Arguments About Values

Bart Verheij, Artificial Intelligence, University of Groningen

July 11, 2013

Abstract

Arguments are more or less persuasive depending on the values their outcomes promote and demote. This idea, closely related to Perelman's ideas work on argumentation, has been formally modeled by Bench-Capon in his well-known Value-Based Argumentation Frameworks, based on Dung's abstract argumentation. In this paper, the question is addressed how arguments about which values are promoted and demoted can be modeled. A proposal is made to model Value-Based Argumentation Frameworks in DefLog, a language that extends Dung's abstract argumentation, while keeping central definitions such as preferred and stable extensions. Since DefLog allows the modeling of support of support, attack of support, support of attack, and attack of attack, it becomes possible to model arguments about whether a value is promoted or demoted in case an argument is accepted.

1 Introduction

Trevor Bench-Capon is a respected and prolific author with a broad range of interests. Many will recognize the experience that a topic one just becomes excited about, turns out to have been addressed by someone before, sometimes already a long time ago. In the areas of *Artificial Intelligence and Law* and *Argument and Computation* chances are high that this someone is Trevor. I propose to dub this the *Bench-Capon effect*, in analogy with its philosophical counterpart, the Aristotle effect, that the major questions of philosophy have also been addressed by Aristotle.

For me personally a recent example of the Bench-Capon effect - there are more - is the investigation of argumentation in connection with quantitative techniques.[1] I was aware of Trevor's emphasis on the fact that there are not only all-or-nothing factors in case-based argumentation, but also gradual dimensions going from one extreme via intermediate values to another extreme (e.g., [4, 6]). In the context of a case about the possible voidance of an employee's dismissal, an example factor is whether the employee has forged a diploma or not, and

[1]See [23], and also the recently started NWO-funded project on forensic Bayesian networks 'Designing and Understanding Forensic Bayesian Networks with Arguments and Scenarios'; `www.ai.rug.nl/~verheij/nwofs/`.

an example of a dimension is the degree of damage caused by the employee (cf. examples used by [18]). In the mentioned work, the emphasis is on a discrete spectrum of qualitative degrees. For instance, Bench-Capon and Sartor [6] follow the well-known work by Berman and Hafner [7], when they discuss a series of cases concerning the hunting of wild animals, about the legal issue whether the loss of game should be compensated or not. In the proposal by Bench-Capon and Sartor, the dimension of control can take on values *no-contact, seen, started, wounded, mortally-wounded, captured*, each value of the dimension expressing a gradually higher level of control of the hunted animal. There is also a hint of quantitative approaches, when remarks are made on the extension of the qualitative version of gradual dimensions to a quantitative version. Preliminary experiments related to Thagard's theory of explanatory coherence [19, 9] are reported in [5].

All this I knew. My surprise came when I discovered Trevor's paper 'Neural networks and open texture' [1],[2], in which an attempt is made to learn the rationale underlying legal cases, given only a set of decided cases. By the use of neural network technology, this work has a profoundly numeric, quantitative flavor, while retaining the connection to qualitative reasoning. The chosen domain concerns a welfare benefit to be paid to senior citizens visiting their hospitalized spouse, for which a set of decided cases was constructed artificially. The variables are chosen so as to provide a representative set of possibilities: The variable 'sex' is for instance represented as a Boolean, while the combined capital resources of the spouses are modeled using a numeric variable with a threshold. The paper draws partly positive and partly negative conclusions: Bench-Capon writes (p. 292): 'Neural networks are capable of producing a high degree of success in classifying cases in domains where the factors involved in the classification are unknown.' But he also finds that the patterns in a given set of decisions do not determine the rules that led to the decisions. (See also my discussion of this paper in [3], Section 5.6.)

