Generating Intentions Through Argumentation

Katie Atkinson, Trevor Bench-Capon and Peter McBurney,
Department of Computer Science,
University of Liverpool, UK.
{k.m.atkinson,tbc,p.j.mcburney}@csc.liv.ac.uk

ABSTRACT

In this paper we consider how a BDI agent might determine its best course of action. We draw on previous work which has presented a model of persuasion over action and we discuss a formalism based upon this model which extends the BDI agent architecture to include the notion of value functions. This formalism will allow BDI agents to reason and argue about practical action, in accordance with our underlying model.

CATEGORIES AND SUBJECT DESCRIPTORS

I.2.0 **[Artificial Intelligence]:** General: *Philosophical Foundations*. I.2.4 **[Artificial Intelligence]:** Knowledge Representation Formalisms and Methods: *Representation Languages*. I.2.11 **[Artificial Intelligence]:** Distributed Artificial Intelligence: *Intelligent Agents*. I.2.11 **[Artificial Intelligence]:** Distributed Artificial Intelligence: *Multiagent Systems*. I.2.11 **[Artificial Intelligence]:** Distributed Artificial Intelligence: *Languages and Structures*.

GENERAL TERMS

Design, Languages, Theory.

KEYWORDS

Argumentation, Practical Reasoning, Persuasion, BDI Agents.

1. MOTIVATION

The ability to reason effectively about what is the best or most appropriate course of action to take in a given situation is an essential activity for an agent. However, practical reasoning — reasoning about action — has not received, in either Computer Science or Philosophy, the attention that has been given to reasoning about belief. In this work we provide an account of practical reasoning for agent systems by proposing a formalism for BDI agents to reason effectively and argue with themselves or other agents about practical action. Our account is based on previous work which has made use of an argument scheme and associated critical questions. We now take this general model of persuasion over action and tailor it for use by BDI agents.

One influential approach to practical reasoning in informal logic

has been given by Walton [6] which regards practical reasoning as a species of presumptive argument. Given such an argument, we have a presumptive justification for performing the action. This presumption can, however, be challenged and withdrawn. Subjecting our argument to appropriate challenges is how we hope to identify and consider the alternatives that require consideration, and determine the best choice for us, in the particular context. Because the challenges are, in principle open ended, the process of justification does not end, and discussion can always be re-opened. Walton uses the notion of an argument scheme to present an argument giving a presumption in favour of its conclusion. Whether this presumption stands or falls depends on satisfactory answers being given to the critical questions associated with the scheme, posed in the particular situation.

The primary argument scheme for practical reasoning given in [6] is the *sufficient condition scheme*:

> W1   G is a goal for agent a
> Doing A is sufficient for agent a to carry out G
> Therefore agent a ought to do A.

Walton associates four critical questions with the scheme:

CQW1:   Are there alternative ways of realising G?
CQW2:   Is it possible to do A?
CQW3:   Does agent a have goals other than G which should be taken into account?
CQW4:   Are there other consequences of doing A which should be taken into account?

We believe this argument scheme and its critical questions both need some elaboration, since the notion of a goal is ambiguous. An action may be justified in terms of:

- its direct consequences (the effects of the action),

- a state of affairs following from the direct consequences, which the action was intended to realise (the goal of the action),

- the underlying social value promoted by performing the action, so as to realise the goal (the purpose of the action).

Walton does also not consider all the problems with soundness of W1, presupposing that the second premise is to be understood in terms of what an agent knows or reasonably believes. In [3] an argument scheme is proposed which extends Walton's scheme by unpacking his goal into *direct consequences*, *goals* and *purposes* and also makes explicit the factual context. This extended argument scheme is as follows:

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

Additionally [3] makes Walton's argument scheme more precise by giving relatively formal definitions of the critical questions associated with it. This allows for the separation of conflicting parts of the argument that can be resolved through verification of objective facts, from the parts which are a matter of subjective choice for each individual participant.

