Computational Support for Public Debate and Policy Justification

Katie Atkinson*

Abstract. This paper investigates how computer support can be enlisted to help the Government collect and analyse public opinions regarding the justification of policies. To do this we show how a link can be established between two existing systems for eDemocracy: the first, an implemented online tool based upon a specific model of argument for gathering the public's views about a particular topic; the second, a method to enable autonomous software agents to reason about justifications for action based upon their subscription to social values. We show how the output from the first of these systems can be used as input to the second and how the information gathered can be reasoned about through computer support.

1 Introduction

Communication between the Government and the people has altered from traditional methods of letter writing and discussions at council meetings, to email contact and web interactions. With the development of safe and efficient web-based services governments have the opportunity to exploit the benefits of new computer technologies to provide accessible, efficient and useful systems through which democracy can be effectively conducted. The work presented in this paper aims to contribute to some of these objectives. We explore how a specific computer system implemented to facilitate eDemocracy can be integrated with autonomous agent systems used to reason about the justification of arguments concerning actions. Section 2 gives a brief overview of a tool, named the PARMENIDES system, developed to foster public debate on a particular political issue. Section 3 describes an approach to argument representation for dealing with reasoning about action, which can be deployed in autonomous software systems. In section 4 we show how a link can be established between the systems described in the previous two sections and illustrate the approach with a short example. Finally, in section 5 we offer some concluding remarks.

2 The PARMENIDES Discussion Forum

Numerous computer systems have been developed which aim to facilitate the online conveyance of democracy, e.g., the DEMOS system [6]. This paper focuses on one particular system – Atkinson *et al's* PARMENIDES (Persuasive ARguMENt In DEmocracieS) system – designed to encourage public participation and debate regarding proposed Government policies. PARMENIDES is described in [1] and it can be used at: `http://www.csc.liv.ac.uk/~katie/Parmenides.html`.

* Department of Computer Science, University of Liverpool, Liverpool, L69 3BX, UK. `katie@csc.liv.ac.uk`

The idea of the system is to enable members of the public to submit their opinions about the Government's justification of a particular action. In the prototypical version the subject dealt with (which is used for illustrative purposes only) is the 2003 war in Iraq, with the particular question under scrutiny being: "Is invasion of Iraq justified?". One of the key features of PARMENIDES is the underlying model upon which it is based, as the tool is intended as an implementation to exploit a specific representation of persuasive argument. In recent years numerous approaches, such as those described in [7] and [8], have advocated the use of web-based discussion boards as a useful way of encouraging and supporting eDemocracy. Although such discussion boards can indeed encourage participation and debate, they generally provide no structure to the information gathered. This issue is addressed in the PARMENIDES system as it is implemented upon a firm model of argument, which is transparent to the user, but provides structure to the responses submitted. The background of this model of argument is as follows.

In [2] Atkinson *et al.* describe an argument scheme that makes explicit the different elements that comprise a justification of an action. Argument schemes hail from the argumentation and informal logic literature with notable advocates of the approach being Katzav and Reed [5], and Walton [9], among others. Such argument schemes are used to classify different types of argument that embody stereotypical patterns of reasoning. In [9] Walton uses such schemes to embody a form of presumptive reasoning. His account views an argument scheme as embodying a presumption in favour of the conclusion. Whether this presumption stands or falls can be tested through posing critical questions associated with the scheme. In order for the presumption to stand, satisfactory answers must be given to any such questions that are posed in the given situation. Atkinson *et al.'s* argument scheme, called AS1, is an extension to one of Walton's schemes – the sufficient condition scheme for practical reasoning. AS1 is stated as follows:

AS1 In the current circumstances R
we should perform action A
to achieve new circumstances S
which will realise some goal G
which will promote some value V.

