Behavior-Based Control System in MultiAgent Domain

Stéphane Calderoni
IREMIA - Université de la Réunion
15, Avenue René Cassin - BP 7151
97715 Saint-Denis messag cedex 9 - France
e-mail: calde@univ-reunion.fr
tel/fax: (+262) 93 82 [87/60]

1 The framework

This short contribution presents a classifier-based control system for adaptive behavior synthesis in multi-agent domain, regulated by a reinforcement learning algorithm. This control system is defined by a set of behavioral rules, a credit assignment system and a genetic algorithm. The genetic algorithm we use is common to the most of algorithms presented in the evolutionary litterature, consequently, we chose to limit this report to the description of the chromosome structure and the credit assignment system.

2 The chromosome structure

Each rule R_i is a coupling of a predicate P_i and a task T_i in such a way that, if P_i is *true*, then T_i is considered as applicable. P_i is a multivariable boolean function that combines the stimuli perceived from the environment in a tree where nodes are boolean operators and leaves are stimuli. In such a way, complex environmental situations may be expressed, even from a reduced set of recognizable stimuli. When stimuli are sensed from the environment, they are propagated through the boolean tree, from the leaves, and filtered by the boolean operators so as to be collected as a single boolean value in P_i. This value then makes the corresponding task T_i applicable or not. Thus, P_i. constitutes a necessary precondition for the application of the rule R_i. As for T_i, it is defined as an ordered sequence of basis actions, which constitutes the basis skills of the agent, namely its behavioral primitives. The association of P_i and T_i thus forms the chromosome R_i.

3 The credit assignment system

Actually, each behavioral rule R_i is provided with both an activation level α_i and an inhibitory threshold σ_i to determine its applicability. Thus the task T_i is considered as activable if and only if $\alpha_i > \sigma_i$.

More generally, each rule R_i should be supplied with an additional set of features in order to provide the agents with the ability of estimating their performance by themselves, from satisfaction or disapointment indicators. Consequently, a rule R_i can be defined as following: $R_i = [P_i \longrightarrow T_i, \alpha_i, \sigma_i, \pi_i], i \in I\!N$. α_i is the activation level of T_i, σ_i is the inhibitory threshold of T_i, and π_i is a set of the following form: $\pi_i = \{\eta_j : \phi_j \longrightarrow (r_j^+, r_j^-)\}$. η_j is a rule of reinforcement associated to the task T_i, ϕ_j is a predicate that indicates to the control system when to apply either the positive reinforcement r_j^+ (when ϕ_j is $true$) or the negative one r_j^- (when ϕ_j is $false$). Then we can define the local reinforcement function z_i associated with the task T_i as following: $z_i(t) = \sum_{j=1}^{j=n_i} r_j^+(t) + r_j^-(t)$ where n_i is the cardinal of π_i. Consequently the global reinforcement function Z_i of the task T_i is defined, at the date θ, by: $Z_i(\theta) = \sum_{t=0}^{t=\theta} z_i(t)$

4 Implementation

The control system which has been presented in this short report has been implemented as an extension of the MUTANT platform described in [Calderoni and Marcenac, 1998].

References

[Calderoni and Marcenac, 1998] Calderoni, S. and Marcenac, P. (1998). Mutant: a multiagent toolkit for artificial life simulation. In Singh, M., Meyer, B., Gil, J., and Mitchell, R., editors, *Proceedings of the 26^{th} International Conference on Technology of Object-Oriented Languages and Systems (TOOLS-USA-98)*, pages 218–229, Santa Barbara, USA. IEEE Computer Society Press.