

$a$  relates to value  $v$  if accepting  $a$  promotes or defends  $v$ : the value in question is given by  $val(a)$ . For every  $a \in AR$ ,  $val(a) \in V$ .

The set  $P$  of audiences is introduced because, following Perelman, we want to be able to make use of the notion of an audience. We see audiences as individuated by their preferences between values, since if there is agreement on the ranking of values, there will be agreement on which attacks succeed<sup>4</sup>. We therefore have potentially as many audiences as there are orderings on  $V$ . We can therefore see the elements of  $P$  as being names for the possible orderings on  $V$ . Any given set of arguments will be assessed by an audience in accordance with its preferred values. We therefore next define an audience specific value-based argumentation framework,  $AVAF$ :

**DEFINITION 3.** An audience specific value-based argumentation framework (AVAF) is a 5-tuple:

$$VAF_a = \langle AR, attacks, V, val, Valpref_a \rangle,$$

where  $AR$ ,  $attacks$ ,  $V$  and  $val$  are as for a  $VAF$ ,  $a$  is an audience,  $a \in P$ , and  $Valpref_a$  is a preference relation (transitive, irreflexive and asymmetric)  $Valpref_a \subseteq V \times V$ , reflecting the value preferences of audience  $a$ . The  $AVAF$  relates to the  $VAF$  in that  $AR$ ,  $attacks$ ,  $V$  and  $val$  are identical, and  $Valpref$  is the set of preferences derivable from the ordering  $a \in P$  in the  $VAF$ .

Our purpose in extending the  $AF$  was to allow us to distinguish between one argument attacking another, and that attack succeeding, so that the attacked argument is defeated. We therefore define the notion of *defeat for an audience*:

**DEFINITION 4.** An argument  $A \in AF$  defeats<sub>a</sub> an argument  $B \in AF$  for audience  $a$  if and only if both  $attacks(A, B)$  and  $\neg valpref(val(B), val(A))$ .

Note that an attack succeeds if both arguments relate to the same value, or if no preference between the values has been defined. If  $V$  contains a single value, or no preferences are expressed, the  $AVAF$  becomes a standard  $AF$ . In practice we expect the number of values to be small relative to the number of arguments. Many practical disputes can in fact be naturally modelled using only two values. Note that defeat is only applicable to an  $AVAF$ : defeat is always *relative to a particular audience*. We write  $defeats_a(A, B)$  to represent that  $A$  defeats  $B$  for audience  $a$ , that is  $A$  defeats  $B$  in  $VAF_a$ .

We next define the other notions associated with an  $AF$  for a  $VAF$ ,

**DEFINITION 5.** An argument  $A \in AR$  is acceptable-to-audience- $a$  (acceptable<sub>a</sub>) with respect to set of arguments  $S$ , (acceptable<sub>a</sub>( $A, S$ )) if:

$$(\forall x)((x \in AR \ \& \ defeats_a(x, A)) \rightarrow (\exists y)((y \in S) \ \& \ defeats_a(y, x))).$$

**DEFINITION 6.** A set  $S$  of arguments is conflict-free-for-audience- $a$  if

$$(\forall x)(\forall y)((x \in S \ \& \ y \in S) \rightarrow (\neg attacks(x, y) \vee valpref(val(y), val(x)) \in valpref_a))).$$

DEFINITION 7. A conflict-free-for-audience-a set of arguments  $S$  is admissible-for-an-audience-a if

$$(\forall x)(x \in S \rightarrow \text{acceptable}_a(x, S)).$$

DEFINITION 8. A set of arguments  $S$  in a value-based argumentation framework  $AF$  is a preferred extension for-audience-a (preferred<sub>a</sub>) if it is a maximal (with respect to set inclusion) admissible-for-audience-a subset of  $AR$ .

Now for a given choice of value preferences  $Valpref_a$  we are able to construct an  $AF$  equivalent to the  $AVAF$ , by removing from  $attacks$  those attacks which fail because faced with a superior value.

