



**HAL**  
open science

## Model Based Critique of Policy Proposals

Adam Z. Wyner, Katie Atkinson, Trevor Bench-Capon

► **To cite this version:**

Adam Z. Wyner, Katie Atkinson, Trevor Bench-Capon. Model Based Critique of Policy Proposals. 4th International Conference on Electronic Participation (ePart), Sep 2012, Kristiansand, Norway. pp.120-131, 10.1007/978-3-642-33250-0\_11 . hal-01541479

**HAL Id: hal-01541479**

**<https://inria.hal.science/hal-01541479v1>**

Submitted on 19 Jun 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Model Based Critique of Policy Proposals

Adam Z. Wyner, Katie Atkinson, and Trevor Bench-Capon

Department of Computer Science, Ashton Building, University of Liverpool,  
Liverpool L69 3BX, United Kingdom

**Abstract.** Citizens may engage with policy issues both to critique official justifications, and to make their own proposals and receive reasons why these are not favoured. Either direction of use can be supported by argumentation schemes based on formal models, which can be used to verify and generate arguments, assimilate objections etc. Previously we have explored the citizen critiquing a justification using an argumentation scheme based on Action-based Alternating Transition Systems. We now present a system that uses the same model to critique proposals from citizens. A prototype has been implemented in Prolog and we illustrate the ideas with code fragments and a running example.

## 1 Introduction

Citizens contact their representatives about policies with a number of objectives. They may be seeking a justification of a current or proposed policy; they may wish to object to a current or proposed policy; or they may wish to make policy proposals of their own. Argumentation can support e-participation systems designed to meet any of these [8]. In the first case a simple statement is all the response that is needed. For the second case it is necessary to identify, make precise, and attempt to meet, these objections. These situations are addressed in work such as [7]. The third case is, however, different. Here what is needed is to obtain a precise, well formulated proposal from the citizen and then to identify its flaws and explain why the current policy is preferred. It is this situation that we will address in this paper. Our approach makes use of argumentation schemes, which hail from the informal logic literature and capture stereotypical patterns of reasoning that can be expressed as arguments. Every argumentation scheme is associated with a set of characteristic critical questions, that are used to identify possible challenges to instantiations of the argumentation scheme. Such challenges can be made against the different elements of the argumentation schemes and the relations between them. The particular argumentation scheme [1] that is of relevance to our prototype is one that enables reasoning about, and justification of, actions. Some examples of the critical questions related to the argumentation scheme for practical reasoning are: *Are the current circumstances as stated? Does the action have the consequences stated? Does the goal promote the value?*

We will describe how the argumentation scheme and critical questions of [1] can be used as the basis of a system to help a citizen to form a justification of

a policy proposal and automatically provide a systematic and thorough critique of that position. The proposed system will be complementary to previous work using this argumentation scheme, notably [7] and [11], in which the system proposed a justification and requested a critique from the user. In this paper the roles are reversed, so that we solicit a justification from the user and the system will provide the critique. We will illustrate our approach with a running example, relating to the use of speed cameras originally developed in [2], and some fragments of a Prolog program that we have produced to make our definitions operational.

The argumentation scheme in [1] proposes an action based on an understanding of the current situation, the consequences of actions and the desire to promote particular social values. These social values exist in a preference order that may vary from person to person according to their tastes and aspirations. The scheme is stated in [1] as the practical reasoning argument scheme labelled *AS1*:

*Premise 1:* In the current circumstances S (source)

*Conclusion:* We should perform action A

*Premise 2:* Which will result in new circumstances R (result)

*Premise 3:* Which will realise goal G

*Premise 4:* Which will promote some value V.

In this scheme, S is what is true before the action A, R is what is true after the action, G is the desirable consequence and V the reason why G is desirable.

