Honda Research Institute EU
Honda Research Institute Europe GmbH
https://www.honda-ri.de/

Influence of reward system on the building of abstractions

Inna Mikhailova, Christian Goerick

2011

Preprint:

This is an accepted article published in Int. Conf. on Epigenetic Robotics (EpiRob). The final authenticated version is available online at: https://doi.org/[DOI not available]

Influence of reward system on the building of abstractions

Inna Mikhailova, Christian Goerick
Honda Research Institute Europe GmbH,
Carl-Legien-Strasse 30,
63073 Offenbach / Main, Germany
Email: inna.mikhailova@honda-ri.de

I. Introduction

The integration of multiple sub-systems is a key to building of an autonomous complete system. In this contribution we consider the interfaces between reward system and abstraction system. Both systems were thoroughly investigated but to our knowledge mostly isolated from each other ([1]). Our analysis leads to categorization of different types of abstractions based on the influence of the reward and not on the content of representation. With help of gained overview we show how the integration of different abstraction types supports the ongoing developmental process.

II. Categorization of abstraction types

The appropriate knowledge representation is one of the hardest problems in the actual research on cognitive systems [2]. Here we do not ask what is the best way of knowledge representation, instead we make an overview on different possibilities used in praxis and look at them from the perspective of a complete system that monitors its reward acquisition.

We clarify first what we mean by abstractions. The system can observe its current state in sensorimotor space as well as causality either between time-delayed measurements (predictive model) or between measurements in different sensory channels (associative model). The usage of an associative model can be seen as a process in time, thus both types of causality can be formalized as expectations. As the observation space O is huge, most implementations aim at clustering and discretization of this space: $O \approx \bigcup_{d=1}^{N_d} O_d$, where O_d are clusters in one of described above spaces (sensory, motor, sensorimotor, or causality) and N^d is the number of necessary clusters. This is how the abstractions, i.e. action or perception primitives, come into play. We use the term abstractions for this initial stage in difference to the term symbol that implies existence of symbol-symbol links and grammar.

A. O_d^{stat}: statistical structuring of the system-environment interaction

The observations done by the system obviously depend on the behavior of the system [3]. If the behavior follows the reward optimization, then the abstraction building is indirectly influenced by the reward. This indirect statistical influence can be made more explicit, for example [4] refines the discretization O_d if entropy of the behavior choice is too high.

B. O_d^{dis}: discrimination with respect to reward.

The reward or value function can itself be a part of observation. Then the observation space can be directly segmented according to the amount of the reward [5]. Alternatively the reward or value function can be used as a teaching signal for clustering without being a part of representation as discussed in [6].

C. O_d^{opt}: optimization of abstractions with respect to reward

Optimization is a qualitatively different way how the reward can influence the building of abstractions, [7]. The authors assume that there exists an initial clustering of the sensorimotor flow into the predictive models and corresponding inverse models to produce a desired action. As a result of the reiteration of action optimization and differentiated model update, each pair of prediction and inverse models get specialized to a particular context. The sensorimotor space becomes segmented into a set of optimally tuned controllers for particular contexts.

D. $O_d^{opt^2}$: optimization of abstractions with respect to internal processing on the abstraction layer.

Finally, abstractions can be introduced as a result of the optimization of the operations to be executed on the abstraction layer, e.g. optimization of planning [8] or optimization of search [9].

E. Summary

Figure 1 shows schematically different types of abstractions discussed above. A natural question arises whether we should prefer one type to another. We believe that the discussion about the best suited abstraction type is contra-productive. Instead we investigate how one abstraction type supports the building of another type.

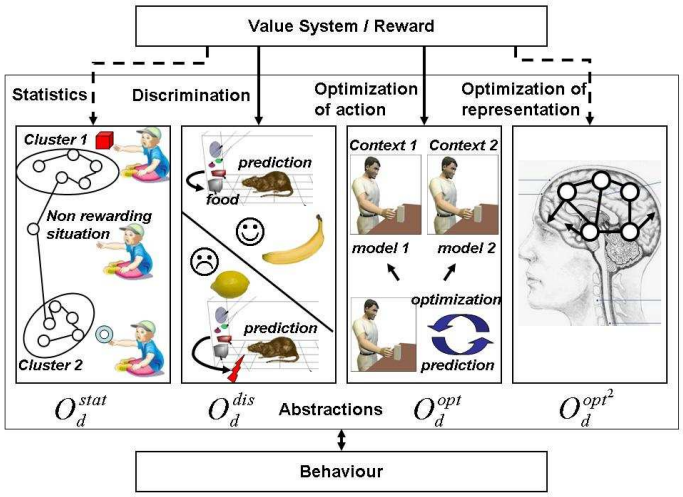

Fig. 1. Different types of abstractions according to the influence of the value system. Solid lines show direct influence and dashed lines indirect one. See the text for more details.