Thus for any  $AVAF$ ,  $vaf_a = \langle AR, attacks, V, val, Valpref_a \rangle$  there is a corresponding  $AF$ ,  $af_a = \langle AR, defeats \rangle$ , such that an element of  $attacks$ ,  $attacks(x, y)$  is an element of  $defeats$  if and only if  $defeats_a(x, y)$ . The preferred extension of  $af_a$  will contain the same arguments as  $vaf_a$ , the preferred extension for audience  $a$  of the  $VAF$ . Note in particular that if  $vaf_a$  does not contain any cycles in which all arguments pertain to the same value,  $af_a$  will contain no cycles, since the cycle will be broken at the point at which the attack is from an inferior value to a superior one. Because multiple preferred extensions can only arise from even cycles, and empty preferred extensions only from odd cycles, both  $af_a$  and  $vaf_a$  will have a unique, non-empty, preferred extension for such cases. Cases of even and odd cycles are illustrated in Figures 1 and 2.

Suppose we have a  $VAF$  with  $k$  values in  $V$ .  $P$  will contain factorial  $k$  distinct audiences, each of which will have a corresponding preferred extension  $preferred_i$ . Now for a given argument,  $A$ , in the  $VAF$  there are three possibilities:

- $A$  is every  $preferred_i$ . In this case  $A$  is accepted by every audience, and so cannot be rejected by adopting a particular ranking of values. We therefore say that  $A$  is *objectively acceptable*.
- $A$  is in at least one  $preferred_i$ , but not every  $preferred_i$ . In this case  $A$  will be acceptable to some audiences but not others, and so can either be accepted or rejected by choosing an appropriate value order. In this case we say that  $A$  is *subjectively acceptable*.
- $A$  is in no  $preferred_i$ . In this case  $A$  will be rejected by every audience. We say that such an  $A$  is *undefensible*.

That all three such cases can arise in quite simple frameworks can be shown by considering cycles which are at the root of the complexity problems in a standard  $AF$ . We will first consider cycles abstractly, so that we can see the

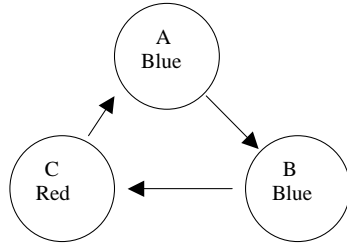


Figure 1. Three arguments cycle.

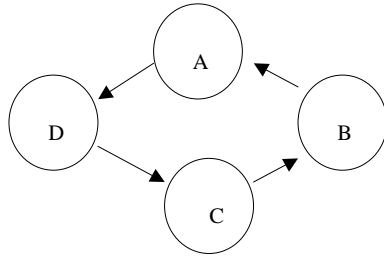


Figure 2. Cycle with four Arguments.

mechanics of *VAFs*. A concrete example is given in Section 4. First consider a three cycle containing arguments relating to two distinct values:

Suppose that C relates to one value (call it Red) and A and B relate to another value (call it Blue). There will be two audiences, and two preferred extensions. Suppose Blue is preferred to Red. Now the attack of C on A will fail, and A will be available to defeat B. Thus  $preferred_{blue}$  will contain  $\{A,C\}$ . Now suppose that Red is preferred to Blue. Now the attack of B on C will fail and C will defeat A, and hence defend B, giving  $preferred_{red}$  as  $\{B,C\}$ . Thus C, the argument related to the lone Red value, is objectively acceptable: it is in the preferred extension for both possible audiences. Which of the other two arguments is accepted will depend on the audience: hence they are both subjectively acceptable.

Now consider a cycle, again with two values, but with an even number of arguments.

