Legal Knowledge and Information Systems
F. Bex and S. Villata (Eds.)
IOS Press, 2016

doi:10.3233/978-1-61499-726-9-101

Arguments for Ethical Systems Design

Bart VERHEIJ [1]
Artificial Intelligence, University of Groningen, The Netherlands

AbstractToday's AI applications are so successful that they inspire renewed concerns about AI systems becoming ever more powerful. Addressing these concerns requires AI systems that are designed as ethical systems, in the sense that their choices are context-dependent, value-guided and rule-following. It is shown how techniques connecting qualitative and quantitative primitives recently developed for evidential argumentation in the law can be used for the design of such ethical systems. In this way, AI and Law techniques are extended to the theoretical understanding of intelligent systems guided by embedded values.

1. Introduction

Once Artificial Intelligence was science fiction, and the study of ethical AI could be left to creative speculation in novels and films. A good example of a fictional ethical system appears in Verhoeven's 1987 film Robocop, where the choices of a cyborg police officer are guided by three 'prime directives':

1. Serve the public trust;
2. Protect the innocent;
3. Uphold the law.

These directives—inspired by Asimov's 1942 Three Laws of Robotics—guide Robocop's behavior, but the plot involves several twists where ethical choices based on Robocop's personal values must be made.

Today Artificial Intelligence is a science with real life applications, and the investigation of ethical AI should be done systematically by scientists and engineers. Autonomous systems for driving and warfare must do the right thing in complex, unforeseeable situations. The design of social media asks for a careful balance between what is good for users and for businesses. The invention of virtual currencies and related blockchain-based technology inspires the automation of trust mechanisms in finance and other businesses.

Advanced intelligent techniques operate in problem domains that involve the complex ethical decision-making that people perform routinely everyday. And even though we make many mistakes—often enough with extremely bad consequences—, humans outperform all other natural and artificial systems in real-life ethical decision making. Only we can choose our actions while carefully

[1]Email: b.verheij@ai.rug.nl.

considering the context, taking human values into account, and following normative rules.

The state-of-the-art in artificial systems with ethical impact in use today is what have been called *implicit ethical systems* (Moor, 2006): they are limited by design to performing only the right kind of behavior. Think of an ATM that is carefully designed in order to give money only to the person entitled to receiving it. When Silicon Valley speaks of ethical system design today, this typically concerns such implicit ethical systems (see, e.g., the interesting and relevant TEDx talk by Tristan Harris[2] on systems incorporating human values).

In contrast, Moor speaks of *full ethical systems* when they have an embedded explicit ethical model. Such a model allows a system to make autonomous judgments it can justify, in the face of conflicting ethical considerations. There is a slow shift of attention towards full ethical systems, but the technological hurdles are huge and require fundamental research (see also Broersen, 2014, who emphasises responsibility in intelligent systems).

We study ethical decision making using an argumentation perspective, focusing on three themes:

Context-dependence *An ethical system's decisions depend on the circumstances.* What counts as a good decision in one situation, may not be good in another, similar situation. Similarities and differences between the circumstances of situations can determine what counts as a good decision. For instance, when driving a car, an abrupt stop can be a good choice to avoid a collision in front of you, but maybe not when someone is close behind you.

Value-dependence *An ethical system's decisions depend on values embedded in the system.* A system's decisions are not determined by the external circumstances alone. There is room for discretionary choices depending on the values embedded in the system. For instance, when driving a car, some base their choices more on speed, others more on safety.

Rule-dependence *An ethical system's decisions depend on rules embedded in the system.* A system's decisions are typically not made on a case by case basis, but follow rules. For instance, when driving in a suburban area, as a rule you reduce your speed. It does not matter much which suburban area you are in, not even whether you have been there before.

