Appendix 1:
Dynamic Programming Equations

Balancing sampling and specialization:
An adaptationist model of incremental development

Willem E. Frankenhuis; Karthik Panchanathan
UCLA Department of Anthropology, 341 Haines Hall,
Box 951553, Los Angeles, CA, 90095-1553, USA.

State variables

$X_0(t) \equiv$ the number of cues sampled indicating a *world–0* at time t

$X_1(t) \equiv$ the number of cues sampled indicating a *world–1* at time t

$Y_0(t) \equiv$ the number of steps towards the *world–0* phenotype at time t

$Y_1(t) \equiv$ the number of steps towards the *world–1* phenotype at time t

Where t denotes the current time period. The organism makes decisions in time period 1 to time period 20. T denotes the end of development, $t = 21$.

For a given time period t, these state variables take on the values x_0, x_1, y_0, y_1.

Parameters

$P(w_0) \equiv$ the probability of being born into *world–0* across evolutionary time

$P(w_1) \equiv$ the probability of being born into *world–1* across evolutionary time

$$P(w_0) + P(w_1) = 1$$

$P(c_0|w_0) \equiv$ the probability of a cue indicating *world–0* when in *world–0*

$P(c_1|w_1) \equiv$ the probability of a cue indicating *world–1* when in *world–1*

$$P(c_0|w_0) + P(c_1|w_0) = 1$$
$$P(c_1|w_1) + P(c_0|w_1) = 1$$

We assume that these cue validities are symmetric in the following way:

$$P(c_0|w_0) = P(c_1|w_1)$$

Posterior beliefs

After each cue sampled, the organism updates its belief about the state of the world. These beliefs are computed by Bayes' theorem, using the evolutionary prior probabilities $\{P(w_0), P(w_1)\}$, the cue validities $\{P(c_0|w_0), P(c_1|w_1)\}$, and the set of cues the organism has sampled thus far $\{x_0, x_1\}$.

Letting *D* denote the set of cues sampled, where $D = \{x_0, x_1\}$, we can compute the probability of observing this set of cues:

$$P(D) = P(D|w_0)P(w_0) + P(D|w_1)P(w_1)$$

We can compute the probabilities of obtaining D in either *world–0* or *world–1* by:

$$P(D|w_0) = \binom{x_0 + x_1}{x_0} P(c_0|w_0)^{x_0} P(c_1|w_0)^{x_1}$$
$$P(D|w_1) = \binom{x_0 + x_1}{x_0} P(c_1|w_1)^{x_1} P(c_0|w_1)^{x_0}$$

Next, we denote the posterior beliefs, which will be functions of the sampled cues:

$b_0(D) \equiv$ the posterior belief of being in *world–0*

$b_1(D) \equiv$ the posterior belief of being in *world–1*

Because there are only two states of the world, the organism's beliefs in these two states must sum to one, $b_0(D) + b_1(D) = 1$.

Using Bayes' theorem, we can compute these posterior beliefs:

$$b_0(D) = \frac{P(D|w_0)P(w_0)}{P(D|w_0)P(w_0) + P(D|w_1)P(w_1)}$$

$$b_1(D) = \frac{P(D|w_1)P(w_1)}{P(D|w_0)P(w_0) + P(D|w_1)P(w_1)}$$

Fitness of mature phentoypes

The developing organism makes decisions to sample or specialize in each time period. There are no fitness consequences to these decisions *during* development. Instead, fitness is accrued *after* development, determined by the correspondence between the mature phenotype and environmental state.

Let $\phi(x_0, x_1, y_0, y_1)$ denote the expected lifetime fitness of a mature organism, having sampled a set of cues $\{x_0, x_1\}$ and developed a phenotype $\{y_0, y_1\}$.

The expected lifetime fitness (ϕ) is the sum of two products: the probability of being in *world–0* (b_0) multiplied by the fitness associated with the realized *world–0* phenotype (y_0), added to the probability of being in a *world–1* (b_1) multiplied by the fitness associated with the realized *world–1* phenotype (y_1).