Assume again that we have two values, Blue and Red. There are three possible distinct cases: three arguments of one value and one of the other (e.g., blue, blue, blue, red - three reds and a blue is equivalent); two of each arranged alternately (e.g., blue, red, blue, red), and two connected followed by two connected (e.g., blue, blue, red, red). Each case will have two preferred extensions depending on which value is preferred. The result is summarised in Table I.



*Table I.* Preferred extensions for cases of four arguments cycle

Case	Red preferred	Blue preferred
A Red, B, C, D Blue	{A, C}	{B, D}
A and C Red, B and D Blue	{A, C}	{B, D}
A and D Red, B and C Blue	{A, C}	{A, C}

The interesting case is the third one: here one pair of arguments, {A,C} is objectively acceptable, and one pair, {B,D}, indefensible. The reason is that A and C will defeat the arguments they attack since they have the same value: D, since defeated by A, is not available to defeat C, even when Red is the preferred value, and B, since defeated by C, is not available to defeat A, even when Blue is the preferred value.

The above are just some examples of how we can exploit values to resolve situations about which we cannot say anything in a standard AF: in fact we can say quite a lot more about the properties of cycles in Value-Based Argumentation Frameworks, even where we have an arbitrary number of values (for example, only arguments attacked by arguments with the same value can be indefensible). For more details and formal proofs, see Bench-Capon (2003a). For the moment the important thing to note is that the cycle construct, which posed difficulties in the standard AF, leading to empty preferred extensions or unmotivated choices, enables us in a VAF either to establish that an argument is objectively acceptable/indefensible, or else to motivate its acceptance in terms of value preferences.

Tractability problems also arise with standard Argumentation Frameworks when they are seen as the basis for dialogue. These problems include, in addition to problems arising from cycles, problems arising from the length of some argument chains, and difficulties in choosing which of several attackers to deploy. Again these problems can be reduced by taking the values promoted by arguments into account.

For example, in a line of argument, once an argument has been attacked by an argument with a different value, no arguments relating to the original value can have any effect. Either the original value is stronger than its successors, and so can defend itself against them, or it is weaker and so no argument relating to it can defeat the successors. This consideration can both reduce the length of lines of argument that need to be considered, and can eliminate certain attackers, easing the choice problem. The choice problem is also helped by the fact that an attacker with the same value can potentially establish objective acceptance, and so should be tried before an attacker with a different value, which cannot aspire to do more than establish subjective acceptance.

Cycles also behave differently in dialogue contexts when values are considered. Consider Figure 1 with C relating to Red and A and B relating to Blue. In a dialogue game based on a standard AF, such as that proposed in Dunne and Bench-Capon (2003), if we claim that C, A becomes unavailable, because attacked by C. So now B can be used to defeat the claim. In a dialogue based on a VAF, in contrast (e.g. the dialogue game with values proposed in Bench-Capon 2003b), we can keep A available by preferring Blue to Red. Now when B is played to attack C – which reveals that the opponent must indeed prefer Blue to Red, since otherwise the attack would be unavailing - we can defeat it by playing A, thus ensuring the acceptability of the original claim. Note that it is the *opponent's* value preferences that must be used to persuade: in fact I may accept A because I prefer Red to Blue, but I cannot hope to persuade my opponent on the basis of my choice of value preferences. The ability that values give to deploy arguments that would be unavailable in games based on the standard framework is rather important, since without it the decisive line of argument often cannot be played.

We can illustrate this with an example, derived from Christie (2000). Hal, a diabetic, loses his insulin and can save his life only by breaking into the house of another diabetic, Carla, and using her insulin. Consider Figure 1, but interpret the arguments as follows:

- A. Hal must replace Carla's insulin once the emergency is over.
- B. Hal must not take Carla's insulin because it is Carla's property.
- C. Hal can take the insulin as otherwise he will die.

