Adaptación a Propiedades Estadísticas de Sucesos Biológicamente Importantes

Arturo Bouzas

Universidad Nacional Autónoma de México

Resumen

El presente capítulo presenta las ventajas conceptuales de entender el comportamento

En 1968 George Reynolds publicó el libro “A primer of operante conditioning”, su propósito era presentar de una forma concisa lo que hasta ese momento eran considerados los principios básicos del análisis experimental del comportamiento (AEC). El propósito del presente capítulo es presentar un bosquejo del contenido de ese libro si se escribiera el día de hoy, cuales serían los principios básicos núcleo de nuestra disciplina y cuales preguntas aun sin respuesta serían importante atender. Como cualquier disciplina la nuestra se transforma en respuesta a nueva evidencia, a la disponibilidad de nuevas metodologías y de análisis de datos y por su contacto con esquemas conceptuales de disciplinas cercanas. El mensaje principal del capítulo es que contrario a lo que se piensa en ciertos ámbitos tanto dentro como fuera del AEC, el estudio del comportamiento es una área madura con una enorme vitalidad y en constante transformación. El esquema conceptual brillantemente resumido por Reynolds (19..) tiene hoy una cara significativamente diferente {KilleenPeterR 2014}{Baum 2012} Staddon, 2016). El capítulo esta organizado de la siguiente forma. Primero revisaremos los cambios en el contexto filosófico que han facilitado un pluralismo en las formas de explicar el comportamiento. Un segundo apartado presentará argumentos a favor del estudio del comportamiento como una adaptación a la estructura estadística de su entorno. En el tercer apartado presentaremos los principales elementos de los modelos de integradores que subyace a los modelos de reforzamiento de la asignación de crédito que serán el tema del último apartado.

1. Cambios en el contexto empírico y conceptual del AEC.

Justo alrededor de los años que se publicó el libro de Reynolds se reportaron un gran número de fenómenos empíricos que cuestionaban varios de los principios más centrales del AEC, entre ellos destacan por su impacto sobre el desarrollo del área los resultados encontrados en programas concurrentes de refuerzo (Herrnstein, 196), los de bloqueo en protocolos de condicionamiento clásico (Kamin, 196), los del protocolo “aleatorio” (Rescorla, 196) los de automoldeamiento (Brown y Jenkins, 196), los de refuerzo de Premack (196), los de psicofísica del tiempo (Church, 19), y los obtenidos en protocolos de evitación desarrollados por Hineline and Herrnstein (Herrnstein, 19). Esta literatura empírica cristalizo en dos grandes modelos matemáticos alrededor de los cuales hasta la fecha ha girado buena parte de la investigación empírica y conceptual del comportamiento adaptable (Herrnstein, 1970; Rescorla y Wagner, 1969). Cabe destacar dos artículos teóricos (Bolles, 1972 y Bindra, 1974) que desde diferentes perspectivas cuestionaron la ley del efecto y pueden verse como antecedentes de los modelos recientemente propuestos por Baum {KilleenPeterR 2014}{Baum 2012}) y por Killeen (201) y la literatura sobre aprendizaje por refuerzo basado en modelos ().

Otro conjunto de desarrollos que impactaron al AEC se dieron en la Filosofía de la Ciencia y la consecuente apertura a esquemas conceptuales en otras disciplinas. La teoría psicológica surge tomada de la mano de la filosofía de la ciencia, la cual antes de los años 1960s sostenía que las teorías eran entidades lingüísticas (proposiciones) y que su relación con el mundo empírico era vía reglas de correspondencia entre un lenguaje teórico y lenguaje observacional. Dentro de esta perspectiva, los problemas de la “verdad” y la “referencia” eran el principal punto de discusión. Artículos sobre el “operacionalismo” y los argumentos sobre “variables interventoras” y “constructor hipotéticos” poblaron las revistas psicológicas.

