Value Based Reasoning and the Actions of Others

Katie Atkinson and **Trevor Bench-Capon** [1]

Abstract. Practical reasoning, reasoning about what actions should be chosen, is highly dependent both on the individual values of the agent concerned and on what others choose to do. We discuss how value based argumentation about what to do can be performed without making assumptions about the preferences of the other agents. We then show how expected utility calculations relate to the value-based argumentation approach, and express the reasoning as arguments and objections, so that they can be integrated value-based practical reasoning. We illustrate our discussion with examples of value based reasoning in public goods games as used in experimental economics and present an initial evaluation of the approach in terms of these experiments.

1 Introduction

A key difference between theoretical reasoning (reasoning about what is the case) and practical reasoning [35] (reasoning about what to do) is the *direction of fit* [36]. Whereas in theoretical reasoning an agent is trying to fit its beliefs to the world, in practical reasoning an agent is choosing an action intended to fit the world to its desires. For theoretical reasoning, there is only one, shared, world, and so agents should tend to agree, but desires will legitimately differ from agent to agent and so practical reasoning depends on the subjective aspirations and desires of the *individual* agent. Agents may even be in conflict, so that they attempt to bring about different worlds. The conclusions are therefore legitimately subjective, and disagreement is both rational and to be expected. Acceptance of an argument as to what to do depends not only on the argument itself - for it must, of course, be a sound argument - but also on the audience to which it is addressed [33]. This notion of audience was computationally modelled in [21] and made more formal in Value-Based Argumentation Frameworks (VAFs) [9]. VAFs are an extension of the abstract Argumentation Frameworks (AFs) introduced in the seminal paper of Dung [13]. In a VAF arguments are associated with the social (i.e. *not* numeric) values[2] their acceptance promotes or demotes. Different audiences can now be characterised by the ordering they place on these values. Whereas in an AF an argument is defeated by any attacking argument, in a VAF an argument is *defeated for an audience* by an attacker only if the value associated with the attacking argument is ranked at least as highly by that audience. In this way different audiences will accept different sets of arguments (preferred semantics [13] is used to determine acceptance), and, as is shown in [9], provided the VAF contains no cycles in the same value, there will be a unique non-empty preferred extension. Thus, use of VAFs provides a way of explaining (and computing) the different arguments accepted by different audiences. Value Based Reasoning has been used as the basis of practical reasoning ([19], [2], [23], [15], [41], [12]) and applied in particular areas such as law ([7], [26], [20]), e-democracy ([11], [45]), policy analysis ([38]), medicine, ([4]), experimental economics ([8]), rule compliance ([10]), decision support ([28]) and even ontology alignment ([39], [32]). Complexity results for VAFs were established in [14] and [29].

1.1 An Argumentation Scheme for Value-Based Practical Reasoning

The application of the preferences of an audience, expressed as an ordering on values, to practical reasoning requires the generation of the arguments and identification of the values associated with them. The proposal made in [3] was to use an argumentation scheme (now included in the compendium of argumentation schemes collected in [44]) justifying an action in terms of the values it promotes. The scheme appears in [3] as:

> In the current circumstances *R*, I should perform action *A*, to bring about new circumstances *S*, which will achieve goal G and promote value *V*.

We will henceforth refer to this scheme as *Practical Reasoning Argumentation Scheme* (PRAS). Like all argumentation schemes, PRAS establishes its conclusion only presumptively [42] and can be challenged using what [42] and [44] call *critical questions*. Thus an argument using PRAS can be challenged by claims against its soundness such as: that the current state is different, that the action is not possible, that the action will reach a different state, fail to achieve its goal or fail to promote its value. It can also be challenged on the basis of the desirability of the action: that it will also demote values and these values are more important, or that alternative actions promote values that are more important. This second group of objections is what gives room for subjectivity arising from different value orderings so that, as Searle puts it in [36]:

> Assume universally valid and accepted standards of rationality, assume perfectly rational agents operating with perfect information, and you will find that rational disagreement will still occur; because, for example, the rational agents are likely to have different and inconsistent values and interests, each of which may be rationally acceptable

In [2] seventeen different critical questions were identified that could give rise to objections to, and counter-arguments against, instantiations of PRAS.

1.2 Computational Realisation of this Scheme

In order to make this approach computable, it is necessary to provide an underlying representation of the world and how it can be

[1] Department of Computer Science, University of Liverpool, email: tbc@csc.liv.ac.uk

[2] *Values* are the aspirations or the purposes an agent might pursue, such as liberty, equality, fraternity, wealth, health and happiness.

affected by the actions of agents. State Transition Diagrams (STDs) are a natural choice for this, since they can represent the world as a set of states, and actions as the transitions between them. In open agent systems, however, the outcome of an action may well depend on what the other agents in the situation choose to do. Thus an individual's choice does not necessarily determine the state that will be reached. To account for this, open agent systems should model transitions as the *joint actions* composed of the the individual actions of all the agents relevant to the situation[3]. A suitable variant of STDs for use in open agent systems is *Action-based Alternating Transition Systems* (AATS), introduced in [46], which have joint actions as their transitions. AATS are formally based on Alternating-time temporal logic [1]. The basic AATS was augmented in [2] to allow the labelling of the transitions with the values promoted and demoted by that transition (AATS+V) and AATS+Vs were used to provide the underpinning semantical structure for the approach to practical reasoning set out in that paper[4]. Given a representation of the problem situation as an AATS+V, the discovery of arguments, counter arguments and objections can be implemented in the manner of [47]. A database containing tables for the states, joint actions and transitions of the AATS+V is created to hold the problem information and then instantiations of PRAS and challenges to those instantiations can be found by fairly simple queries to that database. For example there will be an instantiation of PRAS if there is a transition from the current state which promotes a value.

Three stages in practical reasoning are identified in [2]:

- **Problem formulation**: essentially the construction of an AATS+V for the particular problem situation;
- **Epistemic stage**: this involves determination of the current state and the joint action that will result from the choice of a particular individual action by the agent concerned;
- **Option selection**: the arguments generated from the AATS+V are formed into a VAF and their acceptability status determined according to the preferences of the agent concerned.

While problem formulation and the identification of the current state can be resolved using normal theoretical reasoning techniques, and the option selection stage is carried out using value-based reasoning base on VAFs, the determination of the joint action is less clear and will be the topic of this paper. The essential problem is that in order to know what it is best to do, it is necessary to anticipate what the other agents will do, since this will critically affect what results from our own actions. But since this reasoning will depend on the beliefs, aspirations and preferences of these other agents, this will require a number of assumptions which are often difficult to justify to be made. For example, agents which adopt the naive approach of assuming that others will be like themselves, tend to perform badly in practice [17].

