

Norms and Extended Argumentation Frameworks

Trevor Bench-Capon

Department of Computer Science, University of Liverpool
Liverpool, UK

Sanjay Modgil

Department of Computer Science, King's College London
London, UK

ABSTRACT

This paper discusses reasoning about norms using Extended Argumentation Frameworks. We argue that norms emerge from a process of value based practical reasoning. We model practical reasoning using Action Based Alternating Transition Systems with Values, with the resulting arguments evaluated in an Extended Argumentation Framework, which permits arguments to attack attacks as well as arguments, and so provides a natural way of representing exceptions. Following our consideration of how norms are justified, we consider their application. Our representation of norms in EAFs enables us to offer an accessible account of permissions and exceptions. We illustrate our approach with an example relating to some US Supreme Court Fourth Amendment cases.

CCS CONCEPTS

• Applied computing → Law.

KEYWORDS

norms, exceptions, values, extended argumentation frameworks

ACM Reference Format:

Trevor Bench-Capon and Sanjay Modgil. 2019. Norms and Extended Argumentation Frameworks. In *Seventeenth International Conference on Artificial Intelligence and Law (ICAIL '19)*, June 17–21, 2019, Montreal, QC, Canada. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3322640.3326696>

1 INTRODUCTION

State transition diagrams are commonly used in multi-agent systems to represent the actions available to the agents, and the norms applicable to them (e.g. [4]). One common way of representing prohibitions (e.g. [1]) is to remove the transition representing the forbidden action. But this makes norms implicit and, worse, prevents any possibility of violation, which is central to any discussion of norms [14]. Essentially this way of representing norms reduces *should not* to *can not*. Our approach, following [8] is to use the transition diagrams to generate value based arguments, as in [2]. In this way norms can be seen to emerge from the process of practical reasoning, reasoning about what actions should be chosen, and can then be represented as arguments for and against the actions they concern. We use Extended Argumentation Frameworks (EAF) [17] to represent and evaluate the relevant arguments, to enable

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ICAIL '19, June 17–21, 2019, Montreal, QC, Canada

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-6754-7/19/06...\$15.00

<https://doi.org/10.1145/3322640.3326696>

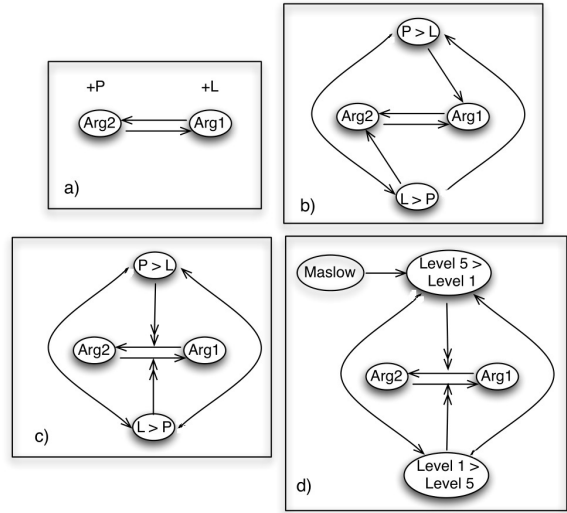


Figure 1: a) Basic VAF for Summer b) Basic VAF for Summer with arguments expressing value preferences. c) Basic EAF for Summer. d) EAF for Summer with Maslow.

reasoning about both what the norms should be, and, once the norms have emerged, how they should be applied. EAFs not only enable a more expressive and elegant representation of preferences, but also provide a very natural representation of exceptions, which are of great importance in any discussion of legal norms.

Our approach is based on [8]. The idea is that norms emerge through a process of value based practical reasoning [2]. The reasoning may be performed by an individual, citizens of a state, members of a social community. etc, and may give rise to prudential, legal, social and moral norms as appropriate. The example in [8] is taken from [11] and is based on a scenario taken from the parable of the *Ant and the Grasshopper*¹. In that scenario, in Summer the agents have a choice of whether to work, and build up a stock of food, or to play, when they will have no food for the Winter. In the fable, the grasshopper plays while the ant works, and when Winter comes asks the ant for food, but is refused and dies.

In Summer there are two arguments:

- Arg 1: You should work to avoid demoting life
- Arg 2: You should play to promote pleasure

These arguments conflict as shown in Figure 1(a).