Trevor's work is highly cited, and his most cited work is on value-based argumentation [2].[3] In this work, it is argued that disagreements cannot always be solved conclusively, and that hence the role of persuasion and of the values promoted by arguments should be emphasized. Here Trevor follows Perelman. In [2], argument attack succeeds depending on the preferences of the values promoted. As a model, Bench-Capon proposes an extension of Dung's abstract argumentation frameworks [11], called value-based argumentation frameworks.

In value-based argumentation frameworks, it is statically modeled which arguments promote or demote which values. As a result, there is no reasoning *about* which values are promoted or demoted. That is the question addressed in the present paper: How can arguments about which values are promoted and demoted be modeled?

An example can be found in the Dutch debate about the gradual elimi-

[2]Trevor himself played a pivotal role in this discovery: it was on a long list of possible papers to be discussed in [3].

[3]Google Scholar's author profile for Trevor J.M. Bench Capon (accessed July 9, 2013) records 6877 citations, 3278 of which since 2008, an h-index of 41, and 425 citations for [2].

nation of home mortgage interest deduction. That political debate can only be understood when considering arguments about which values are promoted or demoted, and which of these are preferred. For instance, for a laissez-faire oriented liberal, the current system of interest deduction would be seen as demoting the value of a free house market, thereby artificially inflating house prices. However, the Dutch political party that one would expect to be closest to laissez-faire liberalism (the VVD, currently providing the prime minister), has always fiercely opposed the gradual elimination of the interest deduction system. A key reason for the VVD seems to be that the system tends to especially benefit high-income citizens, and hence has as a side effect that it softens the high tax rates for higher incomes. So a VVD liberal will argue that the interest deduction system promotes the value of keeping taxes low for higher incomes.

In the following, Trevor's proposal about value-based argumentation is discussed (Section 2). Then the DefLog theory of structured arguments is presented (Section 3), followed by a proposal for modeling Trevor's value-based argumentation in DefLog (Section 4). Using a characteristic property of DefLog, namely that it models arguments about attack and support, a discussion of the modeling of arguments about values follows (Section 5).

2 Value-based argumentation

In AI & Law, Berman and Hafner [7] may have been the first to emphasize the need for the computational modeling of the role of the values and goals underlying legal decisions. Bench-Capon [2] builds on their work, but also pays tribute to the ground-breaking rhetorician and dialectician Perelman, who wrote (with Olbrechts-Tyteca):

> 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. ([12], as quoted in [2])

Hence, it is not to be expected that actual arguments about a decision to be taken can be conclusively settled. Bench-Capon focuses on a decision by a judge on a case where two sides argue for their position, and writes that the 'arguments [of both sides] may all be sound. But their arguments will not have equal value for the judge charged with deciding the case: the case will be decided by the judge preferring one argument over the other.' As a result, mathematical proof is not an appropriate model of the kind of arguments used when taking decisions, as acceptance of such arguments depends on the values of the audience addressed.

Bench-Capon [2] presents a formal account of what is needed for the modeling of the role of values in argumentation. In the account, he uses Dung's abstract argumentation frameworks [11] as a starting point. Mathematically,

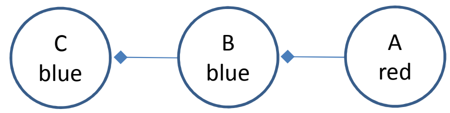

Figure 1: A value-based argumentation framework with two values (adapted from [2], p. 437)

Dung's abstract argumentation frameworks are directed graphs, where the nodes represent arguments, and the directed edges argument attack. Bench-Capon's central notion is that of a value-based argumentation framework (Definition 5.1). Formally, a value-based argumentation framework, or VAF for short, is an abstract argumentation framework in which each argument has an associated (abstract) value. The intended meaning of the value associated with an argument is that the value is promoted by accepting the argument. For instance, when parliament discusses a policy about raising taxes, the argument may focus on the values of social equality and of enterprise. Accepting a tax raise will demote the value of enterprise and promote the value of social equality. Bench-Capon also defines audience-specific argumentation frameworks (Definition 5.2). In an audience-specific argumentation framework, the values have an associated preference relation, that express the preferences of some audience. For instance, the Labour Party may prefer the value of social equality, and the Conservative Party that of enterprise.