III. Incremental development

Ongoing development supposes that the system can build up on its abilities by creating favorable interaction, by skill optimization, and control reorganization in the sense of usage of learned regularities by higher metal functions (simulation, planning, attention). We agree with [1] that one of systems reward functions should evaluate the quality of interaction with environment. Then with help of simple abstractions of type O_d^{stat} and O_d^{dis} that segments the observation according to this reward, the system can optimize its interaction with environment as to support the learning of expectation models. These in turn can be used as initialization of optimization process that leads to abstractions of type O_d^{opt}. The learned expectations can be used for the transition from reactive to anticipative control [10]. The optimization of higher mental function related to the anticipation give rise to the type $O_d^{opt^2}$. For the stability reasons it is important that more tight coupling of abstractions to behavior, e.g. O_d^{opt}, occurs on later stages. The results of testing these ideas on real-world applications are summarized in [11].

IV. Conclusion

In this contribution we categorized different types of abstractions with respect to interfaces to reward system. The influence of reward on abstractions building is an important part of symbol grounding. Further we showed how one abstraction type supports the building of another in the sense of ongoing development. We hope that this contribution encourages more intensive discussion on integration of multiple types of learning started long ago in [12].

References

[1] P.-Y. Oudeyer and F. Kaplan, "What is intrinsic motivation? a typology of computational approaches," *Frontiers in Neurorobotics*, vol. 1, no. 6, 2007.

[2] G. Pezzulo and C. Castelfranchi, "The symbol detachment problem," *Cognitive Processing*, vol. 8(2), pp. 115–131, 2007.

[3] M. Lungarella and O. Sporns, "Information self-structuring: Key principle for learning and development," *International Conference on Development and Learning*, vol. 0, pp. 25–30, 2005.

[4] S. Hart, S. Ou, J. Sweeney, and R. Grupen, "A framework for learning declarative structure," in *Robotics: Science and Systems - Workshop on Manipulation for Human Environments.* Philadelphia, Pennsylvania, August 2006.

[5] H. Ishiguro, R. Sato, and T. Ishida, "Robot oriented state space construction," in *IEEE/Robotics Society of Japan Int'l Conf. Intelligent Robots and Systems*. IEEE, 1996.

[6] E. Körner and G. Matsumoto, "Cortical architecture and self-referential control for brain-like processing in artificial neural systems," *Artificial life and robotics*, vol. 2, no. 3, pp. 170–178, 1998.

[7] M. Haruno and M. Kawato, "Heterarchical reinforcement-learning model for integration of multiple cortico-striatal loops; fMRI examination in stimulus-action-reward association learning," *Neural Networks*, vol. 19, pp. 1242–1254, 2006.

[8] M. Toussaint and A. J. Storkey, "Probabilistic inference for solving discrete and continuous state markov decision processes," in *ICML '06: Proceedings of the 23rd international conference on Machine learning*. New York, NY, USA: ACM, 2006, pp. 945–952.

[9] J. Eggert and H. Wersing, "Approaches and challenges on the road towards a cognitive vision system," in *Creating Brain-like Intelligence 2007*, B. Sendhoff, E. Koerner, O. Sporns, H. Ritter, and K. Doya, Eds. Springer Verlag, LNCS Series, 2007.

[10] C. Balkenius, "Natural intelligence for artificial creatures," *Lund University Cognitive Studies*, vol. 37, 1995.

[11] I. Mikhailova, *System design for autonomous open-ended acquisition of new behaviors.* Sdwestdeutscher Verlag fr Hochschulschriften, 2010.

[12] C. Balkenius, "Biological learning and artificial intelligence," *Lund University Cognitive Studies*, vol. 30, 1994.