Semantic Models for Policy Deliberation

Katie M. Atkinson, Trevor J.M. Bench-Capon, Dan Cartwright and Adam Z. Wyner
Department of Computer Science
University of Liverpool
Liverpool, UK
{k.m.atkinson,tbc,dan,azwyner}@liverpool.ac.uk

ABSTRACT

Semantic models have received little attention in recent years, much of their role having been taken over by developments in ontologies. Ontologies, however, are static, and so have only a limited role in reasoning about domains in which change matters. In this paper, we focus on the domain of policy deliberation, where policy decisions are designed to change things to realise particular social values. We explore how a particular kind of state transition system can be constructed to serve as a semantic model to support reasoning about alternative policy decisions. The policy making process includes stages that support the construction of a model, which can then be exploited in reasoning. The reasoning itself will be driven by a particular argumentation scheme for practical reasoning, and the ways in which arguments based on this scheme can be attacked and evaluated. The evaluation provides alternative policy positions. The semantics underpin a current web-based implementation, designed to solicit structured feedback on policy proposals.

1. INTRODUCTION

At one time, in the 1980s, so-called deep conceptual models, which we refer to here as semantic models, were seen as key to the future of AI and Law. The most prominent advocate was Thorne McCarty who as early as 1983, wrote **?**:

> the most critical task in the development of an intelligent legal information system, either for document retrieval or for expert advice, is the construction of a *conceptual* model of the legal domain.

The felt need for semantic models was not confined to legal applications, but was a widespread reaction to certain difficulties encountered with rule based expert systems, such as brittleness, difficulties of knowledge acquisition, unsatisfactory explanation facilities, lack of verifiability and difficulties of maintenance and reuse. This led to work on model based representations of the domain (*model* in the sense of representing components, their relations, and constraints rather than in the sense of an assignment of truth values to a set of

variables) coupled with reasoning from first principles, in particular qualitative reasoning. Good examples of this style of work are **?** and **?**. McCarthy developed a number of tools to support legal conceptual modelling including his treatment of deontic notions **?**, culminating in his Language for Legal Discourse **?**. Similar considerations motivated the isomorphic modelling of legal sources **?**. By the early nineties, however, Gruber had introduced the notion of ontologies **?**, and this idea rapidly took root as a software engineering solution to the problems of rule based knowledge based systems, which retained their advantages and did not require the computationally intensive solutions associated with model based expert systems. Ontologies also became established as the answer to these problems for legal information systems, **?** and **?**, and deep conceptual models faded out of the mainstream. This was also true in other areas of AI, where the increasing demand for logical rigor has led to much work based essentially on logic programmming techniques, resulting in rule based representations, such as BDI in the agent's field, becoming the dominant paradigm many areas (e.g. for agents, see **?**). Meanwhile progress in domains such as vision and speech have been achieved using non-symbolic methods. Ontologies are now established as a central topic in AI and Law, with many papers and regular workshops. Ontologies, however, are best suited to dealing with a snapshot of static knowledge, not knowledge of dynamic domains, where change is essential.

In our domain of interest, policy deliberations, notions of change are fundamental. By policy deliberation, we mean those discussions between government officials and interested members of the public, e.g. unions, business people, voluntary organisations, and individuals, about what policies ought to be made into legislation, where the policies concern the actions that ought to be taken to attain desired objectives. Policy deliberation can occur in other contexts as well; in general, in a policy deliberation the point is to consider ways of moving from the current situation to one which is more desirable in certain defined respects. For such reasoning the underlying semantic model must encompass different states of affairs, and the causal relationships between them.

In this paper, to underpin deliberation on, and justification of, choice of actions, we provide a semantic structure designed to support reasoning about the actions of groups of agents. With this, we use an argumentation scheme designed to support practical reasoning to justify policies. This scheme requires the basic semantic structure to be augmented with social values. The argumentation scheme is formally expressed in terms of the semantic structure. The context in which the policy is made is then modelled as a semantic structure of the required type, and alternative instantiations of the scheme create arguments for and against particular policies. These

form a value based argumentation framework (VAF) **?**, which may be represented as a graph. Evaluating the VAF results in alternative policy positions. The whole process can be seen as a formal basis for generating policy proposals for the Parmenides tool **?**, used to solicit structured feedback on policy justifications. We will illustrate the approach by using a domain previously discussed in **?** and **?**.

In the next section we will introduce our formalism, the *Argumentation Scheme for Practical Reasoning* (PRAS argumentation scheme) **?** and the underlying semantic structure *Action-Based Alternating Transition Systems* (AATS) **?**. In section **??**, we construct a sequence of AATSs to act as semantic models of our context. These models are used in section **??** to identify candidate policies, their justifications and possible objections. These are organised as a VAF in section **??** and the competing proposals are evaluated against different factual assumptions and value priorities to produce defensible policy positions. The analysis is then considered with respect to tools for policy modelling, in particular *Parmenides*, in section **??**. We conclude with some future directions.

2. BACKGROUND

There are two elements to the formal background: an argumentation scheme for practical reasoning, which can be used to justify actions intended to achieve social ends, and an underlying structure, an Action-Based Alternating Transition System (AATS), which can provide semantics for such a scheme and its associated critical questions.

2.1 Argumentation Scheme for Practical Reasoning

In **?** an argumentation scheme and associated critical questions were given to enable agents to propose, attack and defend justifications for action. This argumentation scheme follows Walton **?** in viewing argumentation as presumptive justification - *prima facie* justifications of actions can be presented as instantiations of the argumentation scheme, and then critical questions characteristic of that scheme can be used to identify ways of challenging these justifications. However, **?** gives no formal semantics for its argumentation schemes, and these are needed for use in computational systems. We first give the scheme informally, then in section **??**, give a formal expression in terms of the semantical structure introduced in section **??**.

The argumentation scheme is an extension of Walton's *sufficient condition scheme for practical reasoning* **?** stated as follows:

PRAS1

In the current circumstances R
We should perform action A
Which will result in new circumstances S
Which will realise goal G
Which will promote some value V.

This extended scheme can be viewed as a 'positive' justification in that it proposes an action that would promote some value which is desirable for the agent. PRAS1 can also be stated in a 'negative' version, PRAS2. Where a value is demoted we have a reason to refrain from an action to avoid a state containing particular features (which we continue to call a "goal", although we wish to avoid rather than attain it) which demotes a value. This negative version is stated as follows:

PRAS2

In the current circumstances R
We should not perform action A
Which will avoid new circumstances S
Which would realise goal G
Which would demote some value V.