In [2] sixteen critical questions associated with AS1 are identified. This allows us to question the presumptions made in the argument scheme in order to consider all alternative options and thus choose the "best" action based upon the justifications. Also worthy of note is that each of the critical questions falls into one of three distinct categories which relate to the nature of the attack: issues relating to the beliefs as to what is the case; issues relating to desires as to what should be the case; and issues relating to representation concerning the language being used and the logic being deployed in the argument. How the attacks are resolved depends upon the categories into which they fall [3]. The full list of critical questions associated with AS1 are:

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

2.   APPLICATION TO BDI AGENTS

In [1] we present a formalism based upon the model of practical reasoning detailed above, which specifies how BDI agents can effectively reason about the best course of action to take in a given situation. This formalism provides a set of pre-conditions to specify when an agent can construct a position based upon its beliefs, the set of actions available for performance, the agent's desires, and its values. Current BDI architectures do not use the notion of values, and so we have extended the architecture to include values which provide justifications for the agent's choice of intentions, based upon its beliefs and desires. We then specify a full set of pre-conditions which when satisfied allow agents to attack a justification for action by posing any of the critical questions from our theory against the position. Once all the possible presumptive justifications and their critiques have been identified, the defensible course of action can be reasoned about using a Value-Based Argumentation Framework [4], which is an extension to Dung's argu-

mentation framework [5]. Such frameworks are intended to be used as part of the agent's filtering process to enable it to choose which intentions to adopt from the set of candidate actions generated in the previous phase.

We have applied this proposal to three different application domains: medicine, law and politics. The medical application describes a multi-agent system for reasoning about the treatment of a patient. The agents bring different information, representing different areas of medical knowledge and facts about individuals, and also represent different policies and perspectives (such as cost, efficacy, safety, etc.) relevant to the problem. A central agent performs the argumentation by coordinating these contributions and forming them into arguments to come to a decision. Presumptive arguments are generated to identify and critique all the possible treatments for the particular patient. These options are then represented as a VAF to determine the preferred treatment. The critical point of the reasoning involved here concerns the uncertainly of effects of different treatments. As individuals respond differently to the same treatment, the coordinating agent must gather as much information as possible about the patient and the likely effects of drugs on them. It can then assess the risks and benefits by taking multiple perspectives into account, in order to choose the best treatment.

In the legal application we use our approach to model the reasoning in a much studied property law case by reconstructing it in terms of our account. We model the different views of the case using BDI agents, describe the resulting argumentation frameworks, and relate them to the case's original majority and dissenting opinions. An important feature of this application is that the reasoning produces three related levels of VAF. Frameworks represent factual and normative levels and a level connecting the two, conclusions at one level forming premises at the next, offering some insight into the relation between factual and deontic notions in legal reasoning.

Finally, we have illustrated our model using a recent political debate involving the Government's justification of a proposed action. Our example models the viewpoint of the Government in putting forward their position on the issue and some attacks that this justification was subjected to by members of Parliament and the public. Again different viewpoints are modelled by different BDI agents. The distinctive feature of this application is that it requires the modelling of the extent to which the values involved are promoted.

3.   REFERENCES

[1] K. Atkinson, T. Bench-Capon, and P. McBurney. Attacks on a presumptive argument scheme in multi-agent systems: pre-conditions in terms of beliefs and desires. Technical Report ULCS-04-015, Department of Computer Science, University of Liverpool, UK, 2004.

[2] K. Atkinson, T. Bench-Capon, and P. McBurney. Justifying practical reasoning. In F. Grasso, C. Reed, and G. Carenini, editors, *Proc. CMNA 2004*, pp 87–90, Valencia, Spain, 2004.

[3] K. Atkinson, T. Bench-Capon, and P. McBurney. A dialogue game protocol for multi-agent argument for proposals over action. *Journal of Autonomous Agents and Multi-Agent Systems*, 2005. *In press.*

[4] T. Bench-Capon. Persuasion in practical argument using value based argumentation frameworks. *Journal of Logic and Computation*, 13 3:429–48, 2003.

[5] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357, 1995.

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