1.3 Modelling the Values of Others

One approach, common in classical economics, is to see agents as consistently rational and narrowly self-interested agents who usually pursue their subjectively-defined ends optimally. John Stuart Mill [25] put it thus when describing "economic man" (sometimes called *homo economicus*):

> [Economics] is concerned with him solely as a being who desires to possess wealth, and who is capable of judging the comparative efficacy of means for obtaining that end.

Game Theory [27] also takes a single measure of utility expressed as a payoff matrix, which has become a very widespread basis for the design of multi-agent agent systems [31]. This approach has led to some insights, and provided the foundation for much elegant mathematics, but unfortunately does not provide a satisfactory explanation of the way in which humans behave in practice. And of course, if we are deciding what to do, we much cannot expect others to behave as they *should*, so even if this was a good *normative* theory, we would still need an adequate *descriptive* theory.

That others cannot been seen in this way is well demonstrated by a number of experiments carried out in behavioural economics. These experiments are carried out, using a variety of public goods games, to test the theory that behaviour can be predicted using the assumptions of classical economics and game theory. There are valuable meta studies, in particular for the Dictator Game [16] and the Ultimatum Game [30] and [22]. The findings suggest that the canonical model is followed only very rarely. Thus in [22] we read:

> in addition to their own material payoffs, many experimental subjects appear to care about fairness and reciprocity, are willing to change the distribution of material outcomes at personal cost, and are willing to reward those who act in a cooperative manner while punishing those who do not even when these actions are costly to the individual

Even in the Prisoner's Dilemma [34], where defection is clearly the dominant strategy, we find a tendency to deviate from it [6]. In [40], the emergence of norms and conventions is discussed in terms of the Prisoner's Dilemma, and some of the other characteristics influencing behaviour, such as empathy, trust and *esprit de corps* are cited as ways in which these norms can be formed. The role of punishment is explored in [24]. What all these meta studies show is

- The canonical model used in classical economics, game theory and many multi agent systems is not adequate to explain the behaviour encountered in experimental studies;
- There is a significant amount of inter-cultural variation, suggesting that the established values of subjects is carried forward into these experiments;
- There is also a significant amount of intra-cultural variation, suggesting that the behaviour of individuals cannot reliably be predicted solely on the basis of their cultural background.

Our view is that by putting the subjective ordering of values to the fore, value based reasoning can provide a fruitful way of exploring these issues. This was borne out by the examination of the Dictator and Ultimatum games in [8]. There, however, like all approaches based on [2], the reasoning about what others would do relied too heavily on unjustifiable assumptions about the values they would use, and how they would order them. Our objectives in this paper are threefold:

- to take account of the actions of others in the framework of value-based practical reasoning without requiring assumptions about the beliefs and preferences of other agents;
- to do so in a manner compatible with the results of game theory and multi-criteria utility (e.g., [37], [41]) while explicitly allowing for subjectivity and altruism;
- to be able to express the reasoning in the form of arguments and objections so as to facilitate integration with value-based practical reasoning.

[3] This is an important difference from classic planning systems such as STRIPS[18].

[4] To aid readability, formal definitions are collected as an Appendix at the end of the paper. AATS+V are defined in Definitions 1 and 2.

2 The Games

In this section we describe two games used in experimental economics. We will not consider the Dictator Game here, because although as shown in [16] and [8] it is amenable to analysis in terms of value-based reasoning, there is only one decision maker, and so the need to anticipate the actions of others, which is the aspect in which we are interested here, does not arise. We will therefore only consider the Ultimatum Game and the Prisoner's Dilemma in this paper.

2.1 The Ultimatum Game

In the Ultimatum Game the first player is given a sum of money and told that he may offer some of it to the second player. Once the proposer has made an offer the respondent may choose to accept the offer, or reject it, in which case both players receive nothing. Whereas traditional game theory would suggest that the proposer would make the smallest offer possible and the respondent would accept it, experiments do not support this. The meta-analysis of 37 papers reported in [30] found that

> that on average the proposer offers 40% of the pie to the responder. ... On average 16% of the offers is rejected. ... We find differences in behavior of responders (and not of proposers) across geographical regions.

It may well be that regions (at least at the country or even continent level used in [30]) do not provide the best explanation for different behaviours, being themselves large and often culturally heterogeneous. Another study [22], based on small-scale, homogeneous societies, found the different cultures more predictive:

> Among the Achuar, Ache and Tsimane, we observe zero rejections after 16, 51, and 70 proposer offers, respectively. Moreover, while the Ache and Achuar made fairly equitable offers, nearly 50 percent of Tsimane offers were at or below 30 percent, yet all were accepted. Similarly, Machiguenga responders rejected only one offer, despite the fact that over 75 percent of their offers were below 30 percent. At the other end of the rejection scale, Hadza responders rejected 24 percent of all proposer offers and 43 percent of offers at 20 percent and below. Unlike the Hadza, who preferentially rejected low offers, the Au and Gnau of Papua New Guinea rejected both unfair and hyper-fair (greater than 50 percent)

Two aspects of the societies concerned, namely the amount of cooperation found in the general economic activity of the society and the extent to which market exchanges were a feature of daily life, were found to be explanatory in [22]

> the Machiguenga and Tsimane rank the lowest; they are almost entirely economically independent at the family level and engage rarely in productive activities involving more than members of a family. By contrast, the Lamelara whale-hunters go to sea in large canoes manned by a dozen or more individuals. ... The Machiguenga show the lowest cooperation rates in public-good games, reflecting ethnographic descriptions of Machiguenga life, which report little cooperation, exchange, or sharing beyond the family unit.

In contrast, the Lamelara have the highest mean offer (58%) and a zero rejection rate. As shown in [8], this can can be explained by differing values and preferences amongst the participants, with the ordering emerging from their everyday activities being applied in the games. The game was analysed in [8], with the following six values:

- Proposer's Money (M1): Promoted by acceptance of an offer to a degree inversely related to the size of the offer and demoted if the offer is rejected;
- Respondent's Money (M2): Promoted by acceptance of an offer, to a degree related to the size of the offer;
- Generosity (G): Promoted for the proposer by giving away a reasonable amount of money;
- Equality (E): Promoted by both participants receiving the same amount;
- Proposer's Contentment (C1): Promoted by the acceptance of a low offer (did not offer too much) and demoted by the rejection of a low offer (did not offer enough), or by the rejection of a good offer, since the respondent would be considered unreasonable;
- Respondent's Contentment (C2): Promoted by accepting a good offer and demoted by accepting a low offer.

