



## FENS Forum 2010 - Amsterdam

- Posters: to be on display from 8:00 to 13:15 in the morning and from 13:30 to 18:45 in the afternoon. Poster sessions run from 09:30 to 13:15 in the morning and from 13:30 to 17:30 in the afternoon. A one hour time block is dedicated to discussion with the authors (authors should be in attendance at their posters as from the time indicated.)
  - For other sessions, time indicates the beginning and end of the sessions.
- 

**First author** Friedrich, Johannes (poster)

Poster board B17 - Mon 05/07/2010, 14:30 - Hall 1

Session 103 - Synaptic plasticity 3

Abstract n° 103.17

Publication ref.: *FENS Abstr.*, vol.5, 103.17, 2010

**Authors** Friedrich J., Senn W. & Urbanczik R.

**Address** Dept. of Physiology, University of Bern, Bern, Switzerland

**Title** Spatio-temporal credit assignment in population learning

**Text** Learning by reinforcement is important in shaping animal behavior. But behavioral decision making is likely to involve the integration of many synaptic events in space and time. So in using a single reinforcement signal to modulate synaptic plasticity a twofold problem arises. Different synapses will have contributed differently to the behavioral decision and, even for one and the same synapse, releases at different times may have had different effects.

Here we present a plasticity rule which solves this spatio-temporal credit assignment problem in a population of spiking neurons. The learning rule is spike time dependent and maximizes the expected reward by following its stochastic gradient. Synaptic plasticity is modulated not only by the reward but by a population feedback signal as well. While this additional signal solves the spatial component of the problem, the temporal one is solved by means of synaptic eligibility traces. In contrast to temporal difference based approaches to reinforcement learning, our rule is explicit with regard to the assumed biophysical mechanisms. Neurotransmitter concentrations determine plasticity and learning occurs fully online.

Further, it works even if the task to be learned is non-Markovian, i.e. when reinforcement is not determined by the current state of the system but may also depend on past events.

The performance of the model is assessed by studying three non-Markovian tasks. In the first task the reward is delayed beyond the last action with non-related stimuli and actions appearing in between. The second one involves an action sequence which is itself extended in time and reward is only delivered at the last action, as is the case in any type of board-game. The third is the inspection game that has been studied in neuroeconomics. It only has a mixed Nash equilibrium and exemplifies that the model also copes with stochastic reward delivery and the learning of mixed strategies.

**Theme** B - Excitability, synaptic transmission, network functions  
Synaptic plasticity - Spike-timing dependent plasticity

---