

Persuasive Political Argument

Katie Atkinson and **Trevor Bench-Capon** and **Peter McBurney**

Department of Computer Science,

University of Liverpool,

Liverpool L69 3BX, UK.

{k.m.atkinson,tbc,p.j.mcburney}@csc.liv.ac.uk

Abstract

In this paper we discuss how a computational version of argumentation involving practical reasoning can be applied to the domain of e-democracy. We discuss our previous work which proposed an argument scheme and associated critical questions to make use of presumptive reasoning in order to justify a proposal for action. We explain how this proposal can be made computational for use by BDI agents and we illustrate the approach with an example application of a recent political debate involving the Government's justification of a proposed action. We then go on to discuss some interesting issues arising from the example regarding the strength of the arguments and how they can be accrued to put forward a more convincing case for justifying the action. We also examine how this application compares with a previously implemented system which makes use of the same underlying theory.

1 Introduction

Debates about proposed actions are an inherent feature of a democratic society. Not only do government ministers debate the issues of the moment amongst themselves, but members of the public are also encouraged to engage in debates of current affairs. The passing of laws and motions is subject to open-ended discussion and questioning before issues are put to a vote. In discussing the issues the audience must be allowed explore the reasons behind justifications presented to them, as well as being able to pose critical questions to the proponents of arguments in order to consider all aspects of the case. One of the most widely debated, and perhaps most controversial debates to feature in the recent international political agenda has been the invasion of Iraq. Not only did this issue spark debate at national levels, it also received a great amount of time and attention at an international level. Debates of such importance require clarity about the issues and any arguments advanced by parties need to withstand critical arguments in order to be justifiable. In this paper we seek to take a previous theory developed by the authors of persuasion over action [2] and apply it to this particular scenario to show how it can be used by autonomous agents to effectively aid such reasoning. Section 2 of the paper recapitu-

lates our general approach to practical reasoning through the use of an argument scheme and critical questions. This section then goes on to provide definitions and preconditions for instantiating the argument scheme and challenging instantiations of it for use in multi agent systems based on the Belief-Desire-Intention (BDI) model [9]. We then briefly explain how the presumptive arguments and attacks generated are resolved through organisation into a Value-Based Argumentation Framework (VAF)[4], which enables the calculation of the dialectical status of the various arguments, with respect to the participants in the debate. Section 3 shows how this general theory is applied to the particular example. Our approach models the various participants in the debate as different agents. These agents subscribe to individual beliefs, goals and values, and therefore can represent the different views that can be brought to bear on the problem. Here we will show the relations between these these views and how the arguments can be evaluated through the use of VAFs. In Section 4 we will discuss the notion of accrual and how unacceptable arguments can become acceptable for a particular agent through the support of cumulative arguments, which may individually be unacceptable. In Section 5 we will briefly compare the model of the debate presented here with a system previously implemented by the authors [3] which deals with the same political issue. Finally, in Section 6 we will offer some concluding remarks.

2 General Approach to Practical Reasoning with Agents

One way of addressing practical reasoning is as a species of presumptive argument: given an argument, we have a presumptive reason for performing the action. This presumption can, however, be challenged and withdrawn. Subjecting our argument to appropriate challenges is how we hope to identify and consider the alternatives that require consideration, and determine the best choice for us, in the particular context. One account of presumptive reasoning is in terms of argument schemes and critical questions, as given in [8]. The idea here is that an instantiated argument scheme gives a presumption in favour of its conclusion. Whether this presumption stands or falls depends on satisfactory answers being given to the critical questions associated with the scheme.

In [2] we presented a scheme for practical reasoning. Our

scheme extends the Sufficient Condition Scheme of Walton [8] by unpacking his notion of a goal into three elements: the state of affairs brought about by the action; the goal (the desired features in that state of affairs); and, the value (the reason why those features are desirable). Our scheme is expressed as follows:

AS1: In the current circumstances R
 we should perform action A
 to bring about new circumstances S
 which will realise goal G
 which will promote value V

This presumptive justification may be attacked in a large variety of ways, as described in [7], as the various elements and the connections between them are open to question, and additionally there may be alternative possible actions, and side effects of the proposed action. We will define the subset of attacks used in the example of this paper later on in this section.

The computational setting for our approach is a multi agent system, in which the agents form intentions based on their beliefs and desires. This is essentially the standard BDI agent model [9], except that we make a small extension by associating each desire with a value, the reason why it is desirable.