Our specific focus concerns the comparison of values and its role in decision making. Values are typically studied using either qualitative or quantitative modeling primitives. For instance, values are modeled as a qualitative logical property that can either be promoted or demoted when a decision is made (as, e.g., in value-based argumentation frameworks by Bench-Capon, 2003). Alternatively, values are handled using quantitative numeric properties such as the probability that a consequence follows and the utility of a decision (as, e.g., in expected utility theory). In recent research on evidential argumentation (Verheij, 2014, 2016b; Verheij et al., 2016), techniques have been developed for the connection between qualitative or quantitative modeling primitives. In this paper, these techniques are applied to the comparison of values in ethical decision making, emphasising

[2] `www.ted.com/talks/tristan_harris_how_better_tech_could_protect_us_from_distraction`.

the role of context-dependence, value-dependence and rule-dependence. In this way, we provide a perspective on ethical decision making as value-guided argumentation.

2. Formalism

The following formal perspective has been developed in recent research on evidential argumentation, in order to bridge between qualitative or quantitative modeling primitives, in particular arguments, scenarios and probabilities (Verheij, 2014, 2016b; Verheij et al., 2016), building on (Verheij, 2010, 2012). In subsequent sections, we show how the formalism also can be put to work for ethical decision making and its context-dependence, value-dependence and rule-dependence.

2.1. General Idea

The formalism models arguments that can be presumptive (also called ampliative), in the sense of logically going beyond their premises. Against the background of classical logic, an argument from premises P to conclusions Q goes beyond its premises when Q is not logically implied by P. Many arguments used in practice are presumptive. For instance, the prosecution may argue that a suspect was at the crime scene on the basis of a witness testimony. The fact that the witness has testified as such does not logically imply the fact that the suspect was at the crime scene. In particular, when the witness testimony is intentionally false, based on inaccurate observations or inaccurately remembered, the suspect may not have been at the crime scene at all. Denoting the witness testimony by P and the suspect being at the crime scene as Q, the argument from P to Q is presumptive since P does not logically imply Q.

For presumptive arguments, it is helpful to consider the *case made by the argument*, defined as the conjunction of the premises and conclusions of the argument (Verheij, 2010, 2012). The case made by the argument from P to Q is $P \wedge Q$, using the conjunction of classical logic. An example of a non-presumptive argument goes from $P \wedge Q$ to Q. Here Q is logically implied by $P \wedge Q$. Presumptive arguments are often defeasible (Pollock, 1987; Toulmin, 1958), in the sense that extending the premises may lead to the retraction of conclusions.

Figure 1 shows two presumptive arguments from the same premises P: one supports the case $P \wedge Q$, the other the case $P \wedge \neg Q$. The $>$-sign indicates that one argument makes a stronger case than the other, resolving the conflict: the argument for the case $P \wedge Q$ is stronger than that for $P \wedge \neg Q$. The figure also shows two presumptions P and $\neg P$, treated as arguments from logically tautologous premises. Here the presumption $\neg P$ makes the strongest case when compared to the presumption P. Logically such presumptions can be treated as arguments from logical truth $\top$. The arguments make three cases: $\neg P$, $P \wedge Q$ and $P \wedge \neg Q$ (Figure 2). The size of their areas suggest a preference relation.

The comparison of arguments and of cases are closely related in our approach, which can be illustrated as follows. The idea is that a case is preferred to another case if there is an argument with premises that supports the former case more

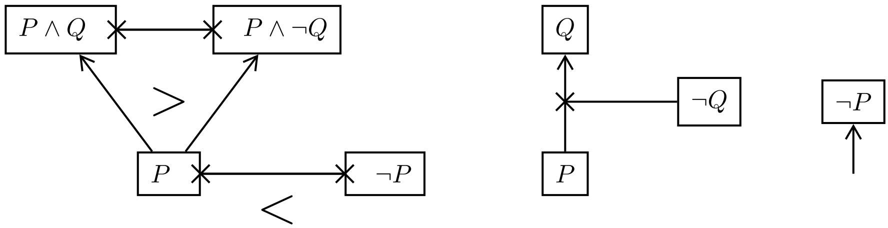

Figure 1. Some arguments

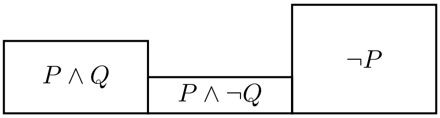

Figure 2. Some cases

strongly than the latter case. Hence, in the example in the figures, $\neg P$ is preferred to both $P \wedge Q$ and $P \wedge \neg Q$, and $P \wedge Q$ is preferred to $P \wedge \neg Q$. Conversely, given the cases and their preferences, we can compare arguments. The argument from P to Q is stronger than from P to $\neg Q$ when the best case that can be made from $P \wedge Q$ is preferred to the best case that can be made from $P \wedge \neg Q$.