In Bench-Capon's model, an argument A *defeats* an argument B *for an audience a* if A attacks B and the value associated with B is not preferred to the value associated with A for audience a. An attack succeeds for instance when the arguments promote the same value, or when there is no preference between the values. Bench-Capon continues to define variations of Dung's notions of argument acceptability, admissibility and preferred extension, relative to audience attack.

Bench-Capon uses a VAF with two values *red* and *blue* as an example (Figure 1). In the underlying abstract argumentation framework, there are three arguments A, B and C. A attacks B and B attacks C. The argumentation framework has a unique preferred extension (which is also grounded and stable), in which A and C are accepted and B rejected. In the framework, accepting the argument A promotes the value *red* and accepting the argument B or C promotes *blue*. Hence, for an audience preferring *red* to *blue*, defeat for the audience coincides with the underlying attack relation. In the corresponding unique preferred extension, A and C are accepted and B rejected, modeling that an audience preferring *red* accepts A and C, and rejects B. However, for an audience preferring *blue*, A does not defeat B. So for such an audience, A can exist side by side B, while B defeats C. This models that, for an audience preferring the value *blue*, A and B are accepted and C is rejected.

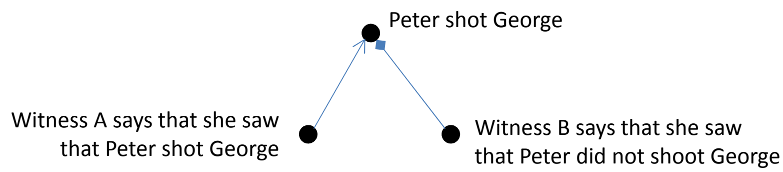

Figure 2: Support and attack

Bench-Capon illustrates value-based argumentation by considering the case of a diabetic who almost collapses into a coma by lack of insulin, and therefore enters the house of another diabetic, and uses her insulin. Bench-Capon models the roles of the value of protection against property right infringement and of saving one's life.

3 Structured arguments: the DefLog approach

Dung's abstract argumentation frameworks [11] have had an immense impact on the formal study of argumentation, as evidenced by its exceptionally high number of citations.[4] The paper focused on argument attack and its properties, uncovering the mathematical intricacy of the concept. Also a new puzzle arose, namely how argument attack is related to argument support. Dung himself contributed especially to one approach for the combination of attack and support, namely assumption-based argumentation [10]. A related influential approach is the ASPIC+-framework [15], which can be regarded as a modernized version of the model developed in close collaboration by Prakken and Sartor [16, 17], after incorporating the experiences in the European ASPIC project.

The formal argumentation model that combines support and attack used in this paper is DefLog [21]. We discuss a number of characteristic properties.

1. DefLog models both support and attack between arguments.

This first property is illustrated in Figure 2. The claim that Peter shot George is supported on the grounds of the testimony by witness A who says she saw that Peter shot George. The claim is also attacked by a testimony since witness B says she saw that Peter did not shoot George. DefLog uses a logical language with two connectives: a conditional for the representation of support combined with a negation operator for the representation of attack.

Modeling the arguments in Figure 2 requires two DefLog sentences:

$$a \to p; b \to \times p. \tag{1}$$

[4] Google Scholar records 1915 citations (July 9, 2013).