A and B are based on property and C on necessity to preserve life. If C is attacked by B, it can be defended by a preference for necessity, but this will not persuade someone with a preference for property. This person can be persuaded by attacking B with A. Christie, however, who does prefer necessity to property, wishes to deny that Hal must replace the insulin, on the grounds that he should take it even if he is too poor to compensate. This involves attacking A with C, which he is allowed to do since A is not defeated by B given his preferences. Thus both will be persuaded that Hal can take the insulin, but will remain in disagreement as to the obligation to compensate.

#### 4. Application to example cases

In this Section I will apply the above analysis to a much discussed set of cases (e.g., Bench-Capon and Rissland 2001; Bench-Capon and Sartor 2001; Berman and Hafner 1993; Prakken 2000). I will base my discussion on the representation of these cases as a Dung-style Argumentation Framework

given in Bench-Capon (2002). We will consider three cases, all concerning the disputed pursuit of wild animals. The facts of the three cases are:

- *Keeble v. Hickergill* (1707). This was an English case in which Keeble owned a duck pond, to which he lured ducks, which he shot and sold for consumption. Hickergill, whose land adjoined Keeble’s, out of malice, scared the ducks away by firing guns. The court found for Keeble.
- *Pierson v. Post* (1805). In this New York case, Post was hunting a fox with hounds. Pierson intercepted, killed and carried off the fox. The court found for Pierson.
- *Young v. Hitchens* (1844). In this English case, Young was a commercial fisherman who spread a net of 140 fathoms in open water. When the net was almost closed, Hitchens went through the gap, spread his net and caught the trapped fish. The case was decided for Hitchens.

The Argumentation Framework for *Pierson* is shown in Figure 3. In each case the diagrams below show only the subgraph representing the particular interpretation under discussion.

The claim that the original pursuer had a right to the animal is challenged by the fact that it had not been caught (A), which is challenged by the argument that there was every prospect of catching it (B). Authorities are then cited to the effect that pursuit is not enough (C, taken from Justinian), but that bodily seizure is not necessary (D, taken from Barbyrac). Finally E states the need for possession to give a clear line. Two values seem to be involved here: the need for clear law which motivates A, B, C and E, and the desire to encourage useful and pleasurable activity which motivates D. The status of the claim turns on the resolution of the two argument cycle  $D \rightarrow E \rightarrow D$ , which represents the conflicting views of the two cited authorities. To resolve this the court must choose which value is to be preferred, whether the less clear situation where the animal has not caught can be tolerated in

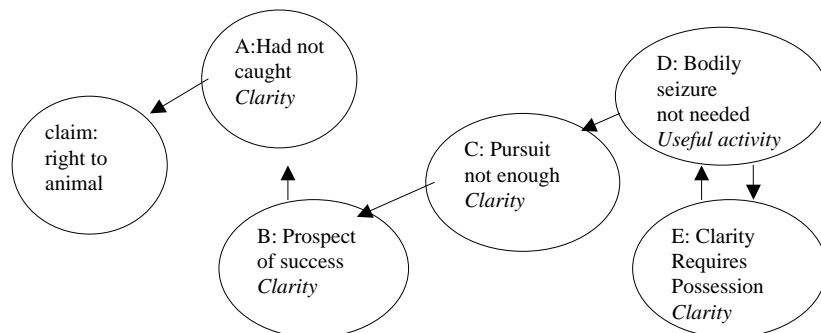


Figure 3. AF for *Pierson v. Post*.

order to encourage the activity. In the decision, preference was shown for clarity, although the dissenting opinion opted for the other preference.<sup>5</sup>

In *Keeble* we can see two different lines of reasoning. The first is shown in Figure 4. Here we introduce an argument (F) based on a new value, the right to enjoy one's property, which is challenged with another argument based on clarity (G) to the effect that the animals are not confined. Now assuming that the preference of clarity over encouraging a valuable activity remains, we can choose that the plaintiff wins by ranking property rights higher than clarity. This is what was decided in this case as I read it.

