# Appendix A.
# List of all symbols and functions used

| Variable | Description | Value |
|---|---|---|
| ***Markov decision process variables*** | | |
| $CY_t$ | The $t$'th cycle. Each cycle consists of a resource encounter followed by an extrinsic event | |
| $Q$ | The set of all possible resource qualities | Normal distribution range: [-20,20] Step size: 0.2 |
| $\Pr(q\|\mu_{resource}, \sigma_{resource})$ | The probability distribution over possible resource qualities, given the mean and standard deviation of resource quality in the environment. | |
| $E$ | The set of all possible extrinsic event values | Normal distribution Range: [-20,20] Step size: 0.2 |
| $\Pr(e\|\mu_{extrinsic}, \sigma_{extrinsic})$ | The probability distribution over possible extrinsic event values, given the mean and standard deviation of extrinsic events in the environment. | |
| $B$ | The set of all possible somatic states an agent can be in | Range: [0,10] Step size: 0.2 |
| $b_t$ | An agent's somatic state at the start of cycle $t$ | |
| $b_0$ | The somatic state when starting the first cycle. | |
| $b_{t=\infty}$ | The somatic state at the end of life. | |
| ***Priors, cues, and posteriors*** | | |
| $c$ | The cost (i.e., reduction in somatic state) of sampling a single cue | 0.2 |
| $c_-$ | A negative cue | |
| $c_+$ | A positive cue | |
| $n_{c-}$ | The number of negative cues observed during the current resource encounter | |
| $n_{c+}$ | The number of positive cues observed during the current encounter | |
| $D$ | The set of all cues observed during the current resource encounter | |
| $\vec{D}$ | A set-of-sets. Specifically, the set of all possible sets of observed cues | |
| $\mathrm{B}(q\|D)$ | The posterior belief that a resource quality is $q$ after an agent observed the set of cues $D$ | |
| $\mathrm{B}(q\|\emptyset)$ | The prior belief that a resource quality is $q$. | |
| $\mathrm{N}(c_-\|q)$ , $\mathrm{N}(c_+\|q)$ | The non-normalized cue reliability; the probability of sampling a negative or positive cue, respectively, given the resource quality is $q$. These values do not sum to 1. | |
| $\Pr(c_-\|q)$ , $\Pr(c_+\|q)$ | The normalized cue reliability: the probability of sampling a negative cue or positive cue, respectively, given that the resource quality is $q$. These probabilities do sum to 1. | |
| $\Pr(c_-\|D)$ , $\Pr(c_+\|D)$ | The probability that sampling results in a negative or positive cue, respectively, after having observed the set of cues $D$ . | |
| $ca$ | The cue accuracy: a standard deviation that determines how reliable cues are | 15 |

| | | |
|---|---|---|
| $\Pr(D\|q)$ | The probability of observing a set of cues, $D$, when the resource quality is $q$ | |
| ***Utility and fitness*** | | |
| $O_{immediate}(CY_t\|b_t)$ | The immediate expected outcome (i.e., change in somatic state) of the $t$'th cycle given that an agent starts that cycle with a somatic state of $b_t$ | |
| $O_{total}(CY_t\|b_t)$ | The total expected outcome of the $t$'th cycle – the sum of all future changes in somatic state, given the outcome of the current cycle | |
| $U(CY_t\|b_t)$ | The expected utility of the outcome of a cycle, given that the cycle starts with $b_t$ – that is, the expected lifetime change in somatic state including the current cycle. | |
| $\lambda$ | The discount rate | 0.95 |
| $\omega(b_{t=\infty})$ | The fitness function associated with having a somatic state $b_{t=\infty}$ at the end of life. | |
| $\alpha$ | The shape of the fitness function | 0.4 (diminishing fitness returns); 1 (linear fitness returns); 2.5 (increasing fitness returns) |
| $\beta$ | A normalizing constant for the fitness function | |
| ***Environmental variables*** | | |
| $\mu_{resource}$ | Mean resource quality, a kind of harshness | Range: [-5,5]<br>Step size: 0.5 |
| $\sigma_{resource}$ | Standard deviation of resource quality, a kind of unpredictability | Range: [0,8]<br>Step size: 0.5 |
| $\mu_{extrinsic}$ | Mean extrinsic event quality, a kind of harshness | Range: [-5,5]<br>Step size: 5 |
| $\sigma_{extrinsic}$ | Standard deviation of extrinsic event quality, a kind of unpredictability | Range: [0,8]<br>Step size: 4 |
| $\rho$ | Interruption rate – the rate at which resources become unavailable after accepting. A kind of unpredictability | Range: [0,0.5]<br>Step size: 0.25 |
| $env$ | An environment. Environments are vectors of size 5 containing a $\mu_{resource}$, $\mu_{extrinsic}$, $\sigma_{resource}$, $\sigma_{extrinsic}$, and $\rho$ | |
| $ENV$ | The set of all environments | |
| ***Within cycle actions, states, and policies*** | | |
| $S$ | The set of all possible states | |
| $S^{start}$ | The set of all possible starting states | |
| $S^{transition}$ | The set of all transition states an agent enters after accepting or rejecting. A terminal state denotes the end of a current cycle, and the start of the next cycle | |
| $S^{final}$ | The set of all final state. Final states are the terminal states of the first cycle. | |
| $S^{sample}$ | The set of all non-terminal 'sampling' states in which the best action includes sampling | |