PRAS2 can thus be used to argue in terms of avoiding some undesirable effect, rather than achieving some positive effect, on our social values.

Associated with both these schemes are critical questions that can be used to identify arguments that attack the arguments which are generated by instantiations of the schemes. Sixteen critical questions associated with each of these schemes are given in **?**. We next present AATSs and show how the PRAS argumentation schemes can be given a formal semantic interpretation using AATSs, so that the schemes can be used in computational systems.

2.2 Action-Based Alternating Transition Systems

In order to be able to automate reasoning based on the argumentation schemes given above, they need to be grounded on some well-defined representation. In **?** such a formal structure was presented by describing PRAS1 in terms of an Action-based Alternating Transition System (AATS). AATSs were originally presented in **?** as semantical structures for modelling game-like, dynamic, multi-agent systems in which the agents can perform actions in order to modify and attempt to control the system in some way. These structures are thus well suited to serve as the basis for the representation of arguments about action. Below we re-capitulate the definition of the components of an AATS **?**.

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

- Q is a finite, non-empty set of *states*;

- $q_0 \in Q$ is the *initial state*;

- $Ag = \{1,\ldots,n\}$ 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 α 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);

- Φ 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.

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 is a tuple $\langle \alpha_1,...,\alpha_k \rangle$, where for each α_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 α_j and $\alpha_{j'}$ in j_{Ag} that belong to the same Ac_i. The set of all joint actions for the set of agents Ag is denoted by J_{Ag}, so $J_{Ag} = \prod_{i \in Ag} Ac_i$. Given an element j of J_{Ag} and an agent $i \in Ag$, i's action in j is denoted by j_i.

A model $\mathcal{M}$ of a policy domain is defined as an AATS, giving instances for each component of the structure - the states, agents, actions, transitions, and propositions. We indicate the models with $\mathcal{M}_1, \ldots, \mathcal{M}_n$. As our understanding of the policy domain increases or changes, we refine, extend, or revise the AATS. It is in this respect that, with the addition in the next section, we provide semantic models which support policy deliberation.

2.3 Extending AATSs with Values

To represent the practical reasoning argumentation scheme using AATSs, we need to extend the AATS structure to enable the representation of values along the lines of **?**. We introduce a set V of values[1]. As given by δ, every transition between two states either promotes, demotes, or is neutral with respect to each value. Additionally, as noted in **?**, some actions promote or demote values by their very performance. This is captured using the function ϵ.

- 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 $v_u \in V$.

- $\epsilon : J \times V \rightarrow \{+, -, =\}$ is a *valuation function* which defines the status (promoted (+), demoted (–) or neutral (=)) of a value $v \in V$ ascribed to performance of an action: $\epsilon(j_x, v_y)$ labels the transition made using j_x with one of $\{+, -, =\}$ with respect to the value $v_y \in V$.

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 $(n + 10)$ tuple $S = \langle Q, q_0, Ag_1, ..., Ag_n, Ac_i, \rho, \tau, \Phi, \pi, V, \delta, \epsilon \rangle$. We continue to use $\mathcal{M}_1, \ldots, \mathcal{M}_n$ to indicate models built from the AATS+V components.

2.4 Restating the Scheme

Given the AATS+V structure, we re-state PRAS1 in terms of it (PRAS2 is similar, but we omit the formal version for reasons of space). This gives us:

PRAS1 (restated)

The initial state $q_0 = q_x \in Q$,
Agent $i \in Ag$ should participate in joint action
$j \in J_{Ag}$ where $j_i = \alpha_i$,
Such that $\tau(q_x, j)$ is q_y,
Such that $p_a \in \pi(q_y)$ and $p_a \notin \pi(q_x)$,
Such that for some $v_u \in V$, $\delta(q_x, q_y, v_u)$ is + or $\epsilon(j, v_u)$ is +.

[1] These were indexed to agents in **?**, but this is not necessary here.

With the formal structure presented above, we first build a sequence of AATS+Vs to supply policy models and then generate arguments for and against policy proposals from the models using the practical reasoning argumentation scheme.

3. THE POLICY MODELLING CYCLE

The formal machinery using the AATS+V is intended to provide the basis for the specification of semantic models which enable arguments about policy proposals to take place. We now consider the general process of policy making, and show how policies can be modelled and argued about, using a running example. The example is an issue in UK Road Traffic policy, previously used as an e-participation example in **?** and **?**. 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 which we will consider is how to reduce road deaths.

The starting point of policy making is when a policy issue on a particular topic is identified and the relevant governing body wishes to launch a consultation to solicit views on the issue. Since there is no specific commitment to a particular action at this stage, a *Green Paper* on the issue will be released publicly. The Green Paper is intended to encourage debate, with a view to interested parties, such as unions, pressure groups, think tanks, companies, universities etc, putting forth their views and comments on the issue, which they submit as formal responses. Considering our running example, the Green Paper would solicit opinions on the issue of what to do to reduce road deaths.

At this deliberative stage of the process, typically a wide range of proposals is put forward representing the different perspectives of different parties with different expertise, interests and values on the issue. For these to inform policy making, the relevant government Department must analyse them to identify relevant facts, theories, interests and values, trying to synthesise them into some coherent form which can provide the basis of deliberation as to the policy to recommend in the subsequent *White Paper*. A White Paper sets out a concrete policy intended to form the basis of legislation and its justification. Again comments are sought from interested parties on the White Paper, but now with this rather specific focus. In short, when moving from the Green Paper to the White Paper, the government department tries to make sense of the alternative views submitted to try to produce a coherent picture of the domain of interest. Of course, this sense-making is not done using a formal apparatus. We argue, however, that such sense-making could be facilitated by formally representing the alternative views as AATS+V models, then reasoning with these models using argumentation schemes. This would clarify the alternative positions on the policy, force reconciliation of any incompatible views, and provide an integrated summary of the consultation. In the next section, we first articulate the process of constructing AATS+Vs, and then provide a sequence of models for our example.