Which is chosen will depend on the value preferences. If we suppose a preference for life over pleasure, Arg1 will be chosen,

¹One of Aesop's Fables, numbered 373 in the Perry Index.

since Arg2 will be attacked by this preference. The contrary preference would attack Arg1 and so Arg 2 would be chosen. One can express these attacks in a basic Value Based Argumentation Framework (VAF), as shown in Figure 1(b). However, Extended Argument Frameworks (EAFs) [17] provide a more elegant representation [5] (shown in Figure 1(c)) as EAFs allow attacks not only on arguments, but also on attacks. The preferences more properly attack not the arguments, but the attacks between arguments. Thus, if for example, given the preference for Life, some other argument emerged so as to directly attack Arg1, then Arg2 would now be acceptable (i.e., justified) in the EAF in Figure 1(c), whereas Arg2 would remain unjustified in the VAF in (b).

Definitively choosing a winning argument does, however, rely on appealing to a 'subjective' audience encoded in the alternative value orderings. An EAF that only includes the argument (audience) $L > P$, and that therefore excludes $P > L$, will render Arg1 as uniquely acceptable (winning). On the other hand a community of sybarites might well feel that a short and merry life is better than a long life, and so choose to play (so that the EAF excludes $L > P$ and includes $P > L$). We can however attempt to provide objective groundings for the preferences by talking not of *values* but of *needs*. If we adopt Maslov's hierarchy of needs [16] and prefer the more basic needs, we can justify choosing $L > P$ since life relates to level 1 needs and pleasure of this sort to level 5 needs. Now, as shown in Figure 1(d) Maslow will attack the preference for pleasure, and Arg1 will be the only acceptable argument.

We now follow Hare [13] to move from a one-off argument to a norm. Hare argues that, in a novel situation, we need to reason from first principles, but having arrived at a conclusion we "crystallise it into a not too specific or detailed form, so that its salient features may stand out and serve us again in a like situation without so much thought." In other words, having identified a moral principle or norm it can be applied without reference to the arguments which originally gave rise to it.

This means that the argument can be encapsulated as the norm:

- N1: It is obligatory to work in Summer.

Note that it is this kind of thinking that underlies Rawls' Theory of Justice [18]: under his notions of the *veil of ignorance* and the *difference principle* citizens will choose to adopt N1 [9].

If we move forward to the Winter, the ant has a choice of whether to give or refuse food to the grasshopper. Although giving would save the grasshopper's life, it has been shown in [15] that if violations of norms are not punished, normative collapse is inevitable. Therefore a second norm is needed to protect N1:

- N2: It is forbidden to give food in Winter.

This will ensure that violations of N1 are punished, or else licence sanctions against the violators of N2 and enable a stable situation. A more detailed discussion of the emergence of these norms can be found in [9].

2 REPRESENTING NORMS AND EXCEPTIONS

Once we have decided what the norms should be, we need to represent them so that they can directly influence the choices of agents, rather than requiring them to derive the norms from first principles every time. The arguments used in the frameworks so far have

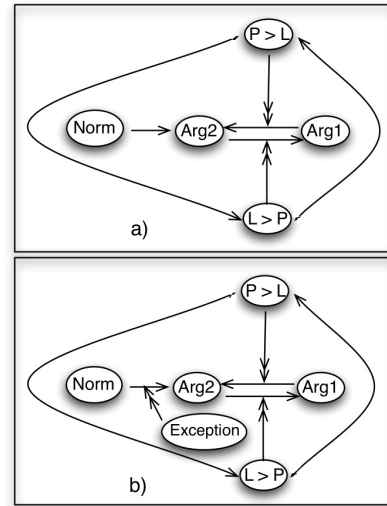


Figure 2: a) EAF in which Arg1 encodes the action enjoined by Norm. b) EAF with norm and exception.

been of the form *citizens should perform a particular action in order to fulfill some need*. Promulgations of norms are essentially the actions available to the state to influence behaviour of its citizens. The actions available to the citizens are (in State Transition terms) to follow transitions. In many multi-agent systems (MAS) which use norms (such as [1]) norms are represented by simply removing the transitions corresponding to prohibitions. Obligations (if represented) are represented by removing all transitions from a state *except* the one that is obligated. Permissions seem rather mysterious on this approach, since the existence of the transition makes them permissible without a norm². The real problem, however, with removing transitions in this way is that it removes any possibility of violation, while the point of representing norms is precisely so that violations can be reasoned about [14]³.