2.2. Case Models and Arguments

We now formalize case models and how they can be used to interpret arguments. The formalism uses a classical logical language L generated from a set of propositional constants in a standard way. We write $\neg$ for negation, $\wedge$ for conjunction, $\vee$ for disjunction, $\leftrightarrow$ for equivalence, $\top$ for a tautology, and $\bot$ for a contradiction. The associated classical, deductive, monotonic consequence relation is denoted $\models$. We assume a finitely generated language. First we define case models, formalizing the idea of cases and their preferences. The cases in a case model must be logically consistent, mutually incompatible and different; and the comparison relation must be total and transitive (hence is what is called a total preorder, commonly modeling preference relations (Roberts, 1985)).

Definition 1. A *case model* is a pair $(C, \geq)$ with finite $C \subseteq L$, such that the following hold, for all φ, ψ and $\chi \in C$:

1. $\not\models \neg\varphi$;
2. If $\not\models \varphi \leftrightarrow \psi$, then $\models \neg(\varphi \wedge \psi)$;
3. If $\models \varphi \leftrightarrow \psi$, then $\varphi = \psi$;
4. $\varphi \geq \psi$ or $\psi \geq \varphi$;
5. If $\varphi \geq \psi$ and $\psi \geq \chi$, then $\varphi \geq \chi$.

The strict weak order $>$ standardly associated with a total preorder $\geq$ is defined as $\varphi > \psi$ if and only if it is not the case that $\psi \geq \varphi$ (for φ and $\psi \in C$). When $\varphi > \psi$, we say that φ is (strictly) preferred to ψ. The associated equivalence relation $\sim$ is defined as $\varphi \sim \psi$ if and only if $\varphi \geq \psi$ and $\psi \geq \varphi$.

Example 2. Figure 2 shows a case model with cases $\neg P$, $P \wedge Q$ and $P \wedge \neg Q$. $\neg P$ is (strictly) preferred to $P \wedge Q$, which in turn is preferred to $P \wedge \neg Q$.

Next we define arguments from premises $\varphi \in L$ to conclusions $\psi \in L$.

Definition 3. An *argument* is a pair (φ, ψ) with φ and $\psi \in L$. The sentence φ expresses the argument's premises, the sentence ψ its conclusions, and the sentence $\varphi \wedge \psi$ the *case made* by the argument. Generalizing, a sentence $\chi \in L$ is a *premise* of the argument when $\varphi \models \chi$, a *conclusion* when $\psi \models \chi$, and a *position* in the case made by the argument when $\varphi \wedge \psi \models \chi$. An argument (φ, ψ) is *(properly) presumptive* when $\varphi \not\models \psi$; otherwise *non-presumptive.* An argument (φ, ψ) is a *presumption* when $\models \varphi$, i.e., when its premises are logically tautologous.

Note our use of the plural for an argument's premises, conclusions and positions. This terminological convention allows us to speak of the premises $\mathtt{p}$ and $\neg\mathtt{q}$ and conclusions $\mathtt{r}$ and $\neg\mathtt{s}$ of the argument $(\mathtt{p} \wedge \neg\mathtt{q}, \mathtt{r} \wedge \neg\mathtt{s})$. Also the convention fits our non-syntactic definitions, where for instance an argument with premise χ also has logically equivalent sentences such as $\neg\neg\chi$ as a premise.

Coherent arguments are defined as arguments that make a case that is logically implied by a case in the case model. Conclusive arguments are defined as coherent arguments with the property that each case that implies the argument's premises also implies the argument's conclusions.

Definition 4. Let $(C, \geq)$ be a case model. Then we define, for all φ and $\psi \in L$:

$(C, \geq) \models (\varphi, \psi)$ if and only if $\exists \omega \in C$: $\omega \models \varphi \wedge \psi$.