We use the following definitions taken from our general account which is specified in [1]:

2.1 Definitions

Definition 1: *The Beliefs of an Agent.* The beliefs of an Agent J is a four tuple $\langle W_J, A_J, D_J, V_J \rangle$ where,

W_J represents beliefs of Agent J about the world;
 A_J represents beliefs of Agent J about actions;
 D_J represents beliefs about the desires of Agent J;
 V_J represents beliefs about the values of Agent J;

Definition 2: *Beliefs about the World.* The beliefs about the world of Agent J is a set of triples $\langle p, cert_{p,J}, t \rangle$ where,

p is a proposition; $cert_{p,J} = -1 \leq cert_{p,J} \leq 1$; t is a time.

We interpret this as J has $cert_{p,J}$ regarding p at time t . If $cert_{p,J} = -1$, J believes p to be definitely false, if $cert_{p,J} = 1$, J believes p to be definitely true, and if $cert_{p,J} = 0$, J has no opinion as to the truth of p .

Let M denote the set of all agents in the system and T the set of all times.

The set P denotes the set of all p such that $\langle p, cert_{p,J}, t \rangle \in W_J$ for some agent $J \in M$ and some time $t \in T$.

Definition 3: *Beliefs about Actions.* The beliefs about action of Agent J is a set of triples $\langle a, Pre_{a,J}, Post_{a,J} \rangle$ where,

action a is an action; $Pre_{a,J}$ is a set of pairs $\langle p, threshold_{p,J} \rangle$ and $Post_{a,J}$ is a set of pairs $\langle p, truth_{p,J} \rangle$,

$-1 \leq threshold_{p,J} \leq 1$, and $-1 \leq truth_{p,J} \leq 1$.

$Pre_{a,J}$ is a set of preconditions for *action a* recognised by agent J. The interpretation is that J believes that *action a* can be performed at t if all elements of $Pre_{p,J}$ are satisfied with respect to W_J at t .

$\langle p, threshold_{p,J} \rangle$ is satisfied with respect to W_J if $\langle p, cert_{p,J}, t \rangle$ and if $threshold_{p,J} > 0$, then $cert_{p,J} \geq threshold_{p,J}$, else if $threshold_{p,J} < 0$, $cert_{p,J} \leq threshold_{p,J}$. J believes that if *action a* is performed at t , then for all $\langle p, truth_{p,J} \rangle \in Post_{a,J}$, $\langle p, truth_{p,J}, t+1 \rangle$ will be an element of W_J .

$W_{J,a}$ is the state of the world that J believes will result from performing *action a*.

Additionally, J may *assume* that *action a* can be performed at t if all elements of $Pre_{a,J}$ can be *assumed to be satisfied* with respect to W_J at t . $\langle p, threshold_{p,J} \rangle$ can be assumed satisfied with respect to W_J if $\langle p, cert_{p,J}, t \rangle$ and if $threshold_{p,J} > 0$, then $cert_{p,J} \geq 0$ and if $threshold_{p,J} < 0$, $cert_{p,J} \leq 0$.

The set A denotes the set of all actions such that $\langle a, Pre_{a,J}, Post_{a,J} \rangle \in A_J$ for some agent $J \in M$.

Definition 4: *Desires of an Agent.* The desires of an Agent J is a set of pairs $\langle d, Cond_{d,J} \rangle$ such that,

d is a desire and $Cond_{d,J}$ is a set of pairs $\langle p, threshold_{p,J} \rangle$. The interpretation is that J believes that d is satisfied at t if $Cond_{d,J}$ is satisfied with respect to W_J at t . The notions of satisfaction and assumed satisfaction for $Cond_{d,J}$ is the same as that for $Pre_{a,J}$.

The set D denotes the set of all desires such that $\langle d, Cond_{d,J} \rangle \in D_J$ for some agent $J \in M$.

Definition 5: *Values of an Agent.* The values of an Agent J is a set of triples $\langle v, d, prom_{v,J} \rangle$ such that,

v is a value,
 d is a desire,
 $prom_{v,J}$ a number $-1 \leq prom_{v,J} \leq 1$, representing the degree to which the satisfaction of d promotes v .

The set V denotes the set of all values such that $\langle v, d, prom_{v,J} \rangle \in V_J$ for some agent $J \in M$.

Definition 6: Let $satA(\text{Formula}, W_J)$ be true if Formula can be assumed to be satisfied with respect to W_J .

Let $satS(\text{Formula}, W_J)$ be true if Formula can be satisfied with respect to W_J .

Now J has a presumptive argument for *action a* at time t if:

there is an $\langle a, \text{Pre}_{aJ}, \text{Post}_{aJ} \rangle \in A_J$ such that:
 $\text{satA}(\text{Pre}_{aJ}, J)$ at t ;
 $\text{satA}(\text{Cond}_{dJ}, J)$ at $t+1$ and
 Cond_{dJ} will be satisfied at $t+1$ with respect to W_J ;
there is a $\langle v, d, \text{prom}_{vJ} \rangle$, such that $\text{prom}_{vJ} > 0$.