3.1 Constructing Semantic Models of Policies

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 Φ with which we can identify the possible member states of Q. Since if there are n elements in Φ there may be 2^n elements in Q, it is desirable to keep Φ as small as possible and only include propositions if they are definitely relevant to the problem. Given Φ, we can constrain the size of Q by identifying logical relationships between members of Φ, such that for $p_1, p_2 \in \Phi$, $\neg(p_1 \wedge p_2)$.

We need to give the set of agents, Ag, the actions they can perform, and any values inherently promoted or demoted by the performance of the action. Again, in order to keep the number of joint actions within reasonable bounds, we will need to be as frugal as possible in including agents and actions: n agents, each with m actions, give rise to n^m potential joint actions. Finally, we need a transition matrix expressing ρ, τ, δ and ϵ. 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 that the preconditions for the joint action are satisfied, and comprises 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.

Returning to our running example, we suppose that we have received responses from which we can extract intuitive rules; in turn, we use these rules to construct the various components of the AATS+V. As there may be alternative responses, we create alternative models, incrementally (in our example) building a complex model which represents the sum of the policy deliberation. Once we have a final model, we can turn to considering how to argue with it using PRAS1 and PRAS2.

First Model - $\mathcal{M}_1$. For example, one response to the Green Paper issue put forward by those concerned about road safety might be that we install and operate speed cameras at strategic points. The speed cameras automatically photograph speeding cars, and the photographs are subsequently used to identify the car and issue speeding tickets to the drivers; we will use the installation of the cameras to refer to this overall process. There is evidence from other countries and pilot studies that this measure can be effective. So we might propose the following as the intended meaning of the response: *The Government should install speed cameras to reduce road deaths, which will promote the value of Life.* However, we want to argue about policy using our practical reasoning argumentation scheme, which explicitly references circumstances and consequences. The response just given is elliptical, having only the action and the value. So to be compatible with PRAS1, we need to add the circumstances (that road deaths are too high, and that speeding is rife), and a consequence (that there will be fewer accidents and so fewer deaths). There is still some magic here, however: it is not the speed cameras themselves that reduce the accidents: the belief is that speed cameras will cause motorists to observe the speed limits, that observing speed limits will reduce accidents, and this will lead to fewer deaths.

Let us represent this response in the rules SR1-5, which articulate various aspects which may otherwise be implicit; we use these rules to provide the model in AATS+V. For example, SR1 says that in the current state, if the action *Government does not introduce speed cameras* is done, then the action *Motorists do not cut their speed* is also done; in other words, if the Government does nothing, Motorists don't change their behaviour. This rule constrains the joint actions that are admissible in the model. SR2 says that if the action *Motorists cut their speed* is done relative to a state, then in the next state, there is no speeding; this is a constraint on the states reached by joint actions of which this is a component. Finally, SR3 is a constraint on states. Note that these constraints may vary from model to model; here we are modelling the initial response about the government's introduction of speed cameras. As additional responses are contributed, the model will change accord-

ingly. Finally, our representation is a fragment of a fully specified model, which would require completion. This process of completion forces us to think further about the context.

SR1 $G_0 \to M_0$, where G_0 means the *Government does not introduce speed cameras*, and M_0 means *Motorists do not cut their speed*.

SR2 $M_1 \to \odot \neg s$, where M_1 means the *Motorists cut their speed*, s is *there is excessive speeding*, and $\odot$ is the "next state" operator. That is, if motorists cuts their speed is part of a joint action, then $\neg s$ will be true in the state reached.

SR3 $\neg s \to \neg a$, where a is *there is an excessive number of accidents*, and $\neg a \to \neg r$, where r is *there are excessive road deaths*, therefore, $\neg s \to \neg r$. The first premise expresses the constraint that there will be no decrease in accidents without a decrease in speeding; the second premise expresses a constraint between accidents and deaths; the conclusion is that there is a constraint between speeding and deaths.

SR4 $(r \wedge \odot \neg r) \to +L$, where L means *life*, so +L means that life is promoted. This expresses the notion: If q is a state of the AATS+V in which r holds and there is a transition to a state in which $\neg r$ holds, then that transition is labelled with +L.

SR5 $(\neg r \wedge \odot r) \to -L$, which means that a transition from a state with $\neg r$ to a state with r demotes life.

From this conceptualisation of the problem, we present our first model, the AATS+V $\mathcal{M}_1$:

- $Q = \{q_0, q_1, q_2\}$. SR3 means that one potential state is in fact impossible;

- $Ag = \{G, M\}$, where G is the government and M is motorist;[2]

- $Ac_G = \{G_1, G_0\}$, which are the actions the government does or does not perform, respectively. $Ac_M = \{M_1, M_0\}$, which are the actions the motorist does or does not perform. here G_1 is *operate speed cameras*, and M_1 is *cut speed*.

- $\Phi = \{r, s, \neg r, \neg s\}$. While we informally also have a, we assume, for the example, that a and r can be taken as equivalent, since accidents and deaths are correlated as in SR3;

- $V = \{L\}$;

- δ is such that $\delta(q_x, q_y, L) = +$, if r holds in q_x and $\neg r$ holds in q_y; $-\neg r$ holds in q_x; and r holds in q_y, and = otherwise. This applies the evaluation of SR4;

- π is a function such that $\pi(q_0) = \{r, s\}$, $\pi(q_1) = \{\neg r, s\}$, and $\pi(q_2) = \{\neg r, \neg s\}$, given SR3;

- J_{Ag}, the set of all joint actions, is $\{j_0, j_1, j_2\}$, where j_0 is $< G_0, M_0 >$, j_1 is $< G_1, M_0 >$, j_2 is $< G_1, M_1 >$. Given SR1, we have assumed that Motorists do not cut their speed if the Government does not install speed cameras.