In this scheme the notion of a goal has been separated into three distinct elements: states of affairs (the effects of actions), goals (the desired features in those states of affairs) and values (the reasons why those features are desirable). The distinction between states and goals is made to represent the difference between effects of actions which the agent wishes to attain, and the effects which follow from an action but are not necessarily desired. Values in turn are different from goals as they provide the actual reasons for the desire to achieve a goal. Values denote some actual descriptive social attitude/interest which the proponent of an action may or may not wish to uphold or subscribe to. Thus, values are not just a qualitative measure of a state, but they provide more subjective reasons as to why states of affairs are desirable or undesirable. Additionally, values relate states of affairs, since a given state of affairs may be desirable through promoting several values, and a given value can be promoted by several states of affairs. Instantiations of argument scheme AS1 provide *prima facie* justifications of proposals for action. Associated with this scheme are sixteen different critical questions that challenge the presumptions in instantiations of AS1. These critical questions are:

CQ1: Are the believed circumstances true?
CQ2: Assuming the circumstances, does the action have the stated consequences?
CQ3: Assuming the circumstances and that the action has the stated consequences, will the action bring about the desired goal?
CQ4: Does the goal realise the value stated?
CQ5: Are there alternative ways of realising the same consequences?
CQ6: Are there alternative ways of realising the same goal?
CQ7: Are there alternative ways of promoting the same value?
CQ8: Does doing the action have a side effect which demotes the value?
CQ9: Does doing the action have a side effect which demotes some other value?
CQ10: Does doing the action promote some other value?
CQ11: Does doing the action preclude some other action which would promote some other value?
CQ12: Are the circumstances as described possible?
CQ13: Is the action possible?
CQ14: Are the consequences as described possible?
CQ15: Can the desired goal be realised?
CQ16: Is the value indeed a legitimate value?

Given this argument scheme and critical questions, debates can then take place between dialogue participants whereby one party attempts to justify a particular action, and another party attempts to present persuasive reasons as to why elements of the justification may not hold or could be improved. This structure forms the underlying model of PARMENIDES. In the prototypical version a justification upholding the action of invading Iraq is presented to users in the form of AS1. Users are then led in a structured fashion through a series of web pages that pose the appropriate critical questions to determine which parts of the justification they agree or disagree with. Once a critique of the initial justification has been given, users can then state their own full justification of any action they believe should be proposed, regarding the topic. Users are not aware (and have no need to be aware) of the underlying structure for argument representation, but it is imposed on the information they submit. The Government is thus able to collect information that has been structured in a clear and unambiguous fashion from a system which does not require users to gain specialist knowledge before being able to use it. All responses submitted are written to a back-end database to gather information as to which points of the argument are more strongly supported than others. The Government can then analyse this information to review public support of its case and perhaps revise or change its justification to make the policy more amenable to public support.

This brief description of the PARMENIDES system is intended as an overview of the tool. We now show how argument scheme AS1 and its critical questions can be employed in an alternative application making use of autonomous software agents.

3 Reasoning About Action Using Autonomous Agents

In [2] Atkinson *et al.* formally describe how their argument scheme and critical questions can be transformed into a computational account for use in software systems consisting of autonomous agents based upon the popular belief-desire-intention (BDI) ar-

chitecture. They provide pre-conditions to describe how an agent can construct a position based upon its beliefs, the actions available for performance, the agent's desires, and its values. As standard BDI architectures do not use the notion of values, the architecture is extended to include values which provide justifications for the agent's choice of intentions, based upon its beliefs and desires. Further pre-conditions then specify when agents can attack a justification for action by posing any of the critical questions against the position. The output of this process is a set of presumptive arguments[1] plus attacks on them. Resolution of a chosen course of action is then done by organising the arguments and attacks into a Value-Based Argumentation Framework (VAF) [3]. Argumentation Frameworks (AFs) were introduced by Dung in [4] as an abstract method for evaluating arguments against how well they defends themselves against attack and defeat from other arguments. VAFs provide an extension to Dung's AFs to accommodate different audiences with different values and interests. The key elements of Dung's AFs are *preferred extensions* (PEs). A PE is a subset of the arguments in the framework which satisfies the following properties: the PE is conflict free, in that no argument in the PE attacks any other argument in the PE; the PE is able to defend every argument in the PE against attacks from outside the PE, in that every argument outside the PE which attacks an argument in the PE is attacked by some argument in the PE; the PE is maximal, in that no other argument can be introduced to the PE without either introducing a conflict or an argument that cannot be defended against outside attacks. Thus, in Dung's AFs an argument is always defeated by an attacker, unless that attacker can itself be defeated. In a VAF, however, attack is distinguished from *defeat for an audience* whereby strengths of arguments for a particular audience are compared with reference to the values to which they relate. This allows a particular audience (characterised by a particular preference ordering on the values) to choose to reject an attack, even if the attacking argument cannot itself be defeated. This is on the proviso that the audience ranks the purpose motivating the attacked argument (the value cited in its justification) as more important than that motivating the attacker. The notion of attack in Dung's PE is then replaced by the notion of defeat for an audience to get the PE for that audience. VAFs are represented as directed graphs, with nodes representing arguments (labelled with an argument identifier and the value promoted by the argument), and edges representing attacks between arguments. Once the appropriate VAFs have been constructed, the dialectical status of the arguments can be calculated and the justified action for each audience determined. Atkinson *et al.* have demonstrated the entire approach in an example from the legal domain [2]. In the next section we show how a link can be made between PARMENIDES and the account described here to show how computer support can aid the democratic debating process, whilst accounting for public opinions.

