A biologically plausible model of the learning rate dynamics

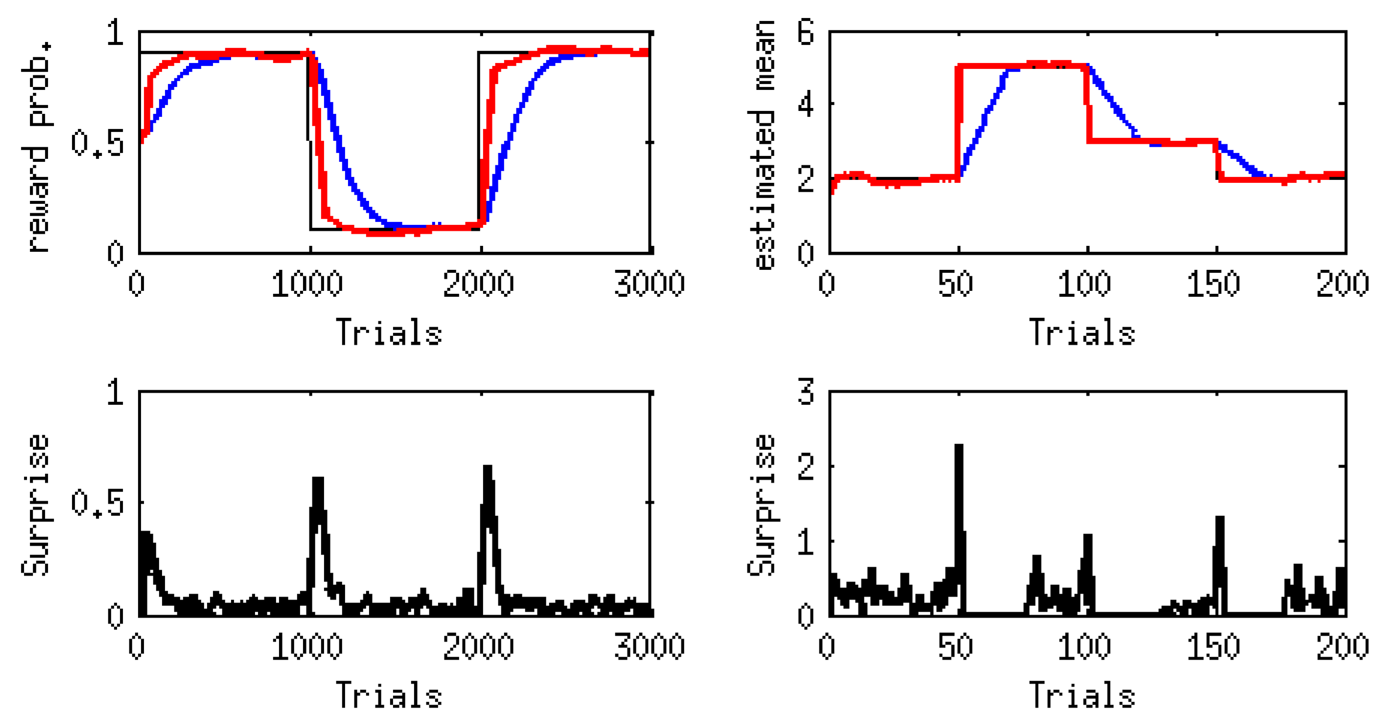
Mohammad Javad Faraji[1], Kerstin Preuschoff[1], and Wulfram Gerstner[1]

1. School of Life Sciences, Brain Mind Institute & School of Computer and Communication Sciences, EPFL, Switzerland

Abstract

We address how surprise affects learning. We propose a model of the learning rate dynamics suitable for learning in stable and unstable environments. This model is based on uncertainty and surprise measures. We further show that surprise, in principle, can improve learning by modulating the learning rate, regulating exploration-exploitation trade-off, and generating new environmental states.