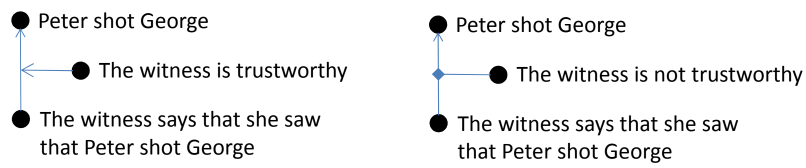

Figure 3: Arguments about support

Here a expresses the claim about the testimony by A, b the claim about the testimony by B, and p the claim that Peter shot George. The conditional $a \to p$ expresses that given A's testimony, it is supported that Peter shot George. The sentence $b \to \times p$ expresses that given B's testimony, it is attacked that Peter shot George. It depends on whether a and b are assumed what can be said about the justification of p. When only a is assumed as a premise, there is only an argument for p, making p justified. When only b is assumed as a premise, there is only an argument against p, and p is not justified. When both a and b are assumed as premises, there is one argument for p and one argument against p. As no information is available that solves the conflict between the testimonies, it is undetermined whether Peter shot George. In the DefLog formalization, this corresponds to the fact that the theory consisting of the four sentences $a \to p$, $b \to \times p$, a, and b does not have a stable extension in which p has a value.

2. DefLog models arguments about the issue whether an argument supports or attacks another argument.

The second property is illustrated in Figure 3, where arguments are provided about whether a witness testimony that Peter shot George supports the claim that Peter shot George. On the left-hand side, it is argued that the testimony indeed supports the claim since the witness is trustworthy. In contrast, on the right-hand side, it is attacked that the testimony supports the claim since the witness is not trustworthy. The right-hand side argument is an example of what since Pollock [13, 14] is called undercutting attack.

Modeling the arguments on the left-hand side in DefLog requires a nested conditional:

$$t \to (w \to p). \qquad (2)$$

Here w expresses the witness testimony about the claim p that Peter shot George, and t the trustworthiness of the testimony. The nested conditional expresses that, given the trustworthiness of the witness, it is supported that, given the testimony, the claim is supported. Assuming t as a premise, it is justified that $w \to p$, i.e., that, given the testimony, the claim is supported. Assuming also w as a premise, the claim p that Peter shot George is justified.

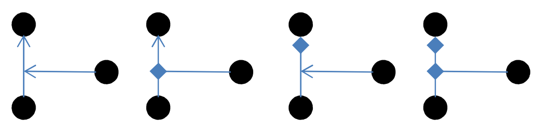

Figure 4: The four ways of arguing about support and attack in DefLog

Modeling the undercutter on the right-hand side, requires this nested conditional:

$$n \to \times(w \to p). \tag{3}$$

Here n expresses that the witness is not trustworthy. The nested conditional expresses that if the witness is not trustworthy it is attacked that, given the testimony, the claim is supported. The examples in Figure 3 can be adapted for arguments about attack, e.g., by considering a witness who says that she saw Peter *not* shoot George.

Figure 4 gives an overview of the four ways of arguing about support and attack in DefLog: support of support, attack of support, support of attack, and attack of attack.

3. In DefLog, the arguments supporting and attacking claims are expressed by sentences.

The third property has already been illustrated in the examples above. Every node in the graphs corresponds to a sentence. Also the arrows in the diagrams correspond to sentences, then with a conditional structure. The point is made since, following Dung [11], one influential way of modeling argumentation considers the argument nodes as structured arguments, in particular as (defeasible) Modus Ponens-based derivations. The attack relation is then considered to be between such derivations. Such a formalization separates an argument support layer from an argument attack layer in the model, which are then later reconnected when considering argumentative warrant (e.g., [15]). Instead, in DefLog's sentence-based approach, claims about support and attack are treated on a par. Layers that more strongly separate support and attack are not needed. In DefLog, sentences express prima facie assumptions that are assigned a justification status in light of the other available information: justified, defeated, or undetermined (cf. labeling approaches to argumentation with labels 'in', 'out' and 'undecided').

4. DefLog extends the expressiveness of Dung's abstract argumentation frameworks, generalizing its preferred and stable semantics.

