Chapter 8

Adaptive cognitive systems

Timo Honkela, Krista Lagus, Ville Könönen, Ann Russell, Mikaela Klami, Tiina Lindh-Knuutila, Matti Pöllä, Jaakko Väyrynen, Kevin Hynnä

8.1   Introduction

Our research on cognitive systems focuses on modeling and applying methods of unsupervised and reinforcement learning. The general aim is to provide a methodological framework for theories of conceptual development, symbol grounding, communication among autonomous agents, agent modeling, and constructive learning. We also work in close collaboration with other groups in our laboratory, e.g., related to multimodal environments.

An important part of our acitivity has been the active role in organizing international scientific events:

- International and Interdisciplinary Conference on Adaptive Knowledge Representation and Reasoning (AKRR'05) [2],

- Symposium on Adaptive Models of Knowledge, Language and Cognition (AMKLC'05) [12], organized in conjunction with AKRR'05 conference, and

- Workshop on Reinforcement Learning in Non-Stationary Environments in conjunction with the 16th European Conference on Machine Learning (ECML'2005) [9].

International and Interdisciplinary Conference on Adaptive Knowledge Representation and Reasoning, AKRR'05, was raising awareness of adaptive approaches to knowledge representation and reasoning. The power of the adaptive systems lies in the fact that they enable computers to adapt to the needs of individuals, groups, enterprises and organizations in the changing world. There were two special symposia in the conference that provide a focused view on their topics: Adaptive Models of Knowledge, Language and Cognition (AMKLC'05) and Knowledge Representation for Bioinformatics (KRBIO'05). The goal of the reinforcement learning workshop in ECML'2005 was to foster co-operation inside European reinforcement learning research community and raise international visibility of European reinforcement learning research.

8.2 Emergence of cognitive and conceptual representations

Conceptual modeling is a task which has traditionally been conducted manually. In artificial intelligence, knowledge engineers have written descriptions of various domains using formalisms based on predicate logic and other symbolic representations such as semantic networks and rule-based systems. As modern related topics, the Semantic Web and knowledge representation formalisms like eXtendable Markup Language (XML) can be mentioned.

Philosophical considerations

The traditional symbolic approach has concentrated on the linguistic domain. Therefore, the models often lack the connection to the perceptual domain. It has been assumed that knowledge can be represented as propositional structures that are based on static shared concepts. It has been commonplace to assume that there is a rather one-to-one correspondence between words and concepts. Moreover, it is assumed that a concept refers unambiguously to a number of distinct objects or events in the reality. The individual differences are assumed to be small and explained as errors.

In radical constructivism (consider, e.g., [8, 15]), it is pointed out that cognitive agents construct their description of the world, and this description consists of constructed categories such as objects and events along with their associated subcategories. Each of those constructions is subjective but at the same time their formation is based on the interaction with other agents as well as artefacts that reflect the structural characteristics of the constructions of other agents. It should not be be taken as a fact that only the rules or principles observed in the past shall apply to the future. Constructive learning involves qualitative restructuring and modification of internal knowledge representations, rather than just accumulation of new information in memory. [5]

The static nature of the information systems in general makes them also prone to be "incompatible with the reality". One reason is that the domain of use is changing. Another, more profound reason is that human beings have individual conceptual systems, gained through constructive learning processes. A conceptually static and coarse-grained information system matches with our conceptual systems only partially. This misfit may lead to errors or unjustified procedures. Therefore, it appears necessary that any information system should be adaptive in order to be able to deal with the variety of conceptual construction and in order to be able to conduct meaning negotiations.[3]