The position is expressed as:

In circumstances r , where each $r \in R$ is the first term,
in each element of Pre_{aJ} ,
Performing *action a*,
Will result in s , where each $s \in S$ is the first term in
each element of Post_{aJ} ,
Which will realise d ,
Which promotes v .

On the basis of these definitions we can state the following preconditions for attacking such presumptive arguments. Each of these attacks may be associated with a critical question, as described in [2]. For each attack we give the critical question from which it derives and a natural language expression of the attack. The complete set of preconditions for attacks is in [1]: here we give only those used subsequently in the paper.

Source CQ: Are the believed circumstances true?

Attack 1a: *Pre-conditions for A_K to make an attack:*

$\text{satA}(\text{Pre}_{aK}, W_K)$ and,
not $\text{satS}(\text{Pre}_{aK}, W_K)$.

Argument: p may not be true.

Source CQ: Assuming the circumstances are true, does the action have the stated consequences?

Attack 2a: *Pre-conditions for A_K to make an attack:*

$\text{satA}(\text{Post}_{aK}, W_K)$ and,
not $\text{satS}(\text{Post}_{aK}, W_K)$.

Argument: *action a* may not have the desired consequences.

Source CQ: Assuming the circumstances are true and the actions has the stated consequences, will the action bring about the desired goal? (CQ3).

Attack 3a: *Pre-conditions for A_K to make an attack:*

for no $\langle d, \text{Cond}_{dK} \rangle$ does $\text{satA}(\text{Cond}_{dK}, W_{Ka})$ hold.

Argument: the state of affairs resulting from performing the action will not bring about the goal.

Source CQ: Does the goal realise the value intended? (CQ4).

Attack 4a: *Pre-conditions for A_K to make an attack:*

$\langle v, d, \text{prom}_{vK} \rangle$ and,
 $\text{prom}_{vK} \leq 0$.

Argument: the goal may not promote the value.

Source CQ: Are there alternative ways of realising the same goal?

Attack 6: *Pre-conditions for A_K to make an attack:*

$\text{satA}(\text{Pre}_{bK}, W_K)$ and,
 $\text{satA}(\text{Cond}_{dK}, W_{Kb})$ and $b \neq a$.

Argument: there is an alternative action which will realise the same goal.

Source CQ: Does doing A have a side effect which demotes the value V ?

Attack 8: *Pre-conditions for A_K to make an attack:*

$\text{satA}(\text{Cond}_{eK}, W_{Ka})$, $e \neq d$ and,
 $\langle v, e, \text{prom}_{vK} \rangle$ and,
 $\text{prom}_{vK} < 0$.

Argument: *action a* has a side effect which satisfies an alternative desire, which demotes the value.

Source CQ: Does doing A have a side effect which demotes some other value? (CQ9).

Attack 9: *Pre-conditions for A_K to make an attack:*

$\text{satA}(\text{Cond}_{eK}, W_{Ka})$, $e \neq d$ and,
there is a w , $w \neq v$ such that $\langle w, e, \text{prom}_{wK} \rangle$ and,
 $\text{prom}_{wK} < 0$.

Argument: *action a* has a side effect which satisfies an alternative goal, which demotes some other value.

Source CQ: Would doing A promote some other value? (CQ10).

Attack 10: *Pre-conditions for A_K to make an attack:*

$\text{satA}(\text{Cond}_{eK}, W_{Ka})$, $e \neq d$ and,
there is a w , $w \neq v$ such that $\langle w, e, \text{prom}_{wK} \rangle$ and,
 $\text{prom}_{wK} > 0$.

Argument: *action a* has a side effect which satisfies an alternative goal, which promotes some other value.

Now given a set of agents and a situation in which an action must be chosen we can first instantiate a number of presumptive justifications for each agent, and then see which attacks the agents can make on these justifications. The result will be a set of arguments and a set of attack relations between them, providing the key elements for an argumentation framework

[6]. Moreover, since the arguments produced by instantiating AS1 are associated with values, we can see this as a VAF [4], and use the procedures in that paper to calculate the dialectical status of the arguments with respect to the different audiences represented by the different agents. This process will be illustrated by a detailed working through of our example debate in the section below.

3 Political Example

The example we will use is based on the debate which took place in 2003 as to whether the UK should go to war with Iraq (set before Iraq was invaded). This was perhaps one of the most widely debated issues of recent years, and disagreement as to the motives and justification of the action taken remains to this day. Our example will model the viewpoint of the Government in putting forward its position on the issue and some of the attacks that this justification elicited from members of Parliament and the public.