Alternatively, we can represent the reasoning of Berman and Hafner (1993) in Figure 5, which does not consider the ownership of the land relevant. Here we introduce an argument (H) based on a fourth value, the economic usefulness of the activity. If we prefer that value to clarity, then A is defeated and the claim rendered acceptable, without the need to consider the other attacks on A.

In both versions of the argument, the plaintiff wins, despite the availability of the line of argument of *Pierson* because a value is preferred to clarity, which means that the attack on A deriving from an alternative clear criterion need no longer be defended against: it appears that there are areas in which it is felt right for the law to intervene, and defending property rights and the pursuit of livelihood both seem plausible candidates for such intervention.

The third case is *Young vs Hitchens*. Here we can represent the line of argument as in Figure 6. This introduces two additional arguments, I and J, based on the value of protecting economic activity, one pointing to the fact that the defendant was in competition with the plaintiff, and the other suggesting that the competition was not fair. At this point we have two four argument cycles. One,  $A \rightarrow J \rightarrow I \rightarrow H \rightarrow A$ , contains three arguments of one

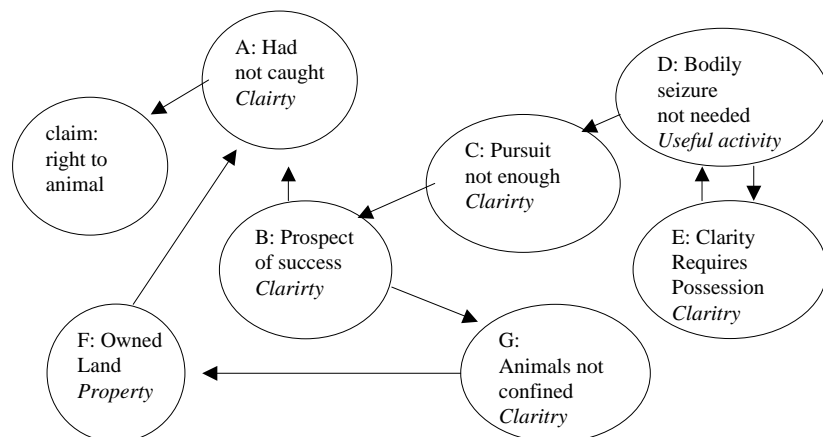


Figure 4. *Keeble v. Hickergill*, version 1.

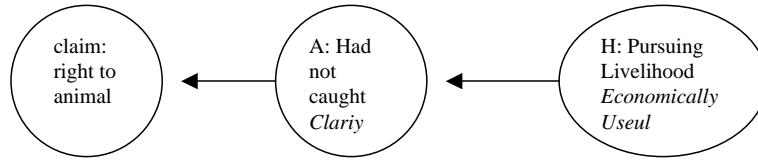


Figure 5. Keeble v. Hickergill, version 2.

value and one of another. But we can take from the second interpretation of *Keeble* that encouragement of economic activity is preferred to clarity, and hence we will accept J and H, finding for the plaintiff. But, in fact, we do not need to rely on *Keeble* and this value preference, because the other four cycle,  $A \rightarrow J \rightarrow I \rightarrow B \rightarrow A$ , contains two arguments with each of the two values in a sequence which renders B and J objectively acceptable. To avoid this consequence the defendant must break these four cycles: this is done by introducing the argument (K) that it is not proper for the court to rule on what counts as unfair competition, based on the value of maintaining the proper separation of powers between legislature and judiciary. Essentially this argument maintains that it is the prerogative of the legislature to rule on what counts as unfair competition, and that anything should be permitted if not explicitly forbidden by statute. Since the defendant won, we can assume that this argument was accepted, and this value was preferred even to the encouragement of economic activity, although we may think that a different view might have been taken at other times and in other places.

