ϵ-optimal AIS is computable

Daniel Filan, daniel.filan@anu.edu.au

First, we define the AIS model by putting the speed prior into the AIXI model:

$$\dot{a}_k = \arg\max_{a_k} \lim_{m\to\infty} \max_a \sum_e (\gamma_k r_k + \cdots + \gamma_m r_m) S(ae_{k:m}|ae_{<k})$$

where the γ_t are discount factors such that

$$\sum_{t=1}^{\infty} \gamma_t < \infty$$

Here, we show that AIS is computable to within ϵ accuracy: that is, that the value gained by this approximation may be up to ϵ less than the value that AIS expects to get.

To do this, note that since S and rewards are bounded, there exists some $n > k$ such that

$$\begin{aligned}\lim_{m\to\infty} \max_a \sum_e (\gamma_n r_n + \cdots + \gamma_m r_m) S(ae_{k:m}|ae_{<k}) &\leq C \sum_{t=n}^{\infty} \gamma_t \\ &\leq \epsilon\end{aligned}$$

Therefore, we can simply compute

$$\dot{a_k} = \arg\max_{a_k} \max_a \sum_e (\gamma_k r_k + \cdots + \gamma_n r_n) S(ae_{kn} : ae_{<k})$$

and the value gained by $\dot{a_k}$ can be no less than ϵ less than the value gained by the optimal policy. Therefore, AIS is computable to within ϵ accuracy, assuming that $\sum_{t=n}^{\infty} \gamma_t$ is computable.