[2]Where *motorist* is an abstraction to use the 'collective' interpretation of 'motorist'.
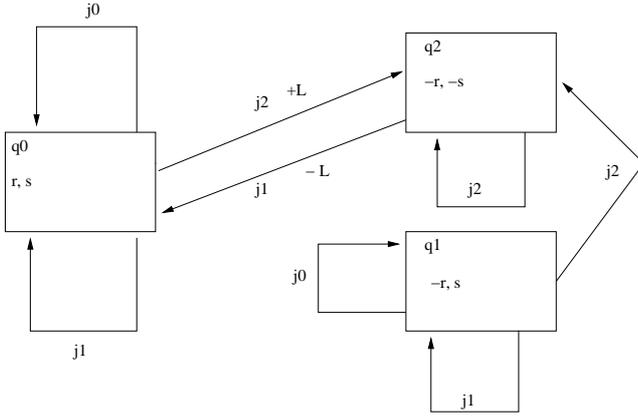
Figure 1: $\mathcal{M}_1$

The model $\mathcal{M}_1$ also requires the functions ρ and τ. However, given SR1-5 along with the rest of $\mathcal{M}_1$, we can express these as in a transition matrix shown in Table **??**: an entry in a cell indicates the preconditions for the joint action are satisfied; the first argument is the state reached if that joint action is executed, the second is the set of values promoted, and the third is the set of values demoted; where no value is promoted or demoted, we have _; *null* means the preconditions of one or more of the component actions cannot be satisfied, so that joint action is not possible in that state. This is true of j_0 in our example: in the case where the speed cameras have succeeded in reducing speeds, it is assumed that the Government will continue to operate them, so that only the joint actions containing G_1 are possible in q_2.

Table 1: Joint Actions in $\mathcal{M}_1$

	j0	**j1**	**j2**
q0	$\langle q0,\_,\_\rangle$	$\langle q0,\_,\_\rangle$	$\langle q2,+L,\_\rangle$
q1	$\langle q1,\_,\_\rangle$	$\langle q1,\_,\_\rangle$	$\langle q2,\_,\_\rangle$
q2	null	$\langle q0,\_,-L\rangle$	$\langle q2,\_,\_\rangle$

Put together this gives the AATS+V shown as $\mathcal{M}_1$ in Figure **??**.

Second Model - $\mathcal{M}_2$. As the consideration of Green Paper responses continues, we find people who dispute $\neg s \rightarrow \neg r$, the constraint expressed in SR3. These people dispute that excessive speeding is a factor in deaths, so that the previously impossible state with $[\neg r \wedge s]$ is now possible, which we will refer to as q_3. In order to represent that the effect of M_1 is not determinate, we introduce a third agent N, which is usually termed *nature*, and distinguish two joint actions containing M_1, depending on whether nature cooperates (N_1), meaning that a reduction in speed has the desired effect on deaths, or nature does nothing (N_0), in which case a reduction in speed does not have the desired effect.

To the constraints SR1-5 we add or substitute the following:

SR6 $N_0 \rightarrow \odot r$. Any joint action containing N_0 reaches a state with r;

SR7 $N_1 \rightarrow \odot(\neg s \rightarrow \neg r)$ (replacing SR3, which now is conditional on N_1.)

Having formally specified $\mathcal{M}_1$, we can use it to assimilate other views and build more complex models. We don't formally specify all the components each time, only those which are relevant, assuming all else is unchanged. To accommodate the new rules, SR6 and SR7, we have a second model $\mathcal{M}_2$ just like $\mathcal{M}_1$ except that $Ag = \{G, M, N\}$ and $Ac_N = \{N_0, N_1\}$. In joint actions without M_1, nature does nothing. So we now have four joint actions, the additional action reflecting the two possible effects of M_1: j_0 is $< G_0, M_0, N_0 >$; j_1 is $< G_1, M_0, N_0 >$; j_2 is $< G_1, M_1, N_0 >$, and j_3 is $< G_1, M_1, N_1 >$. Referring again to Figure **??** suitably modified, j_2 leads to q_3 (where SR6 applies), and j_3 leads to states where SR7 applies, e.g. q_2.

The new argument was intended as an objection to speed cameras. We could, however, respond to this by saying that even if compliance with speed limits did not have a significant effect on accidents, it would still be worthwhile, since it would mean that there was increased compliance with the law, and that this is a value in itself. We can then label transitions with $+C$, where C represents the value of compliance, in accordance with the constraint SR8.

SR8 $(s \wedge \odot \neg s) \rightarrow +C$. Any transition from a state with excessive speeding to one without promotes C;

Not every submission need lead to a change in the model. For example, it might be argued that speed cameras will have no effect unless the speeding motorists identified are prosecuted and punished, so that there are no joint actions with both G_1 and M_1, unless some third agent, such as the police, is considered. Here, however, there is no need to change the model. The idea of applying sanctions to the wrongdoers identified was at least implicit in our notion of G_1 in that actions by the Government are "supported" by a system of sanctions on wrongdoers; here we can modify, or rather make explicit, aspects of the *interpretation* of this action but leave the *model* unchanged.

Third Model - $\mathcal{M}_3$. Here we add some additional aspects, considering the cost of the proposal and an alternative proposal involving education. Speed cameras cost money, and there is only a limited budget available for improving road safety. We then need to consider monetary matters. This will relate to a value B, which is demoted if the budget is exceeded and promoted if there is a surplus. Assuming we do have money to spend, we can cost our plan and interpret G_1 as being the introduction of such speed cameras as the budget will allow. Where cameras are installed according to budget, then G_1 is neutral with respect to B; thus the transition to q_2 will be neutral with respect to B. If, however, motorists fail to respond to the deterrent effect of the cameras, continue to speed, and pay the fines, then because we can easily identify and prosecute the speeders, income from fines will be greater than expected and the expenditure will be recouped. Thus, the transition back to q_0 involving G_1 will promote B, meaning that there is justification for G_1, even if this is the joint action that results.

SR9 $[G_1 \in j_x \wedge \odot s] \rightarrow +B$ (if introducing speed cameras does not lead to a cut in speeding, budget is improved).