Modulating learning rate by surprise

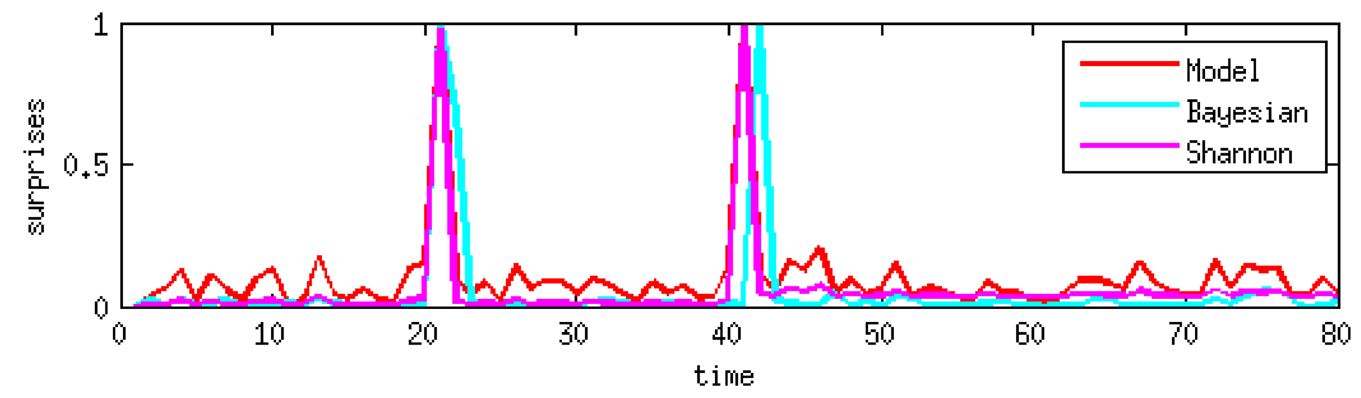


The agent estimates the probability of reward delivery in a reversal task (upper-left). We altered a standard SARSA learning algorithm (blue line) such that when the agent detects an unexpected event beyond the stochasticity of the environment, the ensuing surprise signal (bottom-left) temporarily accelerates learning leading to **more accumulated reward** for the surprise-based reinforcement learner (red line). In the dynamic decision making task (upper-right), the learner observes samples from a Gaussian distribution with varying mean (black line). Modulating the learning rate by surprise signals (bottom-right) leads to **faster detection of change points** (red line) than that in the SARSA model (blue line).

Learning rate dynamics

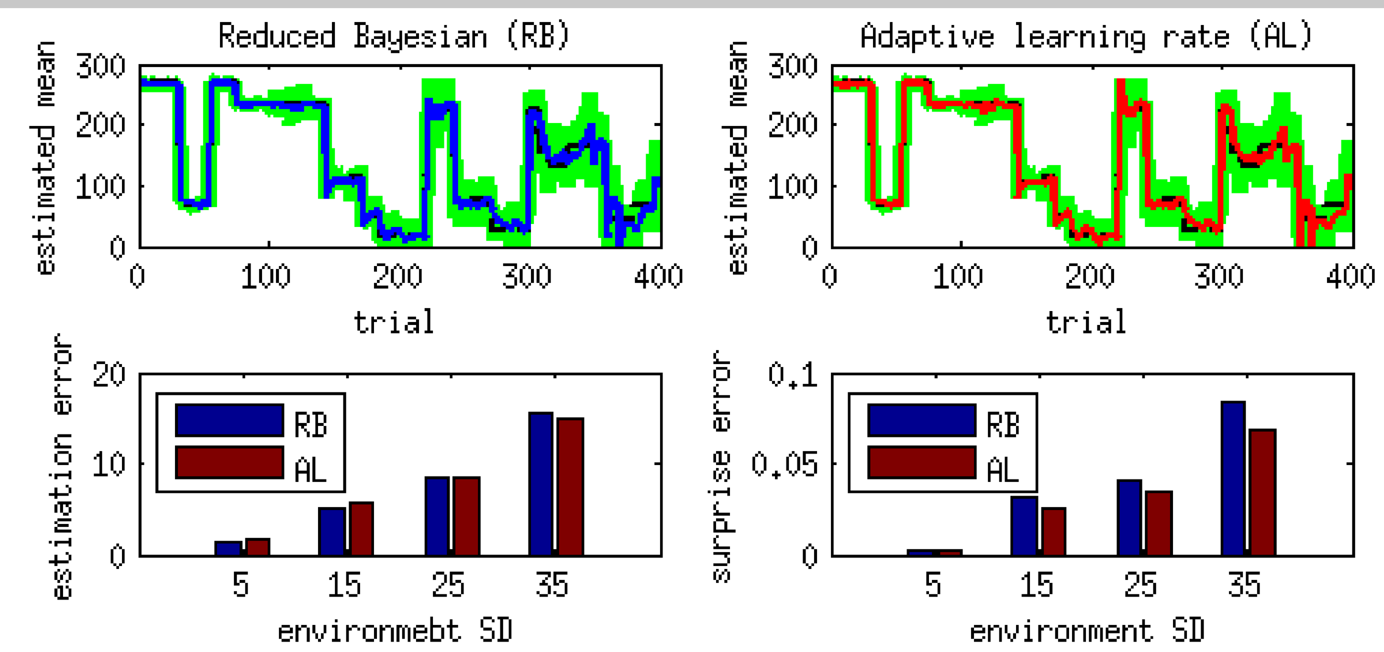
Reward prediction error $\delta_n = r_n - \hat{\mu}_{n-1}$ and the estimated risk $\hat{\sigma}_n$ of the environment are used to measure surprise $S_n = f(|\delta_n|/\hat{\sigma}_{n-1})$. The dynamics of the learning rate α is then controlled by surprise S_n and the level of relative uncertainty $U_n = u_n^2/(u_n^2 + \hat{\sigma}_n^2)$. The estimation uncertainty u_n^2 determines variation of the estimated mean reward, i.e., $u_n = std(\hat{\mu})$. Finally, $\dot{\alpha} = -U_n\alpha + S_n$.