The transition diagram for the Ultimatum Game used in [8] is given in Figure 1. This considers the actions as happening serially, so that the joint actions have two stages. Whilst this makes the interaction, where values are promoted and demoted, more explicit, here we prefer to combine the actions, The proposer may make a very high (vho) offer (more than 50%), an equal (eo) offer (=50%), a fair (fo) offer (40-50%), or a low (lo) offer (less than 40%)). The respondent may accept or reject, giving 8 joint actions. j_1 is {vho,accept}, j_2 is {vho,reject} and so on. The AATS state records the money for each participant, and two flags, indicating whether the partcipants are content. Most important are the values promoted and demoted by the joint actions. These are shown in Table 1.

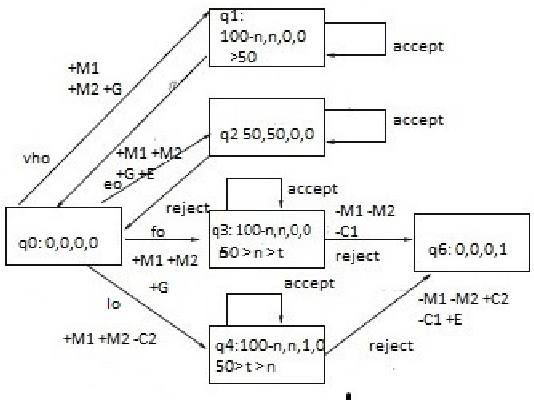

Figure 1. AATS for Ultimatum Game from [8]

Table 1. Value Promotion and Demotion in the Ultimatum Game

Joint Action	Proposal	Response	Promoted	Demoted
j1	vho	accept	M1,M2,G, C2	E
j2	vho	reject	G	M1
j3	eo	accept	M1,M2,G,C2	
j4	eo	reject	G	M1
j5	fo	accept	M1,M2	E
j6	fo	reject		M1
j7	lo	accept	M1,M2,C1	E,C2
j8	lo	reject		M1,C1

2.2 Prisoner's Dilemma

In this very well known game [34], widely used in discussions of norm emergence such as [40] and [6], both players may either cooperate or defect. Mutual cooperation results in a pay off of 3 to each player, mutual defection a payoff of 1 to each player, and if one cooperates and the other defects the defector receives 5 and the cooperator receives 0. The "correct" strategy is to defect since that gives a better payoff whichever move the other makes (is the *dominant* strategy). Also it is not a zero-sum game: collective utility is maximised by mutual cooperation. Here too, experiments find that the game-theoretic choice is not always made in practice. As explained in [40] conventions to encourage mutual cooperation often emerge or are devised. An example used in [40] is a military situation where much effort is made to build up trust and loyalty to create an *esprit de corp* in a regiment so that members will cooperate rather than defect, feeling that they are able to rely on their comrades, and in turn reluctant to let their comrades down. The conventions are often reinforced by punishing defectors [24]. Again there seem to be additional values considered by participants. Here we use the following values:

- *Player Money (M1 and M2)*: promoted if a player's payoff is greater than 1 (which is the least that can be ensured), and demoted if it is less that 1.
- *Player Guilt (G1 and G2)*: demoted if player defects and the other player cooperates
- *Player Self-Esteem (S1 and S2)*: demoted if player 1 (or 2) cooperates and player 2 (or 1) defects: since the player may feel that they should have known better.

In this game there are four joint actions which promote and demote values as shown in Table 2. Note that mutual defection provides a baseline, neither promoting nor demoting any values, since it can always be achieved or bettered.

Table 2. Value Promotion and Demotion in the Prisoner's Dilemma

Joint Action	Player 1	Player 2	Promoted	Demoted
j1	C	C	M1,M2	
j2	C	D	M2	M1,S1,G2
j3	D	C	M1	M2,S2,G1
j4	D	D		

3 Justification of Actions

The current approach to reasoning about the actions of others based on [2] and used in [8] is:

1. Select a desirable transition based on the values it promotes and demotes.
2. Argue for the individual action performed by the agent in the joint action corresponding to that transition.
3. Consider objections based on the other agents choosing different actions and so causing different joint actions to be performed.
4. Attempt to rebut these objections because:
   (a) The values promoted and demoted by the alternative transition are acceptable.
   (b) It is considered that the other agents will not act in this way.

Whereas 4a can be resolved on the basis of the preferences of the agent concerned, 4b, which is very often needed, requires more assumptions about the other agents than can be really justified.

In previous treatments based on such transition diagrams and using PRAS (e.g. [8]) we would get arguments such as *we should cooperate to promote M1* which would be challenged with objections such as *but player 2 might defect which would demote M1*. Now if M1 is the most important value for Player 1, then the objection will succeed, unless cooperation can be assumed. If M1 is the only value considered, defection is *dominant*, giving a better outcome whatever the other player chooses. Only if other values are considered will Player 1 choose cooperation. For example, M2 might be rated as highly as M1 (perhaps Player 2 is Player 1's child, or a close colleague), or a clear conscience is regarded as more important than money, in which case Guilt must be considered, The arguments are, however, really for a *particular* transition (joint action), with the agent's own action justified in virtue of its appearance in the transition: the objections are available because other joint actions contain the same individual action. Better would be an argument for the individual action itself, not the joint action and its corresponding transition. This will require us to look at the *set* of transitions containing the action. In the Ultimatum Game suppose that $prob(jointaction)$ is the probability of *jointaction* being performed when the agent concerned chooses some particular individual action. Now the values will be expected to be promoted and demoted according to the probability of the second player's response, as shown in Table 3, and so expected utility can be calculated, obviating the need to assume that the other will perform a particular action.