For an alternative action, suppose there is a submission by a group who believes that introducing speed cameras will not reduce road
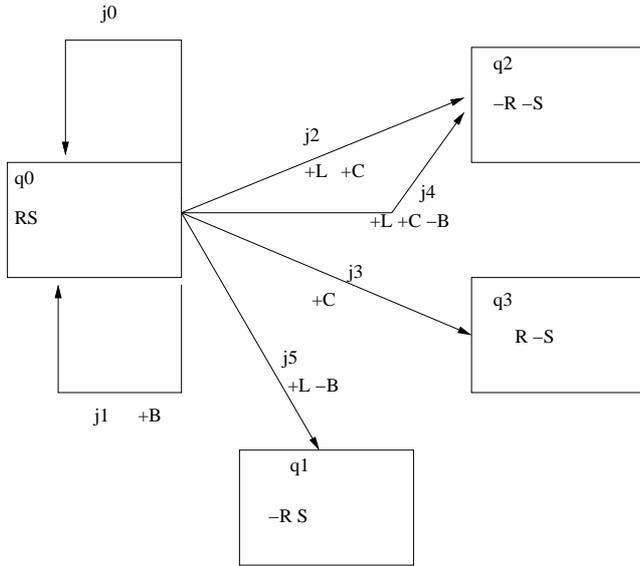
Figure 2: $\mathcal{M}_3$

Figure 3: $\mathcal{M}_4$

deaths, but are very much in favour of reducing these deaths. They may argue that some other action is required to be effective. For example, if we were to educate drivers, so that they were better aware of the effects of speed, and better able to handle their vehicles at speed, then we would expect to reduce accidents, and hence deaths. Thus the government's education of drivers, G_2, would, it is argued, lead to a reduction in deaths whether or not speeding decreased, since motorists who continue to speed are better able to control their cars. The only problem is that education is more expensive than cameras, does not give rise to any revenue stream, and so this proposal would be over budget, demoting B.

SR10 $G_2 \in j_x \rightarrow \odot \neg R$. Uses the ϵ function to indicate that G_2 will always reduce deaths;

SR11 $G_2 \in j_x \rightarrow -B$. Uses the ϵ function to indicate that G_2 will always exceed the budget.

G_2 can also be used to represent any other Government actions which it is claimed will lead to a reduction in accidents but which will exceed the budget, such as deploying increased numbers of traffic police to catch speeders. Note, however, that we now need to distinguish between speeding and accidents, in contrast to the presumption in $\mathcal{M}_1$, and so require the fourth state where speeding continues, but deaths decrease, reachable by the joint action educating motorists who continue to speed.

This gives the AATS+V shown in Figure **??**, where j_4 is education and reduced speed and j_5 is education and no reduced speed (but better driving). The figure only includes transitions from the initial state. $\mathcal{M}_3$ can be seen as representing the response to the Green Paper of several bodies incorporating scientific evidence, representations from motorists, and other pressure groups, and Treasury constraints.

Fourth Model - $\mathcal{M}_4$. Finally, we will consider responses to the Green Paper that are representative of arguments from Civil Liber-
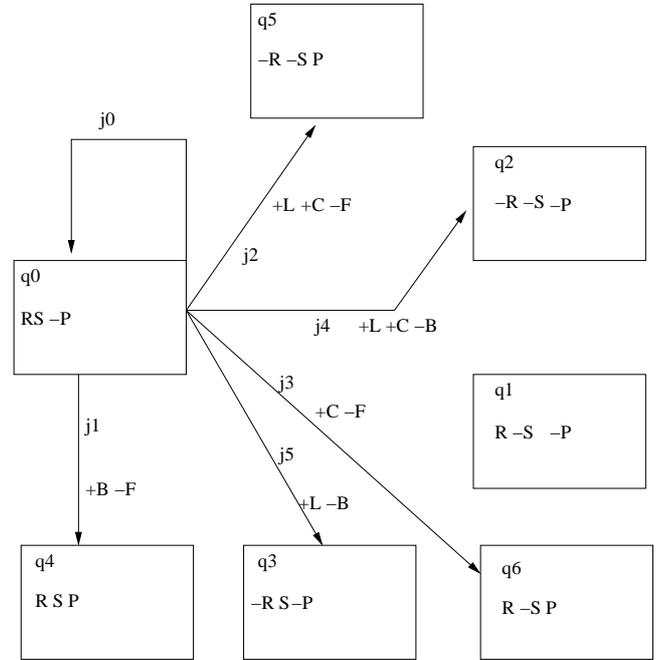
ties pressure groups. They argue that speed cameras, by revealing the location and movements of citizens, represent an unacceptable intrusion of privacy. This requires a new proposition p to represent the existence of the speed cameras making an excessive intrusion on privacy. This will be accompanied by an additional value, F representing civil liberties.

This requires an extension to the model: adding p splits every state reachable by G_1 into two: where p is false (essentially the existing states which consider p to be irrelevant), the transition to a state where p is true demotes F; where p true, transition to a state where $\neg p$ will promote F.

This adds the constraints:

SR12 $\neg p \wedge \odot p \rightarrow -F$. Intruding on privacy with cameras demotes civil liberties;

SR13 $p \wedge \odot \neg p \rightarrow +F$. Undoing SR12.

The row for q_0 of the final transition matrix which embodies SR1-13 is shown in Table **??** and the resulting AATS+V appears in Figure **??**. For clarity, and because we take it as undisputed that the initial state is q_0, only the transitions from q_0 are shown in that diagram.

Table 2: $\mathcal{M}_4$ **transition matrix.**

	j0	**j1**	**j2**
q0	$\langle q0,\_,\_\rangle$	$\langle q0,+B,-F\rangle$	$\langle q5,+L+C,-F\rangle$
	j3	**j4**	**j5**
q0	$\langle q6,+C,-F\rangle$	$\langle q2,+L+C,-B\rangle$	$\langle q3,+L,-B\rangle$

Although we ended our discussion after considering four alternative proposals, expressed in four semantic models each represented

as an AATS+V, it could well be that many more points of view were put forward at the Green Paper stage from which additional models could be constructed. Green Papers may well generate a large number of responses and it may be necessary to filter these to rule out proposals that do not meet some necessary constraints. For example, an alternative approach to improving road safety could be proposed through the introduction of more transport police. Whilst this may be an effective action to achieve the goals, it is unlikely that, for financial and other reasons, sufficient additional police could be recruited, a fact known to the policy analyst. Furthermore, other departments within the police force are always competing for scarce resources and an increase in traffic police would not feature as high on the list of priorities as would an increase in police numbers for other departments such as counter-terrorism or street crime. Thus, any such opinions put forward could be considered and immediately discounted as inappropriate by the policy analyst. Effectively, the policy analyst is deciding that such an action is simply not possible. Finally note that each element in the model and each rule in SR1-13 can be justified in terms of some statement taken from the responses to the Green Paper.

When the response period for the Green Paper closes, the opinion gathering ends and the policy analyst can then focus on the proposal to be chosen as the preferred option to be set out in the White Paper, forming the next part of the process.