In [1] seventeen critical questions, which can be used as a basis to attack justifications using *AS1*, are given. A dialogue protocol based on this scheme will be very complicated, as demonstrated by the protocol for a similar scheme using sixteen of the critical questions given in [3]. As is noted in that paper, the protocol offers too much choice to be effective in practice. We must therefore look for a simpler way of proposing, defending and critiquing such justifications. Moreover, in addition to deploying the variety of available moves, difficulties can also arise when interpreting the critical questions. Although they can be understood informally, with the reader supplying an appropriate context, they are not always entirely precise. For example, one critical question against *AS1* is whether the current circumstances hold. As we will see in section 3, this may be asked either because the questioner believes the action not to be currently possible, or because the questioner believes that the action will have different consequences given the circumstances that are actually currently true. To resolve these problems of vagueness and ambiguity, it is necessary to anchor the questions in a well defined semantical structure. This was done in [1], which used Action-based Alternating Transition Systems (AATS) [10] for this purpose.

Presenting justifications based on this scheme and receiving critiques was the subject of [11]. Here we will describe how an AATS and the argumentation scheme and critical questions of [1] can be used to *accept* a justification and then automatically *provide* a critique. We believe allowing citizens to express their own ideas, and to receive feedback on possible problems with them, is

an important part of engaging citizens in policy debates, and an important supplement to justifications of actual policy. Interacting with policy proposals in this way will help citizens to better understand why certain proposals were adopted and others rejected.

Section 2 will define an AATS, give a concrete instantiation of the AATS for our example, and give the Prolog predicates required to realise it. Section 3 will explain how the justification is elicited from the user, and the various steps of the critique. Section 4 will give an illustration using our example. Section 5 will give a brief discussion and some concluding remarks.

## 2 Model

An AATS is a formal structure that enables representation of states in an agent system and actions that cause transitions between states. Due to their precise formalisation, AATSs can be used in automated systems where agents need to have an explicit representation of states, actions and transitions. AATSs are useful for our tool because they enable us to map out in precise detail the space of a policy debate. As defined in [10], the AATS made no reference to values. In order to adapt it for use with AS1, [1] extended the structure to include labels on the transitions indicating which values are promoted and which demoted by following a transition. The extension of the original specification of an AATS to accommodate the notion of values is an *Action-based Alternating Transition System with Values* (AATS+V), defined as a  $(n + 9)$  tuple  $S = \langle Q, q_0, Ag, Ac_1, \dots, Ac_n, \rho, \tau, \Phi, \pi, V, \delta \rangle$ , where:

- $Q$  is a finite, non-empty set of *states*;
- $q_0 \in Q$  is the *initial state*;
- $Ag = \{1, \dots, n\}$  is a finite, non-empty set of *agents*;
- $Ac_i$  is a finite, non-empty set of actions, for each  $i \in Ag$  where  $Ac_i \cap Ac_j = \emptyset$  for all  $i \neq j \in Ag$ ;
- $\rho : Ac_{Ag} \rightarrow 2^Q$  is an *action pre-condition function*, which for each action  $\alpha \in Ac_{Ag}$  defines the set of states  $\rho(\alpha)$  from which  $\alpha$  may be executed;
- $\tau : Q \times J_{Ag} \rightarrow Q$  is a partial *system transition function*, which defines the state  $\tau(q, j)$  that would result by the performance of  $j$  from state  $q$  – note that, as this function is partial, not all joint actions are possible in all states (cf. the pre-condition function above);
- $\Phi$  is a finite, non-empty set of *atomic propositions*; and
- $\pi : Q \rightarrow 2^\Phi$  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  $q$ .
- $V$  is a finite, non-empty set of values.
- $\delta : Q \times Q \times V \rightarrow \{+, -, =\}$  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  $v_u \in V$ .