3.1 Context

In our reconstruction of the arguments we will use seven different agents to represent the different views put forward by the parties involved. Firstly we have four agents advocating the action of invading Iraq but for different, though sometimes overlapping reasons. We will refer to these agents as: G, representing an agent named George; T, representing an agent named Tony; D, representing an agent named Donald; and C representing an agent named Colin. We will also make use of three agents who oppose the action of invading Iraq, again for different reasons. We will refer to these agents as: M, representing an agent named Michael; R, representing an agent named Robin; and J, representing an agent named Jacques.

We begin by instantiating our agents with the appropriate beliefs, desires and values, in accordance with the definitions given in Section 2.

We use six possible propositions about the world to describe the given situation and these are as follows:

- P1: Saddam has weapons of mass destruction (WMD).
- P2: Saddam is a dictator.
- P3: Saddam will not disarm voluntarily.
- P4: Saddam is a threat to his neighbours.
- P5: Saddam is defying the UN.
- P6: Saddam is running an oppressive regime.

Our agents differ quite widely as to which propositions are believed true. Each agent subscribes to the propositions as shown in Table 1 with 1 representing belief in the proposition, -1 representing disbelief in the proposition and U representing unknown to show that the agent has subscribed to neither belief nor disbelief in the proposition.

We now identify the desires and values. For definition 4 we need to identify a set of desires for the agents, and give conditions under which the agents will accept that these desires are realised. Definition 5 requires us to associate these

Table 1: **Propositions about the World**

| Agent | P1 | P2 | P3 | P4 | P5 | P6 |
|-------|----|----|----|----|----|----|
| G | 1 | 1 | 1 | 1 | 1 | 1 |
| T | 1 | 1 | 1 | U | U | 1 |
| D | 1 | U | 1 | 1 | U | U |
| C | U | 1 | U | U | U | 1 |
| M | -1 | 1 | -1 | -1 | -1 | 1 |
| R | U | 1 | U | U | U | 1 |
| J | U | 1 | U | -1 | U | 1 |

desires with a value, and a degree to which the satisfaction of the desire promotes the value. For now we list the set of desires, conditions and values in Table 2 below and we will go on to discuss degrees of promotion in the next section.

Table 2: **Possible Desires and Values in the Debate**

| No. | Desire | Value | Condition to be satisfied |
|-----|-------------------------|----------------------|----------------------------------|
| 1 | No WMD | World Security | Iraq has no WMD |
| 2 | No dictator | World Security | Saddam deposed |
| 3 | Democracy in Iraq | Human Rights | Saddam deposed |
| 4 | International agreement | Good World Relations | All allies agree with the action |
| 5 | No human casualties | Respect for Life | No war |

Based upon the beliefs and desires given in the above tables, each agent can provide one or more instantiation of AS1. The figures presented below give two argumentation frameworks to show the views of the agents. Initially we can see that there are two values involved in the debate: ‘world security’ and ‘human rights’. The Government’s argument provides two major justifications (which are instantiations of AS1) which endorse the same action of invading Iraq. However, each justification provides different reasons and promotes different values, even though both justifications endorse the same action. We can therefore construct two argumentation frameworks to show the instantiations of AS1 representing the Government’s justifications and the attacks on these instantiations that can be made by the opposing agents. These attacks instantiate the remaining three values. We present these argument schemes, frameworks and attacks below. In the figures, nodes represent arguments. They are labelled with a description of the argument, and on the right hand side, with letters representing the agents introducing the argument. Arcs are labelled with the number of the attack they represent. We then summarise what can be deduced from each framework.

3.2 Argument Based on Threat

Firstly we present in Figure 1 argument scheme Arg1 showing the Government’s first justification of the action and the attacks made on it by opposing agents:

Arg1

R1: Saddam has WMD, Saddam is a dictator, Saddam will not disarm voluntarily, Saddam is a threat to his neighbours, Saddam is defying the UN.

A1: we should invade Iraq

S1: which would get rid of the WMD and depose the dictator

G1: so this will remove the threat that Saddam poses to his neighbours and assert the authority of the UN

V1: which will promote world security.