4. USING THE MODEL TO CRITIQUE A POLICY PROPOSAL

Returning to our example, we have reached the stage when some particular policy proposal needs to be chosen and justified to form the basis of a White Paper. When considering which proposal to present, the analyst will need identify candidate proposals, the arguments in their favour, the objections to them, and any further rebuttals of these objections. In doing so, a number of arguments and the attack relations between them will be identified. Having produced this value based argumentation framework (VAF)**?**, the arguments need to be to evaluated, so that sets of acceptable arguments - perhaps relative to some factual assumptions and some prioritisation of values - identified. Each such set will represent a *policy position*, and justify some line of action. Examining the AATS+Vs will enable an analyst first to identify what actions can promote values and so be considered candidate proposals, and to identify the attacks that can be made on each of the proposals put forward, and the counterattacks on these attacks. The framework can then be resolved into admissible sets of arguments (sets which, given an ordering on values, is free from internal conflict, and able to defeat any argument outside the set which attacks any of its members). VAFs and their evaluation are fully and formally described in **?**.

4.1 Generating the Arguments

Thus, now we have a model, we can use it to generate arguments and attacks on them. These arguments may include those advanced in the responses to the Green Paper used in the construction of the model, but we can now approach the task in a more organised fashion. We first establish that the initial state, q_0 is where speeding is excessive, and deaths are excessive, but privacy is respected. Of course, this could be challenged, but for this paper we will assume that all are agreed that this is the current situation. The submissions to the Green Paper are likely to contain sufficient evidence and support for these statements.

We now identify proposals by finding instantiations of PRAS1. Be-

cause all accept q_0 as the current situation, we will omit reference to it in the arguments here. We will have one instantiation of PR for every transition from q_0 which promotes a value. Thus:

PR1 We should perform G_1 to reach q_5 to promote L (this was the argument, we started from)

PR2 We should perform G_1 to reach q_5 or q_6 to promote C (this argument is new)

PR3 We should perform G_1 to reach q_4 to promote B (again this argument is new)

PR4 We should perform G_2 to reach q_2 or q_3 to promote L (proposed in a submission)

PR5 We should perform G_2 to reach q_2 to promote C (new)

We also have arguments based on the negative version PAS2, for every transition which demotes a value:

NPR1 We should not perform G_1 to avoid q_5 and q_6 since this would demote F

NPR2 We should not perform G_2 to avoid q_2 and q_3 since that would demote B

These arguments attack one another in a variety of ways. In **?** seventeen ways of attacking an instantiation of PR were given and formalised in terms of an AATS+V. These attacks were related to three stages of the practical reasoning process: the *problem formulation* stage, which disputes features of the AATS+V, the *epistemic* stage which queries either the current state, or the assumed action of another agent, and the *option selection* stage which considers alternative proposals to discover which is best. Questions about the formulation of the AATS+V were discussed in the previous stage, and we added propositions, actions and values to an adequate, but incomplete, model like that shown in Figure 1, to yield the additional transitions and states shown in Figure 3. For example, we initially assumed in $\mathcal{M}_1$ that cutting speeds would result in fewer deaths. When this assumption was challenged we introduced an agent N, so that we could represent situations in which this assumption did not hold. The final model, $\mathcal{M}_4$, is intended to represent our considered view, and discussion on these lines is, for the present, closed. Epistemically, we may agree for the present that q_0 is the initial state. The behaviour of the motorist is a source of epistemic questions and that motorists may perform M_0 instead of M_1 represents an attack on PR1, PR2 and PR5. Similarly nature may or may not reduce accidents in response to speed, representing an attack on PR1, but no challenge to PR4, since it is held that the increased skills of motorists will reduce accidents and deaths even if they continue to speed. Finally PR3 assumes that motorists will do M_0 even when cameras are introduced. This gives rise the three epistemic objections.

Ob1 Motorists may choose M_0 not M_1: attacks PR1, PR2 and PR5.

Ob2 Reducing speed may not reduce accidents and deaths. Attacks PR1.

Ob3 Motorists may choose M_1 not M_0: attacks PR3.

Only PR4 is immune to epistemic objections, mainly because it succeeds in promoting the desired value whatever motorists and nature do. In order to answer Obj1-3 we need to step outside of our model. For example we may quote the experience of other countries, or of pilot studies, to support our view that motorists will (or will not) reduce their speed in the presence of cameras. Accident statistics may be used to establish the proportion of accidents, and of fatal accidents, in which speeding was considered a crucial element. Much of the information we need will be available from the responses to the Green Paper. Considerations such as these - and the argumentation which supports them - takes us outside of deliberation. We are no longer concerned with *choice*, but with scientific questions of fact or causality (physical and psychological). Such questions therefore need to be resolved for the purposes of continuing the deliberation by agreeing to take a view on the facts, or by obtaining sufficient additional information to resolve the questions.

We now reach the final stage, when we weigh the merits and demerits of competing options. One source of attack is that a value is demoted: thus NPR1 attacks PR1, PR2 and PR3, and NPR2 attacks PR4 and PR5. Another source of attack, giving rise to symmetric attacks, is an alternative way of promoting the same value: thus PR1 and PR4 mutually attack, and PR2 and PR5 mutually attack. Finally we have different actions promoting different values: PR1 and PR5 and PR2 and PR4 mutually attack in this way. Finally we can have attacks which question the motive put forward: if PR1 is advanced to justify speed cameras, some may argue that the real expectation is that q_4 will be reached and that the real motive is the save money, rather than lives. This, however, does not challenge the action, but the justification, and we will not include these attacks here. We can now evaluate the arguments.

5. EVALUATING THE ARGUMENTS

We can organise these arguments in a Value based Argumentation framework (VAF). VAFs are formally defined in **?**. Here we may say that an argument is defeated if it is attacked by an undefeated argument *unless* it relates to a value preferred to that of its attacker. Factual arguments defeat all value based argument: practical arguments have the value they promote or demote. A set of arguments is mutually acceptable (*admissible*) if none of them defeat a member of the set, and some member of the set can defeat every argument attacking a member of the set. A maximal admissible set is a *preferred extension*. The VAF for $\mathcal{M}_4$ is shown diagrammatically in Figure 4.

