Fuzziness in Decision-theoretic Planning

Jean-Marc Guinnebault
IRISA
Campus de Beaulieu
F-35042 Rennes Cedex
email: jguinneb@irisa.fr

Abstract

Imperfect knowledge of the world is one of the main issues AI systems have to deal with. In recent years there has been much interest in coping with such imperfections in planning systems. However, most works only emphasize one side of the problem, namely uncertainty. We argue that vagueness (or imprecision) is a fundamental issue which planning systems should also take into account . With this aim in view, we report a work in progress on the extension of a well-known decision-theoretic planning framework to fuzzy variables. We claim that, not only could this extension deal with imprecision and produce more robust reactive plans, but also, it could help to reduce the complexity of the original approach.

1 Introduction

Imperfection is a common issue which any human being can more or less easily deal with in his everyday life. On the contrary, mechanical agents encounter intractable difficulties to take it into account. Imperfection in our knowledge can be roughly divided into three main kinds [Bouchon, 1995]:

a) uncertainty, which refers to a doubt about the validity of statements. This can be caused by the unreliability of information.

b) imprecision (or vagueness), which refers to the difficulty of describing statements, either because numerical knowledge is not well defined, or because some vague terms are used.

c) incompleteness, which refers to the lack of knowledge or to partial knowledge about some features of the environment.

Those different kinds of imperfection are not independent of each other. Incompleteness implies uncertainty, some cases of imprecision can be associated with incompleteness and produce uncertainty and so on.

In recent years, there has been much interest in dealing with uncertainty in planning; See [Kushmerick *et al.*, 1995] for a probabilistic planner based on SNLP, [Dean *et al.*, 1995] for algorithms based on dynamic programming for Markov Decision Processes, and [Thiébaux *et al.*, 1995], [Hanks *et al.*, 1994] for other examples.

We argue that none of those planning systems explicitly take imprecision into account, although human beings and sensors can hardly ever provide us with a better world state description than a vague one. So as to deal with both uncertainty and imprecision, we therefore propose a fuzzy framework based on Markov Decision Processes. By considering unpredicted world states which are close to predicted ones, this framework leads to more robust reactive plans.

This paper report a work in progress in this direction, and is organized as follows. We first give a brief overview of decision-theoretic planning based on Markov Decision Processes (MDPs). We then extend this framework to fuzzy propositions in order to deal with imprecision. We finally discuss the benefits of this extension with respect to reducing the complexity of the original approach.

2 Markov Decision Processes

Following the work based on Markov Decision Processes [Howard, 1960], [Dean *et al.*, 1995], the entire environment is modeled as a stochastic automaton. A Markov decision process can be described [Dean *et al.*, 1995] as a tuple $\langle \mathcal{S}, \mathcal{A}, T, R \rangle$, where

- $\mathcal{S}$ is a finite set of states of the world;
- $\mathcal{A}$ is a finite set of actions;
- $T : \mathcal{S} \times \mathcal{A} \rightarrow \Pi(\mathcal{S})$ gives the transition model of the environment; we write $T(s_1, a, s_2)$ for the probability that the world will make a transition from state s_1 to state s_2 when action a is chosen.
- $R : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$ is the reward function specifying the instantaneous reward that the agent derives from being in each state.

A policy π is a mapping from $\mathcal{S}$ to $\mathcal{A}$, specifying an action to be taken in each situation. An environment combined with a policy for choosing actions yields a Markov chain.

Given a policy π and a reward function R, the value $V_\pi(s)$ of state $s \in S$ is the sum of the expected values of the reward to be received at each future time step, discounted by how far into the future they occur. That is, $V_\pi(s) = \sum_{t=0}^{\infty} \gamma^t E(R_t)$, where R_t is the reward received on the t^{th} step of executing policy π after starting in state s. The discounting factor, $0 \leq \gamma < 1$, controls the influence of rewards in the distant future. When $\gamma = 0$, the value of a state is determined entirely by rewards received on the next step; we are generally interested in problems with a longer horizon and set γ to be near 1.

A policy π dominates (is better than) π' iff, for all $s \in \mathcal{S}, V_\pi(s) \geq V_{\pi'}(s)$, and for at least one $s \in \mathcal{S}, V_\pi(s) > V_{\pi'}(s)$. A policy is optimal iff it is not dominated by any other policy.

Given a state transition model, a reward function and a value for γ, it is possible to compute the optimal policy using, for instance, the policy iteration algorithm or the value iteration algorithm.

3 Fuzzy States

In order to integrate imprecision into the previously described framework, we consider the world state description as being composed of a number of fuzzy linguistic variables [Zadeh, 1975]. The use of linguistic variables is to represent imprecise or vague knowledge on variables whose precise value might be unknown.

