

POSTER PRESENTATION

Open Access

Biologically plausible reinforcement learning of continuous actions

Jaldert O Rombouts^{1*}, Pieter R Roelfsema^{2,3,4}, Sander M Bohte¹

From Twenty Second Annual Computational Neuroscience Meeting: CNS*2013
Paris, France. 13-18 July 2013

Humans and animals have the ability to perform very precise movements to obtain rewards. For instance, it is no problem at all to pick up a mug of coffee from your desk while you are working. Unfortunately, it is unknown how exactly the non-linear mapping between sensory inputs (e.g. your mug on the retina) and the correct motor actions (e.g. a set of joint angles) are learned by the brain. Here we show how a biologically plausible learning scheme can learn to perform non-linear transformations from sensory inputs to continuous actions based on reinforcement learning.

To arrive at our novel scheme, we built on the idea of attention-gated reinforcement learning (AGREL) [1], a biologically plausible learning scheme that explains how networks of neurons can learn to perform non-linear transformations from sensory inputs to discrete actions (e.g. pressing a button) based on reinforcement learning [2]. We recently showed that the AGREL learning scheme can be generalized to perform multiple simultaneous discrete actions [3], and we now show how this scheme can be further generalized to continuous action spaces. The key idea is that motor areas have feedback connections to earlier processing layers which inform the network about the selected action. Synaptic plasticity is constrained to those synapses that were involved in the decision, and it follows a simple Hebbian rule which is gated by a globally available neuromodulatory signal that codes reward prediction errors. In our novel scheme motor units are situated in a population coding layer that encodes the outcome of the decision process as a bump of activations [4]. This contrasts to our earlier work where single motor units code for actions [1,3]. We show that the synaptic updates perform stochastic gradient

descent on the prediction error that results from the combined action-value prediction of all the motor units that encoded the decision. Unlike other reinforcement learning based approaches, e.g. [5], our reinforcement learning rule is powerful enough to learn tasks that require non-linear transformations. The distribution of population centers in the motor layer can also be automatically adapted to task demands, yielding more representational power when actions need to be precise.

We show that the novel scheme can learn to perform non-linear transformations from sensory inputs to motor outputs in a variety of direct reward tasks. The model can explain how visuomotor coordinate transforms might be learned by reinforcement learning instead of semi-supervised learning as used in [6]. It might also explain how humans learn to weigh the accuracy of their movement against the potential rewards and punishments for making inaccurate movements as in the visually guided movement task described in [7].

Author details

¹Life Sciences, Centrum Wiskunde en Informatica (CWI), 1098XG, Amsterdam, The Netherlands. ²Department of Vision & Cognition, Netherlands Institute for Neurosciences, an institute of the Royal Netherlands Academy of Arts and Sciences (KNAW), 1105 BA, Amsterdam, The Netherlands. ³Department of Integrative Neurophysiology, Centre for Neurogenomics and Cognitive Research (CNCR), VU University, 1081 HV, Amsterdam, The Netherlands. ⁴Psychiatry Department, Academic Medical Center (AMC), Amsterdam, The Netherlands.

Published: 8 July 2013

References

1. Roelfsema PR, van Ooyen A: **Attention-gated reinforcement learning of internal representations for classification.** *Neural Comp* 2005, **17**:2176-2214.
2. Sutton RS, Barto AG: **Reinforcement Learning: an introduction.** MIT Press; 1998.
3. Rombouts JO, van Ooyen A, Roelfsema PR, Bohte SM: **Biologically Plausible Multi-dimensional Reinforcement Learning in Neural Networks.** *ICANN* 2012, 443-450.

* Correspondence: J.O.Rombouts@CWI.nl

¹Life Sciences, Centrum Wiskunde en Informatica (CWI), 1098XG, Amsterdam, The Netherlands

Full list of author information is available at the end of the article

4. Zhang K: Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *J Neurosci* 1996, **16**:2112-2126.
5. Ognibene D, Rega A, Baldassarre G: A model of reaching that integrates reinforcement learning and population encoding of postures. *From Animals to Animats 9* 2006, 381-393.
6. Ghahramani Z, Wolpert DM, Jordan MI: Generalization to local remappings of the visuomotor coordinate transformation. *J Neurosci* 1996, **16**:7085-7096.
7. Trommershäuser J, Maloney LT, Landy MS: Statistical decision theory and the selection of rapid, goal-directed movements. *J Opt Soc Am A Opt Image Sci Vis* 2003, **20**:1419-1433.

doi:10.1186/1471-2202-14-S1-P28

Cite this article as: Rombouts *et al.*: Biologically plausible reinforcement learning of continuous actions. *BMC Neuroscience* 2013 **14**(Suppl 1):P28.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

