Short and long term plasticity as cause-effect hypothesis testing in robotic ambiguous scenarios

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Citation: SOLTOGGIO, A., 2013. Short and long term plasticity as cause-effect hypothesis testing in robotic ambiguous scenarios [poster presentation]. 2nd Bernstein Sparks Workshop. NeuroEngineering the Brain: from Neuroscience to Robotics ...and back, 2-4 December 2013, Ev. Akademie Tutzing, near Munich.

Additional Information:

- This is a poster presentation

Metadata Record: [https://dspace.lboro.ac.uk/2134/16975](https://dspace.lboro.ac.uk/2134/16975)

Version: Accepted for publication

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: [https://creativecommons.org/licenses/by-nc-nd/4.0/](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Please cite the published version.
Short and long term plasticity as cause–effect hypothesis testing in robotic ambiguous scenarios

Andrea Soltoggio

Neurorobotic systems learn with difficulties when a continuous flow of information and delays make the cause–effect relationships ambiguous. The Hypothesis Testing (HT) plasticity proposed in this study models learning dynamics that account for ambiguity in the sensory–motor information flow improving drastically discrimination capabilities and memory capacity with respect to previous models. The new rule models consolidation of long–term memory and helps solve the plasticity–stability dilemma.

Ambiguous cause–effect relationships with delayed rewards

What sequence of stimuli and actions effectively causes a reward? One single instance cannot disambiguate the relationship if delays and asynchronous information flow are present.

Rarely Correlating Hebbian Plasticity (RCHP) (equivalent to R–STDP)

RCHP \( j(t) = \begin{cases} +1 & \text{if } v_j(t) \cdot w_{ji} > \theta_{bi} \\ -1 & \text{if } v_j(t) \cdot w_{ji} < \theta_{ci} \\ 0 & \text{otherwise} \end{cases} \)

Eligibility traces: \( c_{ji}(t+\Delta t) = c_{ji}(t) \cdot e^{-\Delta t / \tau_{el}} + \text{RCHP}_{ji}(t) \)

Modulation: \( d(t+\Delta t) = d(t) \cdot e^{-\Delta t / \tau_{mod}} + \lambda r(t) + b \)

Weight update: \( \Delta w_{ji}(t) = c_{ji}(t) \cdot d(t) \)

Hypothesis Testing Plasticity. Addition of 1) short–term and long–term components 2) hypothesis testing potentiation and depression with negative baseline modulation (term b)

RCHP \( j(t) = \begin{cases} +1 & \text{if } v_j(t) \cdot w_{ji} > \theta_{bi} \\ 0 & \text{otherwise} \end{cases} \)

Eligibility traces: \( c_{ji}(t+\Delta t) = c_{ji}(t) \cdot e^{-\Delta t / \tau_{el}} + \text{RCHP}_{ji}(t) \)

Short–term weight dynamics: \( \dot{w}_{ji}(t) = -w_{ji}(t)/\tau_{st} + m(t) \cdot c_{ji}(t) \)

Consolidation of hypotheses (weights): \( W_{ji}(t) = w_{ji}(t) + w_{ji}^{\prime}(t) \)

HT–plasticity in a network model

Problem
- large input–output flow
- delayed rewards
- ambiguous cause–effect relationships
- no external reward–predictors (i.e. expected average rewards)

Network
- 300 inputs, 30 outputs
- 9000 stimulus–action pairs

Advantages
- short–term components represent hypotheses
- long–term components represent established facts
- HT allows for learning over multiple scenarios
- HT plasticity allows for higher discrimination of true cause–effect relationships

References
- Soltoggio, Lemme, Reinhart, Steil. Rare neural correlations implement robotic conditioning with delayed rewards and disturbances. Frontiers in Neurorobotics (2013)

SIMULATIONS

(a) Memory capacity, preservation of information over different learning scenarios (b) Utility of memory when revisiting a previously learned scenario

NEURO–ROBOTICS APPLICATIONS: combination of skills

Previous experiments with RCHP can now be all combined and expanded thanks to the increased memory capabilities and different learning scenarios

iCub plays Rock–Paper–Scissors

iCub learns the colors

Operant conditioning with reversal learning. Learning from trial and error in human–robot interactions

Uncertain feedback signals in human–robot interaction

Classical conditioning: iCub makes new friends

Conclusion. The HT-plasticity improves both exploration and exploitation capabilities of the network. It increases memory capabilities by preserving established learned relationships, it detects coincidental facts as irrelevant and allows for the combination of more learning scenarios and skills.