ESANN'2000 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 26-28 April 2000, D-Facto public., ISBN 2-930307-00-5, pp. 103-108

Learning of perceptual states in the design of an adaptive wall-following behaviour

R. Iglesias, M. Fernández-Delgado and S. Barro

University of Santiago de Compostela (Spain)
Dept. of Electronics and Computer Science
{rober, delga, senen}@dec.usc.es

Abstract. In this work we propose a new model that aims to overcome some of the limitations that are associated with reinforcement learning. In order to do so, we include not only prior knowledge of the task to be undertaken, by means of the so-called Supervised Reinforcement Learning (SRL), but also the creation and adaptation of the perceptual states of the environment during the learning process itself, by means of the MART neural network.We have tested the application of this model on a wall following behaviour. The results obtained confirm the great usefulness and advantages that are derived from its employment.

1. Introduction

The use of reinforcement learning (RL) in complex or uncertain domains is becoming ever more important, insofar as RL systems are adaptive and self-improving. Nevertheless, there are certain difficulties in the application of RL, such as, for example, the trade-off between the exploration of the actions that can be executed and the convenience of exploiting what the system already knows. A further problem is that associated with the representation of states: RL assumes that the knowledge of the state in which the system finds itself is enough to determine the action to be executed and to predict future behaviour, which is not always the case.

In this work, we propose a new learning model, which we call MART-SRL, with which we will aim to improve the performance of RL by minimizing the influence of the aforementioned two problems. In order to do so, our model unites Supervised Reinforcement Learning (SRL) [3], which is capable of combining RL with the knowledge available on the task to be resolved or the specifications of sub-objectives of interest, with the neural computation paradigm Multichannel Adaptive Resonance Theory (MART) [2] , which we have adapted to allow the adequate construction and dynamic adaptation of perceptual states.

ESANN'2000 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 26-28 April 2000, D-Facto public., ISBN 2-930307-00-5, pp. 103-108

2. The MART-SRL model

As has already been mentioned, with this model the aim is to improve the performance of RL by obtaining more robust behavior, as well as faster and more stable learning. In order to achieve this, we thought useful to consider during the learning process all the information that is available on the task to be carried out, as well as the specifications about the sub-objectives of interest, with the aim of focussing the exploration, or even imposing control decisions in risk situations. On the other hand, the observation of those situations in which the system's behaviour is unstable will enable us to adequately create and update the representative internal states of the environment, which will henceforth be referred to as "perceptual states". This double objective can be seen through the presence of two clearly differentiated blocks, SRL and MART, shown in figure 1.

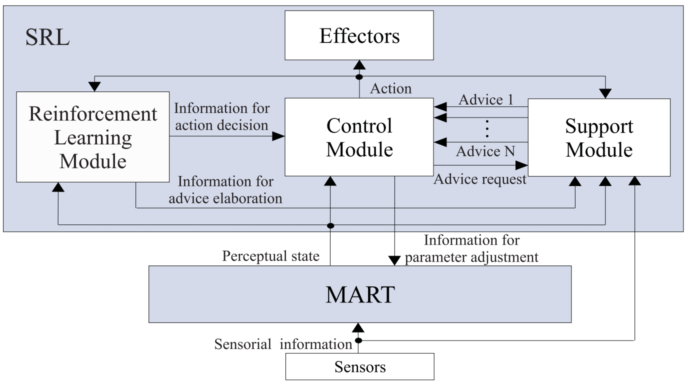

Figure 1: Structure of the model MART-SRL. MART handles recognition, learning and adaptation of perceptual states in the operational environment. In order to do this, it uses both sensorial information and feed-back from SRL.

The structure and *modus operandi* of the system with supervised reinforcement learning (SRL) appear in a detail in [3], due to which, in the present work, we will refer only to certain basic aspects. As can be seen in figure 1, three modules can be discerned within the SRL: the Support Module (SM), the Reinforcement Learning Module (RLM) and the Control Module (CM). The RLM includes a reinforcement learning algorithm whose objective is to determine, in each state, the utility associated with the execution of each one of the possible actions. The SM is made up of several blocks that contain knowledge on the task to be carried out. The objective of these blocks is to indicate to the CM, through pieces of advice, which are the actions whose execution would seem to be the most recommendable. Finally, the CM will assess the information provided by this advice, and will combine it with the information coming from the RLM in order to decide, at each instant, what action needs to be executed.

ESANN'2000 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 26-28 April 2000, D-Facto public., ISBN 2-930307-00-5, pp. 103-108

Each block in the SRL may work with its own representation of states. However, we are primarily concerned with the states that are associated with the RLM, since it is here where useful information is acquired as the system explores the possible actions to be carried out. In learning processes, perceptual states are often defined *a priori*, and in this way the system only has to learn the action associated with each state. This supposes starting from strong conditioning factors and may provoke that a state includes situations which require significantly different actions; this would hinder the convergence of the information stored in the RLM to a satisfactory action policy. In order to mitigate this problem, our proposal includes a neural computational block, MART, which is capable of handling the construction and dynamic adaptation of the different perceptual states that are of interest for a specific application.