AATSs are particularly concerned with the joint actions of the set of agents.  $j_{Ag}$  is the joint action of the set of  $k$  agents  $Ag$ , and can be represented as a tuple  $\langle \alpha_1, \dots, \alpha_k \rangle$ , where for each  $\alpha_j$  (where  $j \leq k$ ) there is some  $i \in Ag$  such that  $\alpha_j \in Ac_i$ . Moreover, there are no two different actions  $\alpha_j$  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  $i \in Ag$ ,  $i$ 's action in  $j$  is denoted by  $j_i$ .

## 2.1 Instantiating the AATS

In this paper we will use the same example as [2], and use the AATS+V developed there. To fully describe a model using the AATS+V we need to specify the various components of the structure. We need the set of propositions  $\Phi$ , combinations of which make up the possible member states of  $Q$ . Given  $\Phi$ , we can constrain the size of  $Q$  by identifying logical relationships between members of  $\Phi$ , such that for  $p_1, p_2 \in \Phi$ ,  $\neg(p_1 \wedge p_2)$ : states containing such combinations will then be impossible and so not be in  $Q$ . We need to give the set of agents,  $Ag$  and the actions they can perform. We need the set of values that may be promoted and demoted by the movement from one state to another. Finally, we need a transition matrix expressing  $\rho$ ,  $\tau$ , and  $\delta$ . This matrix comprises a row for each state in  $Q$  and a column for each joint action in  $J$ . An entry in a cell indicates that the preconditions for the joint action are satisfied. Such entries comprise a triple consisting of the state reached if that joint action is executed, the set of values promoted, and the set of values demoted. These transitions are a reflection of a causal theory, which explains the effects of various actions, and an evaluative theory, which tells us when values are promoted and demoted.

Our example is an issue in UK Road Traffic policy, modelled in [2] and used before that as an e-participation example in [7] and [4]. The number of fatal road accidents is an obvious cause for concern, and in the UK there are speed restrictions on various types of road, in the belief that excessive speed causes accidents. The policy issue that we will consider is how to reduce road deaths. One suggestion would be to deter motorists from speeding by introducing speed cameras, which would greatly increase detection and punishment of speeding offences. Points that might be contested are whether fines are sufficient to deter, and whether speeding is an important factor in road accidents. Additionally there are civil liberties issues associated with the loss of privacy resulting from the increased surveillance. One, more expensive, alternative to speed cameras would be to have a programme of education for motorists that could make them more aware of the dangers of speeding, better able to control their vehicles at speed, or both. This gives the set of propositions as:

$\Phi = \{R, S, P\}$ , where  $R$  is that there are excessive road accidents,  $S$  is that there is excessive speeding and  $P$  that the intrusions on privacy are excessive.

These three propositions give rise to, potentially, eight states. We may, if we wish, exclude one or more of these as impossible. For example if we believe that

it is impossible that there should be a reduction in road deaths without a reduction in speeding, no state with  $\neg R$  and  $S$  would be possible and so such states would not appear in  $Q$ . We also need to identify the current state,  $q_0$ , which we take to be  $\{R \wedge S \wedge \neg P\}$ . The main agents involved are the *Government*, and *Motorists*, considered as a body. In some cases the consequences of an action are indeterminate (or at least cannot be determined using the elements we are modelling). To account for this we introduce a third agent, termed *Nature*. The action ascribed to Nature determines the outcomes of the actions of the other agents, where these outcomes are uncertain or probabilistic. The Government has three actions: introducing speed cameras, educating motorists, or doing nothing. Motorists may reduce their speed or do nothing. Nature has two actions according to which fatal accidents are or are not reduced as a result of the Government and motorist actions. For values we consider, the cost in terms of human life ( $L$ ), compliance with the law ( $C$ ), the financial cost to the Government ( $B$  for budget) and the impact of civil liberties ( $F$  for freedom). Figure 1 shows the transitions from the current state for the six possible joint actions:

- $j_0$  Government does nothing, motorists do nothing and nature does nothing.
- $j_1$  Government introduces cameras, motorists do nothing and nature does nothing.
- $j_2$  Government introduces cameras, motorists reduce speed and nature reduces accidents.
- $j_3$  Government introduces cameras, motorists reduce speed and nature does nothing.
- $j_4$  Government educates motorists, motorists reduce speed and nature reduces accidents.
- $j_5$  Government educates motorists, motorists do nothing and nature reduces accidents.