Sin embargo, atención a la práctica cotidiana de los científicos llevó a los filósofos a atender otras formas de expresar las teorías y su relación con el mundo empírico. Los científicos cada vez más frecuentemente emplean otros “vehículos representaciones” diferentes a los lingüísticos. A estos vehículos los llaman modelos y los hay de una gran variedad. Incluyen en un extremo estructuras físicas como una maqueta, hasta en el otro extremo estructuras matemáticas abstractas. Las nociones de referencia y verdad no parecen aplicables a los modelos, la discusión la ocupa la noción de representación y similitud (Frigg y Hoffman, 2012).

Una característica importante de los modelos es que no tienen que ser una réplica perfecta de todo lo que ocurre en el mundo empírico. Consideré el mapa de una ciudad, este puede tener diferentes usos y dependiendo de estos incluirá o dejará fuera ciertos detalles. Un mapa muy útil para orientarse por sus calles elimina la mayoría de los detalles, el color de las casas por ejemplo, pero mantiene la relación espacial entre las calles, lo que permite usarlo para orientarnos en ellas. Sin embargo, ese mapa no sería útil para capturar la belleza y detalles arquitectónicos de la ciudad.

Al igual que los mapas, los modelos matemáticos muestran la flexibilidad necesaria para representar los diferentes intereses de un científico. Hay modelos que simplemente permiten predecir relaciones de entrada-salida entre variables, como por ejemplo los modelos psicofísicos tradicionales; mientras que otros intentan una representación precisa de lo que ocurre entre la entrada y salida, como por ejemplo los modelos de acumuladores en tareas de decisión perceptual.

Para una ciencia expresada en modelos, la relación entre teoría y el mundo empírico se basa en la noción de representación. La Filosofía de la Ciencia contemporánea ha hecho de la elucidación del concepto de representación uno de sus principales objetos de estudio (Van Fraassen, 2008 ). Atrás de casi todas las teorías de representación está la noción de un morfismo, esto es, la idea que una “buena” representación mantiene en el sistema representacional, las relaciones que son ciertas en el mundo empírico que pretende representar. En este esquema la evaluación de los modelos es una actividad pragmática alejada de la noción de “verdad”. Note que la representación no es la teoría, ésta consiste en especificar qué aspectos del mundo (dominio) pueden ser representados con esa estructura de representaciones particular.

Dentro de la Filosofía de las matemáticas, el enfoque dominante para entender modelos matemáticos es similar al anterior y se le conoce como la teoría del mapeo (Pincok, 2012). Más recientemente, dentro de las matemáticas y en general en filosofía de la ciencia ha habido un movimiento hacia una interpretación de representación en términos de “similitud” (van Fraassen, 2008); con este enfoque es posible hacer sentido de modelos que se sabe son ficticios y que sin embargo representan fenómenos empíricos y resultan muy útiles en el trabajo de la ciencia (para una completa reseña de la literatura filosófica contemporánea sobre modelos ver Frigg and Hoffman, 2012)

Finalmente, el modelo representacional no debe interpretarse como el supuesto de que hay una estructura matemática esperando que alguien encuentre una estructura empírica de la que sea un morfismo de ella. Una parte fundamental del trabajo científico es precisamente el establecimiento de dichas estructuras empíricas. Ello implica una guía teórica previa al proceso de modelase y substancial trabajo empírico.

Por último, debe mencionarse el importante papel que las simulaciones computacionales juegan hoy en día en el desarrollo y evaluación de los modelos científicos (Weisberg, 2013), así como las nuevas formas de analizar estadísticamente a los modelos (Lee, en prensa) .

La naturaleza de la explicación fue el segundo tema que dominó la discusión en la Filosofía de la Ciencia durante el último tercio del siglo pasado, en gran parte por la mayor atención que recibió el trabajo de biólogos, psicólogos y economistas. Hasta antes de ese momento la interpretación dominante de la explicación fue la teoría “nomológica deductiva (ND)” que suponía que explicar consistía en deducir lógicamente de leyes generales los fenómenos de interés. Su problema más notorio resulta más fácil de entender con el siguiente ejemplo paradigmático. La sombra de un asta bandera puede derivarse de su altura pero igualmente, con los mismos principios, la altura del asta bandera puede derivarse a partir del largo de la sombra. Este problema de “simetría” fue devastador para la teoría ND y obligó a considerar el papel que juega la “causalidad” en las explicaciones.