This argument and the attacks that can be made on it by opposing agents, given their beliefs and desires, are represented in Figure 1 below:

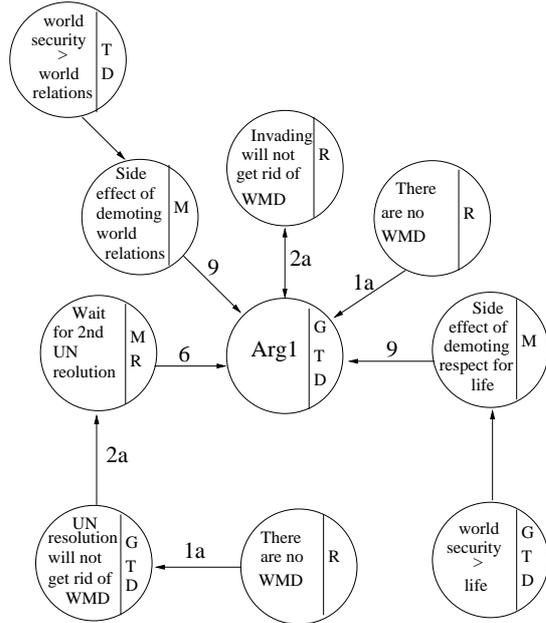


Figure 1. Arg1 and the attacks on it.

Looking at the above argumentation framework we can see that the agents subscribe to the following arguments:

Agents G, T and D all put forward Arg1 to justify the action of invading Iraq. The first challenge to be made on this is executed by agent R who uses attack 1a to deny fact 1 presented in R1 of Arg1. This attack states that agent R does not believe that Saddam has weapons of mass destruction (WMD). Agent R also does not believe that invading Iraq will get rid of the WMD and so he makes attack 2a as well. Agent M points out that there is a bad side effect of the action in that the unconsidered value of good world relations will be demoted due to there being international disagreement about the proposed invasion. This is stated in attack 9. This is counter-attacked by agents T and D who state that they rank the value of world security higher than they rank the value of good international relations. Agents M and R then make a new attack to propose an alternative action to realise the goal. Using attack 6, the alternative action they propose here is waiting for a second UN resolution on the matter. However, this is attacked by all agents supporting Arg1, through attack 2a stating that

this newly proposed action will not achieve the goal, as waiting for a second UN resolution will not get rid of the WMD. This argument is counter-attacked by agent R who again uses attack 1a to state that he does not believe that there are any WMD in Iraq in the first place. The final attack on Arg1 is made by agent M who believes that the action will have the bad side effect of demoting the value of respect for life and he uses attack 9 to state this. However, this is attacked by all proponents of Arg1 through their statement of value preference in which they rank world security as a more important value than respect for life, in this situation.

3.3 Argument Based on Regime Change

Now that all the agents' arguments have been articulated regarding the justification in Arg1, we turn to Arg2 in which agents C and T provide a second justification for the same action:

Arg2

R2: Saddam is running an oppressive regime.

A2: we should invade Iraq

S2: to depose Saddam

G2: which will bring democracy to Iraq

V2: which will promote human rights.

This argument and the attacks on it are represented in Figure 2 below:

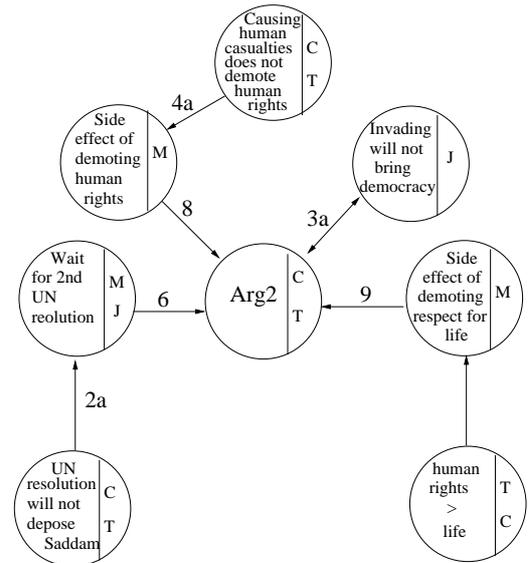


Figure 2. Arg2 and the attacks on it.

Looking at the above argumentation framework we can see that the agents subscribe to the following arguments:

Firstly, we can see from Arg2 that this argument is based on the fact that Saddam is running an oppressive regime and unlike in the last justification no agent disagrees outright with this fact, as we can see from Table 1. So, the first attack made on Arg2 is by agent J who, using attack 3a, states that the action of invading Iraq will not result in democracy being achieved. Agent M then makes attack 8 stating that there

is the side effect on the action of demoting the value of human rights. This is itself attacked by agents C and T who use attack 4a to state that causing human casualties does not demote human rights as they believe that human casualties may be a necessary evil involved in bringing about democracy. Next, as in the previous framework, attack 6 is used to propose the alternative action of waiting for a second UN resolution and this is put forward by agents M and J. As before, this argument is counter attacked using attack 2a to state that this action will not achieve the goal of deposing Saddam and this is put forward by both agents supporting Arg2. However, this time no agent can attack this argument as agent M did in the previous framework, because they all believe the facts upon which the argument is based to be true. The final attack made on Arg2 is attack 9 in which agent M states that the action of invading Iraq again has the side effect of demoting the value of respect for life. Agents C and T both attack this by restating their belief that the value of human rights is more important than the value of respect for life. This concludes the analysis of all the individual attacks used in each framework.