Accidents are always reduced when motorists are educated since either they do not speed, or can control their vehicles better.

## 2.2 Prolog Representation

To represent the AATS+V of Figure 1 in a Prolog program we state facts representing the propositions in  $\Phi$ , the states in  $Q$ , the joint actions  $J$  and the transitions defined by  $\tau$ . Propositions are in fact represented as pairs of literals to allow separate natural language forms for positive and negative occurrences.

```
%literal(id,positive or negative, english text, other text)
literal(1,1,[there,is,excessive,speeding],[ ]).
literal(2,0,[speed,limits,are,generally,obeyed],[ ]).
literal(3,1,[there,are,too,many,road,deaths],[ ]).
literal(4,0,[road,deaths,are,acceptable],[ ]).
literal(5,1,[there,are,unacceptable,intrusions,on,privacy],[ ]).
literal(6,0,[privacy,is,respected],[ ]).
```

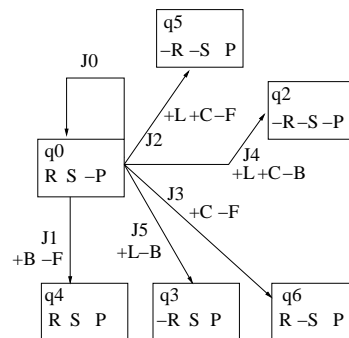
```

%state(id, R, S, P). R S and P where the positive or negative
% literals hold in the state. 0 is the current state
state(0,1,3,6).    state(2,2,4,6).    state(3,1,4,6).
state(4,1,3,5).    state(5,2,4,5).    state(6,2,3,5).

%jointAction(id,government,motorist,nature).
jointAction(j0,[do,nothing],[do,nothing],[there,is,no,effect]).
jointAction(j1,[introduce,speed,cameras],[do,nothing],
             [there,is,no,effect]).
jointAction(j2,[introduce,speed,cameras],[reduce,speed],
             [there,are,fewer,accidents]).
jointAction(j3,[introduce,speed,cameras],[reduce,speed],
             [there,is,no,effect]).
jointAction(j4,[educate,motorists],[reduce,speed],
             [there,are,fewer,accidents]).
jointAction(j5,[educate,motorists],[do,nothing],
             [there,are,fewer,accidents]).

%transition(id, source state, target state,
%joint action, values promoted, values demoted).
transition(1,0,5,j2,[l,c],[f]). transition(2,0,0,j0,[],[]).
transition(3,0,4,j1,[b],[f]). transition(4,0,2,j4,[l,c],[b]).
transition(5,0,6,j3,[c],[f]). transition(6,0,3,j5,[l],[b]).

```



**Fig. 1.** Transitions from  $q_0$

### 3 Critique

The process now is to solicit from the citizen a proposed action and a justification in the form of AS1, and to critique that proposal and its justification. The

critique will proceed through a number of steps, each related to critical questions associated with AS1. In our e-participation setting, citizens will advocate a policy, either as part of a debate as to what should be done, or as an alternative to what the Government is currently doing. In response they will receive a reply suggesting why their proposal is not acceptable to the Government. Some of these responses will indicate errors: the citizen may simply be wrong about a current fact. Others will indicate possible problems, perhaps that a particular value is demoted: the citizen should be alerted to this problem, but may choose to accept it as a price worth paying. Thus some responses will indicate errors, whereas others will provide warnings. The subsections below describe the elements of a critique that can be automatically constructed to provide an instant response to such proposals. Then in section 4 we will step through a complete example to show how these elements fit together to build into a comprehensive whole.