Table 3. Values Promoted and Demoted in the Ultimatum Game

Proposer Action	Promoted	Demoted
vho	G, prob(j1)M1, prob(j1)M2, prob(j1)C2	prob(j2)C1, prob(j2)M1, prob(j1)E
eo	G, prob(j3)M2, prob(j3)C2, prob(j3)M1	prob(j4)C1, prob(j4)M1
fo	prob(j5)M1, prob(j5)M2	prob(j6)M1, prob(j5)E
lo	prob(j7)M1, prob(j7)M2, prob(j7)C1	prob(j7)C2, prob(j8)M1, prob(j8)C1, prob(j7)E

Now we can base arguments on the complete set of transitions containing an action, rather than having to assume an action on the part of the other and then consider objections based on the potential performance of a different action. Several forms of argument are available (our examples assume the context of a persuasion dialogue with the proposer in the Ultimatum Game [43]):

- Where an action is certain to promote a value. E.g. *You should make a very high offer to promote G.*
- Where an action cannot promote a value. E.g. *You should not make a very high offer as that cannot promote C1.*
- Where an action can promote a value. E.g. *You should make a fair offer as this can promote M1.*
- Where an action can demote a value. E.g *You should not make a low offer as that will risk demoting C1.*

The third and fourth forms will have variants, if we can say something about the relative probabilities of acceptance and rejection. These variants will replace "can" with an indicator of how probable promotion is, such as "very likely", "more likely than not", "may possibly" etc. For example, we know from [30] that a fair offer is much more likely to be accepted than rejected, and so we can say *you should make a fair offer as that is likely to promote M1*, or, since low offers are more likely to be rejected, *you should not make a low offer as there is a substantial risk of demoting M1*.

Similar arguments can be generated for Prisoner's Dilemma. Promotions and demotions of the extended set of values for each action are shown in Table 4. From this table we can generate arguments, as given below.

Table 4. Values promoted and demoted in Prisoner's Dilemma

Proposer Action	Promoted	Demoted
C	M2, prob(j1)M1,	prob(j2)M1, probj(2)S1, prob(j2)G2
D	prob(j3)M1	prob(j3)M2, probj(3)S2, prob(j3)G1

- You should cooperate to promote M2
- You should not cooperate as this risks demoting M1, S1 and G2
- You should defect as this might promote M1
- You should not defect as this risks demoting M2, G1 and S2.

The real advance here over previous work such as [2] is that there is no longer any need to make assumptions about the what the other believes and prefers: the agent can now come to a decision using its *own* relative preferences between values, its own beliefs and the degree of risk it is prepared to take, whilst requiring no additional machinery: it uses only the AATS+V as developed in [2]. This fulfils the first of the objectives identified in section 1.

3.1 More than one other agent

The games discussed above have only one other agent. Of course, in practice there will typically be several, or even very many, agents that can have an influence. For example we might extend the Ultimatum Game so that there are several respondents and acceptance or rejection is determined by a majority, or acceptance may require unanimity. Or we might want to look at a problem such as the free-rider problem, whereby defection pays, unless some proportion of the population defects. For example, a small number of tax avoiders will not affect services, but if there are too many, the state infrastructure will collapse. In other situations there may be a number of agents with a range of, perhaps different, choices. This might, at first sight present a problem, since the number of joint actions rises rapidly: n agents each with m actions give rise to m^n joint actions. But we are not especially interested in details of the joint actions: the point of our approach here is to consider the set of joint actions in which the agent of concern performs a particular action. In the standard value-based approach, as proposed in [2] the values promoted and demoted by a transition are determined by the source and target states. Even where the action performed does affect values, as in [5], so that the intrinsic value of an action can be taken into account, what matters for the agent concerned is its own individual action, and so all transitions between the same pair of states containing that action will promote and demote the same values, as far as that agent is concerned. Thus, for our current purposes, we will consider all joint actions with the same action by the agent concerned leading to the same state to be equivalent, so that consideration can be limited to the different outcomes possible for a given action, irrespective of how many joint actions reach each outcome. Effectively all the other agents can be considered together as a single other. If a majority is required, it does not matter which agents make up that majority; nor does it matter who the other free loaders are provided that there are not too many of them, and so on. Of course, the probabilities may be affected: if we know that only one agent in six will reject a fair offer, then we can be more confident that the larger the number of respondents the more likely is a majority for acceptance, although it is less likely that the offer will be accepted if we require unanimity.

3.2 Preferred Values

If only a single value is recognised as worthy of promotion, the choice is often unproblematic. In the Prisoner's Dilemma, M1 may be promoted and cannot be demoted by defection, M2 is promoted by cooperation, C1 can only be demoted by defection and S1 can only be demoted by cooperation, but in some cases, whether a value is promoted or demoted may depend on what the other agents do. Similarly some combinations of values are unproblematic, but hard choices arise when different values pull us in different directions, because an action may promote one value and demote another, or because values are promoted and demoted to different degrees. In such cases we need to express and quantify our preferences.

3.3 Expected Utilities

We now turn to our second objective. In all value based reasoning it is assumed that an agent is capable of expressing a preference in terms of an ordering on values. However, sometimes quantification of the degree of preference and the degree of promotion is required (e.g. [28]). In PD the payoff matrix gives the degree of promotion e.g. j_1 promotes M1 and M2 to degree 2 etc: (remember that we only count gains in excess of the baseline towards promoting M1 and M2), but to quantify the preference each value must be expressed in terms of a single selected value (M1 is the obvious choice). The valuation is subjective to each agent, but requires reference only to its own preferences. Agent Preferences are defined in Definition 3 in the Appendix. Unlike previous work such as [2] there is no longer any need to make assumptions about the beliefs, domain conceptualisation and preferences of the other: the agent will be able to decide using its *own* relative preferences between values, its own beliefs and, where necessary, the particular degree of risk it is subjectively prepared to accept.

Once the agent preferences have been established, the expected utilities can be calculated as in Definition 4 of the Appendix.

If we apply this to the Prisoner's Dilemma (PD), since there are only two joint actions containing cooperation, $prob(j_2) = 1 - prob(j_1)$. In the traditional PD only the agent's own payoff is recognised as having utility. The utility is the actual payoff minus the guaranteed payoff (i.e. the payoff from mutual defection). For cooperation the utility is 2 when the other cooperates and -1 when the other defects. For defection it is 4 when the other cooperates and 0 when the other defects. The expected utilities for ag cooperating (dark grey) and defecting (light grey) for the various probabilities of the other cooperating are shown in Figure 2.

Suppose, however, that both the values M1 and M2 are recognised in PD, and M2 is weighted at 0.5M1. Now the utility of cooperating when the other also cooperates will be 3M1, and the utility of cooperating when the other defects M1. Similarly we can calculate the expected utility of defecting for the various probabilities of the

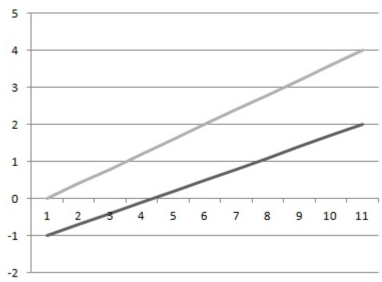

Figure 2. Expected Utilities for M1 only. Dark grey is ag cooperates, light grey is ag defects.

other cooperating. Defecting when the other cooperates yields a utility of 3.5M1, and mutual defection 0 (since this is the base line case, no values are considered promoted). Again the desired joint action is performed when the other agent cooperates. This gives the graph shown as Figure 2a. The crossover is at $prob(j_0) = 0.67$.