Tres interpretaciones de causalidad dominan la literatura: la primera destaca el papel de contrafactuales (qué hubiese pasado en la ausencia del factor explicativo) y es muy común en el trabajo empírico en las ciencias sociales. La segunda destaca el papel de las intervenciones para explicar causalmente la ocurrencia de un suceso o fenómeno; dentro de la Psicología, el AEC ejemplificado por Reynolds es uno de los mejores ejemplos de este tipo de explicación. Strevens (2008) presenta la interpretación causal más elaborada que integra contrafactuales e intervenciones. La teoría asume que las explicaciones buscan encontrar aquellos factores dentro de un mundo de relaciones causales que “hacen una diferencia”. Consiste en “hacer el modelo causal tan abstracto como sea posible, substituyendo las descripciones de las influencias muy específicas por descripciones cada vez más abstractas de las mismas”. Consideré su tratamiento del hecho ampliamente aceptado que muchas explicaciones son idealizaciones en las que la historia causal es incorrecta. De acuerdo a Strevens, estas explicaciones son aceptadas precisamente porque omiten aquello que no hace una diferencia.

No todos los filósofos se sienten cómodos con la noción de causalidad, entre ellos Kitcher (1989), propone la teoría de explicación alternativa de mayor influencia. La teoría de la unificación plantea que las teorías explican a través de unificar un conjunto de observaciones, que quedan subsumidas bajo un conjunto de principios básicos. De acuerdo a él, la ciencia busca “esquemas explicativos” que puedan ser aplicados tan ampliamente como sea posible. La teoría de la selección natural es el mejor ejemplo de un esquema explicativo.

No todos los filósofos se sienten cómodos con la noción de causalidad, entre ellos Kitcher (1989), quien propone la teoría de explicación alternativa de mayor influencia. La teoría de la unificación plantea que las teorías explican a través de unificar un conjunto de observaciones, que quedan subsumidas bajo un conjunto de principios básicos. De acuerdo a él, la ciencia busca “esquemas explicativos” que puedan ser aplicados tan ampliamente como sea posible. La teoría de la selección natural es el mejor ejemplo de un esquema explicativo.

Otros filósofos, aunque aceptan explicaciones de tipo causal, argumentan también que hay otras explicaciones que no son causales. Batterman y Rice (2014) presentan sucintamente esta posición. De acuerdo a ellos, a las diferentes variantes de las explicaciones causales se les puede identificar por incluir como característica fundamental todos los rasgos relevantes causales. Ellos argumentan que hay una clase de explicaciones que los son precisamente porque abandonan la pretensión de exactitud que considera todos los rasgos relevantes. Proponen modelos mínimos que explican, en virtud de un argumento, por qué un número de rasgos son irrelevantes. Los mejores ejemplos de estas explicaciones se encuentran dentro de la biología y la economía, en particular en los modelos de optimización de equilibrio. Estos modelos mínimos se aplican a un rango enorme de instancias con rasgos heterogéneos que pueden ser omitidos y que, sin embargo, dan cuenta de patrones muy generales de comportamiento. Dichos modelos son comunes en el AEC, sin embargo, su estructura no siempre es comprendida, lo que lleva a verlos como modelos provisionales que serán substituidos por modelos “verdaderamente” causales. Por el contrario, los modelos mínimos son robustos en el comportamiento de sistemas a gran escala, a pesar de grandes cambios en los detalles del sistema. Note que estas explicaciones dependen de encontrar patrones muy generales del comportamiento de sistemas muy diversos. En ocasiones, se requiere del genio de alguien como Darwin o Nash para encontrarlos.

Finalmente, el inicio del siglo XXI vio el resurgimiento de otra teoría causal de la explicación, conocida como “nuevo mecanicismo” (Craver, 2014), y que está muy relacionada con la Psicología y las neurociencias. En breve, “el mecanismo para un fenómeno consiste de entidades cuyas actividades e interacciones están organizadas de tal forma que producen el fenómeno”.