Emergence of a shared conceptual system

We studied the emergence of associations between concepts and words. The important questions being how a language learner, or an agent, learns the meaning of new words, and how an agreement on the use of words is reached in a community of agents. The Self-Organizing Map (SOM) were used as a model of an agent's conceptual map, and concepts are seen as areas formed in a SOM based on unsupervised learning. The map may be seen to be an equivalent of a domain in a Conceptual Space[1]. The language acquisition process was modeled in a population of simulated agents by using a series of language games, called the observational games. For the experiments, an agent simulation framework was implemented and tested with different parameters. The results of the experiments verify that the agents learn to communicate successfully and a shared lexicon emerges. [6]

Emergence of word features using ICA

We have studied the emergence of linguistic representations through the analysis of words in contexts using the Independent Component Analysis (ICA). The ICA learns features automatically in an unsupervised manner. Several features for a word may exist, and the ICA gives the explicit values of each feature for each word. In our experiments, we have shown that the features coincide with known syntactic and semantic categories. More detailed description of this research is given in the section on Natural Language Processing in this report.

Similarity of emergent representations

According to the connectionist view, mental states consist of activations of neural units in a connectionist network. We consider the similarity of representations that emerge in an unsupervised, self-organization process of neural lattices when exposed to color spectrum stimuli. Self-Organizing Maps (SOM) are trained with color spectrum input, using various vectorial encodings for representation of the input. Further, the SOM is used as a heteroassociative mapping to associate color spectrum with color names. Recall of association between the spectra and colors is assessed, and it is shown that the SOM learns representations for both stimuli and color names, and is able to associate them successfully. The resulting organization is compared through correlation of the activation patterns of the neural maps when responding to color spectrum stimuli. Experiments show that the emerged representations for stimuli are similar with respect to the partitioning-of-activation-space measure almost independently of the encoding used for input representation. This adds new evidence in favour of the usability of the state space semantics.[11]

Self-refreshing SOM as semantic memory

Natural and artificial cognitive systems suffer from forgetting information. However, in natural systems forgetting is typically gradual whereas in artificial systems forgetting is catastrophic. Methods based on rehearsal and pseudorehearsal have been successfully applied in feedforward networks to avoid catastrophic interference. A novel method based on pseudorehearsal for avoiding catastrophic forgetting in the Self-Organizing Map (SOM) is presented. Simulations results show that the use of pseudorehearsal can effectively decrease catastrophic forgetting. [10]

Simulated emotions in a SOM-based agent model

It is assumed that emotions in a cognitive system have a role in interlinking organism's cognition, needs, goals, motivation and final output behavior. For the purpose of emotion modeling, an earlier SOM-based agent simulation model [4] was simplified in many ways. Proposed emotional model might be classified as a cognition appraisal theory inspired by model considering emotions as emergent labels for the evaluation of prototypical situation or events (modal emotions) rather than basic discrete entities achieved by a response program.[14]

Analysis of interprofessional collaboration

The Self-Organizing Map was used to analyze the online collaborative discourses of an interprofessional team of hospital workers in Toronto area engaged in an 18-month reflective practice and continuous learning project. Preliminary results [13] demonstrate unique

characteristics of the participant group's interactivity that would otherwise remain unidentified using conventional quantitative methods of discourse analysis. The SOM analysis generated a relational profile of participants' reading and linking activity in an online learning environment that not only captures the emergent dynamics of interprofessional collaboration over time, but also highlights individual differences within and between professional groups.

References

[1] Gärdenfors, P. *Conceptual Spaces.* MIT Press, 2000.

[2] Honkela, T.; Könönen, V.; Pöllä, M.; Simula, O. (eds.) *Proceedings of the International and Interdisciplinary Conference on Adaptive Knowledge Representation and Reasoning (AKRR'05).* Espoo, Finland, June 15-17, 2005. Espoo, Finland 2005. 174 p.

[3] Honkela, T. Von Foerster meets Kohonen - Approaches to Artificial Intelligence, Cognitive Science and Information Systems Development. *Kybernetes*, 34(1/2), 2005.

[4] Honkela, T. and Winter, J. *Simulating Language Learning in Community of Agents Using Self-Organizing Maps.* Helsinki University of Technology, Publications in Computer and Information Science Report, 2003.

