Decision-Theoretic Deliberation under Bounded Rationality

Guido Boella
Dipartimento di Informatica
Università di Torino
guido@di.unito.it

Joris Hulstijn
ICS
Utrecht University
jorish@cs.uu.nl

Leendert van der Torre
CWI
Amsterdam
torre@cwi.nl

1 The challenge of decision-theoretic deliberation

Agent theory proposes to model the behavior of complex software systems in terms of mental attitudes like belief, desires, goals, intentions and obligations, ranging from, e.g., the PRS system [7] to the more recent BOID architecture [3]. Decision-theoretic deliberation captures concepts and reasoning mechanisms from agent theory in standard decision-theoretic terms. This is an ambitious enterprise, as it has the objective to bridge the worlds of decision theory and agent theory. Thus far, several partial results on the decision-theoretic characterization have been obtained. The relation between beliefs (as well as defaults) and probabilistic techniques has been studied for some time, there are characterizations of desires and goals in decision-theoretic terms [6], there are various interpretations of obligations and norms, for example as social laws [8], and there are preliminary results on intention [1]. See our comparison paper [4] for an overview.

The most problematic issue in decision-theoretic deliberation is the characterization of intention. Roughly, whereas beliefs have been related to probabilities, desires to utilities, and obligations to social laws, intentions do not seem to have an obvious counterpart in classical game and decision theory. However, most discussions on the popular BDI model have focussed on the role of intention in deliberation [2]. Consequently, we believe that intention is the benchmark example of decision-theoretic deliberation.

The philosophical and logical literature explain why intention is more difficult to capture than other concepts. Intention has been related to choice and commitment, where choice can be interpreted as a decision-theoretic notion, but commitment seems as hard to characterize as intention itself. Moreover, intention has been discussed in the context of planning, and decision-theoretic planning and qualitative decision theory have only been developed recently. Finally, whereas beliefs, desires, and obligations can be characterized by acceptance conditions [10, 11], intention is also characterized by its persistence and reconsideration conditions. More precisely, commitment has been proposed by Bratman [2]'s analysis of intentions as the main component which distinguishes intentions from other motivational attitudes such as goals. The stability of commitment means two things: first, "intentions resist reconsideration and in that sense have inertia" [2, p.30]. Second, maintaining an intention without reconsidering it when the world changes is an irrational behavior. The rationale behind commitment, which is treated by Bratman as a primitive concept, is the bounded rationality of agents (Simon [9]): rationally bounded agents cannot afford continuous reconsideration and revision of their intentions. For similar reasons, [12] propose a meta-deliberation approach to reasoning under bounded rationality.

Our approach to decision-theoretic deliberation – including the characterization of intention – is based on a high level decision-theoretic model of deliberation in which the execution of normal actions and of deliberation actions can both be expressed. We start from Boella's [1] work on intentions in the context of a decision theoretic planner, called DRIPS [5]. This work implicitly contains a notion of intention. However, the definition is hidden in the algorithm executed by the planner. To get results concerning intentions, one would have to do extensive simulations, which are difficult to interpret. Therefore we present in this paper a transparent deliberation model with a decision theoretic interpretation which provides the basis of the study of intentions.

2 Deliberation Model

In defining the internal state of an agent we are inspired by the DRIPS planner. Intentions are defined by [2] as partial plans, and DRIPS is able to model partial plans with associated utility estimates. A state is composed of a set of beliefs B, a set of executable actions A, a set of partial plans P ordered in a plan hierarchy H, an agenda containing a set of commitments C towards goals and partial plans, and two partial functions U^+ and U^- that give for each commitment in C the lower and upper utility bounds in the current state. A set of goals G is defined indirectly as those partial plans that are not a refinement of another partial plan.

In the propositional variables of our language we distinguish decision variables from parameters. Decision variables A can be directly performed by an agent. Abstracting over the problem of the ordering on a sequence of decision variables, we represent a plan recipe as a set of propositional variables; a primitive plan recipe is composed of decision variables only. The plan hierarchy is an AND/OR tree, where partial plan recipes are represented by sets where some decision variables $\{x_1, ..., x_n\} \subseteq A$ have been substituted by a parameter $p \in P$ such that $(p, \{x_1, ..., x_n\}) \in H$.

Definition 1 (States) A state is a tuple $s = \langle B, A, P, H, C, U^+, U^-, G\rangle$, where $B = Lit(X)$ is a set of literals built from propositional variables X, A is a set of actions, P is a set of partial plans, the plan hierarchy $H \subseteq P \times 2^P$ is a relation between partial plans and sets of plans (recipes) that is acyclic, the set of commitments C is a subset of $2^{P \cup A}$, U^+ and U^- are partial functions from commitments C to real numbers such that for $c \in C$ $U^+(c) \geq U^-(c)$ and represent the time-discounted utility of the outcome of the best and worst primitive plan subsumed by c in the hierarchy H. If $(p, \{p_1, \ldots, p_n\}) \in H$ then doing all of $p_1, \ldots, p_n$ is a way to see to it that p. Since H is a relation, there may be several ways to decompose p. Expression $(p, \{x_1, \ldots, x_n\}) \in H^*$ denotes that $\{x_1, \ldots, x_n\}$ is a primitive plan for p, where $\{x_1, \ldots, x_n\} \subseteq A$ and H^* is the transitive closure of H. Goals G are partial plans that do not have a parent in the hierarchy H.

