Surprise Minimization as a Learning Strategy

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

1. Brain Mind Institute & School of Computer and Communication Sciences, EPFL, CH.
2. Geneva Finance Research Institute (GFRI), University of Geneva, CH.

Abstract

Surprise (feeling caused by something *unexpected*) is one of the crucial yet largely undetermined factors that affects learning and attention. We propose a novel measure for calculating surprise which combines the advantages of existing surprise measures. Further, we propose a principle of (future) surprise minimization that can be employed as a learning strategy suitable for learning within changing environments.

Existing Surprise Measures

- Although surprise is ubiquitous, it is difficult to quantify. Surprise is *subjective* and occurs whenever there is *uncertainty*.
- Assume the world is modeled by a generative distribution $\mathbf{p(x|\theta^*)}$ where unknown θ^* governs the world and $\mathbf{x}$ denotes the random variable corresponding to outcomes $\mathbf{X}$. In Bayesian framework, our uncertainty about the world is modeled by a prior $\pi(\theta)$. How likely data $\mathbf{X}$ is to be generated by our internal model of the world is $\mathbf{Z(X)} = \int_\Theta \mathbf{p(X,\theta)\, d\theta}$, where $\mathbf{p(X,\theta) = p(X|\theta)\pi(\theta)}$.
- Two existing measures of surprise:
  - **Shannon surprise** [1]: $\mathbf{S_{Sh.}(X) = -\ln Z(X)}$.
  - **Bayesian surprise** [2]: $\mathbf{S_{Bay.}(X) = D_{KL}}\big[\pi(\theta|\mathbf{X}) \;||\; \pi(\theta)\big]$.
- It is worth noting that by simply averaging Shannon surprise and Bayesian surprise over the distribution $\mathbf{Z(x)}$ of outcomes we can arrive at information theoretic definitions such as *entropy* $\mathbf{H(x)}$ or *mutual information* $\mathbf{I(\theta,x)}$.

Our Proposed Surprise Measure

- Shannon and Bayesian methods are different but complementary approaches for calculating surprise. The former is about *data* and the latter is about a *model*. We propose *average Shannon surprise*

$$\mathbf{S(X;\pi)} = -\int_\Theta \pi(\theta)\; \mathbf{\ln p(X|\theta)\, d\theta}, \tag{1}$$

in which we could calculate surprise prior to any update of our belief about θ and we (implicitly) incorporate features of both Shannon and Bayesian surprise measures, because

$$\mathbf{S(X;\pi) = -\ln Z(X) + D_{KL}}\big[\pi(\theta) \;||\; \pi(\theta|\mathbf{X})\big]. \tag{2}$$

Surprise Minimization Principle

- Surprise is *informative* because it drives attention and modifies learning. If learning is affected by surprise, a repetition of an unexpected event for a second time is perceived less surprising than the first time it is observed. This fact can be employed as a learning strategy that we call surprise minimization principle.
- We define a *learning rule* $\mathbf{L}$ as a mapping function that maps a prior belief $\pi_0(\theta)$ to a posterior belief $\mathbf{q(\theta)} \in \mathcal{F}$ after receiving data $\mathbf{X} \in \mathcal{X}$, i.e.,

$$\begin{aligned} \mathbf{L} &: \mathcal{X} \times \mathcal{F} \to \mathcal{F} \\ \mathbf{q} &= \mathbf{L[X,\pi_0]}. \end{aligned} \tag{3}$$

- We then define a class $\mathcal{L}$ of *plausible learning rules* for which the *posterior average Shannon surprise* $\mathbf{S(X;q)}$ of a new piece of data $\mathbf{X}$ is *at most as surprising as* its corresponding *prior average Shannon surprise* $\mathbf{S(X;\pi_0)}$, i.e.,

$$\mathcal{L} = \{\mathbf{L : S(X;q) \le S(X;\pi_0),\; q = L[X,\pi_0]}\}. \tag{4}$$

- Likelihood maximization which returns the Dirac delta posterior $\mathbf{q(\theta) = \delta(\theta - \hat\theta)}$, where $\hat\theta = \mathbf{arg\,max_\theta\, p(X|\theta)}$, maximally reduces the posterior average Shannon surprise $\mathbf{S(X;q)}$. This of course imposes a huge modification of the prior belief π_0.