This case illustrates the worth of the Value-Based Argumentation approach in explaining decisions quite nicely: it captures the full power of the plaintiff’s argument – but for K his claim is objectively acceptable – whilst also

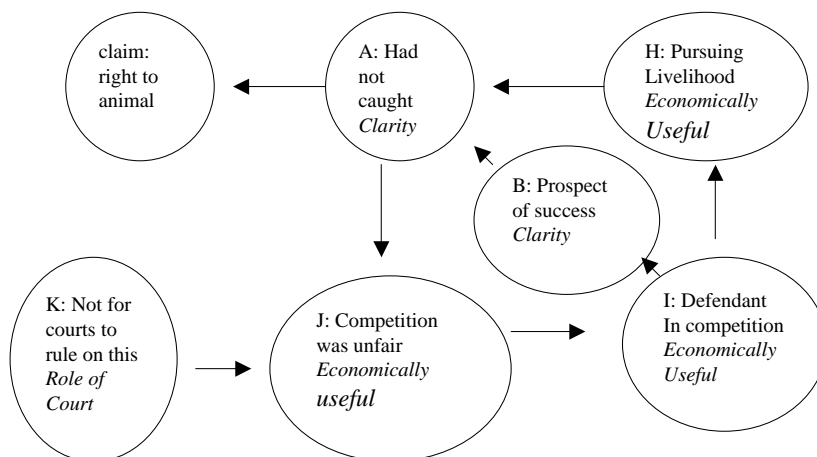


Figure 6. Young v. Hitchens as AF.

explaining the defendant's victory in terms of a particular value preference, turning on a view of the proper role of the courts.

From this body of cases – given the decisions that were made – we can see a partial order on values emerging: that the role of the courts is preferred both to property and economic activity, both of which are preferred to clarity, which is preferred to socially useful activity. But all the cases turn on ranking values: all the decisions are only subjectively acceptable, and a different ordering on values could have overturned any of them.

### 5. Origins of value preferences

At this point we may ask where these value preferences come from, and how agreement can be sufficiently close for these preferences to form an acceptable basis for law. Here I will point to three sources of these preferences: the prevailing legal culture, the prevailing social customs, and the prevailing ideology.

George Christie (2000) reviewed in Bench-Capon (2001) applies the approach of Perelman to comparative law, and in the process identifies a number of dimensions along which legal cultures can differ. These include:

- Narrow versus Broad Interpretation of statute texts: some jurisdictions strive after a precise understanding of the original interpretation of the drafters, whereas others allow much more freedom of interpretation;
- Hierarchical versus co-ordinate officials: some cultures organise their officials into hierarchies whereas others allow their officials more autonomy;
- Reactive versus active state: some cultures see the role of Government as responding to disputes, whereas others see Government as an instrument of social change pursuing its own agenda;
- Common good versus Individual Goods: some cultures see the common good as distinct from the aggregation of individual goods whereas others do not;
- Generality versus the Particular: some jurisdictions attempt to enunciate general principles and to subsume a particular case under them, whereas others concentrate on the facts of the case in hand and do not attempt to generalise further;
- Discretion versus Consistency: when a judge is allowed greater discretion it becomes hard to enforce consistency. Either consistency or the flexibility to fit individual situations may be valued more highly;
- Rights versus Privileges: some cultures see law as conferring privileges and others as making rights explicit.

At any given time a given jurisdiction will occupy a particular position with respect to these various dimensions, and all those involved will typically be aware of what this position is. Change is relatively slow: some dimensions such as the attitude towards interpretation may not move at all, whereas others such as the attitude to discretion behave like a pendulum, the bounds of discretion expanding until inconsistency becomes intolerable, and then contracting until the atmosphere becomes too confining. Part of legal training involves absorbing the culture – and values – of the legal system one is being trained in, and thus – within a jurisdiction – there is likely to be a broad consensus on what is “right”. It is this consensus that may explain why values are rarely mentioned in legal decisions, as noted by Branting (2003): they are simply part of the presupposed background, and everyone is well aware of them. Thus while values play an important role in explaining why decisions are what they are, and why decisions differ across jurisdictions, they need not be mentioned so long as the participants in the process all form part of the same, or very similar, audiences. If this is so, values might be expected to surface explicitly only in landmark cases which reflect a shift in the consensus sufficient to alter the understanding of the law.