| | |
|---|---|
| $S^{leaf}$ | The set of all non-terminal 'leaf' states in which the best action does not include sampling |
| $s$ | The current state of an agent, consisting of a somatic state $b$, a set of observed cues $D$, and the number of the current cycle $t$ |
| $s'$ | A successor state of an agent; a potential future state that follows the current state of the current cycle $t$ |
| $s_{dead}$ | The state an agent enters when its somatic state reaches 0 |
| $b(s)$ | The somatic state of state $s$ |
| $D(s)$ | The set of observed cues in state $s$ |
| $t(s)$ | The cycle number of state $s$ |
| $A(s)$ | The set of all possible actions by an agent in state $s$ |
| $A^*(s)$ | The set of all possible fitness-maximizing actions by an agent in state $s$ |
| $P(s'\|s,a)$ | The transition function. Contains the probability of moving from state $s$ to state $s$' after taking action $a$ |
| $\hat{P}(s'\|s,a)$ | For an infinite number of agents following the same policy, the proportion of agents that transition from $s$ to $s$' |
| $successorState(s,q,e)$ | The state that an agent will be in after it has received a resource of quality $q$ and extrinsic event $e$ when it is in state $s$. This state is always an transition state |
| $parents(s)$ | The parents of state $s$. A parent is a state $s$' from which it is possible to reach $s$ |
| $prop(s)$ | For an infinite number of agents following the same policy, the proportion of agents that visit state $s$ |
| $O_{immediate}(s_{transition}\|s_{start})$ | The expected change in somatic state (i.e., the immediate outcome) when an agent starts the cycle in $s_{start}$ and ends in $s_{transition}$. |
| $U(s_{transition}\|s_{start})$ | The expected utility when an agent starts the cycle in $s_{start}$ and ends in $s_{transition}$.. That is, the expected lifetime change in somatic state including the current cycle. |
| $U(s\|a)$ | The expected utility of being in state $s$ and performing action $a$ |
| $U(s)$ | The expected utility of being in state $s$, assuming that an agent selects the action with the highest utility in that state |
| $\pi$ | A policy that maps states to actions |
| $\pi^*$ | The optimal policy; the policy that maximizes fitness |

***Value iteration***

| | |
|---|---|
| $v_u^k(s_{transition})$ | The utility value function for cycle $k$. Provides the expected outcome of all future $k$ cycles, given that an agent ends the current cycle in state $s_{transition}$ |
| $v_u^*(s_{transition})$ | The optimal utility value function. The expected outcome of all future cycles given that the agent follows the optimal policy and ends the current cycle in state $s_{transition}$ |
| $v_f^*(s_{final})$ | The optimal fitness value function. The expected fitness at the end of life given that the agent follows the optimal policy and ends the first cycle in state $s_{final}$ |

| | | |
|---|---|---|
| $\delta(v_u^{k=x}, v_u^{k=x-1})$ | Delta: the largest absolute difference between the utility of starting states with the same somatic state after iterating one more cycle | |
| $\epsilon$ | The convergence criterion: the largest delta of significance. If delta is lower than epsilon the algorithm has converged. | 0.001 |