5.1 Resolving Epistemic Questions

On the left of the diagram are the two epistemic questions that need to be resolved. In default of anything better let us assume that, on the best evidence available, it is reasonable to expect that motorists will in fact reduce their speed, and that reducing speed will indeed lessen accidents and deaths. Having resolved these two cycles, we have answered the attacks from Ob1 and Ob2, while Ob3 is no longer attacked and will defeat PR3. When arguments are defeated, we can remove them and their attacks (and attacks on them) from the VAF to obtain the simpler VAF, as shown in Figure 5. Note that if we had made different assumptions about the epistemic questions a different VAF, and ultimately a different position, would result from this simplification. When an argument is not defeated, but its attack is resisted by a preferred argument, we mark it as *ineffective*. We cannot ignore it, since we have no argument to defeat it, but we will not act upon it. There are no such arguments as yet, since we have not yet exercised preferences, but only chosen between different factual assumptions.
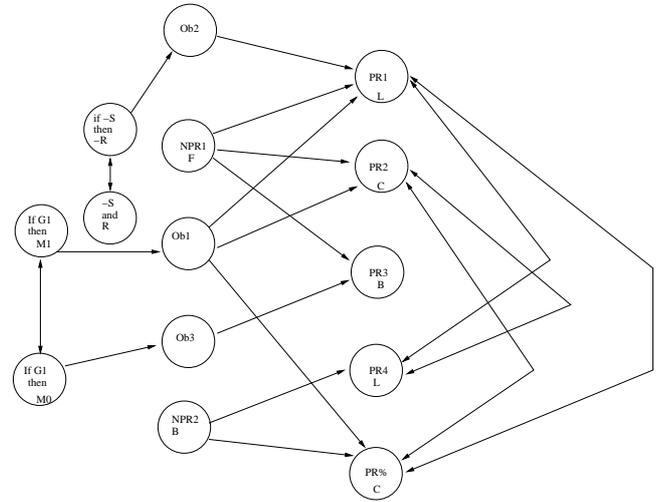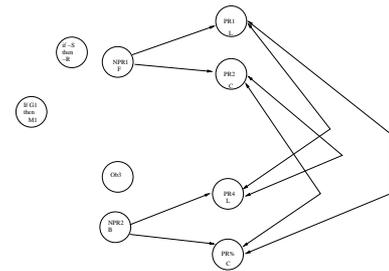

Figure 4: VAF1

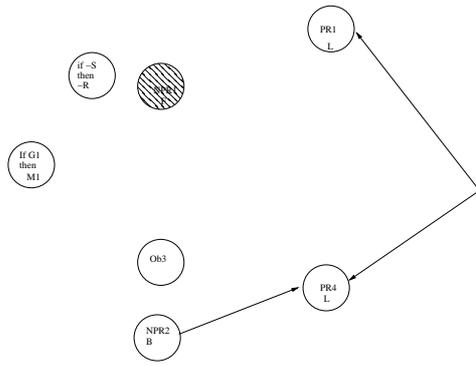**Figure 5: VAF after resolving epistemic questions**

Figure 6: VAF after adopting values $L > F > C$

5.2 Consideration of Preferences

We next consider the two negative arguments based on PRAS2; once we have reached Figure 5 by resolving the epistemic questions, these are unattacked. These arguments will therefore succeed in defeating the arguments they attack unless the value of the attacked argument is preferred to that of the attacker. For NPR1 we must therefore consider Privacy/Freedom against Life to resolve the attack on PR1, and against Compliance to resolve the attack on PR2. A reasonable order would seem to be $L > F > C$: saying that intrusion on privacy is a necessary evil to save lives, but would not be acceptable simply to ensure compliance with speed limits without other gains. NPR1 thus becomes ineffective, which we show in the diagram by shading the argument node. This yields the VAF in Figure **??**.

The final question to resolve is whether PR4 can be accepted given NPR2: that is can we prefer L to B? Unfortunately we are regarding budget as a hard constraint and so we must answer that $B > L$. This means that PR4 falls, leaving a preferred extension for an audience with of $B > L > F > C$ comprising: the two factual assumptions, that motorists will reduce their speed, and that less speed means fewer accidents and deaths; the accepted course of action to install cameras to save lives; and two other considerations, that privacy must unfortunately be lessened (represented by the undefeated but ineffective argument), and that budgetary constraints preclude education as an alternative (represented by Obj3). Of course similar reasoning with different assumptions and different value orders would produce different results. If we assumed that motorists would continue to speed with the same value order, we would still install the cameras, but this time on the basis of PR3. If we made the original assumptions but used the value order $B > F > L > C$, we could do nothing, since we would have no way of saving lives without infringing privacy that we could afford, and if we had the value order $F > B > L > C$, we would educate motorists rather than install cameras.

Finally, if money was available so that it was possible to prefer L to B, we would have two equally valued arguments, PR1 and PR4, neither attack except by each other. In this case we should be inclined to choose PR4, since this would mean that the undefeated NPR1 would no longer have to coexist with an argument it attacks, so that it no longer need be regarded as ineffective. In this way we are able to respect the value of privacy, even though F is not preferred to L.

Considerations of these varied alternatives allows us to see how the policy positions favoured depend very critically on how we rank values: the acceptability of a proposal will often depend on whether the public mood has been correctly judged in this respect.

5.3 Adopting a Proposal

Of course many of the arguments could have been discovered without the AATS+V, and indeed were used when designing the AATS+V. The models we construct using the AATS+V, however, allow the systematic, and in principle automatic, detection of all arguments, and a procedure for resolving them. This procedure separates questions of fact, which need to be resolved outside of the deliberation, or carried into the deliberation as assumptions, from matters of choice, based on value preferences.

Once a policy has been chosen, the justification can be made in terms of factual and behavioural assumptions and in terms of value preferences. The preferred extension records the complete position, including the factual assumptions made, and any arguments rendered ineffective through a preference choice. Such ineffective arguments are points where the position is particularly vulnerable, since the argument cannot be defeated. As well as the position, the AATS+V models themselves can be offered for inspection. Someone confronted with the above reasoning might, for example suggest that the correct thing would be to abolish speed limits altogether (introducing an additional transition), on the grounds that speed did not lead a reduction in accidents (questioning an assumption). This additional transition reaches the currently unreachable q_1 and promotes C while demoting nothing. On the basis of this assumption, and a preference for either C or F over L, we would choose this new course of action. Both extreme libertarians and those who felt that disregard of motoring crimes led to disrespect for law in general, would wish to support this course of action, provided they could accept the factual assumption required.