We then say that the argument from φ to ψ is *coherent* with respect to the case model. We define, for all φ and $\psi \in L$:

$(C, \geq) \models \varphi \Rightarrow \psi$ if and only if $\exists \omega \in C$: $\omega \models \varphi \wedge \psi$ and $\forall \omega \in C$: if $\omega \models \varphi$, then $\omega \models \varphi \wedge \psi$.

We then say that the argument from φ to ψ is *conclusive* with respect to the case model.

Example 5. (continued from Example 2) In the case model of Figure 2, the arguments from $\top$ to $\neg P$ and to P, and from P to Q and to $\neg Q$ are coherent and not conclusive in the sense of this definition. Denoting the case model as $(C, \geq)$, we have $(C, \geq) \models (\top, \neg P)$, $(C, \geq) \models (\top, P)$, $(C, \geq) \models (P, Q)$ and $(C, \geq) \models (P, \neg Q)$. The arguments from a case (in the case model) to itself, such as from $\neg P$ to $\neg P$, or from $P \wedge Q$ to $P \wedge Q$ are conclusive. The argument $(P \vee R, P)$ is also conclusive in this case model, since all $P \vee R$-cases are P-cases. Similarly, $(P \vee R, P \vee S)$ is conclusive.

The notion of presumptive validity considered here uses the idea that some arguments make a better case than other arguments from the same premises. More precisely, an argument is presumptively valid if there is a case implying the case made by the argument that is at least as preferred as all cases implying the premises.

Definition 6. Let $(C, \geq)$ be a case model. Then we define, for all φ and $\psi \in L$:

$(C, \geq) \models \varphi \rightsquigarrow \psi$ if and only if $\exists \omega \in C$:

1. $\omega \models \varphi \wedge \psi$; and
2. $\forall \omega' \in C$: if $\omega' \models \varphi$, then $\omega \geq \omega'$.

We then say that the argument from φ to ψ is *(presumptively) valid* with respect to the case model. A presumptively valid argument is *defeasible*, when it is not conclusive.

3. Dependence on Contexts, Values and Rules

We now discuss the examples used in the introduction to illustrate the context-dependence, value-dependence and rule-dependence of ethical decision-making.

Example 7 (Context-dependence)**.** Context-dependence was illustrated with the example that there is a sudden risk of collision while driving on the highway, an abrupt stop can be a good idea, but not when there is someone close behind you. Then it is better to slow down by careful braking. A case model $(C, \geq)$ for this example consists of three cases:

Case 1: CONTINUE-DRIVING $\wedge$ ¬ABRUPT-STOP $\wedge$ ¬CAREFUL-BREAKING
Case 2: ¬CONTINUE-DRIVING $\wedge$ ABRUPT-STOP $\wedge$ ¬CAREFUL-BREAKING $\wedge$ RISK-OF-COLLISION
Case 3: ¬CONTINUE-DRIVING $\wedge$ ¬ABRUPT-STOP $\wedge$ CAREFUL-BREAKING $\wedge$ RISK-OF-COLLISION $\wedge$ SOMEONE-CLOSE-BEHIND
Case 1 > Case 2 > Case 3

Case 1 is the normal situation of continuing to drive. It is the maximally preferred case, hence is the default situation:

$(C, \geq) \models \top \rightsquigarrow$ CONTINUE-DRIVING

It holds that RISK-OF-COLLISION presumptively implies ABRUPT-STOP, but not when also SOMEONE-CLOSE-BEHIND. Formally:

$(C, \geq) \models$ RISK-OF-COLLISION $\rightsquigarrow$ ABRUPT-STOP
$(C, \geq) \not\models$ RISK-OF-COLLISION $\wedge$ SOMEONE-CLOSE-BEHIND $\rightsquigarrow$ ABRUPT-STOP
$(C, \geq) \models$ RISK-OF-COLLISION $\wedge$ SOMEONE-CLOSE-BEHIND $\rightsquigarrow$ ¬ABRUPT-STOP

Example 8 (Value-dependence)**.** Value-dependence was illustrated with some drivers valuing speed more highly, and others safety. Assuming that maximizing the values of speed and safety are competing purposes to strive for (while driving), we can consider the following three cases in a case model.