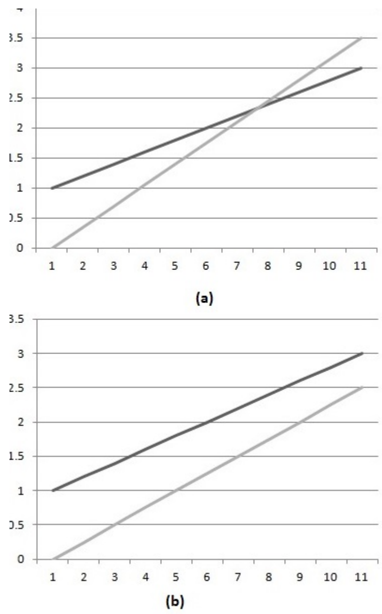

Figure 3. Expected Utilities for (a) M2 = 0.5M1 and (b) M2 =0.5M1 and G = M1. Dark grey is ag cooperates, light grey is ag defects.

If we now add in the value of Guilt (with a weight of 1), which gives a negative utility when an agent defects and the other cooperates, we get the expected utilities shown in Figure 2b.

These three figures represent the three possibilities. In Figure 1, which shows the traditional PD, we find that defection *dominates* cooperation: the expected utility is higher for every value of $prob(j_0)$. Therefore defection is the preferred action, whatever the probability of the other cooperating. In Figure 2b the reverse is true: the inclusion of the additional values means that cooperation dominates defection. In Figure 2a, there is a crossover, at $prob(j_0) = 0.7$, so that for high probabilities of cooperation, defection is preferred, but for low levels, the utility afforded to the payoff received by the other makes cooperation preferred.

3.4 Arguments in Prisoner's Dilemma Using Expected Utilities

Our third objective is addressed by producing arguments based on the expected utilities. These different possibilities mean that several types of argument can be based on the expected utilities. Our examples are expressed in terms suitable for a persuasion dialogue (not between the PD participants, but between a participant and advisor).

1. With your value preferences, you should C (respectively, D) since the expected utility is always greater than any alternative
2. With your value preferences, you should C (respectively, D) since the expected utility is always positive
3. With your value preferences, you should C (respectively, D) since the expected utility is greater than the alternative when the probability of cooperation is greater (less) than P.

Of these (1) is appropriate when the action advocated is dominant, and is the strongest of the three. Argument (2) is rather weak: although the expected utility is always positive, the proposed action can be dominated by the alternative for some (or even all) values of $prob(j_1)$. It may, however, be useful if we wish to reach the target state in order to enable some more beneficial action, since it indicates that no harm is done, and so can be used to rebut objections. The argument shows that we suffer no loss, although there is an opportunity cost. Argument (3) can be effective provided we can give reasons to suppose that probability of cooperation is in the desired range.

A dialogue arising from using (1) for defection might run:

- Since you value M1 and M2 equally, you should C since the expected return is always greater than the alternative.
- this overvalues M2.
- even if M2 is only worth 70% of M1, the expected utility is always greater than the alternative.
- But even 70% overvalues M2
- Even if M2 is only worth half M1, a less than 0.6 probability of cooperation will mean cooperation has the higher expected utility. Moreover the expected utility of cooperation is still always positive.

In the course of the dialogue, the very strong argument of type (1) has become untenable, but a combination of arguments of types (2) and (3) remain potentially persuasive. Here we are producing argumentation dialogues (albeit not yet expressed in a formal dialogue model) which explore the sensitivity to the assessment of the relative valuations, and the sensitivity to the estimates of cooperation. These dialogues do not require any knowledge about the other, but if such information is available these dialogues provide a context in which it can be deployed by constraining the range for the probability of cooperation. For an example based on (2):

- Since you value M2 at 50% of M1, you should C since the expected return is always positive
- But with these values, D gives a better return unless the probability of cooperation is worse than 0.6.

This objection could be reinforced with reasons to suppose it likely that the other will cooperate (family member, team member or similar, or experimental results, if appropriate results are available). Note, however, that these are also reasons to increase the valuation of M2 relative to M1.

The above arguments can, if desired, be presented as argumentation schemes in the manner of [44]. For example the scheme based on (1) above:

- **Values Premise**: V is the set of values considered to be relevant by ag
- **Weighting Premise**: The relative valuation of the members of V given by ag is a set of $\langle value, relativeweight \rangle$ pairs
- **Joint Action Premise**: $\{j_0, j_1, ...j_n\}$ is the set of joint actions J in which ag performs α
- **Expected Utility Premise**: $eu_{ag}(\alpha, prob(j_0))$ returns the expected utilities of agent ag performing α for values of $prob(j_0)$ $0 \leq prob(C) \geq 1$ where j_0 is the desired joint action.
- **Dominance Premise**: $eu_{ag}(\alpha, j_0) \geq eu_{ag}(\beta, j_0)$ for any alternative action β available to ag, for all values of $prob(j_0)$; where j_0 is the joint action compliant with the action of ag.
- **Conclusion**: ag should perform α

This scheme would be associated with critical questions such as: *Are all the members of V relevant? Are any other Values relevant? Are any members of V under or over valued?* These critical questions will have their own characteristic rebuttals: For example the third could be met by *even if the value of v is reduced to n%, the expected utility is always greater than its alternatives.*

3.5 Application to the Ultimatum Game

Similar arguments can be produced for the Ultimatum Game. Different weights for the different values will lead to different arguments being dominant. Also the different actions will promote M1 and M2 to varying degrees. M1 will be promoted most (if accepted) by *lo*, then *fo* then *eo* and least by *vho*, whereas for M2 the reverse will be true. Some examples are given in Table 5.