In this spirit we first represent norms as attacks on prohibited actions, as shown in Figure 2a. Note that we have replaced the reasons for introducing the norm, whether based on Rawl's difference principle, or using preferences based on the Maslow hierarchy, by the argument *Norm*, expressing that *work is obligatory*. Once the norm has been introduced it *replaces* the arguments which gave rise to it: citizens are expected to accept the norm, rather than dispute its justification. As Hare [13] suggests, having developed the moral principle or norm it can be applied without reference to the *arguments* which originally gave rise to it.

The problem with the representation of Figure 2a, as noted in [6] and [7], is that the norm *always* excludes the performance of

²Although deontic logic has always focussed on obligation and permission, in practice prohibitions are in many situations the more usual way of expressing norms (see e.g. the decalogue in 20: 1 -17). This focus on forbidding actions (even when the aim is to avoid undesirable states as in [20]) is widely adopted also in MAS.

³There are sophisticated deontic logic treatments of permissions, such as [12]. Their use of defeaters in defeasible deontic logic to represent permissions has many similarities to the representation of exceptions as attacks on attacks as proposed here.

the action it forbids, whereas it must be recognised that sometimes complying with the norm will lead to undesirable results. For example, some people will be incapable of work, and they should not be punished for not working since it is not their fault. In such circumstances it should be permissible, or even obligatory, to ignore N1. Thus norms are made in the knowledge that they will be violated, and in certain circumstances, the norms admit derogations/exemptions. EAFs provide a very natural way to represent this, by having the exception attack the attack originating from the norm, as shown in Figure 2b. Thus the exception means that either action is permissible, and agents can again choose according to their own value preferences. Note that an exception removes the compulsion on one action, but does not compel the other, which would require the exception to also attack the argument for the action enjoined by the norm. Note also that the norm itself is not defeated: the exception simply means that it is not applied. This means that applying an exception does not *violate* the norm. Rather it encapsulates some first principles reasoning which concludes that the norm should be disregarded in these circumstances.

In [9] a number of reasons for making exceptions to working to produce food are given. These include people incapable of work, people who may fulfill roles other than food production, such as entertainers, law enforcement officers, public officials, and those engaged in approved projects such as community buildings and the like. Some of these exceptions, like incapacity, will be more or less universal, while others will depend on the values of the society concerned. Thus Ancient Athens supported poets and philosophers, whereas Ancient Sparta supported a large military class.

In Figure 2b, the effect of the exception is to accommodate choices determined by the value preferences of the society involved. But, as alluded to above, in other cases the exception will need to rule out other arguments as well as the applicability of the norms. Suppose the exception were to be incorporated in the law as a norm in its own right rather than an exception to another norm:

- N3: Those incapable of work should not work

Now if N3 applies it will replace the exception node and attack both the attack from N1 and the argument to work (Arg1). In this way those incapable of work are not only exempt from the obligation to work, but prohibited from so doing, reflecting the notion that it is not their own choice not to work, and so they should not attract disapproval from able bodied citizens.

If there is an exception the norm can be ignored, but remains in force and is not violated. It may, however, be that in some circumstance the norm *should* be violated [7], [8]. For example, it may be necessary to violate some traffic norms, perhaps to mount the pavement, in order to avoid an accident. If such a situation is sufficiently common, it may become enshrined in law as an exception, or even as an independent norm. If, however, it is a rare situation we will need to return to reasoning from first principles, and consider an argument which attacks and defeats the norm. In this case the defeated norm *will* be violated, but there will be an argument to justify, or at least excuse, the violation. This mechanism may allow for exceptions to emerge, since if the situation is sufficiently common it may well become useful to “crystallise it into a not too specific or detailed form” so that so that it may “serve us again in a like situation without so much thought.”

2.1 Permissions

Permissions have often seemed mysterious in Multi-Agent Systems which represent prohibitions by removing a transition (e.g. [1]). Permissions are there represented by retaining the transition: but then what does permitting an action add? Our framework suggests an answer. There may be several ways of regarding an exception, depending on what it attacks, and whether it is itself attacked by other arguments.

- The exception may attack both the attack from the norm and the action normally enjoined by the norm. For example, if someone is incapable of work is desirable not only that it is permitted to feed them in Winter, but that it is should be obligatory to do so.
- The exception may attack the attack from the norm, but does not attack the enjoined action. Here there is a free choice as to which action to be performed. Thus in figure 3, under the automobile exception to the Fourth Amendment, police are permitted, but not obliged, to search an automobile if there is probable cause.
- The exception may be attacked by an argument *X* which is itself attacked. Now the exception will apply only if reinstated by an attack on *X*, or the attack from *X*. Thus in Figure 3 the automobile exception is attacked by the availability of a warrant, and requires lowered expectations of privacy to reinstate it.
- The exception may be attacked by an argument *X* which is not itself attacked. In this case the exception is only putative and cannot be exercised.