Cuando a los científicos se les toma desprevenidos y se les pregunta cuál es el propósito de su actividad, muchos de ellos dirán que es “explicar” o “modelar” cierto fenómeno. Sin embargo, cuando se les cuestiona acerca de qué entienden por explicar y modelar, encontraremos tantas respuestas como científicos, más aún, frecuentemente hablarán que la ventaja y calidad de su trabajo radica precisamente en el “tipo” de explicación o modelo que proporcionan. Sin embargo, la forma monolítica de entender la ciencia del comportamiento resumida por Reynolds, fue substituida por el pluralismo y riqueza de visiones que hoy caracterizan la actividad científica.

Un ejemplo de este pluralismo se manifiesta de la manera más clara en el reconocimiento de que la explicación es una actividad pragmática donde diferentes demandas de respuesta requieren diferentes tipos de explicaciones (Killeen, : Tinbergen, ). ELABORAR (evolutivas, desarrollo, mecanismos, ). Un ejemplo adicional de este pluralismo más relevante al propósito del capítulo es la distinción que propuso originalmente Marr (19) entre diferente niveles de análisis. De acuerdo a Marr, ….

2. Adaptabilidad del comportamiento

Por buena parte de su historia los psicólogos han debatido acerca de cual es la “clase natural” que conforma su objeto de estudio. Respuestas tales como la mente, la cognición o el comportamiento son tan generales y ambiguas que no es sorprendente que en la practica la ciencia del comportamiento estuvo definido por los protocolos experimentales y sus resultados. El capítulo argumenta que los procesos evolutivos generan una clase natural, el comportamiento adaptable, que es el objeto propio del estudio de la ciencia del comportamiento.

La adaptabilidad del comportamiento no es una construcción social, si no que surge naturalmente de las dos características de la vida de los organismos: metabolismo y reproducción (Godfrey-Smith, ). Los sistemas biológicos gastan energía y requieren de su constante re abastecimiento, los procesos evolutivos a través del éxito reproductivo diferencial de la instancia metabólica (el organismo) van moldeando diferentes formas exitosas de re abastecimiento y de reproducción (comportamiento adaptable).

La adaptabilidad define entonces un sistema empírico, un objeto de estudio, que surge tanto de la propiedad física asociada a la necesidad de recargarse, como al proceso evolutivo asociado a la reproducción. Una condición limitante sobre el re abastecimiento y la reproducción es que los comportamientos que las definen aparte de consumir energía toman tiempo y el tiempo disponible es finito, creando así una restricción lineal en la que los diferentes comportamientos compiten por el tiempo disponible.  
T=t1 +t2+t3.....tn (1)

Los comportamientos y mecanismos adaptables que resultan de los procesos evolutivos son el resultado de las restricciones del organismo y de las propiedades de su entorno asociadas con éxito reproductivo diferencial. Podemos hablar metafóricamente de estas restricciones como un problema de adaptación que resulta en soluciones óptimas o cuasi óptimas, sin embargo no hay que perder de vista que no hay implicación de un agente “adaptándose”, es un proceso sin agencia ni intención que debe verse (y modelarse) como el resultado de la operación de restricciones asociadas con diferentes éxitos reproductivos.

En este nivel de análisis los modelos pueden ser de dos tipos. En el primero el trabajo de un modelo psicológico es operar como ingeniería en reversa. Del comportamiento observado se busca qué restricciones del entorno permitirían entender la evolución (selección) de ese comportamiento (este es básicamente el programa de la Psicología Evolutiva con todas las dificultades que supone). El segundo tipo de modelos inicia con un análisis detallado de las restricciones del entorno ligadas al éxito reproductivo diferencial (el problema de adaptación) y de las restricciones históricas del organismo para encontrar una solución óptima o cuasióptima. Las restricciones del entorno se refieren a las propiedades estadísticas de los sucesos que son relevantes para el éxito reproductivo y que les llamaremos sucesos biológicamente relevantes (SBR). Note que en muchos casos puede haber más de una solución óptima o las restricciones históricas no permiten la solución óptima pero si una mejor solución relativa a las otras posibles. La segunda forma de modelar “optimalizad” es más razonable y ha dominado tanto el AEC como en general a la psicología Experimental (Baum, ; Wei Ma, Gallistel).