3.4 Discussion

We can see from the above analysis that all agents involved in the discussion have different, but sometimes overlapping reasons for their opinions in the debate. In framework 1 we are able to see that agents G, T and D all accept to Arg1 on the basis that they believe Saddam has WMD which he is willing to use to detrimental effect. However, only agents T and D express a value preference of world security over world relations, which they use to defeat attack 9. From this we can see that agent G agrees that there may be the possible side effect of the action of demoting world relations, which is pointed out in attack 9. However, he does go on to defend an attack against the second use of attack 9 by stating that he ranks world security over respect for life. From this we can deduce that agent G only needs to have one goal (as opposed to all goals) of Arg1 satisfied in order to justify the action: asserting the authority of the UN is not important to G. As both agents T and D defend all attacks made by the opposing agents, they require all consequences and goals to be satisfied in order for them to be able to justify the action.

Of the opposing agents in framework 1, agent R rejects Arg1 on the basis the he believes the facts upon which it is based are false i.e. there are no WMD. Agent M rejects the argument on a different basis through reasons that reveal he does not support war rather than refuting the claim that there are WMD. Agent J does not feature in this framework as his views solely relate to arguments about the conveyance of democracy.

In framework 2 we can see that agent T supports this argument in addition to Arg1 and so he is the only agent who sees the need to justify both arguments in order to be able to justify the action. However, agent C also supports Arg2 and as he did not support Arg1 we conclude that he believes Arg2 to be sufficient on its own for the justification of the action.

Of the opposing agents M again reveals his anti-war attitude through the attacks he makes. Agent J disagrees with the result of the action showing that his attitude reflects the

belief that democracy will not be achieved through invasion, which is the main thesis of his position. Agent R does not feature in this scenario as he is only interested in arguments resting on the basis of the evidence of WMD.

4 Accrual and Strength of Argument

In addition to the individual attacks in the frameworks there is also an attack that can be made between the two argument schemes Arg1 and Arg2, and this is attack 10. This is shown below in Figure 3.

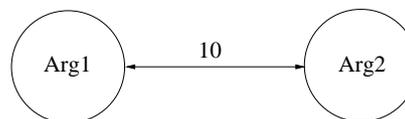


Figure 3. Arg1 and Arg2 mutually attack each other.

An agent making attack 10 does not dispute that the action should be performed, but disputes the motive for performing it. In our example, G and D accept Arg1 but not Arg2 and C accepts Arg2 but not Arg1. Thus G may attack Arg2 by saying that regime change is not a justification for invasion, but removing WMD is, and C may attack Arg1 by arguing the contrary. The case of T is different, since he accepts both arguments. If T accepts that both Arg1 and Arg2 are sufficient to justify invasion, he could be challenged to choose between world security, the value promoted by Arg1, and human rights, the value promoted by Arg2, so as to clarify his “real” reason for advocating invasion. In practice some politicians seemed to be in the position of T, and generally made the removal of WMD their lead justification, although subsequent to the failure to discover WMD, they cite regime change as sufficient in itself. If, however, desires can promote values more or less strongly, it may be that only one of the arguments is sufficient to justify the action. This will then be the “real” reason, and the other argument is superfluous. A different case is where neither argument is sufficient by itself to promote the action. Here attack 10 is inappropriate, since the two arguments are now intended to be mutually supporting and the action is justified only if both arguments stand. An additional precondition for attack 10 is thus that the attacking argument be sufficient to justify the action. The need in some cases to have mutually supporting arguments introduces the notion of accrual. To explore this notion we need to have some mechanism for distinguishing degrees of promotion, and determining when an argument is sufficient to justify the action. The definitions in Section 2.1 allowed for both these ideas, and we illustrate them with an example in the next sub section.

4.1 Degrees of Promotion

The definitions in Section 2.1 represent degrees of promotion as numbers ranging from -1 to 1. This is intended to be flexible enough to accommodate a variety of concrete treatments. Assigning and combining things such as degrees of promotion presents conceptual and practical problems. What we propose here is admittedly rudimentary: we make no claims for its cognitive validity, but use it only to illustrate some points about accrual. We assume that agents can make a subjective assessment of the degree of promotion by responding

to a question such as “from the standpoint of human rights, how important is it that a country be ruled democratically” with a qualitative assessment such as “utterly, very, somewhat, a little, or not at all”. We further assume that satisfying several desires promoting a value will promote that value to a degree greater than is achieved by satisfying only one of them. Our treatment is consistent with these assumptions, but remains rather *ad hoc*. We translate the qualitative assessments into numbers as follows:

- if the agent replies “not at all” the value is *not* promoted and we assign the number 0;
- if the agent replies “a little” the value is *weakly* promoted and we assign 0.3;
- if the agent replies “somewhat” the value is *moderately* promoted and we assign 0.5;
- if the agent replies “very” the value is *strongly* promoted and we assign 0.7;
- if the agent replies “utterly” the value is *fully* promoted and we assign 1.