The Automobile Exception to the Fourth Amendment will be discussed in detail in the next section.

3 EXAMPLE: US FOURTH AMENDMENT AND THE AUTOMOBILE EXCEPTION

We illustrate the use of EAFs to represent norms and exceptions by reference to the US Fourth Amendment. The Fourth Amendment is designed to safeguard the privacy of citizens by guaranteeing:

“The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.”

This amendment has been much discussed in AI since being introduced in [19]. A number of relevant cases are considered in [3]; we will use these cases for our example.

The initial situation can be illustrated by Figure 1c (with Arg2 being *do not search to safeguard privacy rights* and Arg1 being *search to enforce the law* and with P now standing for Privacy and L for Law Enforcement). The matter was felt, however, too important to be left to individual preferences and was resolved by the norm enshrined in the Fourth Amendment. The Fourth Amendment prioritises privacy, and so defeats Arg1. But it also recognises that sometimes searches will be needed, in order to enforce the law in the face of a probable crime, and thus permits searches with

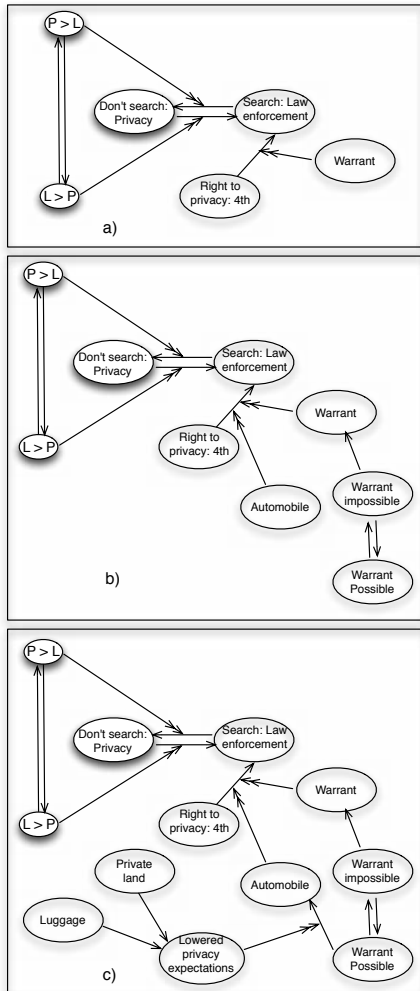


Figure 3: a) Fourth Amendment as EAF. b) After the *Carroll* Case. c) Final Situation of the Automobile Exception.

a warrant, subject to the safeguard that probable cause has been demonstrated. The situation can be shown as the EAF in Figure 3a.

Suppose, however, that it not possible to obtain a warrant because of the urgency of the situation: for example a car suspected of transporting contraband, which is capable of disappearing into a different jurisdiction before a warrant can be sworn. This was the situation in *Carroll v United States*⁴. In that decision it was held that it was lawful to search an automobile, given probable cause and the impossibility of enforcing the law without the power to search.

⁴ *Carroll v United States*, 267 U.S. 132 (1925). Other cases discussed are: *Coolidge v New Hampshire*, 403 U.S. 443 (1971), *Cady v Dombrowski*, 413 U.S. 433 (1973), *Chambers v Maroney* 399 U.S. 42 (1970), *South Dakota v. Opperman*, 428 U.S. 364 (1976), *United States v Chadwick*, 433 U.S. 1 (1977), *Arkansas v. Sanders*, 442 U.S. 753 (1979) and *California v. Carney*, 471 US 386 (1985).

Carroll is the source of what has become known as the *Automobile Exception*. The situation post-*Carroll* is shown in Figure 3b.

One question is whether the possibility of obtaining a warrant defeats the automobile exception. This was not tested in *Carroll*, where it was clearly impracticable to obtain a warrant. Some subsequent cases seemed to suggest that police officers thought if they were dealing with an automobile no warrant was needed, although there were always Justices who seemed to think that if a warrant could be obtained, it should be obtained (as Justice Stewart held in *Coolidge v New Hampshire* “The word “automobile” is not a talisman in whose presence the Fourth Amendment fades away and disappears.”). In *Coolidge* the car was parked in the suspect’s driveway and the suspect was in custody. But in other cases the possibility of obtaining a warrant did not seem to matter: searches were held to be lawful when the car was in a commercial garage (*Cady v Dombrowski*) and even at a police station (*Chambers v Maroney*). The rationale seems to be that an automobile, when not on private land (as in *Coolidge*), carries lowered expectations of privacy because it is liable to be stopped and inspected under traffic legislation (*South Dakota v. Opperman*).