A second source of value preferences lies in social attitudes. Examples abound of how these change over time and come to be absorbed into the law without requiring change in legislation. For example, the role of women has changed greatly since 1945 and this can be seen reflected in decisions concerning such matters as the right to reside in the marital home in domestic violence cases (see Moles 1987) for an excellent discussion of a series of illustrative cases). Again the trend in the UK from a standard six day working week to a five day working week, which has implications for Unemployment and Sickness benefits has been catered for in case law without the legislation being altered. Or one can imagine that accidents involving firearms are viewed differently where guns are ordinary household items from cultures where gun ownership is rare. Society is always in process of social change, and one element of the robustness of law is that it can absorb such changes through the realignment of value preferences.<sup>6</sup>

Finally there is ideology. Many value preferences can be viewed as expressing a contrast between “right” and “left”. An example of this occurred in the UK in the 80s when the prevalence of right wing thinking led to a controversial interpretation of a statute in the House of Lords which ruled an attempt to subsidise public transport in London illegal. Similarly religious ideology can account for different official attitudes to subjects such as contraception and abortion. Again, ideologies change – sometimes swiftly – but at any given time, we can identify a dominant ideological position.

The judge has an important role in the enunciation of value preferences. Judges are steeped in the legal culture of their jurisdiction, and are also aware of social changes, so that they can reflect the currents of the society in which they

move. They do, however, tend to reflect change subject to a certain lag, thus providing an element of predictability, continuity and stability. The movement of ideology is also reflected in the practice of appointing replacements to, for example the Supreme Court, one by one, so as to introduce “new blood” which over time will, if the trend continues, come to dominate.

Values provide us with considerable explanatory power when we wish to consider differences across jurisdictions, and through time within a jurisdiction. They are also a powerful tool in explaining differences between political parties, and often between generations.

## 6. Concluding remarks

In this paper I have argued that it is important to take the notion of an audience seriously if we are to explain and resolve certain kinds of disagreement. I have offered a formalism which allows us to consider the audience, and have shown how this formalism can be used to resolve some disagreements, and to identify the point of contention in others.

I believe that it is essential to recognise this pivotal role of the audience, and that progress in the model of legal argumentation and dialogue requires that we take this seriously. I hope that this paper will draw attention to the issues, and may also provide some foundations for their exploration.

## Notes

<sup>1</sup> This paper is based on an invited talk given at the Ninth International Conference on AI and Law, held in Edinburgh on 27 June, 2003. The talk was, in part, based on work reported in Bench-Capon (2002, 2003a,b). In the development of this work I have been greatly helped by discussions with Paul Dunne, Henry Prakken, Giovanni Sartor and Edwina Rissland.

<sup>2</sup> Of course, this represents something of a simplification, in that it both ignores the possibility that values may be promoted to different degrees, and that the outcome of actions may not be certain, so that the action may not reliably promote the value. In what follows, for the sake of clarity of exposition, I shall treat the promotion of a value as a simple Boolean, and assume that the consequences of an action are known and certain.

<sup>3</sup> Dung’s framework provided the formal basis of the approach used by Prakken and Sartor (1996). Prakken (2000) discusses the extension of their approach to make use of values.

<sup>4</sup> Note that particular individuals are not permanently assigned to some audience. An individual must, for the purposes of a particular dispute be part of some audience, because there is an audience for every possible value order, and a consistent value order is a requirement of rationality. But individuals may change audiences from dispute to dispute, or even during a dispute. Individuals may even enter the dispute undecided as to their value order, and choose which to which audience they belong as the dispute proceeds, as in Bench-Capon (2003b).

