CORRIGENDUM

LEARNING TO SIGNAL WITH PROBE AND ADJUST — CORRIGENDUM

BRIAN SKYRMS

The author wishes to make the following correction.

In 'Learning to Signal with Probe and Adjust', I said: 'Note that this system state enables us to calculate the pay-offs that they got last time they did something' (142). I should have said '...this system state constrains the pay-offs that they got last time they did something to an extent sufficient to establish that signalling systems are the only absorbing states, and that there is a positive path from any state to a signalling system'.

The point is that when the sender pools, pay-offs last time may be underdetermined. We are dealing with a random, time-inhomogeneous Markov chain rather than a time-homogeneous chain. Nevertheless, the proof that Probe and Adjust learns to signal with probability one proceeds just as before.

Given a positive probability path from each state to an absorbing state, there is a maximum path length, n, and a minimum path probability, e. Starting from any state, the probability of not being absorbed after n probes is at most $(1-e)$. After $m*n$ probes, the probability of not being absorbed is $(1-e)^n$. In the limit, the probability of not being absorbed is 0.

REFERENCE

Skyrms, Brian. 2012. 'Learning To Signal With Probe And Adjust.' *Episteme*, 9: 139–50.
doi:10.1017/epi.2012.5.