3. Representation of states in MART-SRL

In this section we describe how MART constructs the representation of perceptual states when it operates in a certain environment. As a detailed explanation of MART appears in [2], in this section we will limit ourselves to dealing with the modifications that have been introduced in order to work in the SRL setting. MART constructs the representation of perceptual states through a process of categorization, which assigns a class to each sensorial pattern of the environment; this class acts as a perceptual state of reference for this pattern. This categorization is carried out based on information coming from various signal channels or groups of sensors, in such a manner that to the advantages derived from simultaneously considering multiple sources of information, one has to add the information that is supplied by a local assessment of each channel and its subsequent integration or global evaluation.

The operation of MART consists of t wo stages: classification and analysis. The former determines which of the learned classes is the one that best identifies the current environment. This class will enable the SRL to decide on the action to be carried out and, later, to observe the results obtained. From the sensorial input pattern and the information fed back from the SRL (figure 1), in the analysis stage the parameters associated to the classes that best identify the sensorial pattern are updated, and, if it is deemed necessary, a new class can be created to represent it henceforth. Between the classification and analysis stages it is possible that there will be a time delay, which responds to the need to observe what happens in the immediate future, with the aim of extracting conclusions and updating that which has been learned. We now go on to describe both stages in some detail.

Classification stage. As has previously been mentioned, MART generalizes the behaviour of ART networks [1] contemplating the possibility that the sensorial input pattern may be obtained from multiple information channels (I denotes the number of these channels). In the same way, each perceptual state, or class that MART learns, will have a projection for each one of these channels. As the analysis of the environment progresses, the channel-class credits

ESANN'2000 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 26-28 April 2000, D-Facto public., ISBN 2-930307-00-5, pp. 103-108

x_{ik} $(i = 1, \ldots, I)$ will reflect which are those channels that supply the most valuable information on each one of the classes that have been learned. These credits are adjusted dynamically according to the sensorial patterns that are assigned to each class. On the basis of this information it is possible to obtain a measurement of similarity s_k between the input pattern and any class k. MART will identify a sensorial input pattern by its two most similar classes k_1 and k_2. In order to determine which of these classes corresponds to the perceptual state, the class credit c_k is taken into account, by means of which the SRL assesses to what degree the actions associated to this class are satisfactory. A vigilance ρ_{ij} is also considered; this measures the relative size of class i with regard to class j. From the quantities $p_{k_1} = f(c_{k_1}, s_{k_1}, \rho_{k_1 k_2})$ and $p_{k_2} = f(c_{k_2}, s_{k_2}, \rho_{k_2 k_1})$, MART selects as being representative of the environment the class with the highest p_k, which is the one that is passed on to the SRL. If $p_k > 1$ it will be said that class k reaches resonance with the input pattern.

Analysis stage. In this stage the results of the previous one are analyzed, evaluating to what degree the perceptual state selected in that stage has been suitable. The result of this analysis will serve to update the perceptual states constructed by MART, thus improving its perception of the environment. Due to the changes in the information learned by MART, which are realised in the time elapsed between the pattern classification and the pattern analysis stages, it is necessary to reassess the similarities s_k between the classes and the sensorial input pattern, advising the SRL on the two most similar classes k_1 and k_2, and calculating p_{k_1} and p_{k_2}. The conclusions obtained by the SRL on the basis of the behaviour observed in the system will serve to indicate to MART whether resonance with one of these two classes seems to be suitable, in which case the parameters associated to it (projections, channel-class credits, vigilances) will be modified in order to contribute to facilitating this resonance. If, however, none of these classes seems to be acceptable, their parameters are modified with the aim of tending to avoid resonance. It may transpire that, once these changes have been effected, resonance is not reached with any of them $(p_{k_1}, p_{k_2} < 1)$, in which case a new class will be created which, henceforth, will represent the input pattern. In this manner, MART-SRL is able to construct a perceptual representation of the environment in which it is located, detecting the appearance of new situations, learning them, remembering them each time that they appear again and updating them as they develop.

