Interval-based Markov Decision Processes for Planning Interactions in Multi-Agent Systems *

Graçaliz Pereira Dimuro, Antônio Carlos da Rocha Costa
Escola de Informática
Félix da Cunha 412, 96010-000 Pelotas, Brazil
e-mail: {liz,costa}@ucpel.tche.br

Autonomous agents of multi-agent systems [5] often need to cooperate with other agents in order to achieve their goals. Thus, the capacity to have representations of the others agents' behaviors and minds, and to reason about the interactions it has with other agents, can be seen as one of the most important requirements for a social agent. Different approaches to the study of such social reasoning, have lead to various studies of evaluations and values assigned to other agents and interactions. The approach we adopt in our work derives from the initial work introduced in [3, 4], that takes Jean Piaget's theory of social exchanges and qualitative exchange values [2] as a basis. It introduced a system consisting of an algebra of qualitative exchange values, a social reasoning mechanism based on that algebra, and the specification of structures for storage and manipulation of such values. Here, we propose to model the exchange qualitative values as intervals, and thus, to have have an interval algebraic structure regulating the exchange of values in the agents' social interactions. Inspired by [1], we apply techniques from Operations Research to bear on the problem of choosing optimal interaction policies in (possibly, partially observable) stochastic domains by means of Interval-based Markov Decision Processes (IMDP).

An IMDP is defined as a tuple $\langle S, I, T, E, R\rangle$, where:

(a) $S \in \mathbb{IR}^n$ is a finite set of states of values, denoting the ways an agent can evaluate a given social exchange, and represented as tuples of interval exchange values;

(b) I is a finite set of possible exchange steps;

(c) $T : S \times I \to \Pi(S)$ is the state-transition function, giving for each state of values and exchange step, a (possibly imprecise) probability distribution over agent's states of values (we write $T(s, i, s')$ for the probability of ending in state s', given that it is in the state s and performs exchange step i);

(d) $E : S \to \mathbb{IR}$ is the state evaluation function, that evaluates the state of values s (we write $E(s)$ for the evaluation of the agent's state of values s);

(e) $R_E : (S \times I) \to \mathbb{IR}$ is the interval reward function, giving the expected immediate reward gained by the agent after it performed exchange step i in state s, considering a given evaluation function E (we write $R_E(s, i)$ for the expected reward for performing exchange step i in state s).

We observe that in the IMDP there is never any uncertainty about the current state of values – the agent has total knowledge of its exchange values and of the exchange values of the other agents, in each moment. However, a partially observable Markov decision processes (POMDP) may be considered, if the agent is not able to make complete observations of the state of values (see,p.ex., [1]). In both cases, the agent's goal is assumed to be to maintain a given preferred balance of values, by means of maximizing the future reward.

In this paper we outline the IMDP approach, considering the special case of just two agents interacting, and in which the transition probabilities are not imprecise and the reward functions is real valued, $R_E : (S \times I) \to \mathbb{R}$.

References

[1] Kaelbling, L. P., et al. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, **101**, 1998. pp. 99–134.

[2] Piaget, J. *Socialogical Studies*, Routlege, London, 1995.

[3] Rodrigues, M. R., Costa, A. C. R., and Bordini, R.. A System of Exchange Values to Support Social Interactions in Artificial Societes. In *Proc. 2nd. Int. Conf. on Autonomous Agnets and Multiagents Systems*, Melbourne, 2003.

[4] Rodrigues, M. R., and Costa, A. C. R.. Using Qualitative Exchange Values to Improve the Modelling of Social Interactions. In Procedings of 4th Workshop on Agent Based Simulations, Melbourne, *Lecture Notes in Artificial Intelligence*, 2003.

[5] Wooldridge, M. *Introduction to Multi-agent Systems*, Wiley, Chichester, 2002.

*Partially supported by CTINFO/CNPq and FAPERGS.