

Let x_i be options present in the environment; in our scenario, $i \in \{0, 1\}$. Let $q_T[x_i]$ denotes the estimated quality for option i at T , we have:

$$q_T[x_0] = 1 - q_T[x_1]$$

At time T , suppose that the agent senses option x_0 , or received an opinion signal for option x_0 , the agent increments the quality of the option by ϵ , the qualities are then normalized:

$$q_{t+1}[x_0] = \frac{q_t[x_0] + \epsilon}{q_t[x_0] + \epsilon + q_t[x_1]} = \frac{q_t[x_0] + \epsilon}{q_t[x_0] + \epsilon + 1 - q_t[x_0]} = \frac{q_t[x_0] + \epsilon}{\epsilon + 1}$$

$$q_{t+1}[x_1] = \frac{q_t[x_1]}{q_t[x_0] + \epsilon + q_t[x_1]} = \frac{q_t[x_1]}{q_t[x_0] + \epsilon + 1 - q_t[x_0]} = \frac{q_t[x_1]}{\epsilon + 1}$$

Let r_{x_i} be the absolute difference between qualities of the previous and current timestep:

$$r_{x_0} = q_{t+1}[x_0] - q_t[x_0] = \frac{q_t[x_0] + \epsilon - (q_t[x_0] * \epsilon) - q_t[x_0]}{\epsilon + 1} = \frac{\epsilon - (q_t[x_0] * \epsilon)}{\epsilon + 1}$$

$$\implies \frac{d(r_{x_0})}{d(q_t[x_0])} = \frac{-\epsilon}{\epsilon + 1} < 0, \forall \epsilon > 0 \quad (1)$$

$$r_{x_1} = q_t[x_1] - q_{t+1}[x_1] = \frac{(q_t[x_1] * \epsilon) + q_t[x_1] - q_t[x_1]}{\epsilon + 1} = \frac{(q_t[x_1] * \epsilon)}{\epsilon + 1}$$

$$\implies \frac{d(r_{x_1})}{d(q_t[x_1])} = \frac{\epsilon}{\epsilon + 1} > 0, \forall \epsilon > 0 \quad (2)$$

(1) \Rightarrow the higher a quality is, the smaller subsequent increases to it is.

(2) \Rightarrow the lower a quality is, the smaller subsequent decreases to it is.

At time $T + 1$, suppose now that the agent senses/receives signal for option x_1 :

$$q_{t+2}[x_0] = \frac{q_{t+1}[x_0]}{q_{t+1}[x_0] + q_{t+1}[x_1] + \epsilon} = \frac{q_{t+1}[x_0]}{\epsilon + 1} = \frac{q_t[x_0]}{(\epsilon + 1)^2} + \frac{\epsilon}{(\epsilon + 1)^2} \quad (3)$$

$$q_{t+2}[x_1] = \frac{q_{t+1}[x_1] + \epsilon}{q_{t+1}[x_0] + q_{t+1}[x_1] + \epsilon} = \frac{q_{t+1}[x_1] + \epsilon}{\epsilon + 1} = \frac{q_t[x_1]}{(\epsilon + 1)^2} + \frac{\epsilon}{(\epsilon + 1)} \quad (4)$$

Observe that, compared to time T , by $T + 2$, the amount that the quality of the option x_0 sensed at $T + 1$ gets incremented by is $\epsilon + 1$ less than that of the option x_1 sensed at $T + 2$ (4,3) \Rightarrow gives more weight to the more recently sensed/received options.