Case 1: DRIVE $\wedge$ MAXIMIZE-SPEED $\wedge$ ¬MAXIMIZE-SAFETY
Case 2: DRIVE $\wedge$ ¬MAXIMIZE-SPEED $\wedge$ MAXIMIZE-SAFETY
Case 3: ¬DRIVE

The preference relation determines which choice is made. When the two cases are equally preferred, we have that both MAXIMIZE-SPEED and MAXIMIZE-SAFETY are presumptively valid conclusions. When Case 1 is preferred over the other, only MAXIMIZE-SPEED presumptively follows; when Case 2 is preferred, only MAXIMIZE-SAFETY. Formally:

When Case 1 $\sim$ Case 2:
$(C, \geq) \models$ DRIVE $\rightsquigarrow$ MAXIMIZE-SPEED
$(C, \geq) \models$ DRIVE $\rightsquigarrow$ MAXIMIZE-SAFETY.
When Case 1 $>$ Case 2:
$(C, \geq) \models$ DRIVE $\rightsquigarrow$ MAXIMIZE-SPEED
$(C, \geq) \not\models$ DRIVE $\rightsquigarrow$ MAXIMIZE-SAFETY.
When Case 1 $<$ Case 2:
$(C, \geq) \not\models$ DRIVE $\rightsquigarrow$ MAXIMIZE-SPEED
$(C, \geq) \models$ DRIVE $\rightsquigarrow$ MAXIMIZE-SAFETY.

When Case 1 $\sim$ Case 2, it does not presumptively follow that MAXIMIZE-SPEED $\wedge$ MAXIMIZE-SAFETY since the (And)-rule does not hold for presumptive validity. When there is no preference for driving or not-driving, Case 3 is preferentially equivalent to both Case 1 and Case 2 (when they are equivalent) or to the preferred case (when one is preferred over the other).

Example 9 (Rule-dependence)**.** Rule-dependence was illustrated with the reduced speed limit in residential areas. The following case model shows four different suburban areas A, B, C and D and their speed limits.

Case 1: RESIDENTIAL $\wedge$ SPEED-LIMIT-30 $\wedge$ AREA-A
Case 2: RESIDENTIAL $\wedge$ SPEED-LIMIT-30 $\wedge$ AREA-B
Case 3: SPEED-LIMIT-30 $\wedge$ AREA-C
Case 4: SPEED-LIMIT-50 $\wedge$ AREA-D
Case 1 $\sim$ Case 2 $<$ Case 3 $\sim$ Case 4
Background theory: $\neg$(AREA-A $\wedge$ AREA-B) $\wedge$ $\neg$(AREA-A $\wedge$ AREA-C))
$\wedge$ $\neg$(AREA-A $\wedge$ AREA-D)) $\wedge$ $\neg$(AREA-B $\wedge$ AREA-C)
$\wedge$ $\neg$(AREA-B $\wedge$ AREA-D) $\wedge$ $\neg$(AREA-C $\wedge$ AREA-D)
$\wedge$ $\neg$(SPEED-LIMIT-30 $\wedge$ SPEED-LIMIT-50)

The preference relation is meant to suggest that the residential areas A and B are exceptional. A separate background theory sentence is specified that holds in all four cases. It expresses that the four suburban areas are different and that there is only one speed limit. Here SPEED-LIMIT-30 follows presumptively (even conclusively) from AREA-A and from AREA-B. We find that in this case model the rule holds that in residential areas the speed limit is 30 km/h. The rule is both presumptively and conclusive valid:

$(C, \geq) \models$ RESIDENTIAL $\rightsquigarrow$ SPEED-LIMIT-30
$(C, \geq) \models$ RESIDENTIAL $\Rightarrow$ SPEED-LIMIT-30

The reversed rule with antecedent and consequent switched is not presumptively valid (hence also not conclusively):

$(C, \geq) \not\models$ SPEED-LIMIT-30 $\rightsquigarrow$ RESIDENTIAL

4. Discussion

We have studied decision making and its dependence on contexts, values and rules. Contexts are present in our use of formalized cases, that can be considered as representing the relevant properties of a situation, possible or real. The values appear in the preference ordering on the cases in case models. The preferences help to make a choice of maximal value. The role of rules comes about when we consider how case models give rise to notions of presumptively and conclusively valid arguments with a conditional form.