For combinations of two desires promoting the same value, or two arguments proposing the same action we add the relevant numbers and subtract their product: i.e. $combine(a,b)=(a+b)-(a*b)$. This is the formula used for rule combination in MYCIN [5]. Again we make no claims for this other than that it satisfies the desired property of increasing the degree of satisfaction while remaining in range. We also, again quite arbitrarily, take 0.7 as the threshold which must be attained if the action is to be justified.

We now return to the example. Agent T needed to accept both Arg1 and Arg2 to convince him to act. Suppose he sees both arguments as moderately supporting the action, both with value 0.5. Separately neither is a sufficient reason to act, but together they support the action at 0.75, and so exceed the threshold. Note that if one argument had been no more than weakly supportive, the combined value would be only 0.65: some third argument offering at least weak support would be required to reach the threshold. In contrast, agents G and D see Arg1 as sufficient to justify the action on its own, and agent C sees Arg2 as sufficient in itself. For these agents there is a single reason to act and attack 10 will lead them to reject the other argument.

This illustrates one form of accrual, where two distinct arguments are involved. Arg1 illustrates a different form of accrual in that it is based on the satisfaction of two desires promoting the same value. If we separate these into Arg1a, based on the removal of Saddam’s threat, and Arg1b, based on asserting the authority of the UN, we can see that there are various possibilities. It may be that for some agents one of the desires promotes the value sufficiently to support the action on its own. Suppose that, as for agent G in Figure 1, the removal of the threat strongly promotes world security. Then Arg1a provides sufficient justification and T need not defend attacks directed at the part of the goal representing Arg1b. Agents D and T, in contrast, may see both parts of the goal as necessary (because both only moderately promote the value) and so must defend it against all attacks. Thus we can use two distinct types of accrual depending on whether one or more

values are promoted. Different values require two arguments and we can ask which provides the *real* motive to act. One value can be represented as a single instance of the argument scheme. In this case, because only one value is involved, a person who believes that an argument based on one of the desires is sufficient need not reject the argument based on the other desire: the second argument is at worst superfluous and may strengthen the justification for the action.

It is often the case that in everyday practical reasoning we may be unsure as to whether to execute an action or not. However, if we find a separate reason (even if it is weak itself) to execute the action then this can compel us enough to carry out the action. This notion of accrual is also mentioned by Walton in [8] in his argument scheme for the ‘Argument From Sign’. Here he states that the more signs that are brought into the scheme, the more inferences can be drawn from them and the more the case builds up. This is particularly relevant for the political forum as the more evidence that is presented to justify an action, the more likely ministers and the public are to accept the argument and win the support of the Government.

The mechanisms we have used to implement degrees of promotion and strengths of justification have been simple and preliminary. They are intended only to illustrate how our model can capture the useful notion of accrual, and can distinguish different types of accrual. We hope to give a more principled account of this in future work, though the treatment of it may be domain specific and we will look to existing accounts to guide our future work on this area.

We also note that in the definitions we gave in Section 2.1 we include degrees of certainty for the individual elements of a position. For example, an agent will only believe some proposition about the world if it passes a certain threshold for belief. Likewise, a value promotion will only be acceptable to an agent if their belief that the goal actually promotes the value passes the set threshold. We have not explored degrees of certainty of individual elements of a position here but it is something we hope to do in future. This would be particularly applicable to a domain that involves elements of risk, such as the medical one. For example, it may be crucial to have a strong belief that a treatment given to a particular patient will have the desired effect. Or it may be acceptable to give a patient a treatment with side effects, as long as we are sufficiently satisfied that these side effects will not be too detrimental for the patient’s health.

5 Comparison with the PARMENIDES System

In earlier work [3] we presented an implementation of a system named Parmenides, which is based upon the same general theory of persuasion over action that we described earlier in this paper. The idea of Parmenides is to provide a system which makes use of our argument scheme and critical questions to solicit the public’s views on a particular issue. The topic used in Parmenides is the same as that discussed here: the invasion of Iraq. Parmenides provides a simple web-based interface to guide users through the justification of an action as presented by the Government. Users are given the opportunity to disagree with the individual elements of the position

and propose alternatives. These points of disagreement each represent an individual attack from our theory of persuasion. In addition to this, after having critiqued the position presented to them, users can then go on to construct their own position regarding the matter. The responses of the users are written to a database to enable the proponent of the action (in this case the Government) to gather and analyse the information in order to identify what elements of an argument are more strongly supported than others. There are some interesting points to note when comparing Parmenides with the example application presented here.