4 Integrating the Approaches

As PARMENIDES is accessible to the public, it is possible for large amounts of data to be received and stored by the system. Thus, it would be useful to have a mechanism to analyse and reason about the data. The use of software agents can serve this purpose.

[1] It is assumed in [2] that these arguments will be represented in some suitable formal logic, such as propositional logic, amenable to reasoning by a software agent.

The PARMENIDES back-end database stores all critiques and counter proposals supplied by members of the public. Any counter proposal offering a justification for action is decomposed into individual entries that record each of the elements that comprise an instantiation of AS1 (the circumstances, action, consequences including the goal, and, the value promoted). Given this information it is then possible to determine different audiences, based upon their value subscriptions, and ascertain the acceptability and popularity of each action suggested. An example to illustrate this is given below.

On entering the PARMENIDES system the user is presented with the Government's (hypothetical) justification for invading Iraq. This justification is as follows:

In the current situation: Saddam has weapons of mass destruction (WMD), Saddam will not disarm voluntarily, Saddam is running an oppressive regime, Saddam is defying the UN, Saddam is a threat to his neighbours. *The action we should take is*: invade Iraq. *Invading Iraq will*: Remove the WMD, Restore democracy to Iraq, Assert the authority of the UN, Remove the threat Saddam poses to his neighbours, Cause military casualties, Cause civilian casualties. *This will achieve*: Removing WMD will promote world security, Restoring democracy will promote human rights.

As two values are involved in this justification we can split the argument into two separate justifications: one based on the presence of WMD whereby the action of invading will get rid of the WMD, promoting the value 'world security', and, one based on the existence of an oppressive regime whereby invading will dispel the regime, promoting the value 'human rights'. We shall call these two arguments Arg1 and Arg2 respectively. We can then take the individual elements of each justification and instantiate the beliefs, desires, goals and values of a value-enhanced BDI agent (in accordance with the preconditions given in [2]) to represent the views expressed in these justifications. Now, returning to PARMENIDES, suppose that a user disagrees with the justification given in Arg1, as revealed in the summary of responses that is displayed when he has been questioned about his views regarding each element of the justification. As an example, suppose he has registered disagreement with the statement "Saddam has WMD", thus he disagrees with the description of the current situation. In critiquing this element the user (without knowing or needing to know) is posing critical question CQ1. We shall call this attack on the justification 'Attack1' and in order to construct the appropriate VAF to represent this attack, we need to identify the value endorsed by this argument. The attack represents a disagreement as to the facts of the situation, so we associate this argument with the value 'opinion'. Thus, this attack isolates the contentious element of the justification, whilst recognising that this is relative to an *opinion* of a particular opponent. This value would initially be ranked as a weak argument against the justification, but it could gain strength were more users shown to have the same opinion.

We can now instantiate a BDI agent that holds the belief manifest in the attack described above, i.e., that there are no WMD. If we now view the arguments considered so far as a VAF, we have the situation shown in Figure 1:

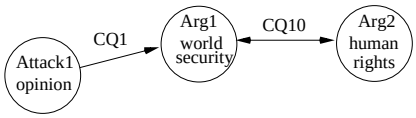

Figure 1. VAF with an attack on a justification.