**Getting the Justification.** To provide a target for the system's critique, the citizen must first instantiate AS1, by giving an action, the current state, the target state to be reached by the action (which will typically assume particular behaviours from the other agents involved), and the values this will promote. To get this information the citizen is first presented with a list of the propositions in  $\Phi$  and indicates which are true and which false using a menu. The user is then prompted for the action the Government should perform. The choice is entered through the use of a menu containing the options in  $Ac_{Gov}$ . The assumed consequences are obtained by presenting again the propositions in  $\Phi$  with a menu. Finally the user is asked to give the most important value promoted by the action, using a menu with the elements of  $V$ .

Note that this form of justification can itself be generated from the Prolog program using (with appropriate instantiations):

```
argumentPro(A,S,R,V):-
  transition(ID,S,R,J,X,_),member(V,X),jointAction(J,A,_,_),
  pp([government,should,A,in,S,to,reach,R,and,promote,V]).
```

where `pp` is simply a pretty print function. Note that if only  $S$  is instantiated in the query, `argumentPro` will generate every possible argument for actions from state  $S$ .

Thus using `transition(3,0,4,j1,[b],[f])` we can instantiate  $A$  to `[introduce, speed, cameras]`,  $S$  to `0`,  $R$  to `4` and  $V$  to `b` and so produce the argument *government should introduce speed cameras in 0 to reach 4 and promote b*. The presentation could further be improved by replacing the state numbers with the text versions of the literals they contain.

We now move through the particular points that make up a comprehensive critique of the justification. The critical question numbers refer to the critical questions presented in [1]: here we use appropriate natural language forms.

**Is the Action Possible? (CQ1, CQ13).** The first thing to check is whether the believed facts are agreed. Thus if the user's answer suggests that literals  $X$ ,  $Y$ , and  $Z$  are believed in the current situation, they should make up the



initial state,  $q_0$  and so  $\text{state}(0, X, Y, Z)$  should be in the program. If this is not so, we have  $\text{state}(0, X1, Y1, Z1)$ , where one or more of  $X1 \neq X$ ,  $Y1 \neq Y$  and  $Z1 \neq Z$ . Thus the user believes we are in state  $q_U$  where  $\text{state}(U, X, Y, Z)$  is in the program. Any disagreement as to the current state should be reported to the citizen. However, disagreement may or may not affect the justification. If  $\text{argumentPro}(A, 0, R, V)$  succeeds, (where the user proposed to do  $A$  to reach  $q_R$  and promote  $V$ ) the advocated action can still be performed in  $q_0$ , and so the disagreement is not material, since the preconditions of the proposed action remain satisfied. The user can therefore be informed that either *I believe the current facts to be  $X1, Y1$  and  $Z1$ , but the action you propose is still possible* or *I believe the current facts to be  $X1, Y1$  and  $Z1$ , and so the action you propose is not possible* depending on whether or not  $\text{argumentPro}(A, 0, R, V)$  succeeds. For example the action *Do nothing* is possible in any state whatsoever.

At this point, if the action is not possible, it is necessary either to terminate the discussion, or to accept what the user says and change which state is taken to be  $q_0$  (and so considered currently true) in the program accordingly. Possibilities also exist for presenting the user with a justification of our beliefs about the current state, perhaps even initiating discussion using a persuasion dialogue in the style of [9] to resolve the differences. We will not go further into these possibilities here, however.

**Can the Action have the Stated Effects? (CQ1, CQ2).** Let the consequences claimed by the user be  $X2, Y2$ , and  $Z2$ , and  $U2$  be the state where they hold. Now the actual consequences of action  $A$  will be given by a transition from state 0, so  $\text{argumentPro}(A, 0, U2, \text{Value})$  would need to succeed if the user's claimed consequences are correct. If this is so the disagreement on the current state does not affect the consequences of the action and so can be disregarded. For example *Introduce Cameras* will reach the same state whether or not privacy is currently respected. Otherwise the user must be informed that *Performing A will not result in  $X2, Y2, Z2$* . Again we could stop here, offer reasons for our position, or engage in a dialogue in an attempt to convince the user of the true consequences of the action proposed.