There were also cases which held it impermissible to search luggage without a warrant, even if the luggage was in an automobile, if the probable cause related only to the luggage and not the automobile as a whole (*United States v Chadwick* and *Arkansas v. Sanders*). The result is the EAF in Figure 3c⁵. Note that the automobile exception does not prevent the police from obtaining a warrant if they choose to do so, it merely frees them to act according to their preferences, We may, however, think that the preferences of the police are such that they are likely to choose to perform a warrantless search if they believe that they are covered by the exception.

In *California v Carney* it was argued that lowered expectations should not apply to a motor home. In the minority opinion Mr Justice Stevens seemed to think that an analogy with luggage might be acceptable.

“The Court in *Chadwick* specifically rejected the argument that the warrantless search was ‘reasonable’ because a footlocker has some of the mobile characteristics that support warrantless searches of automobiles. The Court recognized that ‘a person’s expectations of privacy in personal luggage are substantially greater than in an automobile.’ ... It is perfectly obvious that the citizen has a much greater expectation of privacy concerning the interior of a mobile home than of a piece of luggage such as a footlocker.”

But this was rejected by the majority, who held that while in use as a vehicle a motor home had the lowered expectation of privacy associated with vehicles: its primary purpose at the relevant time was not to contain personal effects. Steven’s argument about motor homes is therefore not included along with luggage and private land in Figure 3c.

The advantages of using an EAF are clear. It enables us to see that the 4th Amendment is never violated, even though there are exceptions to it. Similarly the need for a warrant, if it is possible to

⁵ This situation was modified in *California v. Acevedo*, 500 U.S. 565, where it was held that the presence of suspicious luggage in a vehicle *ipso facto* gave probable cause to search the whole vehicle, effectively excluding the “luggage” argument.

obtain one, is never defeated, but can be ignored where there are lowered expectations of privacy. The lowered expectations do not defeat the argument from the possibility of obtaining a warrant, but block its effect. Note, however, that the presence of the automobile on private land does defeat the argument from lowered expectations rather than blocking its effect: expectations of privacy are lowered when the car is in a public space, but not when it is on private land. Similarly the expectations of privacy relating to luggage, which is typically used to transport personal items, remain even when placed in an automobile (at least until *Avecedo*).

4 CONCLUDING REMARKS

In this paper we began by exploring how norms can emerge through a process of practical reasoning. Once a decision has been agreed, it is encapsulated as a norm so that the argumentation can be applied in future cases, but without the expense of repeating the reasoning. This means that a norm will very often represent the preferences of the individual, social group or society which adopts the norm.

We have distinguished between the notions of *promoting values*, as typically used in previous work such as [2] and the more fundamental notion of *fulfilling needs*. Whereas values are useful in explaining differences between societies and individual behaviours, needs, which are considered universal to all people, are better able to express why the basic norms of various societies are mostly very similar. This universality is necessary if we want to follow the reasoning of Rawls *Theory of Justice* [18], or otherwise avoid moral relativism. We suggested that Maslow's hierarchy [16] would be suitable for this purpose.

Although this will identify basic norms, norms typically require some exceptions to avoid situations in which following them would have undesirable consequences. What exceptions will be produced will depend on the preferences of the society adopting the norm, and may well vary from society to society. We argued that Extended Argumentation Frameworks [17] are particularly suited to representing such exceptions. We distinguish between reasoning to *justify* norms, which requires reasoning from first principles using needs values and state transitions, from reasoning *with* norms, in which the arguments justifying the norm are replaced by the norm itself. The focus then moves from competing preferences to the identification of exceptions. These exceptions are themselves justified in terms of preferences, and we find more inter-societal variation than in norms. Thus while most societies will have a norm requiring people to earn their living, the exceptions to such a norm will vary according to the preferences and ideals of the particular societies. It is to the reasons justifying exceptions that we should look when seeking to identify the values of a society or of groups of citizens.