[5] Honkela, T.; Hynnä, K.; Lagus, K.; Särelä, J. (eds.) *Adaptive and Statistical Approaches in Conceptual Modeling.* Espoo, Finland: Helsinki University of Technology, 2005. (Publications in Computer and Information Science Technical Report A75).

[6] Lindh-Knuutila, T. *Simulating the Emergence of a Shared Conceptual System in a Multi-Agent Environment.* Master's Thesis. Helsinki University of Technology, Department of Electrical and Communications Engineering, Espoo, Finland. 2005.

[7] Manning, C.D. and Schütze, H. *Statistical Natural Language Processing.* MIT Press, Cambridge, Massachusetts, 1999.

[8] Maturana, H. R. and Varela, F. J. *Autopoiesis and cognition: The realization of the living.* Reidel, Dordrecht, 1980.

[9] Nowé, A., Honkela, T., Könönen, V. and Verbeeck, K. (eds.) *Proceedings of the Workshop W9 on Reinforcement Learning in Nonstationary Environments.* Porto, Portugal, 2005 (in conjunction with the 16th ECML and 9th PKDD, Oct. 3-7, 2005), Portugal 2005, 81 p.

[10] Pöllä, M.; Lindh-Knuutila, T.; Honkela, T. Self-Refreshing SOM as a Semantic Memory Model. *Proc. of AKRR'05, International and Interdisciplinary Conference on Adaptive Knowledge Representation and Reasoning*, Espoo, Finland, June 15-17, 2005, pp. 171-174.

[11] Raitio, J. Vigário, R., Särelä, J. and Honkela, T. Assessing similarity of emergent representations based on unsupervised learning. *Proc. of IJCNN 2004, International Joint Conference on Neural Networks*, Budapest, Hungary, 25-29 July 2004.

[12] Russell, A.; Honkela, T.; Lagus, K.; Pöllä, M. (eds.) *Proceedings of Symposium on Adaptive Models of Knowledge, Language and Cognition (AMKLC'05).* Espoo, Finland, June 15-17, 2005. Espoo, Finland 2005. 61 p.

[13] Russell, A.; Honkela, T. Analysis of interprofessional collaboration in an online learn-
ing environment using self-organizing maps. *Proceedings of Symposium on Adaptive
Models of Knowledge, Language and Cognition (AMKLC'05)*, Espoo, Finland, June
15-17, pp. 52-57, 2005.

[14] Skripal, P. and Honkela, T. Framework for Modeling Emotions in Communities of
Agents. In: H. Hyötyniemi, P. Ala-Siuru and J. Seppänen (eds.), *Life, Cognition and
Systems Sciences, Symposium Proceedings of the 11th Finnish Artificial Intelligence
Conference*, Finnish Science Center Heureka Vantaa, 1-3 September 2004, pp. 163-
172.

[15] Von Foerster, H. (1981). Notes on an epistemology for living things. *Observing
Systems*. Intersystems Publications, pp. 257-271. Originally published in 1972 as BCL
Report No 9.3., Biological Computer Laboratory, University of Illinois, Urbana.

8.3   Reinforcement learning in multiagent systems

Reinforcement learning methods have attained lots of attention in recent years. Although these methods and procedures were earlier considered to be too ambitious and to lack a firm foundation, they have been established as practical methods for solving, e.g., Markov Decision Processes (MDPs). However, the requirement for reinforcement learning methods to work is that the problem domain in which these methods are applied obeys the Markov property. Basically this means that the next state of a process depends only on the current state, not on the history. In many real-world problems this property is not fully satisfied. However, many reinforcement learning methods can still handle these situations relatively well. Especially, in the case of two or more decision makers in the same system the Markov property does not hold and more advanced methods should be used instead. A powerful tool for handling these highly non-Markov domains is the concept of Markov game. In this project, we have developed efficient learning methods based on the asymmetric learning concept and tested the developed methods with different problem domains, e.g. with pricing applications.