Surprise measures

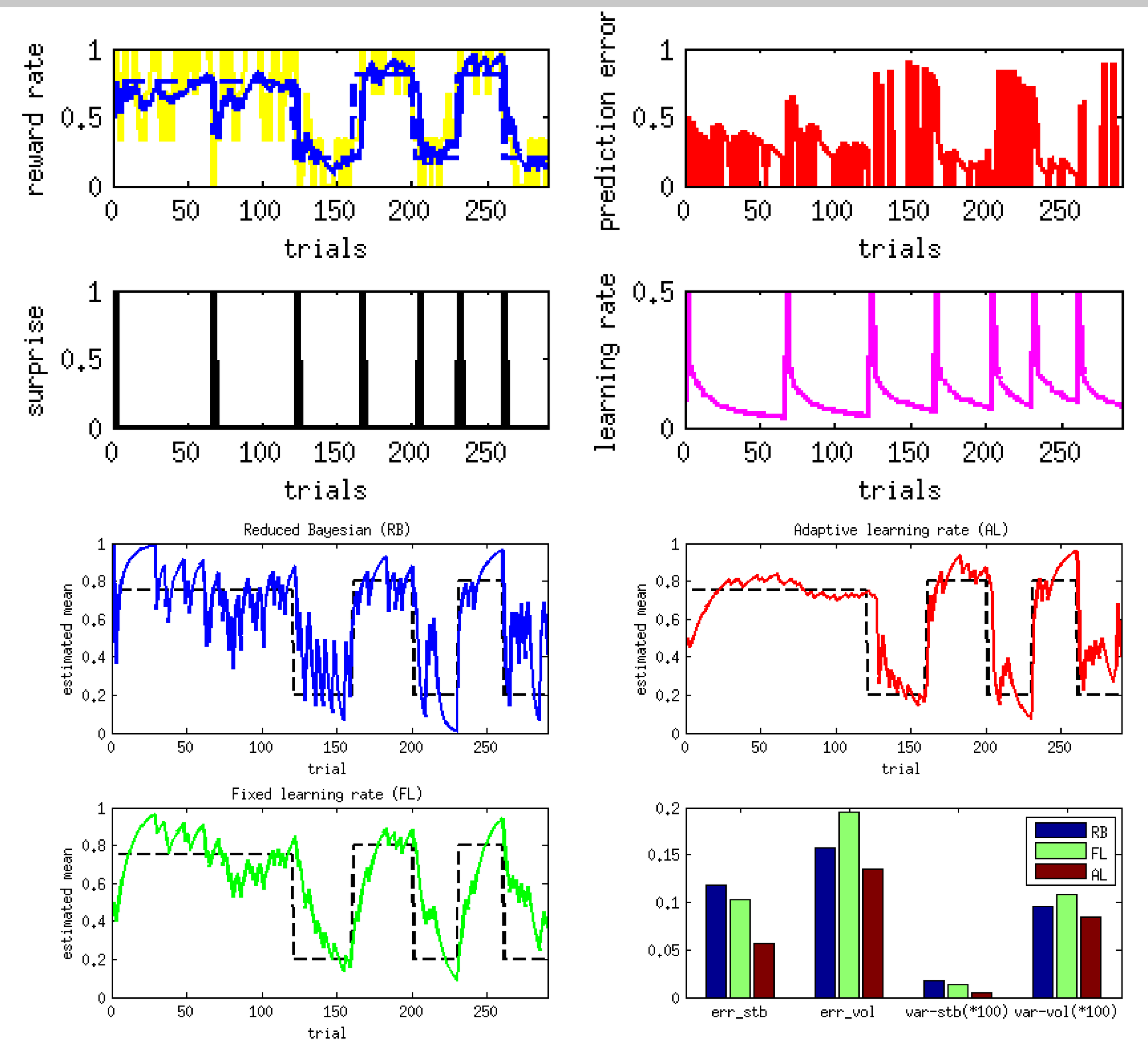


Theses results hold equally well for different surprise measures: Shannon surprise $-\log P(r_n|\hat{\mu}_{n-1}, \hat{\sigma}_{n-1})$, Bayesian approach $D_{KL}[P_{n+1}(\hat{\mu}|r_n)||P_n(\hat{\mu})]$, and model-free $|r_n - \hat{\mu}_{n-1}|/\hat{\sigma}_{n-1}$.