References & Acknowledgment

[1] Shannon, Claude Elwood. "A mathematical theory of communication." ACM SIGMOBILE Mobile Computing and Communications Review 5.1 (2001): 3-55.
[2] Itti, Laurent, and Pierre F. Baldi. "Bayesian surprise attracts human attention." Advances in neural information processing systems. 2005.
[3] Nassar, Matthew R., et al. "An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment." The Journal of Neuroscience 30.37 (2010): 12366-12378.
Funded by the European Research Council (grant no. 268 689).

Constraint Surprise Minimization: SMiLe Rule

- Alternatively, one could limit the search to all the posteriors that are not too divergent from the prior π_0, i.e., all $\mathbf{q}$ for which $\mathbf{D_{KL}[q||\pi_0] \le B}$. This constraint problem can be formulated as follows.

$$\underset{\mathbf{q}\in\mathcal{F}}{\mathbf{arg\,min}}\; \mathbf{S(X;q)} - \frac{\mathbf{1}}{\lambda}(\mathbf{B - D_{KL}[q||\pi_0]}) \tag{5}$$

$$= \frac{\mathbf{p(X|\theta)^\lambda \pi_0(\theta)}}{\mathbf{Z(X;\lambda)}} := \pi_\lambda(\theta|\mathbf{X}) := \mathbf{L_\lambda[X,\pi_0]}. \tag{6}$$

- **S**urprise **Mi**nimizing **Le**arning (**SMiLe**) rule $\mathbf{L_\lambda}$ is reminiscent of Bayes' rule except that the likelihood $\mathbf{p(X|\theta)}$ is modulated by parameter λ.

Theorems

- **Theorem 1:** *A piece of data* $\mathbf{X}$ *has the minimal average Shannon surprise* $\mathbf{S(X;\pi_\lambda)}$ *under the posterior* π_λ*, corresponding to the SMiLe rule (6), among all posterior beliefs* $\mathbf{q} \in \mathcal{F}$ *that are sufficiently similar to the prior* π_0*. Here dissimilarity is expressed as KL divergence and sufficiency is determined by a pre-defined non-negative bound* $\mathbf{B \ge 0}$*, i.e.,*

$$\forall \mathbf{B \ge 0}\; \exists \lambda \ge \mathbf{0} \;\; \mathbf{s.t.} \;\; \big[\mathbf{S(X;\pi_\lambda) \le S(X;q)},\;\; \forall \mathbf{q}, \mathbf{D_{KL}[q||\pi_0] \le B}\big]. \tag{7}$$

- **Theorem 2:** *The SMiLe rule (6) obeys the principle of surprise-minimization for all* $\lambda \ge \mathbf{0}$*. This includes the standard Bayes' rule* ($\lambda = \mathbf{1}$) *as well as the likelihood maximization* ($\lambda \to \infty$)*. That is*

$$\mathbf{L_\lambda} \in \mathcal{L}, \;\; \forall \lambda \ge \mathbf{0}. \tag{8}$$

Simulation of a Dynamic Decision Making Task

- The task is to correctly estimate the true mean of a Gaussian distribution whose underlying mean suddenly changes at some unexpected change-points. In each of four **100**—trial blocks, the standard deviation of the distribution is fixed to **5, 15, 25,** and **35**, respectively [3].
- In case of a fundamental change which could be signaled by an unexpected *surprising* observation, data acquired before the change becomes less informative about the current state of the world. Increasing the influence of surprising outcomes on obtaining posterior belief can be modeled by choosing λ that increases with surprise of each new piece of data.