Table 5. Example value weights and corresponding dominant actions in the Ultimatum Game

M1	M2	G	E	C1	C2	dominant
1	0	0	0	0	0	lo
1	0.3	0	0	0	1	fo
1	0.3	0	1	0	1	eo
1	1	1	0	0.5	0.5	vho
1	0.3	0.7	0	0	0.5	eo/lo
1	0.6	0.4	0	0	0.5	vho/fo

The last two rows give examples of value assignments which produce cross overs. In the penultimate row, at low probabilities of acceptance the best choice is the equal offer: this promotes generosity and avoids angering the other, without sacrificing more money than is necessary to achieve these goals. When the probability of acceptance reaches 0.6 both the fair offer and the low offer take over, with the low offer being slightly preferred. In the final row, the high weight of M2 means that the very high offer is better than the equal offer for low probabilities of acceptance, but the fair offer becomes best for probabilities of acceptance greater than 0.6. When the probability exceeds 0.7, the low offer is also better than the very high offer, but the fair offer remains best.

Finally we have produced some initial results which indicate that the cultural variations encountered in public goods game experiments can be reproduced using suitable value profiles, shown in Table 6. Reproduction of such experimental results will form the basis of our evaluation.

3.6 Evaluation

We offer two aspects of evaluation. Technically, we can ask whether we achieved the objectives set out in section 1. Practically, we can explore the extent to which our proposed approach is able to reproduce the results of empirical studies such as [22].

Three technical objectives were given in Section 1. Our first objective was to accommodate the need to consider the actions of others, while only considering the values, and preferences of the agent concerned, since modelling of others is inevitably unreliable, given the extent of inter- and intra-cultural variation. We have achieved this, using only the structure of AATS+V of [2], by considering all the joint actions containing a given individual action as a set, obviating the need to consider the specific actions performed by others. The second objective was to do this in a way consistent with existing game and multi-criteria utility theory. We have achieved this by relating the value-based approach to expected utilities. The key notion of a dominant action remains, since, if there is a dominant action, the expected utility of the values promoted by that action will always be greater than any alternative. Moreover where an action is not dominant for all probabilities of the other behaving as required, the bounds can be identified, which allows for the sensitivity to the relative weighting of the relevant values, and, where no action is dominant, to the probability of the other performing the appropriate action, to be quantified. To fulfill the third objective, we have given arguments grounded on the expected utilities. Objections can be based on adding, removing or re-weighting values, which can change the dominant action, or restrict its dominance to a certain range of probabilities of the other agents allowing a particular outcome to be reached. Again the required degree of revaluation can be specified.

Whereas the payoffs of game theory are, as is perfectly correct for games which do require firm rules, fixed and unchanging, here the payoffs are subjective with respect to the individual goals and aspirations of the agent concerned, and so can be individually set and made subject to change, possibly as a result of persuasive argument, or of empirical evidence. This means that we can attempt a more practical evaluation in terms of reproducing the results of studies such as [22].

Recall that that study accounted for differences in terms of the degree of cooperation, and degree of commercial exchange found in daily life. We can relate these characteristics to a value profile. Suppose we associate the value of generosity with the cooperative groups such as the whale hunting Lamelara, and the recognition of C2 (the need not to anger the other) with commercial exchange. Ideally we would produce a value profile for each society, and evaluate both proposers and respondents. Such a full study must await future work, but as an encouraging preliminary we offer the results shown in Table 6.

Table 6. Dominant actions for values relating to cooperation and exchange activities

	M1	M2	G	E	C1	C2	dominant
cooperative	0.3	0	1	0	0	0	eo
exchange	0.3	0	0	0	0	1	eo
neither	1	0	0	0	0	0	lo

These results show that these value profiles do indeed correspond to the action choices typical of corresponding societies. Note that it is the equal offer rather than the very high offer that Table 6 predicts for cooperative societies and those accustomed to commercial exchange. This coheres with the highest offers in [22] being 58% and 51%. Similarly the lowest offer of 26% belonged to groups that did not work cooperatively and rarely engaged in commercial exchange, reflected here by a profile which does not recognise either generosity or the feelings of the respondent.

As well as replicating previous studies, we can also perform our own experiments in which the value preferences of the subject are established (e.g. through a questionnaire), and then the behaviour in the games compared with what is predicted by the value profile.

4 Concluding Remarks

Previous work on practical reasoning using value-based argumentation has required assumptions about the values and preferences of other agents which can affect the outcome of an action performed by the reasoning agent. Justification of these assumptions is always difficult, particularly when several other agents are involved, multiplying the alternative actions needing consideration. We have described an approach in which no assumptions need be made about the values and preferences of others: all that is required is that the agent concerned can identify the values it recognises and indicate their relative worth to itself. In some cases success may still depend on what the other does, but this can be assessed using bounds on the probabilities of the alternatives available to the other. In this way we are able to achieve our objectives of allowing arguments which consider the actions of others, but which do not require assumptions about the beliefs and preferences of the others, while remaining consistent with multi-criteria utility theories, and the dominant actions of game theory. Thus we have shown how to:

- Remove the need to speculate on the preferences of other agents;
- Relate the value-based argumentation approach to approaches based on multi-criteria utility and game theory.
- Express reasons based on utility and expected returns as arguments, and objections to them, so that the arguments are genuinely for a particular action by the agent concerned rather than participation in a joint action, as was the case in [2].

We believe that this greatly improves the quality of value-based arguments for particular actions. Note also that the dominance of an action is *dominance for that agent*: it depends on the subjective values and aspirations of the individual agent. Which action is considered dominant by a particular agent or audience will depend on the values recognised, and the relative importance assigned to them, rather than fixed payoffs determined by the game, allowing each agent to set its own objectives. In addition to providing some initial results, we have, for future work, set out how the approach can be more broadly empirically tested using both existing and new experimental studies.

Appendix: Formal Definitions

Definition 1: AATS [46]. An *Action-based Alternating Transition System* (AATS) is an (n + 7)-tuple $S = \langle Q, q_0, Ag, Ac_1, \ldots, Ac_n, \rho, \tau, \Phi, \pi\rangle$, where:

- Q is a finite, non-empty set of *states*;
- $q_0 \in Q$ is the *initial state*;
- $Ag = \{1,...,n\}$ is a finite, non-empty set of *agents*;
- Ac_i is a finite, non-empty set of actions, for each $ag_i \in Ag$ where $Ac_i \cap Ac_j = \emptyset$ for all $ag_i \neq ag_j \in Ag$;
- $\rho : Ac_{ag} \to 2^Q$ is an *action pre-condition function*, which for each action $\alpha \in Ac_{ag}$ defines the set of states $\rho(\alpha)$ from which α may be executed;
- $\tau : Q \times J_{Ag} \to Q$ is a partial *system transition function*, which defines the state $\tau(q, j)$ that would result by the performance of j from state q. This function is partial as not all joint actions are possible in all states;
- Φ is a finite, non-empty set of *atomic propositions*; and
- $\pi : Q \to 2^\Phi$ is an interpretation function, which gives the set of primitive propositions satisfied in each state: if $p \in \pi(q)$, then this means that the propositional variable p is satisfied (equivalently, true) in state q.