**Does the Action Promote the Value? (CQ4).** Even if the action does have the consequences claimed, it may still not promote the desired value. For the citizen's justification to be acceptable, where the citizen claims  $V$  will be promoted,  $\text{argumentPro}(A, 0, U2, V)$  must succeed. If this is not so, then the citizen must be informed that *Although performing A will bring about  $X2, Y2$  and  $Z2$ , the value  $V$  will not be promoted*. For example, the citizen might claim that introducing cameras would reach  $q_5$  and promote budget. Moving to  $q_5$  does not, however, promote budget: that would require that  $S$  remained true.

Assuming, however, that  $\text{argumentPro}(A, 0, U2, V)$  does succeed, the justification will have crossed the first hurdle, and represent an acceptable instantiation of AS1. The critique must now turn to whether the argument is acceptable when set against other objections.

**Are There Negative Side Affects? (CQ8, CQ9).** So far we have established that  $\text{argumentPro}(A, 0, U2, V)$  succeeds. We need, however, to consider

the other consequences of performing the action. Firstly it is possible that the action will demote some values as well as promoting  $V$ . Any such demotion would give us an argument against performing  $A$ . Such arguments can be discovered from the program using

```
argumentCon(A,0,U2,W):-
  transition(ID,S,R,J,_,X), member(W,X), jointAction(J,A,_,_),
  pp([government,should,not,A,in,S,
     to,avoid,R,which,would,demote,W]).
```

If this succeeds with one or more values binding  $W$  we will have one or more grounds for objecting to the original justification. For example the demotion of  $f$  is an argument against reaching  $q_5$  by introducing cameras. Whether the argument is decisive will depend on the user's preferences. Most serious is if  $W = V$ , since the action represents a self-defeating way of advancing this value. So the user would be informed that *You should not do A since this will demote the value you are trying to promote*. If, on the other hand,  $W \neq V$ , the user only needs to be warned about the side effects to consider whether  $V$  is still worth promoting, so the message would be *Performing A will demote values W1 ... Wn. Are you sure it is still worth promoting V?* In this way, an advocate of speed cameras can continue to hold that view, but now does so with the awareness that freedom will be reduced.

**Are there Other Ways to promote the Value? (CQ7).** If there is an action  $B \neq A$  for which  $argumentPro(B,0,Any,V)$  succeeds, then it is possible to promote the desired value with a different action. This different action may be preferable to the chosen action, and so should be drawn to the citizen's attention. For example, if the citizen had suggested education as a way of promoting  $L$ , the critique could draw attention to the possibility of promoting the value by introducing cameras. Thus the system responds *Performing B may also promote V. Are you sure you still wish to perform A?* This might cause the advocates of education to change their mind, particularly since they will have already been alerted to the fact that education will take them over budget.

**Could Other Values be Promoted? (CQ11).** If there is an action  $B \neq A$  for which  $argumentPro(B,0,Any,W)$  succeeds with  $W \neq V$ , then performing  $A$  precludes promoting this other value. Moreover there may be several such values. For our current purposes, we will not object if performing  $A$  promotes  $V$  and other values as well, but we should object if there are values which can be promoted from state 0, but are not promoted by  $A$  taking us to the state proposed by the user. In this case the user should be warned: *If you perform A you will miss an opportunity to promote W1 ... Wn. Do you still want to perform A?*, so that the user is aware of this, possibly preferable, opportunity. Of course, the user may prefer  $V$  and be content with the original choice. For example, an advocate of education to produce better drivers and reach  $q_3$  will be told that  $C$  can be promoted by introducing cameras and reaching  $q_5$  instead, but they may choose to disregard this, if they do not regard  $C$  to be of sufficient importance.

