

SIMULTANEOUS AND SEQUENTIAL CHOICE AS A FUNCTION OF REWARD DELAY AND MAGNITUDE: NORMATIVE, DESCRIPTIVE AND PROCESS-BASED MODELS TESTED IN THE EUROPEAN STARLING (*Sturnus vulgaris*)

MARTIN S. SHAPIRO, STEVEN SILLER, ALEX KACELNIK

2008

Una oportunidad puede rechazarse si aceptarla implica un compromiso temporal que impida aprovechar el resto de la riqueza del ambiente. Se puede modelar como la elección entre la oportunidad actual y otras oportunidades de fondo con las cuales la fuente a analizar compete temporalmente.

La aproximación normativa comienza preguntándose qué resultado sería favorecido por la selección natural, y solo se pregunta después por los mecanismos que le darían lugar. Va de un valor objetivo de recompensa a una regla hipotética de decisión. Se le ha llamado “racionalidad biológica” (Kacelnik, 2006).

La forma en que se modela la asignación de valor a las alternativas varía con cada disciplina. La ecología conductual va del resultado de la selección natural a los mecanismos de decisión. La economía también es normativa, pero no deriva sus predicciones de un principio general como la selección natural, sino de la expectativa de coherencia y la observación de las decisiones. El análisis de la conducta es más extremo en su empirismo y deriva sus modelos de datos, sin referirse a teorías de maximización o coherencia; los modelos se evalúan solo en términos de su ajuste con los datos.

Además de modelar la asignación de valor, es necesario describir cómo ésta se traduce en conducta. La preferencia exclusiva por la alternativa más valiosa difícilmente se da en la realidad.

El propósito del artículo es usar sus datos empíricos para discutir la adecuación y justificación de varios modelos.

PROCEDIMIENTO

Usaron tres tipos de ensayos: *choice*, *no-choice* y *probe*. En *choice* se encendían ambas alternativas; en *no-choice*, solamente una; y *probe* eran ensayos de pico para probar el “conocimiento” de los animales sobre la demora. Las sesiones se componían de 12 bloques

de 3 ensayos: uno no-choice con el símbolo A, uno con B, y uno de elección. Hubo 6 ensayos de probe por sesión distribuidos semialeatoriamente.

Las aves pasaron por 15 “tratamientos” de 6 a 20 sesiones de extensión. Cada tratamiento tenía combinaciones particulares de magnitud y demora.

Había un total de 6 símbolos, cada uno con demoras y magnitudes diferentes (en pellets/demora en segundos, las combinaciones eran 2/5, 4/10, 1/5, 2/10, 1/10 y 2/20). Hacer combinaciones binarias entre todos los símbolos resulta en los 15 “tratamientos”. Nótese que las alternativas forman pares con la misma relación entre magnitud y demora. Se pueden formar tres sets de comparaciones: (1) uno en el que la razón magnitud/demora es la misma entre ambas alternativas, (2) otro en el que A da una razón dos veces mayor que B, y (3) otro en el que A da una razón cuatro veces mayor que B.

Tras cada tratamiento se aplicó una prueba de preferencia sin reforzamiento. Por 5 minutos se dejaba a los animales picar ambas alternativas sin ninguna consecuencia programada y solo se registraban las respuestas, se apagaban las teclas por 5 minutos, se encendían 5 más, se apagaban de nuevo por 5 minutos, y se encendían finalmente 5 minutos más.

RESULTADOS

El procedimiento de pico mostró buena discriminación temporal. Había picos más altos en las alternativas de mayor magnitud, pero estos decaían con mayor rapidez.

Sobre preferencia, hubo diferencias significativas entre sets, lo que indica que mientras aumentaba la diferencia entre las relaciones magnitud/demora, aumentaba la preferencia por la alternativa más rica. Además, se encontró una diferencia dentro del set 1, en el que la relación era la misma para ambas alternativas. Se hicieron pruebas T contra indiferencia: en el set 1 ningún tratamiento resultó en preferencias distintas de la indiferencia; pero en 2 y 3, todos lo hicieron.

Sobre la prueba no-reforzada, se encontraron preferencias similares, pero menos extremas. Eran estadísticamente distintas de las preferencias originales calculadas durante los tratamientos. La diferencia no era atribuible a la extinción, puesto que, aunque la tasa de respuestas a ambas alternativas disminuyó a lo largo de la prueba, la relación entre la tasa de respuestas a ambas alternativas se mantuvo constante.

Las diferencias entre sets para estas preferencias evaluadas con pruebas no-reforzadas fueron mucho menos marcadas que las diferencias reportadas previamente.

Todos los tratamientos de los sets 2 y 3, y un tratamiento del set 1, fueron distintos de la indiferencia.

Se encontró que las latencias de respuesta dependían no solamente de la relación demora/magnitud de una alternativa, sino también de la relación de la alternativa complementaria, incluso cuando ésta no se encontraba presente (en los ensayos *no-choice*). Sucedió el mismo efecto en las pruebas no-reforzadas, a pesar de que en ellas no se registró la latencia sino la frecuencia de respuestas. La frecuencia de respuestas emitidas ante una opción dependía de la alternativa con la que se le comparaba.

DISCUSIÓN

Queda claro que los animales “sabían” la longitud de la demora y eran sensibles a la magnitud. Los *starlings* no mostraron la preferencia casi exclusiva que sería predicha por

políticas de maximización, sino preferencias graduales.