AATSs are particularly concerned with the joint actions of the set of agents Ag. j_{Ag} is the joint action of the set of n agents that make up Ag, and is a tuple $\langle\alpha_1,...,\alpha_n\rangle$, where for each α_j (where $j \leq n$) there is some $ag_i \in Ag$ such that $\alpha_j \in Ac_i$. Moreover, there are no two different actions α_j and $\alpha_{j'}$ in j_{Ag} that belong to the same Ac_i. The set of all joint actions for the set of agents Ag is denoted by J_{Ag}, so $J_{Ag} = \prod_{i \in Ag} Ac_i$. Given an element j of J_{Ag} and an agent $ag_i \in Ag$, ag_i's action in j is denoted by j^i. This definition was extended in [2] to allow the transitions to be labelled with the values they promote.

Definition 2: AATS+V. Given an AATS, an AATS+V is defined by adding two additional elements as follows:

- V is a finite, non-empty set of values.
- $\delta : Q \times Q \times V \to \{+, -, =\}$ is a *valuation function* which defines the status (promoted (+), demoted (–) or neutral (=)) of a value $v_u \in V$ ascribed to the transition between two states: $\delta(q_x, q_y, v_u)$ labels the transition between q_x and q_y with one of $\{+, -, =\}$ with respect to the value $v_u \in V$.

An *Action-based Alternating Transition System with Values* (AATS+V) is thus defined as a (n + 9) tuple $S = \langle Q, q_0, Ag, Ac_1, ..., Ac_n, \rho, \tau, \Phi, \pi, V, \delta\rangle$. The value may be ascribed on the basis of the source and target states, or in virtue of an action in the joint action, where that action has intrinsic value.

Definition 3: Agent Preferences

The preferences of an agent $ag \in Ag$ is the set $O_{ag} = \{\langle v_0 * w_0\rangle, \langle v_1 * w_1\rangle, ..., \langle v_n * w_n\rangle\}$, where $v_0...v_n$ are values and $w_0...w_n$ are weights with $w_0 \geq w_1 \geq ... \geq w_n$.

Using these weights we can calculate the expected utility of agent i performing α. We will assume that if the desired joint action (j_0) does not result from the performance of α the worst case alternative joint action (j_w) will be the one that does result (providing a lower bound). Informally the expected utility of performing α will be the utility of j_0 multiplied by the probability of j_0 plus the utility of j_w (which will often be negative) multiplied by (1 minus the probability of j_0).

Definition 4: Expected Utility of ag performing α in state q_s

- Let $J_\alpha = \{j_0, j_1...j_n\}$ be the set of joint actions in which ag performs α (i.e. $j^{ag} = \alpha$) available in the starting state, q_s.
- Let P_{ag_k} be the values for ag promoted by the performance of $j_k \in J_\alpha$ in q_s. Let D_{ag_k} be the values of ag demoted by the performance of $j_k \in J_\alpha$ in q_s.
- The positive utility for ag, $pu(ag, j_k)$, of the performance of $j_k \in J_\alpha$ in q_s is $\Sigma_{i=0}^{i=n}(v_i * w_i)$ where $v_i \in P_{ag_k}$ and the negative utility for ag, $du(ag, j_k)$, of the performance of $j_k \in J_\alpha$ in q_s is $\Sigma_{i=0}^{i=n}(v_i * w_i)$ where $v_i \in D_{ag_k}$. The utility, $u(ag, j_k)$, for ag of the performance of $j_k \in J_\alpha$ in q_s is $pu(ag, j_k)$ - $du(ag, j_k)$.
- Let U_{ag} be the set of utilities for ag, $\{u_0, u_1...u_n\}$, such that $u_k = u(ag, j_k)$ for $j_k \in J_\alpha$. Let u_w be such that for all $u_i \in U_{ag}$, $u_w \leq u_i$.
- Let $prob(j_0)$ be the probability of j_0 being the joint action performed when ag performs α in q_s.
- Now the expected utility, $eu_{ag}(\alpha)$ for ag of performing α in q_s is $(u(ag, j_0) * prob(j_0)) + (u(ag, j_w) * (1 - prob(j_0)))$.

REFERENCES

[1] Rajeev Alur, Thomas A Henzinger, and Orna Kupferman, 'Alternating-time temporal logic', *Journal of the ACM (JACM)*, **49**(5), 672–713, (2002).

[2] Katie Atkinson and Trevor Bench-Capon, 'Practical reasoning as presumptive argumentation using action based alternating transition systems', *Artificial Intelligence*, **171**(10), 855–874, (2007).

[3] Katie Atkinson, Trevor Bench-Capon, and Peter McBurney, 'Computational representation of practical argument', *Synthese*, **152**(2), 157–206, (2006).

[4] Katie Atkinson, Trevor Bench-Capon, and Sanjay Modgil, 'Argumentation for decision support', in *Database and expert systems applications*, pp. 822–831. Springer, (2006).

[5] Katie Atkinson and Trevor J. M. Bench-Capon, 'Taking the long view: Looking ahead in practical reasoning', in *Computational Models of Argument - Proceedings of COMMA 2014*, pp. 109–120.

[6] Robert Axelrod, 'An evolutionary approach to norms', *American political science review*, **80**(04), 1095–1111, (1986).

[7] Trevor Bench-Capon, Katie Atkinson, and Alison Chorley, 'Persuasion and value in legal argument', *Journal of Logic and Computation*, **15**(6), 1075–1097, (2005).

[8] Trevor Bench-Capon, Katie Atkinson, and Peter McBurney, 'Using argumentation to model agent decision making in economic experiments', *Autonomous Agents and Multi-Agent Systems*, **25**(1), 183–208, (2012).

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

[10] Brigitte Burgemeestre, Joris Hulstijn, and Yao-Hua Tan, 'Value-based argumentation for justifying compliance', *Artificial Intelligence and Law*, **19**(2-3), 149–186, (2011).

[11] Dan Cartwright and Katie Atkinson, 'Using computational argumentation to support e-participation', *Intelligent Systems, IEEE*, **24**(5), 42–52, (2009).

[12] Günther Charwat, Wolfgang Dvořák, Sarah A Gaggl, Johannes P Wallner, and Stefan Woltran, 'Methods for solving reasoning problems in abstract argumentation–a survey', *Artificial intelligence*, **220**, 28–63, (2015).