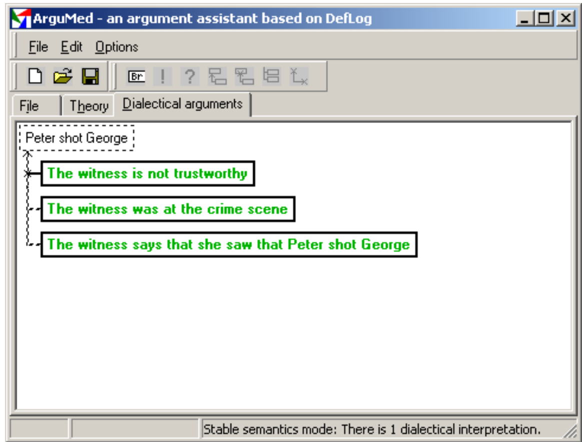

Figure 5: An argument against support in ArguMed

Dung's preferred semantics is based on the notion of admissibility, following the idea that an argument is only acceptable when it can be defended against attacks by other arguments. DefLog's semantics is based on the related notion of dialectical justification, which uses a slightly stronger constraint, namely that an argument is only dialectically justified when it can be defended against all conflicts. Since attacks by other arguments are examples of conflicts, dialectical justification requires, in general, a stronger defense. Dialectical justification implies admissibility, but not in general vice versa. Verheij [21] shows that the two notions coincide when restricting to the expressiveness of Dung's abstract argumentation. To show this, a selection of DefLog's more general theories is made that directly correspond to Dung's abstract argumentation frameworks: these are called Dung theories, defined as DefLog theories that only contain elementary sentences and sentences of the form $\alpha \to \times\beta$, where α and β are elementary. Dung theories have a straightforward translation to a Dung-style abstract argumentation framework, and vice versa: an elementary sentence corresponds to an argument, and a sentence $\alpha \to \times\beta$ to the attack of the argument β by the argument α. Since for Dung theories admissibility and dialectical justification coincide (Proposition 6.1), it follows that the DefLog analogs of preferred and stable extensions correspond exactly to Dung's preferred and stable extensions. It is also explained that when the richer DefLog language is used, admissibility is not a sufficiently strong notion of defense, when one wants to characterize the number of extensions (Theorem 4.3): then the stronger notion of dialectical justification is needed.

5. DefLog has been implemented in the argumentation software ArguMed. ArguMed diagrams and evaluates arguments, and has been assessed in user experiments.

The DefLog language and its formalization have been developed in parallel with a series of implemented argument assistance tools [20, 22]. Figure 5 shows an example of an argument against support in ArguMed (cf. the undercutting argument at the right-hand side of Figure 3). As shown, in ArguMed based on DefLog, reasons for and against conclusions can be composite, i.e., consist of several conjuncts. In Figure 5, the reason for the claim that Peter shot George is that the witness was at the crime scene *and* that the witness says that she saw that Peter shot George. The reason does not justify its conclusion, by the undercutting attack that the witness is not trustworthy.

ArguMed based on DefLog was designed following the experiences in different prototype implementations, and evaluated on the basis of user experiences systematically collected using test protocols. ArguMed implements DefLog's dialectical interpretations, and DefLog's dialectical interpretations are a formally faithful generalization of Dung's semantics (as explained under Property 4 above). As a result, ArguMed based on DefLog is the first argumentation diagramming tool with a formal foundation corresponding to Dung's semantics. Specifically, ArguMed implements the stable semantics when only Dung sentences are used.