Parmenides makes use of a subset of the attacks from our theory, as does the example presented here. These two subsets of attacks are extremely similar with the exception of three additional attacks being used in Parmenides (attacks 7, 11 and 15). The purpose of Parmenides was to build a system in which a user could critique a justification of an action in a particular domain and express their own views in the most complete way possible, using our theory. However, it would be perfectly acceptable to use the extra attacks found in Parmenides in the example presented here: they do not in fact arise because we have limited ourselves to only seven agents, whose beliefs and desires do not happen to satisfy the pre-conditions for these attacks. Thus, it would be possible to reconstruct all the arguments made in Parmenides in the format we have presented in this paper, though this was not our original aim of the exercise.

One of the main motivations of Parmenides was to provide a system which facilitated debate between the Government and members of the public whilst being grounded in a firm model of argument that was transparent to the user. Conversely, the model presented in this paper is intended for use solely by autonomous computer agents. However, we believe that there may be a useful link between the two models. As mentioned earlier in this section, all the information entered into the Parmenides system is stored in a back-end database. Therefore, it would be possible to reconstruct new positions on the issue from the users' responses by introducing agents to represent their views. These new positions could then be used as input to generate presumptive arguments to be used by BDI agents, as demonstrated in this paper. This would allow us gather a wide range of differing views on the topic and evaluate the warrant of each view. As part of the practical reasoning process this would ensure that all possible scenarios have been considered and thus aid us in choosing the best action and justification for the issue in question. This is would be an interesting avenue to pursue between our two systems and is something that we hope to explore in future work.

6 Concluding Remarks

In this paper we have used an approach to modelling practical reasoning and persuasive argument to represent the reasoning in a recent and important international political debate. Our account demonstrates how superficial agreement may conceal subtle but important differences in beliefs and aims. We recapitulated a previous theory of persuasion over action and showed how this can be made computational for use in BDI agents enhanced to deal with the notion of values. We applied

this general theory to the political domain to show how real life issues can be debated with computational agents using our model. We went on to show how such argumentation can make effective use of the notion of accrual of arguments. This concept is an important feature of persuasive debate and we believe that our model is able to give some insight into this phenomenon. Finally, we drew some comparisons between this method of argumentation and a previous system built by the authors based on the same theory of persuasion over action. In future work we hope to apply the methods presented here to different domains, such as the medical one, where decisions are often based on the degree of certainty assumed to be attached to expected effects of actions. This will enable us to explore in more detail the notions of accrual and strength of arguments and how these differ across domains.

Acknowledgements

Katie Atkinson is grateful for support from the EPSRC. Trevor Bench-Capon and Peter McBurney acknowledge partial support received from the European Commission, through Project ASPIC (IST-FP6-002307).

References

- [1] K. Atkinson, T. Bench-Capon, and P. McBurney. Attacks on a presumptive argument scheme in multi-agent systems: pre-conditions in terms of beliefs and desires. Technical Report ULCS-04-015, Department of Computer Science, University of Liverpool, UK, 2004.
- [2] K. M. Atkinson, T. J. M. Bench-Capon, and P. McBurney. Justifying practical reasoning. In F. Grasso, C. Reed, and G. Carenini, editors, *Proc. of CMNA workshop 2004*, pages 87–90, Valencia, Spain, 2004.
- [3] K. M. Atkinson, T. J. M. Bench-Capon, and P. McBurney. Parmenides: Facilitating democratic debate. In R. Traunmüller, editor, *Electronic Government 2004*, LNCS 3183, pages 313–316. Springer, Berlin, 2004.
- [4] T. J. M. Bench-Capon. Persuasion in practical argument using value based argumentation frameworks. *Journal of Logic and Computation*, 13 3:429–48, 2003.
- [5] B. G. Buchanan and E. H. Shortliffe. *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Addison-Wesley, Reading, MA, USA, 1984.
- [6] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357, 1995.
- [7] K. Greenwood, T. Bench-Capon, and P. McBurney. Towards a computational account of persuasion in law. In *Proc. of the Ninth Int. Conf. on AI and Law*, pages 22–31, New York, NY, USA, 2003. ACM Press.
- [8] D. N. Walton. *Argument Schemes for Presumptive Reasoning*. Lawrence Erlbaum Associates, Mahwah, NJ, USA, 1996.
- [9] M. J. Wooldridge. *Reasoning about Rational Agents*. MIT Press, Cambridge, MA, USA, 2000.