**Will the Other Agents Do What they are Supposed To Do? (CQ17).** Finally, suppose that  $argumentPro(A,0,U2,V)$  succeeds with  $jointAction(J,$

A,  $O1$ ,  $O2$ ) but there is also a transition such that `jointAction(J, A,  $O3$ ,  $O4$ )` (where  $O1 \neq O3$  or  $O2 \neq O4$  or both) would move us from state 0 to some new state without promoting V. That is, our chosen action may fail to promote the desired value because one or more other agents do something other than what we had anticipated or hoped for. For example, if motorists do not modify their behaviour when cameras are introduced we will reach  $q_4$  instead of the desired  $q_5$ . Thus in such cases the system should pose the final objection *Are you confident that  $O1$  and  $O2$  will be performed by the other agents?*

## 4 Example

We now put this all together in a full example. Suppose our citizen wishes to propose education as the solution. First the citizen is asked for an opinion on the truth of  $P$ ,  $R$  and  $S$ . Suppose that the user agrees that there is excessive speeding and too many deaths and intrusions on privacy are not at an acceptable level. The citizen is then offered a choice of the three Government actions: suppose `educate motorists` is chosen. When asked for consequences, the user replies  $\{\neg R, S, P\}$ . The user does not believe that education will reduce speeding (which is seen as human nature), but does think that it will reduce fatalities. Finally the user is asked for the value promoted, and chooses  $b$  (budget). The system can now begin the critique.

The user believes the current state to be  $\{R, S, P\}$ , that is  $q_4$ . The proposed action is also possible in  $q_4$ , so the system reports:

*I do not believe that there are unacceptable intrusions on privacy, but the action you propose is still possible*

Suppose that the citizen responds by modifying his original proposal, and concedes that the intrusions of privacy are currently not excessive. The effect of the action is also agreed to take us from  $q_0$  to  $q_3$ . The value, however, produces a disagreement: the user believes that  $b$  will be promoted, whereas the program computes that  $b$  will be demoted. It therefore reports:

*Although educate motorists will bring about there is excessive speeding, road deaths are at an acceptable level and privacy is respected, the value budget will not be promoted*

Moreover since  $b$  is actually demoted the system will add

*You should not educate motorists since this will demote the value you are trying to promote*

There are, however, ways in which  $b$  can be promoted: we can try to reach  $q_4$  by introducing speed cameras. Thus we can say:

*Introducing speed cameras may promote budget. Are you sure you still wish to educate motorists?*

Although the proposed action will promote  $l$  it fails to promote  $c$ . This value can be promoted by introducing speed cameras. The critique continues:

*If you educate motorists you will miss an opportunity to promote compliance with the law by introducing speed cameras. Do you still want to educate motorists?*

The final perspective, reliance on other agents, is not a problem in this instance since the proposed action fails to promote the value anyway.

## 5 Discussion

The above critique covers nine of the seventeen critical questions of [1]. Some relate to the elements of the AATS+V: the propositions and actions included in the model cannot be questioned by the system. Rather we should expect the citizen to object that his position cannot be expressed. In such cases the system is not adequate for the citizen to provide his desired justification. But once the justification has been provided, consensus as to the propositions and actions can be presumed and so related critical questions will not arise. The other missing critical questions all concern goals, which are present in AS1, but have no obvious correspondence to the AATS+V. Such goals are probably best thought of as predicates defined in terms of the propositions in  $\Phi$ . For example one might consider a society to be civilised if road deaths were not excessive and privacy respected, i.e  $\neg R \wedge \neg P$ . Thus the goal of a civilised society could be realised in state 0 by moving to state 2 or state 3. Thinking in terms of a set of basic propositions in  $\Phi$  and a set of intensional definitions allowing for more complex aspirations to be expressed might be useful in a more complex domain. Here, however, we will, like [2], not include intensional definitions and goals in our current critique.