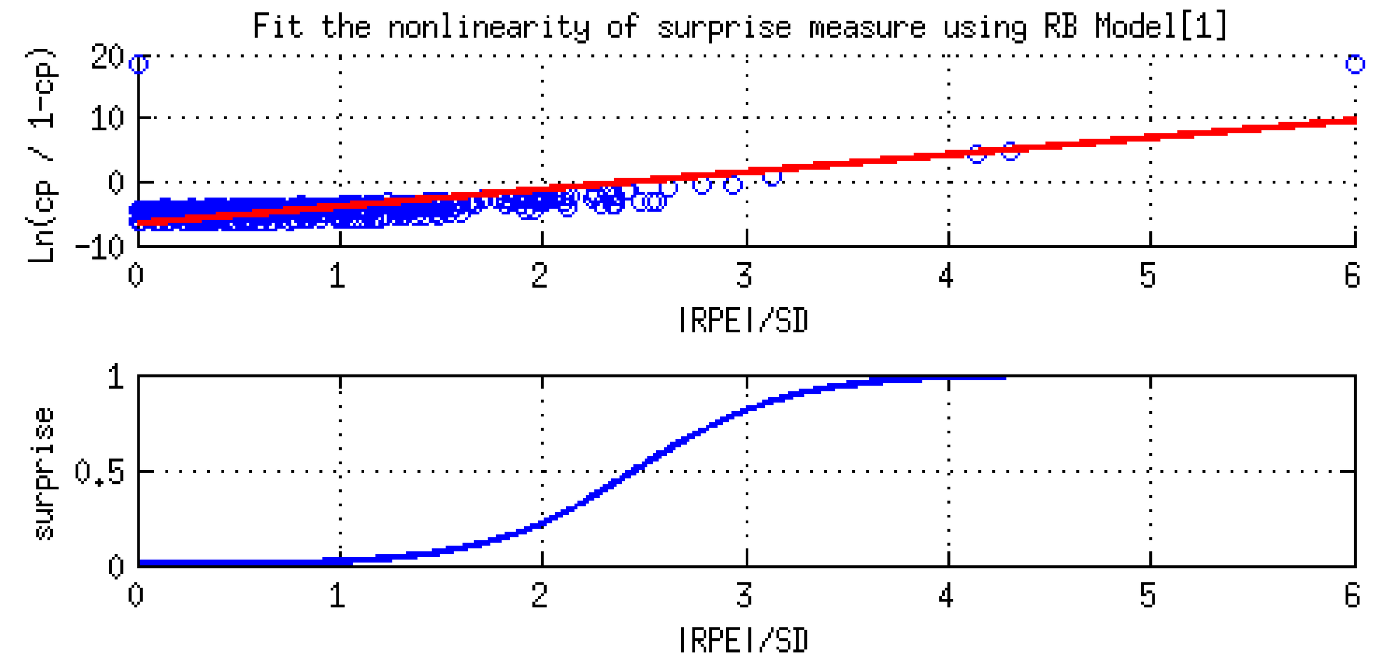
Dynamic decision making task



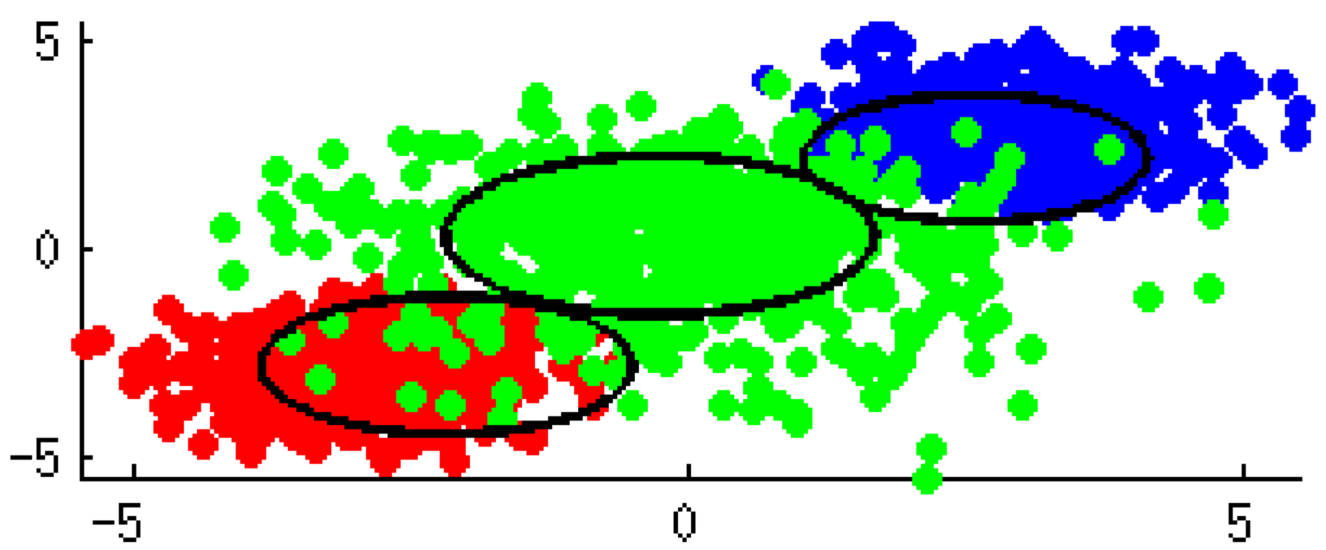
Reversal task



Non-linearity in surprise



Surprise triggers new clusters



A classic K-means clustering algorithm is modified such that if the total number of clusters is initially unknown, the agent (classifier) equipped with surprise is able to add more clusters (black circles) whenever it judges a pattern (colored data point) to be surprising, i.e., a pattern which may belong to none of the existing clusters. It represents an agent who is able to generate (trigger) new states, **an essential feature for learning new environments**.

Conclusion

Humans keep track of uncertainty and surprise measures which control the learning rate dynamics in our simple proposed model. A simple computational model is also suggested for the surprise function. This model works well for both dynamic decision making and reversal tasks. Furthermore, surprise-based SARSA accelerated learning in decision making tasks. A surprise-based clustering algorithm can trigger new clusters if it judges a pattern to be novel.

References

[1] Nassar M, et al. "An approximately Bayesian delta-rule model explains the dynamics of belief updating." J. Neuros. 2010.
[2] Behrens T, et al. "Learning the value of information in an uncertain world." Nat. Neuros. 2007.

Acknoledgment: Research was supported by the ERC (grant no. 268 689, H.S.)