[13] Phan Minh Dung, 'On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games', *Artificial intelligence*, **77**(2), 321–357, (1995).

[14] Paul E. Dunne, 'Tractability in value-based argumentation', in *Proceedings of COMMA 2010*, pp. 195–206, (2010).

[15] Uwe Egly, Sarah Alice Gaggl, and Stefan Woltran, 'Answer-set programming encodings for argumentation frameworks', *Argument and Computation*, **1**(2), 147–177, (2010).

[16] Christoph Engel, 'Dictator games: a meta study', *Experimental Economics*, **14**(4), 583–610, (2011).

[17] Sevan G Ficici and Avi Pfeffer, 'Simultaneously modeling humans' preferences and their beliefs about others' preferences', in *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 1*, pp. 323–330. International Foundation for Autonomous Agents and Multiagent Systems, (2008).

[18] Richard E Fikes and Nils J Nilsson, 'Strips: A new approach to the application of theorem proving to problem solving', *Artificial intelligence*, **2**(3-4), 189–208, (1971).

[19] Artur SD'Avila Garcez, Dov M Gabbay, and Luis C Lamb, 'Value-based argumentation frameworks as neural-symbolic learning systems', *Journal of Logic and Computation*, **15**(6), 1041–1058, (2005).

[20] Matthias Grabmair and Kevin D Ashley, 'Facilitating case comparison using value judgments and intermediate legal concepts', in *Proceedings of the 13th international conference on Artificial intelligence and law*, pp. 161–170. ACM, (2011).

[21] Floriana Grasso, Alison Cawsey, and Ray Jones, 'Dialectical argumentation to solve conflicts in advice giving: a case study in the promotion of healthy nutrition', *International Journal of Human-Computer Studies*, **53**(6), 1077–1115, (2000).

[22] Joseph Henrich, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, and Richard McElreath, 'In search of homo economicus: behavioral experiments in 15 small-scale societies', *The American Economic Review*, **91**(2), 73–78, (2001).

[23] Souhila Kaci and Leendert van der Torre, 'Preference-based argumentation: Arguments supporting multiple values', *International Journal of Approximate Reasoning*, **48**(3), 730–751, (2008).

[24] Samhar Mahmoud, Nathan Griffiths, Jeroen Keppens, Adel Taweel, Trevor JM Bench-Capon, and Michael Luck, 'Establishing norms with metanorms in distributed computational systems', *Artificial Intelligence and Law*, **23**(4), 367–407, (2015).

[25] John Stuart Mill, 'On the definition of political economy, and on the method of investigation proper to it', *Essays on some unsettled questions of Political Economy*, 326, (1844).

[26] Sanjay Modgil and Trevor J. M. Bench-Capon, 'Integrating object and meta-level value based argumentation', in *Proceedings of COMMA 2008*, pp. 240–251, (2008).

[27] Roger B Myerson, 'Game theory: analysis of conflict', *Harvard University*, (1991).

[28] Fahd Saud Nawwab, Trevor J. M. Bench-Capon, and Paul E. Dunne, 'A methodology for action-selection using value-based argumentation', in *Computational Models of Argument: Proceedings of COMMA 2008*, pp. 264–275, (2008).

[29] Samer Nofal, Katie Atkinson, and Paul E Dunne, 'Algorithms for decision problems in argument systems under preferred semantics', *Artificial Intelligence*, **207**, 23–51, (2014).

[30] Hessel Oosterbeek, Randolph Sloof, and Gijs Van De Kuilen, 'Cultural differences in ultimatum game experiments: Evidence from a meta-analysis', *Experimental Economics*, **7**(2), 171–188, (2004).

[31] Simon Parsons and Michael Wooldridge, 'Game theory and decision theory in multi-agent systems', *Autonomous Agents and Multi-Agent Systems*, **5**(3), 243–254, (2002).

[32] Terry R Payne and Valentina Tamma, 'Using preferences in negotiations over ontological correspondences', in *PRIMA 2015: Principles and Practice of Multi-Agent Systems*, 319–334, Springer, (2015).

[33] Chaim Perelman, *The new rhetoric*, Springer, 1971.

[34] Anatol Rapoport and Albert M Chammah, *Prisoner's dilemma: A study in conflict and cooperation*, volume 165, University of Michigan press, 1965.

[35] Joseph Raz, *Practical Reasoning*, Oxford University Press, Oxford, 1979.

[36] John R Searle, *Rationality in action*, MIT press, 2003.

[37] Theo J Stewart, 'A critical survey on the status of multiple criteria decision making theory and practice', *Omega*, **20**(5), 569–586, (1992).

[38] Jasmin Tremblay and Irène Abi-Zeid, 'Value-based argumentation for policy decision analysis: methodology and an exploratory case study of a hydroelectric project in québec', *Annals of Operations Research*, **236**(1), 233–253, (2016).

[39] Cássia Trojahn, Paulo Quaresma, and Renata Vieira, 'An extended value-based argumentation framework for ontology mapping with confidence degrees', in *Argumentation in Multi-Agent Systems*, 132–144, Springer, (2007).

[40] E. Ullmann-Margalit, *The emergence of norms*, Clarendon Press Oxford, 1977.

[41] Thomas L van der Weide, Frank Dignum, J-J Ch Meyer, Henry Prakken, and GAW Vreeswijk, 'Multi-criteria argument selection in persuasion dialogues', in *Argumentation in Multi-Agent Systems*, 136–153, Springer, (2011).

[42] D.N. Walton, *Argumentation Schemes for Presumptive Reasoning*, Lawrence Erlbaum Associates, Mahwah, NJ, USA, 1996.

[43] Douglas Walton and Erik CW Krabbe, *Commitment in dialogue: Basic concepts of interpersonal reasoning*, SUNY press, 1995.

[44] Douglas Walton, Christopher Reed, and Fabrizio Macagno, *Argumentation schemes*, Cambridge University Press, 2008.

[45] Maya Wardeh, Adam Wyner, Katie Atkinson, and Trevor Bench-Capon, 'Argumentation based tools for policy-making', in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law*, pp. 249–250. ACM, (2013).

[46] Michael Wooldridge and Wiebe van der Hoek, 'On obligations and normative ability: Towards a logical analysis of the social contract', *J. Applied Logic*, **3**(3-4), 396–420, (2005).

[47] Adam Wyner, Maya Wardeh, Trevor Bench-Capon, and Katie Atkinson, 'A model-based critique tool for policy deliberation.', in *JURIX*, pp. 167–176, (2012).