4 Value-based arguments in DefLog

The first step towards our proposal to model arguments about values is to show how value-based arguments can be modeled in DefLog. We start with a discussion of Bench-Capon's example [2] of a value-based argumentation framework (VAF) with two values (cf. the example in Figure 1). We propose a translation of a VAF to DefLog using three steps:

1. *Given a VAF, translate the underlying abstract argumentation framework to the corresponding Dung sentences in DefLog.*

   This step models the attack relation between arguments, when the values are ignored. For the example in Figure 1 these sentences are:

   $$A \to \times B; B \to \times C. \tag{4}$$

2. *For each value V promoted by an argument X, add the conditional that expresses the support of V by X.*

   This step models which values are promoted by accepting which arguments. For the example we get:

   $$A \to red; B \to blue; C \to blue. \tag{5}$$

3. *For every pair of arguments X and Y, and pair of values V and W, such that X attacks Y, X promotes V, Y promotes W, and W is preferred to V, add an attack of an attack, namely that the preference attacks that X attacks Y.*

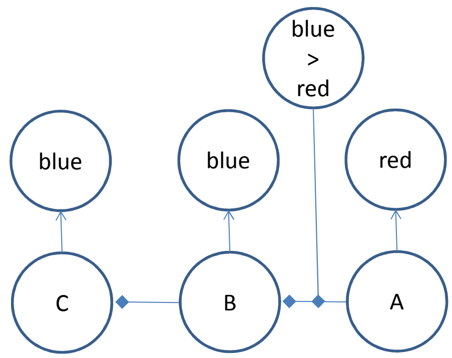

Figure 6: The example of Figure 1; translation to DefLog

This step models that an attack between arguments depends on the value preferences. For the example we get:

$$(blue > red) \to \times(A \to \times B). \tag{6}$$

We make use of a connective $>$ for expressing preferences, that is not a part of DefLog proper. This is unproblematic now that preference statements do not influence the DefLog semantics other than by their use in nested conditionals that result from this translation step. As a result, a sentence $blue > red$ behaves like an elementary proposition, that however by its structure is connected to the sentences $blue$ and red.

This translation of the VAF of Figure 1 is depicted in Figure 6. If we now assume the three arguments A, B, and C, but not the value preference, DefLog evaluates A and C as justified, and B as defeated (by the attack by A). Both values red and $blue$ are justified, expressing that they are promoted: red since A is justified, and $blue$ since C is justified. When we do assume the value preference, DefLog evaluates A and B as justified, as they no longer attack each other, and C as defeated (by the attack by B). Again both values are justified/promoted, but $blue$ is now promoted by B. The results in this DefLog translation correspond to what happens in Bench-Capon's VAF model.

In Figure 7, it is shown how different value preferences, modeling different audiences can be modeled. The two different value preferences indicated exclude each other, representing that they are different choices of an audience. In the added preferential choice $red > blue$, the attack between A and B remains unaffected, and we have that A and C are justified, and B is defeated.

The translations discussed show informally how the apparatus of DefLog can be used to model Bench-Capon's value-based argumentation frameworks.

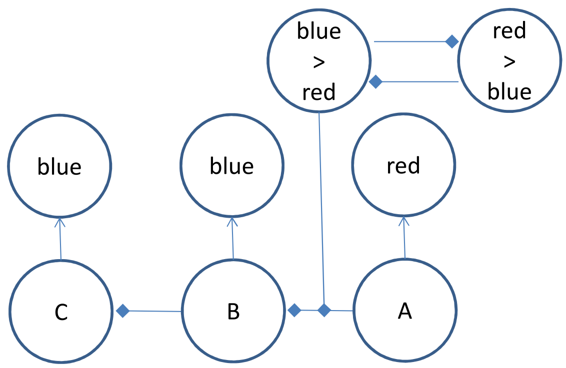

Figure 7: The example of Figure 1; two preference orderings, representing two audiences

A next step would be to evaluate the proposal by developing and proving a precise formal correspondence. Since the translation is so close to Bench-Capon's definitions, and both DefLog and VAFs are close to Dung's abstract argumentation frameworks, formal correspondence results can likely be developed.

5 Arguments about values

The next step of this essay is about the modeling of arguments about values: What if arguments are produced about which values are promoted and which demoted? What if there can be several values promoted or demoted? By the proposal for the modeling of value-based arguments in terms of DefLog's support and attack, as presented in Section 4, the full power of DefLog's nested support and attack model is now available: support of support, support of attack, attack of support, and attack of attack (Figure 4).