<sup>5</sup> I do not here discuss the possibility that value preferences may themselves be argued for. This has been discussed by e.g., Prakken and Sartor (1996) and Hage (2001).



<sup>6</sup> A fuller discussion of how shifts in the relative priority of competing purposes affect the application of precedents can be found in Berman and Hafner (1995). Rissland and Friedman (1995) also discuss how preferences can change in response to external events such as the energy crisis of 1973.

## References

- Bench-Capon, T.J.M. (2001). Review of George C. Christie, *The Notion of the Ideal Audience in Legal Argument*. *Artificial Intelligence and Law* 9: 59–71.
- Bench-Capon, T.J.M. (2002). Representation of Case Law as an Argumentation Framework. In Bench-Capon et al (eds), *Legal Knowledge and Information Systems*, 103–112. IOS Press: Amsterdam.
- Bench-Capon, T.J.M. (2003a). Persuasion in Practical Argument Using Value-Based Argumentation Frameworks. *Journal of Logic and Computation* 13(3): 429–448.
- Bench-Capon, T.J.M. (2003b). Agreeing to Differ: Modelling Persuasive Dialogue Between Parties With Different Values. *Informal Logic*. In Press.
- Bench-Capon, T.J.M. and Rissland, E.L. (2001). Back to the Future: Dimensions Revisited. In *Proceedings of JURIX 2001*. IOS Press: Amsterdam.
- Bench-Capon, T.J.M. and Sartor, G. (2001). Theory Based Explanation of Case Law Domains. In *Proceedings of the Eighth International Conference on AI and Law*, 12–21. ACM Press: New York.
- Berman, D.H. and Hafner, C.L. (1993). Representing Teleological Structure in Case Based Reasoning: The Missing Link. In *Proceedings of the Fourth International Conference on AI and Law*, 50–59. ACM Press: New York.
- Berman, D.H. and Hafner, C.L. (1995). Understanding Precedents in a Temporal Context of Evolving Doctrine. In *Proceedings of the Fifth International Conference on AI and Law*, 42–51. ACM Press: New York.
- Branting, K.L. (2003). An Agenda for Empirical Research in AI and Law. *Papers from the ICAIL 2003 Workshop on Evaluation of Legal Reasoning and Problem-Solving Systems*, Edinburgh.
- Christie, G.C. (2000). *The Notion of an Ideal Audience in Legal Argument*. Kluwer Academic Publishers: Dordrecht.
- Dung, P.H. (1995). On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-person Games. *Artificial Intelligence* 77: 321–357.
- Dunne, P.E. and Bench-Capon, T.J.M. (2003). Two Party Immediate Response Disputes. *Artificial Intelligence*. In Press.
- Hage, J. (1996). Formalizing Legal Coherence. In *Proceedings of the Eighth International Conference on AI and Law*, 22–31. ACM Press: New York.
- Moles, R.N. (1987). *Definition and Rule in Legal Theory*. Blackwell: Oxford.
- Perelman, C. and Olbrechts-Tyteca, L. (1969). *The New Rhetoric: A Treatise on Argumentation*. University of Notre Dame Press: Notre Dame.
- Perelman, Ch. (1980). *Justice, Law and Argument*. Reidel: Dordrecht.
- Prakken, H. (2000). An Exercise in Formalising Teleological Case Based Reasoning. In Breuker, J., Leenes, R. and Winkels, R. (eds) 2000, *Legal Knowledge and Information Systems: Jurix*, 49–57. IOS Press: Amsterdam.
- Prakken, H. and Sartor, G. (1996). A Dialectical Model of Assessing Conflicting Arguments in Legal Reasoning. *Artificial Intelligence and Law* 4 (3–4): 331–368.
- Rissland, E.L. and Friedman, M.T. (1995). Detecting Change in Legal Concepts. In *Proceedings of the Fifth International Conference on AI and Law*, 127–136. ACM Press: New York.