Markov games

With multiple agents in the environment, the fundamental problem of single-agent MDPs is that the approach treats the other agents as a part of the static environment and thus ignores the fact that the decisions of the other agents may influence the state of the environment.

One possible solution is to use competitive multiagent Markov decision processes, i.e. *Markov Games (MGs)*. In a MG, the process changes its state according to the action choices of all agents and can thus be seen as a multicontroller MDP. In Fig. 8.1, there is an example of a MG with three states (s_1, s_2, s_3) and two agents. The process changes its state according to probability $P(s_i | s_1, a^1, a^2), i = 2, 3$, where a^1, a^2 are actions selected by the agents 1 and 2.
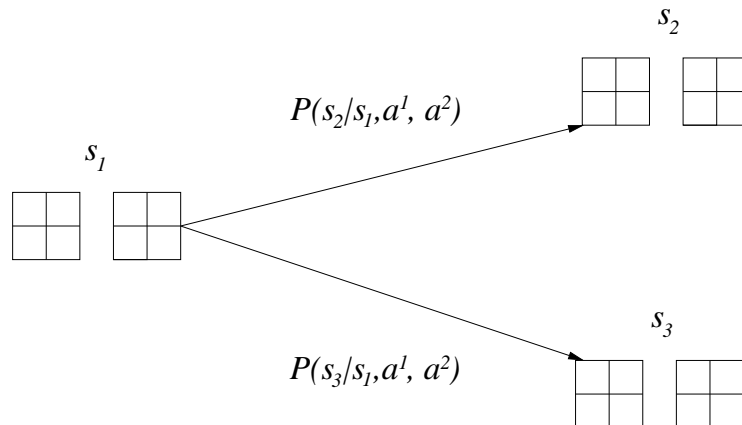

Figure 8.1: An example Markov game with three states.

In single-agent MDPs, it suffices to maximize the utility of the agent in each state. In MGs, however, there are multiple decision makers and more elaborated solution concepts are needed. Game theory provides a reasonable theoretical background for solving this interaction problem. In the single-agent learning, our goal is to find the utility maximizing rule (policy) that stipulates what action to select in each state. Analogously, in a multiagent setting the goal is to find an equilibrium policy between the learning agents.

Practical learning methods

We have concentrated on the case where the state transition probabilities and utility values are not known to the learning agents. Instead, the agents observe their environment and learn from these observations. In general, we use the update rule in the following form:

$$Q_{t+1}^i(s_t, a_t^1, \ldots, a_t^N) = (1 - \alpha_t)Q_t^i(s_t, a_t^1, \ldots, a_t^N) + \alpha_t[r_{t+1}^i + \gamma f(s_{t+1})], \qquad (8.1)$$

where $Q_t^i(s_t, a_t^1, a_t^2)$ is the estimated utility value for the agent i at the time instance t when the system is in the state s_t and agents select actions $a_t^1, \ldots, a_t^N$. r_{t+1}^i is the immediate reward for the agent i and γ is the discount factor. f is the function used to evaluate values of the games associated with states. If a symmetric evaluation function is used, i.e. Nash or Correlated Equilibrium function, the update rule is similar for each agent. In the asymmetric case, there is an ordering (some agents make their decisions prior other agents) among learning agents and thus the learning rules are different on different levels of the corresponding agent hierarchy. Further discussion about symmetric learning methods can be found in [1] and [2]. Respectively, fundamental principles and theoretical analysis of the asymmetric model can be found in [3].

Pricing problem in economy

In this section we provide an example of multiagent reinforcement learning. In the problem, there are two competing agents (brokers) that sell identical products and compete against each other on the basis of price. At each time step, one of the brokers decides its new price based on the opponent's, i.e. other broker's, current price. After the prices have been set, the customer either buys a product from the seller or decides not to buy the product at all. The objective of the agents is to maximize their profits.