The function specifying how degree of phenotypic specialization translates into fitness can take one of three forms: linear, diminishing, and increasing.

Linear returns

$$\phi(x_0, x_1, y_0, y_1) = b_0\ y_0 + b_1\ y_1$$

The maximum fitness is $T - 1 = 20$, as there are 20 developmental decisions.

Diminishing returns

$$\phi(x_0, x_1, y_0, y_1) = b_0\ \alpha(1 - e^{-\beta y_0}) + b_1\ \alpha(1 - e^{-\beta y_1})$$

We set β = 0.2, where β determines the deceleration of the curve. To ensure that the maximum fitness is 20, we set $\alpha = \frac{T-1}{1-e^{-\beta(T-1)}}$.

Increasing returns

$$\phi(x_0, x_1, y_0, y_1) = b_0\ \alpha(e^{\beta y_0} - 1) + b_1\ \alpha(e^{\beta y_1} - 1)$$

We set β = 0.2, where β determines the acceleration of the curve. To ensure that the maximum fitness is 20, we set $\alpha = \frac{T-1}{e^{\beta(T-1)}-1}$.

Decisions during development

In all time periods, except the last one (T), the organism chooses between three options: sample a cue to the state of the world, specialize one increment toward the *world–0* phenotype, or specialize one increment toward the *world–1* phenotype.

The organism chooses the option with the highest expected fitness. In the event of a tie between two or three choices, the organism chooses amongst the alternative optimal choices with equal probability.

Whereas choosing to specialize toward a phenotypic target occurs with certainty, choosing to sample a cue to the state of the world can result in one of two outcomes (c_0 or c_1) depending on the belief of being in *world–0*. Because $b_0(D) + b_1(D) = 1$, we only need to track one posterior belief.

$P_s(c_0) \equiv$ the probability of sampling a cue to *world–0* if the organism samples

$P_s(c_1) \equiv$ the probability of sampling a cue to *world–1* if the organism samples

$$P_s(c_o) = P(c_o|w_o)\, b_0 + P(c_o|w_1)\, b_1$$
$$P_s(c_1) = P(c_1|w_o)\, b_0 + P(c_1|w_1)\, b_1$$

Let $F(x_0, x_1, y_0, y_1, t)$ denote the maximum expected value of the lifetime fitness function based on decisions made between time period t and T, when the organism has thus far sampled cues $\{x_0, x_1\}$ and a phenotype $\{y_0, y_1\}$.

$$F(x_0, x_1, y_0, y_1, t) = \\ \max E \left\{ \phi\big(X_0(T), X_1(T), Y_0(T), Y_1(T)\big) \middle| \begin{matrix} X_0(t) = x_0 \\ X_1(t) = x_1 \\ Y_0(t) = y_0 \\ Y_1(t) = y_1 \end{matrix} \right\}$$

After development, in time period T, the fitness of the mature organism will be:

$$F(x_0, x_1, y_0, y_1, T) = \phi(x_0, x_1, y_0, y_1)$$

In each time period during development ($t < T$), the organism makes the optimal decision (i.e., the decision that maximizes expected lifetime fitness). The fitnesses associated with each choice in time period t are denoted by:

$$F_0(x_0, x_1, y_0, y_1, t) = F(x_0, x_1, y_0 + 1, y_1, t + 1)$$

$$F_1(x_0, x_1, y_0, y_1, t) = F(x_0, x_1, y_0, y_1 + 1, t + 1)$$

$$E\big[F_s(x_0, x_1, y_0, y_1, t)\big] = \begin{matrix} P_s(c_o)F(x_0 + 1, x_1, y_0, y_1, t + 1) + \\ P_s(c_1)F(x_0, x_1 + 1, y_0, y_1, t + 1) \end{matrix}$$

The expected fitness for a given state is the expected fitness associated with the choice resulting in the highest expected lifetime fitness:

$$F(x_0, x_1, y_0, y_1, t) = \max \begin{bmatrix} F_0 \\ F_1 \\ E[F_s] \end{bmatrix}$$