4. Example of application

With the aim of illustrating the possibilities of the MART-SRL model, in this section we will demonstrate its application to the learning of a wall following behaviour in the setting of mobile robotics. The robot that we use is a Nomad200. Apart from other sensors, this robot has 16 evenly-spaced ultrasound sensors that encircle its upper part. In order to follow the wall we only use the 9 sensors positioned frontally and laterally on the robot, on the side closest

ESANN'2000 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 26-28 April 2000, D-Facto public., ISBN 2-930307-00-5, pp. 103-108

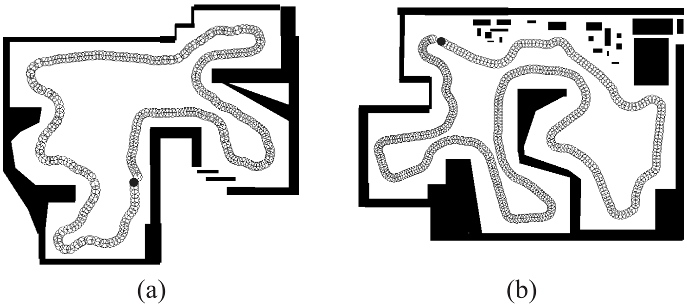

Figure 2: Behaviour learned by the robot in two environments (a) and (b).

to the wall. The information supplied by the sonar sensors is preprocessed before feeding it into MART, obtaining a 7-component vector, which codifies and abstracts the information from the environment into a set of segments according to their orientation and their distance with respect to the robot [3]. The components of this vector $\overrightarrow{V} = (V_1, ...V_7)$ $(V_i \in \{1, 2, \ldots, 15\}, 1 \leq i \leq 7)$ are distributed into 5 groups or channels: thus, the sensorial pattern in the i-th channel is (V_i, V_{i+1}, V_{i+2}).

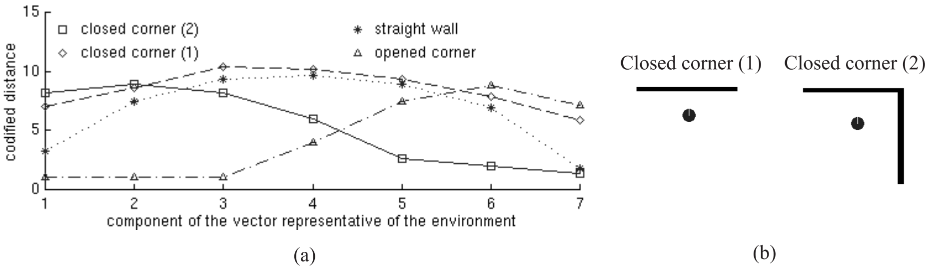

Figure 3: Panel (a) shows some perceptual states learned for the environment of figure 2b for different situacions, some of them shown in (b). Each state is represented using a vector in which each component codifies the distance to the nearest wall segment with a particular orientation, from the frontal wall segments (component 1) to the back ones (component 7). The greater the value of the component, the closer it is to the wall.

Given that the usefulness of SRL to reduce the instability that is associated with the exploration of the possible actions for each state has already been highlighted in other works [3], through these results we will emphasize the capacity of our model for conveniently generating the perceptual states. Starting always from an initial situation in which the action that should be carried out at each instant and the suitable representation of perceptual states are unknown, we have applied our model for the learning of the proposed task in different environments. The results obtained were satisfactory, as can be

ESANN'2000 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 26-28 April 2000, D-Facto public., ISBN 2-930307-00-5, pp. 103-108

seen in figures 2a and 2b. In both environments the final number of perceptual states generated was 13, underlining that both include habitual situations in this t ypeof problems and the behaviour should be valid, even for other environments. Finally, in figure 3 we include an example that illustrates some of the states generated. More specifically, we can see (figure 3a) the representation corresponding to 4 of the classes learned, associated to clearly differentiated situations in the environment. Two of these classes identify a "closed corner" situation, and will be activated when the robot "sees" a close frontal wall (figure 3b), with and without a nearby lateral wall. The other t wo states identify an open corner and a straight wall.

5. Conclusions

In this paper we have presented the utility and advantages derived from the use of a new model, which we call MART-SRL, and which is capable of increasing the performance of basic RL. On one hand we succeed in combining the information learned by the RL with that coming from prior knowledge of the task or sub-objectives of interest. Furthermore, we succeed in constructing, in an adaptive manner and in real time, the representation of perceptual states associated to the environment in which the task has to be learned. We have applied MART-SRL to the learning of wall-following behaviour in mobile robotics. The potential and utility of MART-SRL have been demonstrated by the satisfactory learning of the behaviour pattern without user intervention, starting from an initial state of total ignorance with regard to both the actions to be carried out at every instant and the suitable representation of the perceptual states in the operational environment.

Acknowledgements
This work has been possible thanks to the project XUGA20608B97, and the availability of a Nomad200 mobile robot acquired through an infrastructure project, both funded by the Xunta de Galicia.

References

[1] Carpenter G.A. and Grossberg S. ART2: Self-organizing of stable category recognition codes for analog input patters. *Applied Optics*, Vol. 16, No. 23, December, 1987.

[2] Fernández-Delgado M. and Barro S., "MART: A multichannel ART-based neural network.", *IEEE Tr ansactions on Neural Networks*, Vol. 9, No. 1, pp. 139-150, January 1998.

[3] Iglesias R., Barro S., Regueiro C.V., Correa J. and RodríguezM., "Supervised Reinforcement Learning: A new approach for behaviour learning in mobile robotics", sent to the journal *Robotica*, Cambridge University Press.