Bench-Capon built his value-based argumentation frameworks on top of Dung's abstract argumentation, a natural choice by the innovative technical possibilities allowed by that formalism. Our approach is not based on abstract argumentation, but has been developed in a way to stay close to classical logic and standard probability theory (see Verheij 2012, 2014, 2016a,b). Bench-Capon modeled the promotion and demotion of values as an argument selection mechanism. In our model, the promotion and demotion of values appears in the arguments that are conclusively and presumptively valid given the premises.

Here we have not addressed reasoning about values, as we did in (Verheij, 2013). There we built on an argumentation formalism (DefLog), a model extending Dung's abstract argumentation with support and with support/attack about support/attack by the use of nested conditionals. Here we have not included such reasoning in our discussions. It can be noted that nested conditionals such as $P \to (Q \to R)$ play a role in reasoning that is in relevant ways similar to the conditional with a composite antecedent $P \wedge Q \to R$. Concretely, for the nested conditional and for the conditional-with-composite-antecedent, one expects that when both P and Q hold, R follows. The conditional-with-composite-antecedent has been studied in the present paper, in its presumptive and conclusive forms $P \wedge Q \rightsquigarrow R$ and $P \wedge Q \Rightarrow R$. One idea would be to define $P \rightsquigarrow (Q \rightsquigarrow R)$ and $P \Rightarrow (Q \Rightarrow R)$ as these conditionals-with-composite-antecedent. In collaboration with Modgil, Bench-Capon has developed his value-based argumentation frameworks to the modeling of arguments about value preferences (Bench-Capon and Modgil, 2009; Modgil and Bench-Capon, 2011). In contrast with thise models, the present stays close to logic and probability logic, whereas they work with adaptations of abstract argumentation.

Another kind of model has been developed by Atkinson and Bench-Capon who focused on practical reasoning about which actions to choose (Atkinson and Bench-Capon, 2006, 2007), where they use Belief–Desire–Intention (BDI) modeling, Action-Based Alternating Transition Systems (AATS) and argumentation schemes . These approaches are very relevant for the present work, now that the kind of decision making studied here has close similarities to practical reasoning. However, intentional aspects (associated with BDI modeling), coordination between agents (as studied in AATS modeling) and dialogical themes (as they naturally arise when studying argument schemes and their critical questions) are beyond the scope of the present abstract model.

By the use of case models, the present work has connections to case-based reasoning in the law more generally. For instance, there are clear connections to Rissland and Ashley's work (Ashley, 1990; Rissland and Ashley, 1987, 2002).

The elementary propositions of the logical language used to express the cases in our case models are closely related to their factors, although the latter are pro-plaintiff or pro-defendant, and ours are not. Whether a proposition is pro-plaintiff or pro-defendant would have to be determined on the basis of other information in the case model. For instance, if a factor F is pro-plaintiff P, this can be thought of as the conditional $F \Rightarrow P$ being valid in the case model. Or, allowing for a factor being hypothetically for a side in the debate, $F \rightsquigarrow P$ could be valid. Our approach does not distinguish the dimensionality that come with factors, although dimensions add significantly to the expressiveness and relevance of a set of modeling tools. Since our model is connected to the bridging of qualitative and quantitative modeling primitives, it may be interesting to apply the model here to dimensions. A key difference between Rissland and Ashley's work and the present is that we stay close to logic and probability theory, and develop a theory of conclusive and presumptive validity.

5. Concluding Remarks

The paper started with the ethical dimension of AI, and discussed how advances in technology necessitate that systems develop to full ethical systems, in the sense that they can make decisions while taking the relevant context, human values and normative rules into account. We showed how a formalism developed for bridging qualitative and quantitative primitives in evidential reasoning can be applied to value-guided argumentation grounded in cases.