6. SEMANTIC MODELS AND TOOLS FOR POLICY MODELLING

The process that we have used as the framework for this paper characterises practical reasoning, and hence policy making, as being embodied in the notion of an argumentation scheme, which has subsequently been articulated in a formal manner. It is the semantic structure which underpins the argumentation scheme that acts as our deep model. However, in the context of deliberative democracy, discussions about policy proposals will involve end users not trained in the technicalities of using such a highly formal model, or even able to appreciate the various aspects of argumentation with the precise interpretations required for computational purposes. Whilst we have argued that the underlying semantic model provides a rigorous basis for supporting reasoning about a topic of debate, in order for such a mechanism to be of real practical use, it must be tailored and placed in a suitable context so as to be amenable for use by laypersons. Indeed one of the benefits of the use of argumentation schemes is that they are a natural language representation that capture familiar patterns of reasoning, and so remain relatively easy for humans to understand at an intuitive level. A tool that provides the precise functionality that we require has been set out in **?**. The tool, named Parmenides, presents end users with policy justifications for them to critique in an online medium. A natural question to pose then, is how the formal model we have provided can underpin this tool to support policy critiquing.

An interaction with Parmenides begins with the user being pre-

sented with a policy proposal that is structured in terms of the practical reasoning argumentation scheme[3]. Users are then given the opportunity to systematically critique the constituent parts of the instantiation, which is done by encouraging them them to provide responses posing certain critical questions in order to express any objections they may have to the proposal presented. Users themselves need have no awareness of the critical questions, they simply give yes or no responses to questions generated by Parmenides, depending on whether or not they are satisfied with a particular element of the policy proposal. Parmenides is best used to solicit opinions once a policy proposal has been formulated, which is why a White Paper would be an appropriate stage to represent in the system.

Parmenides will be administrated by a policy analyst who must put forward the chosen proposal. The first issue to thus be addressed is which of the, possibly competing, potential proposals should be chosen as the one to represent the interests of the organisation[4]. It is likely that a number of criteria will be of relevance here; the values of importance to the body the analyst represents would be the most obvious consideration, but there may also be other relevant criteria such as the likelihood of success of the proposed action. Thus, when the policy analyst is deciding which of the competing proposals to commit to, what he is in fact doing is making assumptions about the facts that will be deemed acceptable to those being presented with the proposal, as well as assumptions as to the value ordering they have that will yield the respective acceptable arguments. The decisions about these assumptions then need to be tested with the public, which is done by them critiquing the policy through the Parmenides system. We now briefly describe how this proposal can be put forward for critique within the Parmenides tool.

Upon entering the Parmenides website and choosing to comment upon this particular topic, users will be presented with a natural language version of the policy, plus justification, chosen to address the issue of reducing road deaths. This proposal will be structured to instantiate PRAS1 as given in section **??** of the paper, but the users are not told that such a structure is being imposed upon the information. Users will then be lead through a series of six web pages that question aspects of the justification. Initially a user will be asked the question "Do you agree that the number of fatal road accidents is a cause for concern?", to which he must answer 'Yes' or 'No' by checking the appropriate box. This is effectively posing a critical question that challenges the basic aim of the proposal. The second screen concerns the initial state, and allows the user to suggest any doubts as to the characterisation of the starting point. Obviously a different q_0 would call all the arguments into question. The next screen asks whether the values are worth promoting. Rejecting values would eliminate some arguments. The next screen allows the connection between various goals and values to be questioned, followed by a screen checking the effects of actions, so that the epistemic doubts of section 5.1 can be raised. Finally the users are asked for alternative actions which they might prefer, which would indicate different value priorities from those used in section

5.2. Thus the user is encouraged to dispute the epistemic questions as to the initial state and the states reached by various actions, the existence of values motivating arguments, the labelling of transitions and the value preferences, without any need to be aware of PRAS1 or AATS+V or VAFs, let alone how they have been used to produce the justification.

We can see that the responses that will be supplied by the user will map to the different states and transitions that appear in the AATS+Vs developed during the previous phase of the policy modelling cycle. Thus, prior to the White Paper being released, the analyst can see in advance the potential attacks that could be made against the proposal being presented. As such he has the opportunity to consider the merits of these and find appropriate rebuttals either using counter attacks or value preferences. Parmenides provides a facility for making such justifications for assumptions available to the users, who may then change their minds and retract their objections. It may well be that some of the alternative suggestions offered by respondents are additional to those represented in the underlying AATS+V and again these can be considered, included and evaluated to see whether there is sufficient reason to make corresponding changes to extend the model.

The Parmenides tool not only collects opinions from end users about policy proposals, but it is also equipped with some analysis facilities that enable the site's administrator to see an aggregated view of the responses collected. Statistical data can be viewed that shows which parts of the policy proposal are most agreed or disagreed with and whether the action suggested in the policy is seen as acceptable overall (with the argument evaluation being carried out through a value-based argumentation framework of the sort used in section **??**). This part of the tool is in effect checking when there is majority agreement with the facts assumed and the value ordering applied to the arguments of concern to the debate. Such data can then enable the policy analyst to report and act upon the responses as appropriate.

7. CONCLUSION

In this paper we have constructed a series of AATS+Vs able to serve as semantic models with which we can examine policy proposals. Policies reflect changes which are designed to realise particular social values, and our models are fully able to capture the dynamic and evaluative aspects of problem. In turn, the models are used to generate arguments using a practical reasoning argumentation scheme and associated critical questions, allowing us to identify and evaluate arguments and question elements of the models to produce, critique and choose between alternative policy positions. Finally, the semantics underpin *Parmenides*, a current web-based implementation designed to solicit feedback on particular policy proposals.

In future work, we plan to further elaborate semantic models of policy deliberation, to analyse existing deliberations in terms of the model, and to extend *Parmenides* so as to support these enriched expressions of policies.

8. ACKNOWLEDGMENTS

---

[3]An earlier version of the speed camera debate is implemented in the Parmenides tool and can be viewed at: *http://cgi.csc.liv.ac.uk/~parmenides/speedcam/*

[4]We envisage that the tool would be of use to government bodies, pressure groups, and any other parties who have vested interests in opinion gathering on policy proposals, but for the purposes of this paper we make no commitment as to the party whom the analyst represents.