The decision process is as a Markov decision process whose deliberation actions change the state of the agent - in particular its commitments - or execute actions which affect the real world.

Definition 2 (Decision process) Let a decision process be a tuple $\langle S, DA, F, V\rangle$ where S is the set of states defined above, DA a set of deliberation actions, $F : S \times DA \rightarrow S$ a function that associates with every state s for every deliberation action a a resulting state s', and V a reward function describing the value assigned to a state.

The deliberation actions DA have access to the agent internal state $s \in S$ and modify it by revising its beliefs according to observations, its commitments, by introducing plans for achieving new goals, and by refining partial plans and selecting the more promising ones. Since the utility functions are partial, it is possible that, in order to compare plans, the agent has to compute the new utility of a plan in a new state. The deliberation actions are defined informally as follows for space reasons. Triggers specify when an action is meaningful, while preconditions specify what must be true in order for an action to achieve its postconditions.

1. Evaluation

Trigger: previous action is refinement or revision or observation; new utility estimates are needed given new beliefs B or commitments C.
Pre: agent committed to some plans whose utility has not been computed yet.
Post: upper and lower utilities of committed plans are re-estimated in current state.

2. Refinement & selection

Trigger: committed partial plans are too partial to make a decision.
Pre: agent committed to some partial plans.
Post: agent is committed to all non-suboptimal refinements of partial plans it is committed to.

5. Execution

Trigger: some action can be executed.
Pre: committed to a single primitive plan.
Post: some actions of the primitive plan are executed;
commitments reduced to remainder of the executed primitive plan.

6. Revision

Trigger:	new information, new goals or new plans have become available.
Pre:	-
Post:	new commitments to partial plans whose utility is not known are introduced.

7. Observation

Trigger:	the utility of some partial plan is too uncertain.
Pre:	there is at least a partial plan in the commitments.
Post:	beliefs about the state of the world are updated.

In order to construct a policy for using the Markov decision process, a reward function is defined. It is defined based on the information on the agent internal state, in particular the lower and upper utility bounds of the plans it is committed to. The agent must maximize its utility by executing plans, thus the policy drives the agent's planning activity towards the best solution, taking into account its bounded rationality: refining the current partial solution and computing the utility require time, while the utility of a plan is time discounted; so, the agent has bounded computational resources when it has to build a solution. Note that the deliberation actions of the decision process consider partial solutions which thus have an uncertain outcome, expressed by the lower and upper bound of a utility function. As suggested in [1], the reward function does not only consider the action with the least uncertain outcome, but also the ambiguity of the outcome. For this reason the reward function is defined not only as maximizing the outcome but also at minimizing the uncertainty of the outcome by means of refinement of partial plans.

3 Characterization of intention

In this paper, rather than specifying an algorithm for meta-deliberation as in [12], we model meta-deliberation as a decision process driven by the value of information provided by refining partial plans. Intentions can be analyzed as a side-effect of meta-reasoning on action and deliberation. In general, devising a new plan has to overcome overhead costs. Continuing the old plan is therefore more beneficial, when the initial utility differences are not very large. The stability, observed as the major function of intentions, therefore automatically comes out of a standard decision theoretic planning setting. Intention gets essentially a dynamic semantics: a semantics in terms of the changes to utilities.

References

[1] G. Boella. Intentions: choice first, commitment follows. In *Procs. of AAMAS'02*, pages 1165 – 1166. ACM Press, Bologna Italy, 2002.

[2] M.E. Bratman. *Intention, Plans, and Practical Reason.* Harvard University Press, 1987.

[3] J. Broersen, M. Dastani, J. Hulstijn, and L. van der Torre. Goal generation in the BOID architecture. *Cognitive Science Quarterly*, 2(3-4):428–447, 2002.

[4] M. Dastani, J. Hulstijn, and L. van der Torre. How to decide what to do? *European Journal of Operational Research*, 2003.

[5] P. Haddawy and S. Hanks. Utility models for goal-directed decisiontheoretic planners. *Computational Intelligence*, 14(3):392–429, 1998.

[6] A. Herzig, J. Lang and P. Marquis. Action representation and partially observable planning in epistemic logic. *Procs. of IJCAI'03*, 2003.

[7] A. Rao and M. P. Georgeff. An abstract architecture for rational agents. In *Procs. of KR'92*, pages 439–449, Cambridge, MA, 1992.

[8] Y. Shoham and M. Tennenholtz. On the emergence of social conventions: Modeling, analysis and simulations. *Artificial Intelligence*, 94(1–2):139–166, 1997.

[9] H.A. Simon. A behavioral model of rational choice. *Qualterly J. of Economics*, 69:99–118, 1955.

[10] L. van der Torre and Y. Tan. Rights, Duties and Commitments Between Agents. *Procs. of IJCAI'99*, pages 1239–1244, 1999.

[11] F. Veltman. Defaults in update semantics. *Journal of Philosophical Logic*, 25(3):221–261, 1996.

[12] S. Zilberstein and S. J. Russell. Anytime sensing, planning and action: A practical model for robot control. In *Procs. of IJCAI'93*, pages 1402–1407, Chambery, France, 1993.