

Models for Autonomously Motivated Exploration in Reinforcement Learning*

Peter Auer¹, Shiau Hong Lim¹, and Chris Watkins²

¹ Chair for Information Technology, Montanuniversität Leoben, Austria
<http://institute.unileoben.ac.at/infotech>

² Department of Computer Science, Royal Holloway University of London, UK
<http://www.rhul.ac.uk/computerscience>

Abstract. One of the striking differences between current reinforcement learning algorithms and early human learning is that animals and infants appear to explore their environments with autonomous purpose, in a manner appropriate to their current level of skills. An important intuition for autonomously motivated exploration was proposed by Schmidhuber [1,2]: an agent should be interested in making observations that reduce its uncertainty about future observations.

However, there is not yet a theoretical analysis of the usefulness of autonomous exploration in respect to the overall performance of a learning agent. We discuss models for a learning agent's autonomous exploration and present some recent results. In particular, we investigate the exploration time for navigating effectively in a Markov Decision Process (MDP) without rewards, and we consider extensions to MDPs with infinite state spaces.

Keywords: Reinforcement learning, autonomous exploration, intrinsic rewards.

References

1. Schmidhuber, J.: A Possibility for Implementing Curiosity and Boredom in Model-Building Neural Controllers. In: Meyer, J.A., Wilson, S.W. (eds.) International Conference on Simulation of Adaptive Behavior: From Animals to Animats, pp. 222–227. MIT Press, Cambridge (1991)
2. Schmidhuber, J.: Developmental Robotics, Optimal Artificial Curiosity, Creativity, Music, and the Fine Arts. *Connection Science* 18(2), 173–187 (2006)

* The full version of this paper is published in the Proceedings of the 22nd International Conference on Algorithmic Learning Theory, Lecture Notes in Artificial Intelligence Vol. 6925.