La latencia y frecuencia de respuesta usadas como medidas de preferencia resultaron ser sensibles tanto a la relación magnitud/demora de la alternativa evaluada como a la relación de su complemento. Es de interés la relación cuantitativa de magnitud y demora con la conducta y elección resultantes.

DISCUSIÓN GENERAL

MODELOS DE PREFERENCIA

Agrupar los modelos en normativos, análisis de la conducta (con guía empírica) y basados en procesos.

Modelos normativos

Se basan en la presunción a priori de que hay una variable relacionada con la adaptación (*fitness*) que se debe maximizar. La energía ganada por unidad de tiempo suele tomarse como punto de partida. Se basa en la idea de que todo mecanismo conductual heredable que dé más beneficios deberá tener más éxito evolutivo e invadirá el *gene pool* de la especie. De forma general, estos modelos estarán dados por

$$V_i = \frac{A_i}{T_i}$$

donde V_i es el valor asignado a una alternativa, A_i es la energía neta que otorga, y T_i es el tiempo esperado para obtener la recompensa. En este caso específico, T_i se descompone en intervalo entre ensayos (ITI), latencia (L_i) y demora de reforzamiento (D_i):

$$V_i = \frac{A_i}{ITI + L_i + D_i}$$

Siguiendo este modelo, la mejor estrategia sería mostrar latencia cero y elegir exclusivamente la alternativa con mayor V_i , cosa que no sucede.

Una forma de adecuar el modelo sería agregar una regla que haga no-determinista la relación entre el valor y la elección. Esta regla podría ser la ley de igualación, porque aunque es probabilista, favorece a las alternativas con alto valor, pero sin ser una regla de maximización.

Nota importante: proponen evitar usar parámetros ajustados para mejorar la exactitud de los modelos. En su lugar, plantean una aproximación más conservadora que sea estricta con los modelos.

La perspectiva del análisis de la conducta

En este caso se buscan algoritmos con buen desempeño descriptivo sin referencia a problemas extraconductuales como las consecuencias evolutivas. En lugar de analizar el problema como una cadena infinita de elecciones donde cada una lleva el costo de privar de las demás, el AC ve el problema de forma molecular como una única elección determinada por el valor sumado por las magnitudes y restado por las demoras. De nuevo, la base puede ser la igualación (que es una regularidad conductual observada y no una inferencia normativa). Un modelo adecuado sería, entonces, uno empírico y no normativo, como el modelo de Mazur:

$$V_i = \frac{v_i}{(1 + kD_i)}$$

donde V es el valor de la recompensa demorada, v_i es su valor si no tuviese demora, k es una constante llamada “tasa de decaimiento del valor”, y D es la demora en segundos.

Modelamiento del proceso: elección secuencial

Parte de un razonamiento tanto normativo como empírico. Normativo debido a que reconoce que las elecciones en ambientes naturales no suelen ser entre alternativas concurrentes, sino entre perseguir y no perseguir una presa. Empírico porque se basa en observaciones cualitativas y cuantitativas: (1) contra la lógica de la maximización, los sujetos del experimento mostraron latencias considerables, (2) hay una relación ordenada entre las latencias y las contingencias (disminuyen cuando aumenta la relación entre magnitud y demora), y (3) otros experimentos en el área han indicado que la latencia podría ser una métrica sensible para la preferencia.

Este modelo asume distribuciones de latencias independientes, de las cuales solo la latencia más breve se expresa y “censura” a la otra. Se hacen dos predicciones: (1) se debe poder predecir la elección de ensayos simultáneos con los ensayos de no-elección, y (2) las latencias en ensayos de elección deberían ser menores que en ensayos de no elección.

Para predecir la proporción de elección tomaron la proporción de pares consecutivos de no-elección en los que hubo una menor latencia para la opción A.

CONCLUSIÓN

Los modelos que tratan con el juego entre magnitud y demora suelen tener dos partes: una regla de valor para determinar el valor subjetivo de una alternativa, y una regla de ejecución para traducir ese valor en conducta con variabilidad. Lo más común es suponer que lo estocástico está en la regla de ejecución y no en la de valor.

Todos los modelos hacen predicciones similares, así que la diferencia está en la bondad de ajuste. Es importante no caer en la trampa de creer que un modelo con más parámetros libres es mejor.

El modelo SCM tiene un gran ajuste sin parámetros libres (a menos que se cuente a la latencia como parámetro). Pero no se toma como explicación, sino como motor para generar más preguntas, por ejemplo, si los encuentros sucesivos o simultáneos son mejores candidatos como causantes de los mecanismos encontrados en la actualidad.

El razonamiento del modelo SCM es que la elección simultánea es en gran medida un constructo de laboratorio. Los animales responden a dos variables: la recompensa obtenible por involucrarse con la presa actual, y la recompensa promedio obtenible del resto del ambiente (como el error reducido y el error reducible por parámetro adicional). Así, los animales se tardarían más en aceptar alternativas que no impliquen una gran mejoría con respecto al resto del ambiente.

Hace falta determinar cómo magnitud y demora determinan la latencia. Una buena apuesta parece ser el análisis de la adquisición de fuerza asociativa por cada estímulo de cada alternativa.

Desde un punto de vista descriptivo, los datos apoyan un modelo hiperbólico derivado de forrajeo óptimo pero con la adición de la latencia; y desde un punto de vista normativo, apoyan la noción de que la mejor explicación de los mecanismos de elección es la aceptación y el rechazo de encuentros secuenciales y no la elección entre encuentros simultáneos.