The results are relevant for ethical system design, as one way of looking at ethical system design is as technology that is better suited for who we are as humans. A simple example could be a smartphone that does not make sounds during the times that we are supposed to be sleeping, or better yet: that does not give immediate access to email and facebook during those times. Such interruptions can be fine, and can under circumstances even be rational, but most often it is best to sleep at night. Autonomous driving requires ethical decisions of significantly greater complexity. Always ethical systems should be aware of their relevant context, have embedded values, and use the rules that apply in order to to what is right. Ethical system design is the way of the future, and here some suggestions have been made for their formal foundations.

References

Ashley, K. D. (1990). *Modeling Legal Arguments: Reasoning with Cases and Hypotheticals.* The MIT Press, Cambridge (Massachusetts).

Atkinson, K. and Bench-Capon, T. J. M. (2006). Legal case-based reasoning as practical reasoning. *Artificial Intelligence and Law*, 13:93–131.

Atkinson, K. and Bench-Capon, T. J. M. (2007). Practical reasoning as presumptive argumentation using action based alternating transition systems. *Artificial Intelligence*, 171:855–874.

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

Bench-Capon, T. J. M. and Modgil, S. (2009). Case law in extended argumentation frameworks. In *Proceedings of the 12th International Conference on Artificial Intelligence and Law (ICAIL 2009)*, pages 118–127. ACM Press, New York.

Broersen, J. (2014). Responsible intelligent systems. *Künstliche Intelligenz*, 28:209–214.

Modgil, S. and Bench-Capon, T. J. M. (2011). Metalevel argumentation. *Journal of Logic and Computation*, 21(6):959–1003.

Moor, J. H. (2006). The nature, importance, and difficulty of machine ethics. *IEEE Intelligent Systems*, 2006(July/August):18–21.

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

Rissland, E. L. and Ashley, K. D. (1987). A case-based system for trade secrets law. In *Proceedings of the First International Conference on Artificial Intelligence and Law*, pages 60–66. ACM Press, New York (New York).

Rissland, E. L. and Ashley, K. D. (2002). A note on dimensions and factors. *Artificial Intelligence and Law*, 10:65–77.

Roberts, F. S. (1985). *Measurement Theory with Applications to Decisionmaking, Utility, and the Social Sciences.* Cambridge University Press, Cambridge.

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

Verheij, B. (2010). Argumentation and rules with exceptions. In Baroni, B., Cerutti, F., Giacomin, M., and Simari, G. R., editors, *Computational Models of Argument: Proceedings of COMMA 2010, Desenzano del Garda, Italy, September 8-10, 2010*, pages 455–462. IOS Press, Amsterdam.

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

Verheij, B. (2013). Arguments about values. In *From Knowledge Representation to Argumentation in AI, Law and Policy Making. A Festschrift in Honour of Trevor Bench-Capon on the Occasion of his 60th Birthday*, pages 243–257. College Publications, London.

Verheij, B. (2014). To catch a thief with and without numbers: Arguments, scenarios and probabilities in evidential reasoning. *Law, Probability and Risk*, 13:307–325.

Verheij, B. (2016a). Correct grounded reasoning with presumptive arguments. In Michael, L. and Kakas, A., editors, *15th European Conference on Logics in Artificial Intelligence, JELIA 2016. Larnaca, Cyprus, November 9–11, 2016. Proceedings (LNAI 10021).* Springer, Berlin.

Verheij, B. (2016b). Formalizing correct evidential reasoning with arguments, scenarios and probabilities. In *Proceedings of the Workshop on Artificial Intelligence for Justice (AI4J 2016) at ECAI 2016*, pages 87–95.

Verheij, B., Bex, F. J., Timmer, S. T., Vlek, C. S., Meyer, J. J., Renooij, S., and Prakken, H. (2016). Arguments, scenarios and probabilities: Connections between three normative frameworks for evidential reasoning. *Law, Probability and Risk*, 15:35–70.