Definition 1 *[Zadeh, 1965] Let X be a set. A fuzzy subset A of X is characterized by a membership (characteristic) function $f_A(x)$ which associates with all $x \in X$ a real number in the interval $[0,1]$. The value of $f_A(x)$ is said to represent the "grade of membership" of x in A. A is normalized iff the highest value taken by its membership function is 1.*

Let us now define what a linguistic variable is:

Definition 2 *[Zadeh, 1975] A linguistic variable is a tuple (V, X, T_V), where*

- *V is a variable defined on set X,*
- *$T_V = \{A_1, A_2, \dots\}$ is a set of normalized fuzzy subsets of X that can be used to characterize the value of V.*

For instance a T_V for a linguistic variable representing size might be {huge, big, medium, small, tiny}. Some of the fuzzy subsets of X used in T_V may be ordinary subsets, e.g., singletons of X, which are particular cases of fuzzy subsets with membership function taking value in $\{0,1\}$.

We then define the set $\mathcal{S}$ of world states.

Definition 3 *Let $\{(V_1, X_1, T_{V_1}), \dots, (V_m, X_m, T_{V_m})\}$ be a finite set of linguistic variables. The set $\mathcal{S}$ of world states is the cross product of the T_{V_i}s. An individual (fuzzy) state is described by the vector $\langle v_1, \dots, v_m \rangle$, where each $v_i \in T_{V_i}$.*

As [Goguen, 1975] pointed out, not only can fuzzy logic help to deal with imprecision, but it can also help improving the robustness of computed plans, that is the ability to respond without plan modification to slightly perturbed, or to somewhat inexactly specified situations. In other words, what we want here is to be able to react to unpredicted situations close to predicted ones. With this aim in view, we define the notion of closeness between fuzzy states.

Definition 4 *Let $S_v = \langle v_1, \ldots, v_m \rangle$ be a fuzzy state. We say that the state $S_x = \langle x_1, \ldots, x_m \rangle$ is closer to S_v than $S_y = \langle y_1, \ldots, y_m \rangle$ is, iff*

$$c(S_v, S_x) > c(S_v, S_y)$$

where

- $c(\langle v_1, \ldots, v_m \rangle, \langle x_1, \ldots, x_m \rangle) = g(f_{v_1}(x_1), \ldots, f_{v_m}(x_m))$;
- *f_{v_i} is the membership function (extended to fuzzy values) of fuzzy value v_i of linguistic variable V_i,*
- *g is a conjunction operator.*

Rather than the *min*, we might choose the sum as conjunction operator g.

So when we are facing an unpredicted state, we still have the possibility of searching for the closest predicted state and then of using the action prescribed in this state. Although this action might not perfectly suit the actual situation, we can expect that it will be better than doing nothing at all or executing an arbitrary action.

One might argue that all this can be done without any fuzziness by simply considering closeness between ordinary states. This is perfectly true. Nonetheless, we argue that ordinary descriptions of environments are much too committed and therefore choosing the closest state in such a situation might be nothing but meaningless. We claim that we need vagueness in the state descriptions so that closeness be meaningful and exploitable.

4 Some Hints at Fuzzy Abstraction

Classical algorithms for solving MDPs [Howard, 1960] represent the state transition function as a matrix and the reward function as a vector, and find policies by enumerating and exploring the entire state space. This approach works well in small state spaces, but can very quickly become too complex. [Dean *et al.*, 1995] have addressed this problem by focusing the planner's attention to judiciously chosen larger and larger subsets of the state space. However, for very large state spaces, this approach, as is, is still insufficient: unless the subsets considered be very large, it is likely that we will encounter a state outside these subsets, for which we have no meaningful action. A complementary solution is to exploit internal regularities that allow state spaces to be represented more compactly.

We therefore propose to make use of the fuzziness in our state description in order to build equivalence classes of states. This aggregation is based, not on abstracting variables themselves, but their fuzzy values. By building fuzzy states, we expect to gain in tractability, at the expense of accuracy.

This abstraction process works as follows. Let us assume that we are provided with an abstraction schema for each linguistic variable $V = (V, X, T_V)$ where $T_V = \{A_1, A_2, \ldots, A_m\}$. This schema can be given as a hierarchy of T_V sets, in which an A_i at a given level results from the aggregation of the closest A_i values at the level below. For instance, dark-grey may be aggregated with light-grey to make grey, which might itself be eventually, at an upper abstraction level, aggregated with black to form dark, and so on.