We also noted that the notion of exceptions helped to make sense of the notion of permissions, which has often proved difficult to understand in normative Multi-Agent Systems approaches such as [1], in which explicit permissions often seem to add little. Here we use them to identify exceptions which the community does not wish to enforce or prohibit, but instead leave it to the voluntary choice of citizens, or particular groups, rather than the citizenry as a whole. The key role of permissive exceptions to norms is to exempt citizens from the obligation to punish certain defined classes of

violators, which otherwise is an essential requirement if normative collapse is to be avoided. Thus in the Fourth Amendment cases, the court is required to punish improper search by refusing to admit the evidence obtained thereby, unless it can be shown that an exception such as the Automobile Exception, applied. We used the Automobile Exception to the Fourth Amendment cases to provide a detailed illustration of the use of EAFs to represent norms.

Finally we should note that some laws relate to very specific features of advanced societies (such as traffic laws). Here it may be better to talk of purposes (as in the origin of this way of resolving conflicts, [10]), reserving values to describe the preference manifested by choosing between different purposes.

The chief contributions of this paper are:

- A discussion of how norms can emerge from practical reasoning using value-based argumentation;
- The proposal to use EAFs to represent norms and exceptions, and an illustration of the advantages of so doing.

We believe that EAFs will provide an excellent framework for future representation and discussion of norms.

REFERENCES

- [1] T. Ågotnes, W. van der Hoek, M. Tennenholtz, and M. Wooldridge. Power in normative systems. In *Proceedings of The 8th AAMAS conference*, pages 145–152. IFAAMS, 2009.
- [2] K. Atkinson and T. Bench-Capon. Practical reasoning as presumptive argumentation using action based alternating transition systems. *Artificial Intelligence*, 171(10-15):855–874, 2007.
- [3] T. Bench-Capon. Relating values in a series of Supreme Court decisions. In *Proceedings of JURIX 2011*, pages 13–22. Citeseer, 2011.
- [4] T. Bench-Capon. Transition systems for designing and reasoning about norms. *Artificial Intelligence and Law*, 23(4):345–366, 2015.
- [5] T. Bench-Capon and S. Modgil. Case law in extended argumentation frameworks. In *Proceedings of the 12th International Conference on Artificial Intelligence and Law*, pages 118–127. ACM, 2009.
- [6] T. Bench-Capon and S. Modgil. Rules are made to be broken. *AI4J—Artificial Intelligence for Justice*, pages 9–17, 2016.
- [7] T. Bench-Capon and S. Modgil. When and how to violate norms. In *Proceedings of JURIX 2016*, pages 43–52, 2016.
- [8] T. Bench-Capon and S. Modgil. Norms and value based reasoning: justifying compliance and violation. *Artificial Intelligence and Law*, 25(1):29–64, 2017.
- [9] T. Bench-Capon and S. Modgil. *Representing Norms with Extended Argumentation Frameworks*. Technical Report ULCS-19-001, University of Liverpool, 2019.
- [10] D. Berman and C. Hafner. Representing teleological structure in case-based legal reasoning: The missing link. In *Proceedings of the Fourth International Conference on Artificial Intelligence and Law*, pages 50–59, 1993.
- [11] F. Bex, K. Atkinson, and T. Bench-Capon. Arguments as a new perspective on character motive in stories. *Literary and Linguistic Computing*, 29(4):467–487, 2014.
- [12] G. Governatori, F. Olivieri, A. Rotolo, and S. Scannapieco. Computing strong and weak permissions in defeasible logic. *Journal of Philosophical Logic*, 42(6):799–829, 2013.
- [13] R. Hare. *Freedom and reason*, volume 92. Oxford Paperbacks, 1963.
- [14] A. Jones and M. Sergot. Deontic logic in the representation of law: Towards a methodology. *Artificial Intelligence and Law*, 1(1):45–64, 1992.
- [15] S. Mahmoud, N. Griffiths, J. Keppens, A. Taweel, T. Bench-Capon, and M. Luck. Establishing norms with metanorms in distributed computational systems. *Artificial Intelligence and Law*, 23(4):367–407, 2015.
- [16] A. H. Maslow. A theory of human motivation. *Psychological review*, 50(4):370, 1943.
- [17] S. Modgil. Reasoning about preferences in argumentation frameworks. *Artificial Intelligence*, 173(9):901–934, 2009.
- [18] J. Rawls. *A Theory of Justice*. Harvard University Press, 1971.
- [19] E. L. Rissland. Dimension-based analysis of hypotheticals from supreme court oral argument. In *Proceedings of the Second ICAIL*, pages 111–120, 1989.
- [20] M. Wooldridge and W. van der Hoek. On obligations and normative ability: Towards a logical analysis of the social contract. *Journal of Applied Logic*, 3:396–420, 2005.