A Novel Measure of Surprise with Applications for Learning within Changing Environments

Mohammad Javad Faraji[1], Kerstin Preuschoff[2], and Wulfram Gerstner[1]

1: School of Life Sciences, Brain Mind Institute and School of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.
2: Geneva Finance Research Institute (GFRI) and Swiss Center for Affective Sciences (CISA), University of Geneva, Geneva, Switzerland.

Summary

Surprise is informative because it drives attention [1] and modifies learning [2]. Correlates of surprise have been observed at different stages of neural processing, and found to be relevant for learning and memory formation [3]. Although surprise is ubiquitous, there is neither a widely accepted theory that quantitatively links surprise to observed behavior, such as the startle response, nor an agreement on how surprise should influence learning speed or other parameters in iterative statistical learning algorithms.

Building on and going beyond earlier surprise measures [1, 4–6], we propose a novel information theoretic measure for calculating surprise in a Bayesian framework so as to capture uncertainty of the world as well as imperfections of the subjective model of the world, two important aspects of surprise. The principle of future surprise minimization leads to a learning rule that can be interpreted as a surprise modulated belief update suitable for learning within changing environments.

Importantly, we do not need an assumption about how quickly the world changes. We apply our surprise-modulated learning rule to an exploration task in a maze-like environment and to a dynamic decision making task. Our results are consistent with the behavioral finding that surprising events induce humans and animals to learn faster and to adapt more quickly to changing environments [7–9].

Additional Detail

Information content [5, 10] captures the inherent unexpectedness of a piece of data for a given set of models (uncertainty of the world), whereas Bayesian surprise [1,4] measures the change in belief caused by a new data point (observer dependent). These are two complementary approaches for calculating surprise. The former is about data and the latter is about a model. In our approach both aspects are combined with a third aspect: if we are uncertain about what to expect, receiving a low-probability data sample is less surprising than in a situation when we are almost certain about the world. A surprise minimizing learning (SMiLe) rule is derived by solving a constrained optimization problem defined as follows: the objective is to maximally reduce surprise when facing the same data again in the not so distant future, under the constraint that the posterior belief (after the update step) is not too different from the prior.

The resulting SMiLe rule balances the influence of newly acquired data with prior knowledge where the balance depends on surprise. In case of a fundamental change in the world signaled by surprising samples, data acquired before the change is valued as less informative about the current state of the world. A simultaneous increase of the influence of newly acquired data on learning leads to a fast adaptation of the model to an environmental change. While in a stationary environment our algorithm approaches the known Bayesian update rule, it also allows the model to react to changes in the environment.

In summary, surprising data increases the uncertainty we have about our current model of the world and gives a bigger influence to newly acquired data on belief update. The interaction between surprise and uncertainty is important for modeling the behavior of humans and animals in changing environments. The surprise signal could be broadly transmitted in the brain by a neuromodulator with widespread axonal ramifications (e.g., norepinephrine released from locus coeruleus neurons) and influence synaptic plasticity rules.

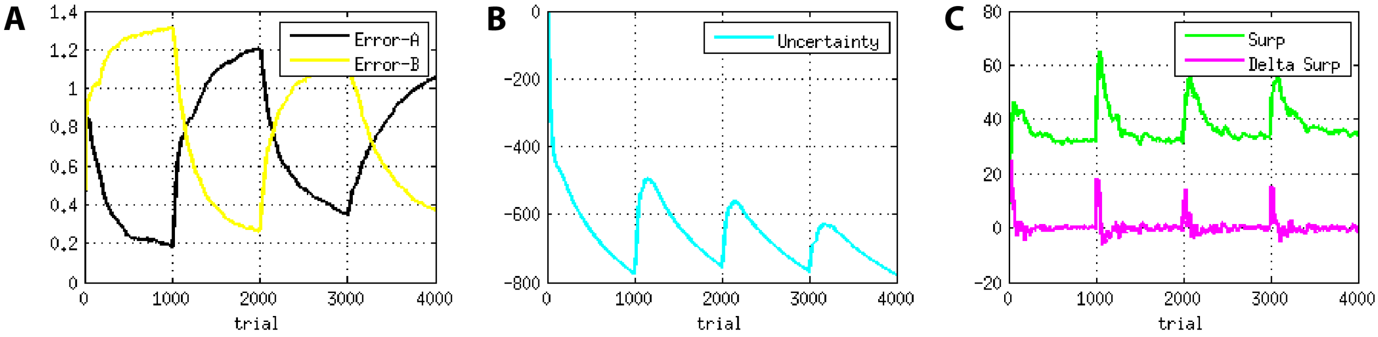

Figure 1: **A.** Average error in estimating the true state transition probabilities in environments A (black) and B (yellow) in a maze exploration task in which a subject (say a rat) is exposed to two $n \times n$ grid worlds A and B. The environment switches between A and B every 1000 trials, starting with A. The surprise signal helps the rat to quickly adapt after each switch. **B.** Model uncertainty (cyan) described as the entropy of the current belief decreases as the rat learns about either A or B except at and shortly after the switching points. **C.** Surprise signal (green), plotted as a running average over 10 trials, decreases as the rat learns more about the environments, except at switching points. Surprise-minus-running-average-of-surprise (amplitudes magnified $\times 20$, magenta) indicates the change points where the model uncertainty increases.

Acknowledgments

This project is funded by the European Research Council (grant agreement no. 268 689).

References

[1] Laurent Itti and Pierre Baldi. Bayesian surprise attracts human attention. *Vision research*, 49(10):1295–1306, 2009.

[2] Wolfram Schultz and Anthony Dickinson. Neuronal coding of prediction errors. *Annual review of neuroscience*, 23(1):473–500, 2000.

[3] Charan Ranganath and Gregor Rainer. Neural mechanisms for detecting and remembering novel events. *Nature Reviews Neuroscience*, 4(3):193–202, 2003.

[4] Pierre Baldi and Laurent Itti. Of bits and wows: a bayesian theory of surprise with applications to attention. *Neural Networks*, 23(5):649–666, 2010.

[5] Claude Elwood Shannon. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(1):3–55, 2001.

[6] Gunther Palm. *Novelty, information and surprise*. Springer, 2012.

[7] John M Pearce and Geoffrey Hall. A model for pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological review*, 87(6):532, 1980.

[8] Paul N Wilson, Patrick Boumphrey, and John M Pearce. Restoration of the orienting response to a light by a change in its predictive accuracy. *Quarterly Journal of Experimental Psychology: Section B*, 44(1):17–36, 1992.

[9] Peter C Holland. Brain mechanisms for changes in processing of conditioned stimuli in pavlovian conditioning: Implications for behavior theory. *Animal Learning & Behavior*, 25(4):373–399, 1997.

[10] Myron Tribus. Information theory as the basis for thermostatics and thermodynamics. *Journal of Applied Mechanics*, 28(1):1–8, 1961.