Now, let s_x and s_y be two states that differ only on one linguistic variable, say V. We may aggregate the two states s_x and s_y into a single state whose linguistic variable V has value A_i, which corresponds to the abstraction level we need to aggregate both variables. This process may, of course, go on iteratively so far as to provide us with very vague state description.

We then propose a simple anytime algorithm that computes a policy for the given problem.

1. generate the fuzziest possible abstraction of the initial description

2. while (not deadline) do

- generate a policy for this description
- refine the description by backtracking to a less fuzzy abstraction

We are currently investigating the problem of the computation of the abstract state transition model T.

5 Related Work

Research on introducing fuzziness into planning is rare. Ours has mainly be motivated by Goguen's paper [Goguen, 1975] on fuzzy robot planning, which is a first attempt to lay foundations for applying fuzziness to highly practical problems of control and communication. More recently, [Saffiotti *et al.*, 1995], have proposed a multi-valued logic approach to integrating planning and control. A classical regression planner is used to generate complex fuzzy controllers from user-provided behaviors schemas. We intend to go one step further, by automatically generating fuzzy policies.

For long [Sacerdoti, 1974], it has been suggested that abstraction is a useful concept for reducing the complexity of real-world problems. Recently, [Nicholson *et al.*, 1994] have proposed an interesting abstraction-based framework for coping with very large stochastic domains. Abstract world views (derivations from an MDP which are MDPs themselves, and which consider only a subset of the possible attributes in the complete world model) are constructed by projecting out some of the dimensions of world state description, so that each state in an abstract view stand for an equivalence class of states in the original world model (See also [Boutilier *et al.*, 1994]). This approach is different from the one we propose, since it deals with variable abstraction, while we suggest making use of abstraction on fuzzy values. We think that our framework might lead to more relevant policies, since we prefer loosing accuracy on variable values, instead of completely omitting these variables.

References

[Bouchon, 1995] B. Bouchon-Meunier La logique floue et ses applications, Editions Addison-Wesley France, 1995.

[Boutilier *et al.*, 1994] C. Boutilier and R. Dearden. Using Abstractions for Decision-Theoritic Planning with Time Constraints, In *Proceedings AAAI (1994)*, pp. 1016–1022.

[Dean *et al.*, 1995] T. Dean, L. P. Kaelbling, J. Kirman and A. Nicholson. Planning under Time Constraints in Stochastic Domains, In *Artificial Intelligence 76 (1995)*, pp. 35-74.

[Goguen, 1975] J. A. Goguen. On Fuzzy Robot Planning, In *Proceedings Fuzzy Sets and their Applications to Cognitive and Decision Processes*, Zadeh, Fu, Tanaka and Shimura Ed., Academic Press (1975), pp. 429-447.

[Hanks *et al.*, 1994] S. Hanks, S. Russel and M. P. Wellman (ed.). Proceedings AAAI Spring Symposium on Decision-theoretic Planning, Stanford University, AAAI Press, 1994.

[Howard, 1960] R. A. Howard. Dynamic Programming and Markov Processes, The MIT Press, Cambridge, Massachusetts (1960).

[Kushmerick *et al.*, 1995] N. Kushmerick, S. Hanks and D. S. Weld. An Algorithm for Probabilistic Planning, In *Artificial Intelligence 76 (1995)*, pp. 239-286.

[Nicholson *et al.*, 1994] A. E. Nicholson and L. P. Kaelbling. Towards Approximate Planning in Very Large Stochastic Domains, In *Proceedings AAAI Spring Symposium on Decision Theoretic Planning (1994)*.

[Sacerdoti, 1974] E. D. Sacerdoti. Planning in a Hierarchy of Abstraction Spaces, In *Artificial Intelligence 5 (1974)*, pp. 115-135.

[Saffiotti *et al.*, 1995] A. Saffiotti, K. Konolige and E. H. Ruspini. A Multivalued Logic Approach to Integrating Planning and Control, In *Artificial Intelligence 76 (1995)*, pp. 481-526.

[Thiébaux *et al.*, 1995] S. Thiébaux, J. Hertzberg, W. Shoaff and M. Schneider. A Stochastic Model of Actions and Plans for Anytime Planning under Uncertainty, In *Int. J. Intelligent Systems*, vol. 10, (2), 1995.

[Zadeh, 1965] L. A. Zadeh. Fuzzy Sets, In *Information and Control (8)* (1965), pp. 338-353.

[Zadeh, 1975] L. A. Zadeh. The Concept of a Linguistic Variable and its Application to Approximate Reasoning, parts 1 and 2, In *Information Sciences (8)*, pp. 199-249, 301-357.