The VAF in Figure 1 has nodes to represent the two arguments promoting different values and it also shows the attack on Arg1 posed by CQ1. Note also that the VAF shows an attack between Arg1 and Arg2 through the use of CQ10. This shows that although both arguments endorse the same action, they each promote different values. Now, suppose the PARMENIDES user has critiqued the original justification but has also gone on to offer an alternative action plus justification which is stored in the PARMENIDES database. Such an alternative, which we will call Arg3, might be as follows:

In the current situation: we do not know whether Saddam has WMD. *The action we should take is:* give weapons inspectors more time to investigate. *This will:* clarify whether the WMD claim is true or not. *This will achieve:* public trust in the facts.

This justification offers an alternative action, incompatible with the original action, which promotes some other value. Thus, it is posing critical question CQ11 and we can instantiate another BDI agent with the beliefs, desires and value cited above. This argument is added to the VAF, as can be seen in Figure 2. Now, the attack of CQ11 on Arg1 would succeed for any audience that ranks the value 'public trust' higher than the value 'world security'. Attack1 would only succeed in defeating Arg1 if it were shown to be an opinion expressed by a sufficiently large number of users, with a threshold for successful defeat perhaps being set by the Government. As yet, Arg2 has no attackers so the action of invasion, for the reasons specified in Arg2, could still be justifiable. However, submissions to PARMENIDES may reveal critiques and counter proposals to Arg2. Suppose a particular user believes that invading Iraq will not restore democracy to the country, i.e., he disagrees with the consequences of the action. This critique makes use of CQ2 and we shall call this attack on the justification 'Attack2'. As in the case of Attack1, Attack2 will also take the value 'opinion'. So, we can instantiate another BDI agent that holds this belief. After having given his critique, a user with such views may also propose an alternative action plus justification, such as the example one given below:

In the current situation: Saddam is running an oppressive regime, Saddam is violating human rights. *The action we should take is:* wait for a second UN resolution on the matter. *This will mean:* unjustified military intervention is not required. *This will achieve:* respect for international law.

As this justification is offering an alternative action, incompatible with the original action, and it promotes some other value, it is again posing critical question CQ11. We will call this Arg4 and we can use another BDI agent to represent this view. Both Attack2 and Arg4 can now be added to the VAF, as shown in Figure 2.

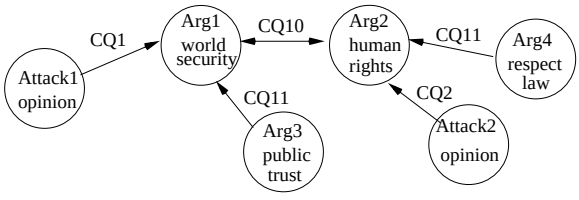

Figure 2. VAF with attacks on Arg1 and Arg2.

In the above scenario the attack of CQ11 on Arg2 would succeed for any audience that ranks the value 'respect for international law' higher than the value 'human rights' (as used in relation to the specific argument). Attack2 would only succeed in defeating Arg2 if it were shown to be an opinion expressed by a sufficiently large number of users.

We have thus shown how both original justifications can be subject to attack through users' critiques (though there are of course further attacks in addition to the examples given here). In order to fully analyse the data submitted to PARMENIDES, all critiques and alternative proposals would need to be represented as VAFs. This would enable the Government to uncover any patterns in the data showing which parts of the justification are most frequently disagreed with, and segment the population according to their values. For example, multiple critiques may reveal that CQ1 is consistently being posed to disagree with the statement 'Saddam has WMD', so the Government may then try to clarify their reasons for endorsing this point. Further critiques may reveal, for example, that CQ9 is consistently used to introduce arguments stating that other values important to members of the public have not been considered by the Government. In this case, the Government may wish to provide justification as to why the values it endorses are the most important ones concerned in the debate. However, to ensure that all opinions have been assessed in relation to each other, the reasoning process should not end here and user-supplied arguments should also be subject to critical questioning. PARMENIDES currently does not provide a facility by which users can critically assess each other's views, though it would be possible to implement this to evaluate all views supplied and identify any conflicts between them. We now provide a short example of this.