In the introduction (Section 1), we discussed the Dutch policy concerning home mortgage interest deduction. It can be argued that such a policy demotes the value of a free house market (V1), and promotes the value of keeping (net) taxes low for higher incomes (V2). A laissez-faire liberal could argue that home mortgage interest deduction demotes the first value, and a VVD-liberal that it promotes the second value. The VVD-liberal can argue against interest deduction demoting the value of a free market by a preference of the value of low taxes for higher incomes to the value of a free house market (though perhaps not too explicitly, because of the risk that the VVD is then more regarded as a party favoring the interests of those who don't need support). A model of the corresponding arguments about values in DefLog is shown in Figure 8.

Note that since DefLog models both support and attack, both the promotion and the demotion of a value can be modeled directly: promotion by support,

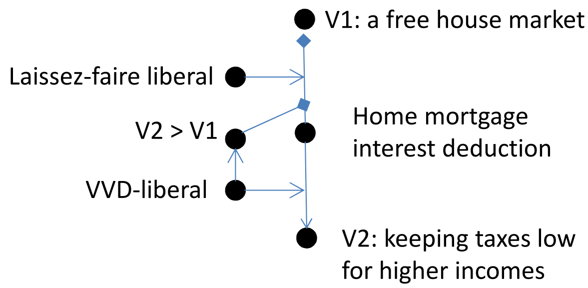

Figure 8: Arguing about values in DefLog

and demotion by attack. The arguments about demotion and promotion of the two types of liberals have been indicated. Also the role of the preference of the VVD liberal is shown; here as an argument against the demotion of the value of a free house market. Their role in determining which policy would be preferred can be modeled as in Section 4. Note also that here the same policy has two associated values, and not just a single value, as in [2].

A side remark: for reasons of national politics, in particular the Dutch necessity of coalition formation, the preferences of VVD-liberals have recently changed slightly, and allowed for a mild reform of the home mortgage interest deduction system.

6 Conclusion: a tribute to Trevor Bench-Capon

I started this text with the explication of an analogy between Aristotle and Bench-Capon. This led me to read the Wikipedia page on the one I know least about.[5] One does not have to read far to find that the analogy is much stronger than just that between the Aristotle and Bench-Capon effects, playfully coined in the introduction. Other matching factors include: productivity; elegant writing; emphasis on dialogue and argument, both in method and in subject matter; emphasis on personal and social values; promoting one's values by living them; the existence of hidden writing gems; being an example to follow and learn from, cf. Aristotle's medieval Muslim nickname المعلم الأول — the First Teacher. The matching factors have a strong correlation to the values that constitute a good life, both in general and in the academy. I am happy that the strength of the value-based analogy makes the following conclusive truth

[5] en.wikipedia.org/wiki/Aristotle; version February 27, 2013.

persuasive. Paraphrasing Cicero on Aristotle:[6]

Flumen orationis aureum fundens Bench-Capon

As the best dialecticians, such as Aristotle, Perelman and Bench-Capon know, but formalists sometimes forget: good form must always follow good content. The beautiful color of Bench-Capon's words is not just in their form, but matches the power of what they express.

References

[1] T. Bench-Capon. Neural networks and open texture. In *Proceedings of the Fourth International Conference on Artificial Intelligence and Law*, pages 292–297. ACM Press, New York (New York), 1993.

[2] T. Bench-Capon. Persuasion in practical argument using value-based argumentation frameworks. *Journal of Logic and Computation*, 13(3):429–448, 2003.

