Ortega and Braun *Complex Adaptive Systems Modeling* 2014, **2**:4
http://www.casmodeling.com/content/2/1/4

Complex Adaptive Systems Modeling
a SpringerOpen Journal

ERRATUM Open Access

Erratum: Generalized Thompson sampling for sequential decision-making and causal inference

Pedro A Ortega[1*] and Daniel A Braun[2*]

*Correspondence:
ope@seas.upenn.edu;
daniel.braun@tuebingen.mpg.de
[1]GRASP Laboratory, Electrical and Systems Engineering Department, University of Pennsylvania, Philadelphia, PA 19104 USA
[2]Max Planck Institute for Biological Cybernetics and Max Planck Institute for Intelligent Systems, Speemanstrasse 38, Tübingen 72076 Germany

Decisions in the presence of latent variables

We correct errors in equations (14), (15) and (19) of the main text.

Equations (14) and (15)

Nature's probability of flipping either coin does not actually depend on the agent's prediction, so we can replace the conditional probabilities $p_0(\theta|x)$ by $p_0(\theta)$. We have then an inner variational problem:

$$\arg\max_{\tilde{p}(\theta|x)} \sum_{\theta} \tilde{p}(\theta|x) \left[-\frac{1}{\beta} \log \frac{\tilde{p}(\theta|x)}{p_0(\theta)} + U(x,\theta) \right] \tag{14}$$

with the solution

$$p(\theta|x) = \frac{1}{Z_\beta(x)} p_0(\theta) \exp\left(\beta U(x,\theta)\right) \tag{15}$$

and the normalization constant $Z_\beta(x) = \sum_\theta p_0(\theta) \exp\left(\beta U(x,\theta)\right)$ and an outer variational problem as described by equation (16) in the main text. Note that deliberation renders the two variables x and θ dependent.

Equation (19)

In the case of $\alpha = \beta$ and uniform prior $p_0(x) = \mathcal{U}(x)$, equation (17) reduces to

$$p(x) = \sum_\theta p_0(\theta) \frac{e^{\alpha U(x,\theta)}}{Z_\alpha}, \tag{19}$$

where $Z_\alpha = \sum_x \sum_\theta p_0(\theta) e^{\alpha U(x,\theta)}$. Note that $e^{\alpha U(x,\theta)}/Z_\alpha$ is in general not a conditional distribution. However, equation (19) can be equivalently rewritten as

$$p(x) = \sum_\theta \frac{p_0(\theta) \sum_{x'} e^{\alpha U(x',\theta)}}{Z_\alpha} \frac{e^{\alpha U(x,\theta)}}{\sum\limits_{x'} e^{\alpha U(x',\theta)}} = \sum_\theta p(\theta) p(x|\theta),$$

where we have expanded the fraction by $\sum_{x'} e^{\alpha U(x',\theta)}$.

This last equality can also be obtained by stating the same variational problem in reverse causal order of x and θ, which is the natural statement of the Thompson sampling problem. The nested variational problem then becomes

$$\arg\max_{\tilde{p}(x,\theta)} \sum_\theta \tilde{p}(\theta)\left[-\frac{1}{\beta}\log\frac{\tilde{p}(\theta)}{p_0(\theta)} + \sum_x \tilde{p}(x|\theta)\left[U(x,\theta) - \frac{1}{\alpha}\log\frac{\tilde{p}(x|\theta)}{p_0(x)}\right]\right]$$

with the solutions

$$p(x|\theta) = \frac{p_0(x)e^{\alpha U(x,\theta)}}{\sum_{x'} p_0(x')e^{\alpha U(x',\theta)}} \tag{i}$$

and

$$p(\theta) = \frac{1}{Z_{\beta\alpha}} p_0(\theta) \exp\left(\frac{\beta}{\alpha}\log\sum_x p_0(x)e^{\alpha U(x,\theta)}\right) \tag{ii}$$

with normalization constant $Z_{\beta\alpha} = \sum_\theta p_0(\theta)\exp\left(\beta/\alpha \, \log \sum_x p_0(x)e^{\alpha U(x,\theta)}\right)$. In the limit $\alpha \to \infty$ and $\beta \to 0$, the Thompson sampling agent is determined by the solutions $p(x|\theta) = \delta(x - \arg\max_{x'} U(x',\theta))$ and $p(\theta) = p_0(\theta)$. Sampling an action from $p(x) = \sum_\theta p(\theta)p(x|\theta)$ is much cheaper than sampling an action from equation (18) because of the reversed causal order in θ and x, which implies that $\beta/\alpha \to 0$ in equation (ii) instead of $\alpha/\beta \to \infty$ as in equation (17).

In the case of $\alpha = \beta$ the solutions for the two different causal orders of x and θ are equivalent. Assuming again a uniform prior $p_0(x) = \mathcal{U}(x)$, we can compute the Thompson sampling agent from equation (i) and equation (ii) for $\alpha = \beta$ to be

$$p(x) = \sum_\theta p(\theta)p(x|\theta) = \sum_\theta \frac{p_0(\theta)\sum_{x'} e^{\alpha U(x',\theta)}}{\sum_{x'}\sum_{\theta'} p_0(\theta')e^{\alpha U(x',\theta')}} \frac{e^{\alpha U(x,\theta)}}{\sum_{x'} e^{\alpha U(x',\theta)}},$$

which is exactly equivalent to $p(x)$ in equation (19). To sample from equation (19), we draw $\theta \sim p_0(\theta)$ and accept $x \sim p_0(x) = \mathcal{U}(x)$ if $u \leq e^{\alpha U(x,\theta)}/e^{\alpha T}$, where $u \sim \mathcal{U}[0;1]$.

Received: 24 June 2014 Accepted: 7 August 2014
Published online: 01 October 2014

References

Ortega PA, Braun DA: **Generalized Thompson sampling for sequential decision-making and causal inference.** *Complex Adaptive Systems Modeling* 2014, **2:**2.

doi:10.1186/s40294-014-0004-x
Cite this article as: Ortega and Braun: Erratum: Generalized Thompson sampling for sequential decision-making and causal inference. *Complex Adaptive Systems Modeling* 2014 **2**:4.