If we examine Figure 2 we can see that it contains Arg4. This argument was constructed from an alternative position to the original justification, as supplied by a user. There are numerous ways in which this could be attacked. For example, the original proponent could counter that the alternative action proposed has side effects which actually demote the value concerned. This could instantiate AS1 with Arg5 as follows:

In the current situation: Saddam is running an oppressive regime, Saddam is violating human rights. *The action we should take is:* wait for a second UN resolution on the matter. *This will mean:* Saddam is allowed to continue his activities. *This will:* demote respect for international law.

This argument is stating that the action proposed has unconsidered consequences which actually demote the value in question ('respect for international law'), and thus it makes use of CQ8. This argument can then be added to the VAF as follows:

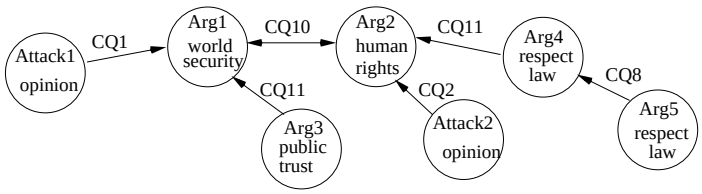

Figure 3. Final VAF.

If the reasoning were to stop here we can see that Arg5 defeats Arg4 for any audience, as the two are motivated by the same value. However, the loss of Arg4 does not

reinstate Arg2 as it could still be defeated by Attack2. Of course, any new arguments introduced into the VAF would be subject to critique and the status of the VAF would need to be updated and re-evaluated accordingly. The example arguments used here are intended to serve as an illustration of the approach, but in practice we envisage the debate encompassing a much larger range of arguments. Once a sufficiently representative number of views had been submitted to PARMENIDES, the Government would then be able to assess the opinions supplied and their relative importance. If the opinions revealed that particular parts of the original justification of the policy in question were viewed as being contentious, then the Government could take measures it deems appropriate to respond to public criticism. This may involve clarification of the facts, release of supporting information, or adjustment to the policy, amongst other things.

5 Concluding Remarks

In this paper we have described how electronic support can aid democratic debate about political issues, by showing how a link can be established between two existing systems that make use of a theory of persuasive argument in practical reasoning. We believe that both the systems described are of value in themselves, as they are based upon a defined method of argument representation. Moreover, once integrated we believe that they have the potential to add further value to domains, such as the political one, where reasoning about and justifying actions is crucial. The approach encourages public participation and it would allow for the gathering of a wide range of differing views on the topic and the evaluation of the warrant of each view. Future work will be to implement and evaluate a large scale version of the approach.

References

1. K. Atkinson, T. Bench-Capon, and P. McBurney. PARMENIDES: Facilitating democratic debate. In R. Traunmüller, editor, *EGOV*, LNCS 3183, pages 313–316. Springer, 2004.
2. K. Atkinson, T. Bench-Capon, and P. McBurney. Arguing about cases as practical reasoning. In *Proc. of 10th Intl. Conf. on AI and Law*, pages 35–44, New York, USA, 2005. ACM Press.
3. T. Bench-Capon. Persuasion in practical argument using value based argumentation frameworks. *Journal of Logic and Computation*, 13 3:429–48, 2003.
4. 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.
5. J. Katzav and C. Reed. On argumentation schemes and the natural classification of arguments. *Argumentation*, 18(2):239–259, 2004.
6. R. Lührs, S. Albrecht, M. Lübcke, and B. Hohberg. How to grow? online consultation about growth in the city of Hamburg: methods, techniques, success factors. In R. Traunmüller, editor, *EGOV*, LNCS 2739, pages 79–84. Springer, 2003.
7. Z. Masters, A. Macintosh, and E. Smith. Young people and e-democracy: Creating a culture of participation. In R. Traunmüller, editor, *EGOV*, LNCS 3183, pages 15–22. Springer, 2004.
8. Ø. Sæbø and H. Nilsen. The support for different democracy models by the use of a web-based discussion board. In R. Traunmüller, editor, *EGOV*, LNCS 3183, pages 23–26. Springer, 2004.
9. D. N. Walton. *Argument Schemes for Presumptive Reasoning*. Lawrence Erlbaum Associates, Mahwah, NJ, USA, 1996.