[3] T. Bench-Capon, M. Araszkiewicz, K. Ashley, K. Atkinson, F. Bex, F. Borges, D. Bourcier, D. Bourgine, J. Conrad, E. Francesconi, T. Gordon, G. Governatori, J. Leidner, D. Lewis, R. Loui, L. McCarty, H. Prakken, F. Schilder, E. Schweighofer, P. Thompson, A. Tyrrell, B. Verheij, D. Walton, and A. Wyner. A history of ai and law in 50 papers: 25 years of the international conference on ai and law. *Artificial Intelligence and Law*, 20(3):215–319, 2012.

[4] T. Bench-Capon and E. Rissland. Back to the future: Dimensions revisited. In *Legal Knowledge and Information Systems. Proceedings of JURIX 2001*, pages 41–52. IOS Press, Amsterdam, 2001.

[5] T. Bench-Capon and G. Sartor. A quantitative approach to theory coherence. In *Legal Knowledge and Information Systems. Proceedings of JURIX 2001*, pages 53–62. IOS Press, Amsterdam, 2001.

[6] T. Bench-Capon and G. Sartor. A model of legal reasoning with cases incorporating theories and values. *Artificial Intelligence*, 150(1):97–143, 2003.

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

[8] J. Blakesley. *A Life of Aristotle, Including a Critical Discussion of Some Questions of Literary History Connected with His Works.* J. and J.J. Deighton, Cambridge, 1839.

[6]'Flumen orationis aureum fundens Aristoteles', in Cicero's Academica. One translation is: Aristotle with his golden river of words (cf. [8], 127).

[9] J. Blakesley. *Coherence in Thought and Action.* The MIT Press, Cambridge (Massachusetts), 2001.

[10] A. Bondarenko, P. Dung, R. Kowalski, and F. Toni. An abstract, argumentation-theoretic approach to default reasoning. *Artificial Intelligence*, 93:63–101, 1997.

[11] P. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357, 1995.

[12] C. Perelman and L. Olbrechts-Tyteca. *The New Rhetoric: A Treatise on Argumentation [English translation of La nouvelle rhtorique: Trait de largumentation].* University of Notre Dame Press, Notre Dame (Indiana), 1958/1969.

[13] J. Pollock. Defeasible reasoning. *Cognitive Science*, 11(4):481–518, 1987.

[14] J. Pollock. *Cognitive Carpentry: A Blueprint for How to Build a Person.* The MIT Press, Cambridge (Massachusetts), 1995.

[15] H. Prakken. An abstract framework for argumentation with structured arguments. *Argument and Computation*, 1(2):93–124, 2010.

[16] H. Prakken and G. Sartor. A dialectical model of assessing conflicting arguments in legal reasoning. *Artificial Intelligence and Law*, 4:331–368, 2007.

[17] H. Prakken and G. Sartor. Modelling reasoning with precedents in a formal dialogue game. *Artificial Intelligence and Law*, 6:231–287, 2007.

[18] B. Roth. *Case-Based Reasoning in the Law. A Formal Theory of Reasoning by Case Comparison. Dissertation.* Universiteit Maastricht, Maastricht, 2003.

[19] P. Thagard. *Conceptual Revolutions.* Princeton University Press, Princeton (New Jersey), 1992.

[20] B. Verheij. Artificial argument assistants for defeasible argumentation. *Artificial Intelligence*, 150(1–2):291–324, 2003.

[21] B. Verheij. DefLog: on the logical interpretation of prima facie justified assumptions. *Journal of Logic and Computation*, 13(3):319–346, 2003.

[22] B. Verheij. *Virtual Arguments. On the Design of Argument Assistants for Lawyers and Other Arguers.* T.M.C. Asser Press, The Hague, 2005.

[23] B. Verheij. Jumping to conclusions. a logico-probabilistic foundation for defeasible rule-based arguments. In L. Fariñas del Cerro, A. Herzig, and J. Mengin, editors, *13th European Conference on Logics in Artificial Intelligence, JELIA 2012. Toulouse, France, September 2012. Proceedings (LNAI 7519)*, pages 411–423. Springer, Berlin, 2012.