Policy proposals are based on a wide variety of different kinds of knowledge including: knowledge as to what is currently the case, knowledge of actions and their effects, awareness of the effect on values, knowledge of what other agents are likely to do and knowledge of preferences between competing values. This diversity is reflected in the range of perspectives from which an action justification can be critiqued. The use of an argumentation scheme such as AS1 enables the critique to systematically explore possible weaknesses with respect to all these aspects. What we have described above enables the systematic critique to be delivered, but while it raises questions, these are not argued for or resolved. We could easily provide simple textual justifications for the position we hold, but deeper exploration would require further argument and dialogue. This, however, would not be a simple extension. All the different kinds of knowledge that underlie the differences raised by the critique will require specific kinds of dialogues if these differences are to be resolved. Some of these dialogue types have been investigated: for example disagreements as to the current facts could be resolved using dialogues such as are presented in [9] and [6], but others have not. While there has been some very preliminary investigation of e.g. arguments between

agents with different preferences [5], much remains to be done for this topic and for dialogues disputing causal theories, agent behaviour and evaluative assessments. Techniques designed for straightforward arguments about the truth of a proposition are unlikely to be entirely suitable for these specialised topics.

The tool we have presented is based upon a precise formal model, but the front-end that the citizen interacts with uses only natural language. Many current tools for e-participation lack a formal and precise underpinning and as such offer little support for automated processes. Our contribution is that we have developed a program through which we can automatically elicit policy proposals and systematically generate a critique from our model. The formal model provides a level of support and consistency that takes our tool beyond the current state of the art. Nevertheless we see the system presented here as providing only the first step in a fully fledged critique system. What we have implemented so far provides organisation and breadth but lacks depth, since it uses, but does not justify, the model. In future work we shall explore support for the dialogue types required to open up these stubs.

## Acknowledgments

This work was partially supported by the FP7-ICT-2009-4 Programme, IMPACT Project, Grant Agreement Number 247228. The views expressed are those of the authors and are not necessarily representative of the project.

## References

1. K. Atkinson and T. Bench-Capon. Practical reasoning as presumptive argumentation using action based alternating transition systems. *Artif. Intell.*, 171(10-15):855–874, 2007.
2. K. Atkinson, T. Bench-Capon, D. Cartwright, and A. Wyner. Semantic models for policy deliberation. In *Proceedings of the 13th ICAIL*, pages 81–90, 2011.
3. K. Atkinson, T. Bench-Capon, and P. McBurney. A dialogue game protocol for multi-agent argument over proposals for action. *Autonomous Agents and Multi-Agent Systems*, 11(2):153–171, 2005.
4. T. Bench-Capon and H. Prakken. A lightweight formal model of two-phase democratic deliberation. In *Proceedings of JURIX 2010*, pages 27–36. IOS Press, 2010.
5. E. Black and K. Atkinson. Choosing persuasive arguments for action. In *Proceedings of AAMAS 2011*, pages 905–912, 2011.
6. E. Black and A. Hunter. An inquiry dialogue system. *Autonomous Agents and Multi-Agent Systems*, 19(2):173–209, 2009.
7. D. Cartwright and K. Atkinson. Using computational argumentation to support e-participation. *IEEE Intelligent Systems*, 24(5):42–52, 2009.
8. A. Macintosh, T. Gordon, and A. Renton. Providing argument support for e-participation. *Journal of Information, Technology and Politics*, 6(1):43–59, 2009.
9. H. Prakken. Formal systems for persuasion dialogue. *KER*, 21, 2006.
10. M. Wooldridge and W. van der Hoek. On obligations and normative ability. *J. Applied Logic*, 3(3-4):396–420, 2005.
11. A. Wyner, K. Atkinson, and T. Bench-Capon. Towards a structured online consultation tool. In *Proceedings of ePart 2011*, pages 286–297, 2011.