Tesauro and Kephart modeled [5] the interaction between two brokers as a single-agent reinforcement learning problem in which the goal of the learning agent is to find the pricing strategy that maximizes its long time profits. Additionally, reinforcement learning aids the agents to prevent "price wars", i.e. repeated price undercutting among the brokers. As a consequence of a price war, the prices would become very small and the overall profits would also be small. Tesauro and Kephart reported very good performance of the approach when one of the brokers keeps its pricing strategy fixed. However, if both brokers try to learn simultaneously, the Markov property assumed in the theory of MDPs does not hold any more and the learning system encounters serious convergence problems. For solving these convergence problems we have modeled the system as a Markov game. In the example depicted in Fig. 8.2, cumulative profits that are averages of 1000 test runs each containing 10 pricing decision for both brokers are plotted against the planning depth (discount factor γ). In this simple example all prices were between $[0, 1]$ and the customer bought the product from the broker with the lowest price. More discussion on the pricing problem can be found in [4].

References

[1] A. Greenwald and K. Hall. Correlated-Q learning. In *Proceedings of the Eighteenth International Conference on Machine Learning (ICML-2003)*, pages 242–249, Washington, DC, 2003.

[2] J. Hu and M. P. Wellman. Nash Q-learning for general-sum stochastic games. *Journal of Machine Learning Research*, 4:1039–1069, 2003.
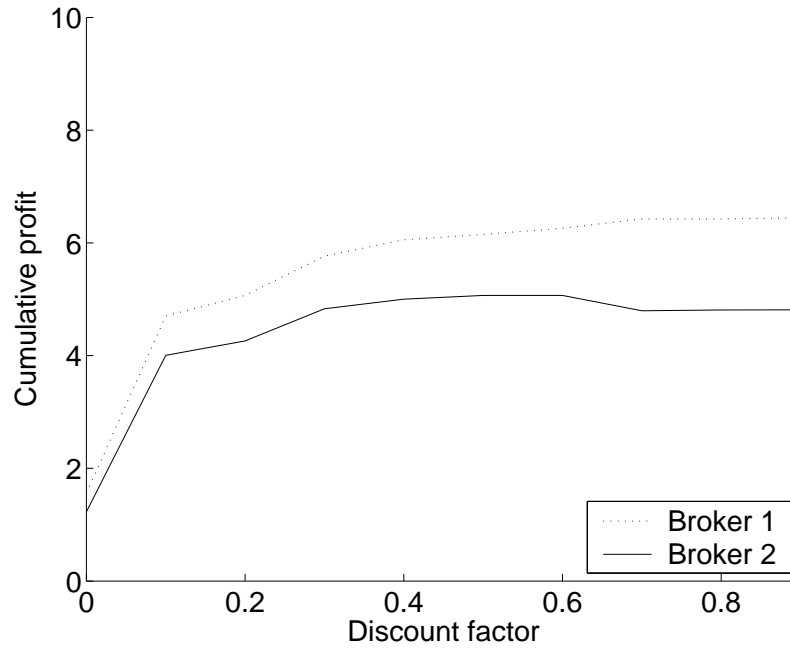
Figure 8.2: Averaged profits in the pricing example. All data points are averages of 1000 test runs each containing 10 pricing decisions for both agents.

[3] V. J. Könönen. Asymmetric multiagent reinforcement learning. *Web Intelligence and Agent Systems: An International Journal (WIAS)*, 2(2):105–121, 2004.

[4] V. J. Könönen. Dynamic pricing based on asymmetric multiagent reinforcement learning. *International Journal of Intelligent Systems*, 21(1):73–98, 2006.

[5] G. Tesauro and J. O. Kephart. Pricing in agent economies using multi-agent Q-learning. In *Proceedings of the Workshop on Game Theoretic and Decision Theoretic Agents (GTDT'99)*, pages 71–86, London, UK, 1999.