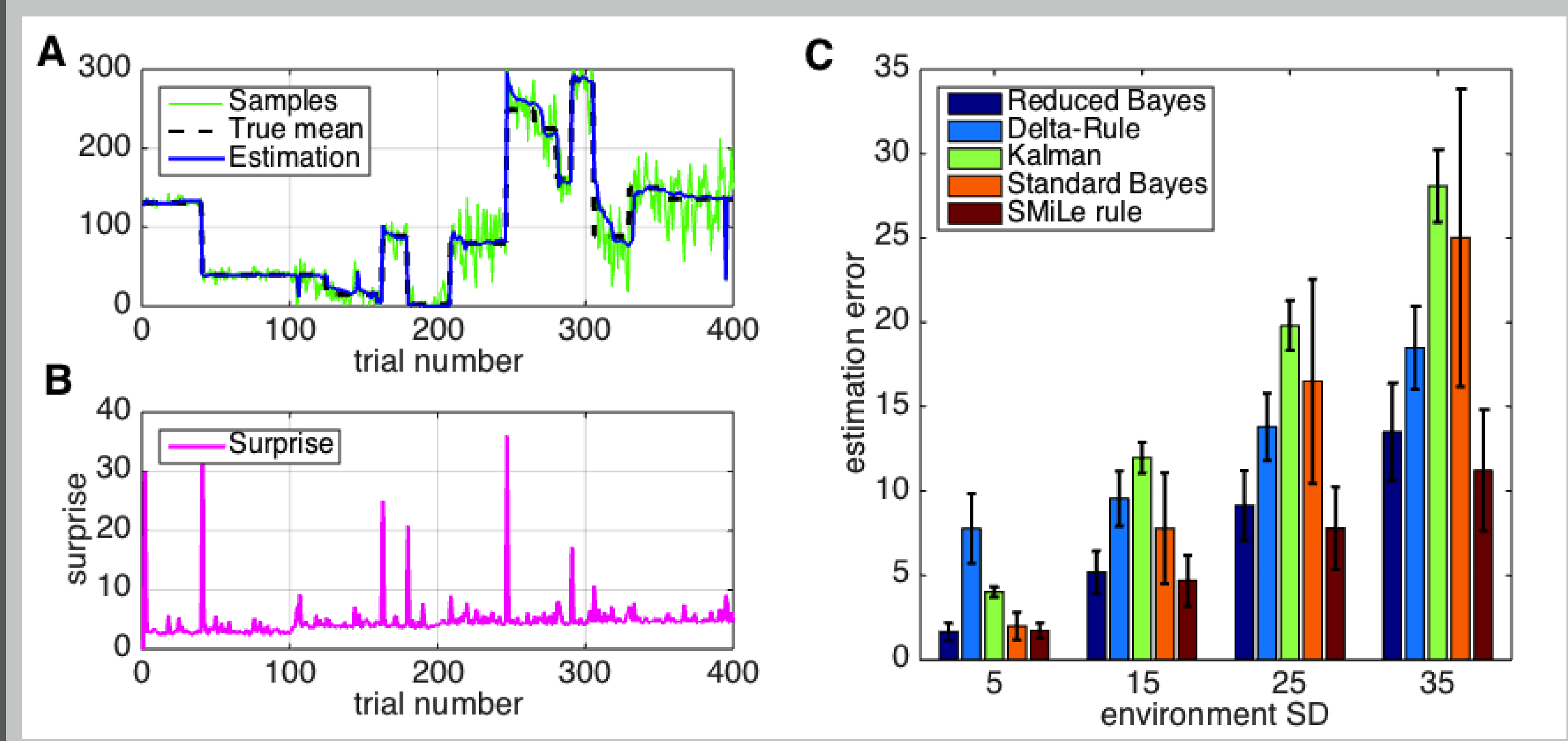

Figure 1: **A**. Samples (green), true mean (black-dashed) and mean estimated by the SMiLe rule (blue) (6), for **400** trials in a dynamic decision making task. **B**. The surprise (magenta) (1) increases whenever the actual mean changes. **C**. The mean and the standard deviation (error bars) of the estimation error (average value of difference between the true mean of the process and the estimation). The mean of the estimation error increases for all the methods as the variability of the samples, environment standard deviation (SD), increases. The lowest mean estimation error within all blocks of the experiment belongs to the SMiLe rule (dark red color).

Take-home Message

- First, the belief update according to the SMiLe rule is a plausible learning rule (especially for learning within changing environments) in the sense that a piece of data becomes less surprising after taking it into account during learning. Moreover, among all other plausible learning rules that return posteriors that are not too divergent from the prior, it *maximizes learning* in the sense that the feeling of unexpectedness is maximally reduced, when facing the same data in the future.

http://lcn1.epfl.ch
mohammadjavad.faraji@epfl.ch