La siguiente pregunta que nos podemos hacer es si se seleccionan, comportamientos o “mecanismos”. Desde luego que si el mecanismo no resulta en comportamiento con éxito reproductivo, este no será seleccionado, sin embargo, el mecanismo que subyace un comportamiento con éxito reproductivo si seleccionado puede ser secuestrado para generar soluciones bajo un conjunto de restricciones diferentes (Houston…). Adicionalmente la estructura biológica metabólica y neuronal puede ser también una importante restricción sobre cuales pueden ser los posibles comportamientos y mecanismos.

El anterior análisis sobre la evolución de comportamiento adaptable ha tenido una gran influencia sobre el tipo de explicaciones y modelos que se pueden desarrollar y que ilustra muy bien la importancia de los desarrollos recientes en la Filosofía de la Ciencia presentados en la introducción. En particular Marr () propuso que las explicaciones en la Psicología pueden tomar tres formas que se complementan mutuamente: la primera conocida como explicación racional consiste en un análisis detallado del problema de adaptación y sus posibles soluciones, implica que para entender un comportamiento es necesario un modelo acabado del problema del que puede ser una solución. El segundo nivel de explicación que se le conoce como algorítmico hace referencia a los posible mecanismos que pueden instrumentar las soluciones óptimas y finalmente un tercer nivel de substrato físico que se refiere a la forma como los algoritmos (mecanismos) se pueden instrumentar en un substrato generalmente neuronal. Estos niveles de explicación no compiten entre si y son un ejemplo de respuestas a preguntas diferentes. Sin embargo, Marr y otros asumen que el nivel racional es prioritario porque impone restricciones acerca de lo que un mecanismo debe lograr y que substrato es necesario.

(Nota: Ponderar poner el ejemplo de memoria.implica conocimiento debe traducirse usarse en acción)

Las explicaciones racionales sostienen que la detección (aprendizaje) de las propiedades estadísticas de sucesos biológicamente importantes no pueden separase de su uso en la distribución óptima de los comportamientos en competencia (acción). Implica por ejemplo que las propiedades de los sistemas de la memoria están afinados a las propiedades estadísticas del uso de los sucesos a recordar posteriormente (Anderson,)

Los organismos están inmersos en un mundo de observaciones en entornos que pueden estar en constante cambio, por ello las explicaciones racionales requieren como un primer paso especificar las propiedades estadísticas de las características del entorno que pueden ser relevantes para el éxito reproductivo.. Cuatro propiedades estadísticas de la disponibilidad de los sucesos biológicamente importantes han dominado el estudio del comportamiento adaptable: tiempo y lugar de ocurrencia, su covarianza con otras características del entorno y con el comportamiento de un organismo.

El supuesto más importante que haremos es que si las consecuencias relevantes en el entorno se distribuyen en ciertos tiempos, lugares y asociadas con ciertas señales y con el comportamiento, un organismo que pueda detectar estas propiedades estadísticas y ajustar su comportamiento a ellas, podrá asignar más óptimamente su comportamiento a las metas en competencia.

Los libros de texto le asignan uno o dos capítulos a cada una de estas restricciones y las abordan como problemas separados. Sin embargo la historia de la Psicología del Aprendizaje se ha dividido en la importancia teórica asignada a estas restricciones. El enfoque asociativo dominante considera las restricciones de covarianza como los elementos básicos del aprendizaje, del cual puede derivarse las adaptaciones a las restricciones temporales y espaciales. Un enfoque alternativo (conocido como representacional y no bien representado en los libros de texto), asume que estas dos últimas restricciones son básicas y deriva de ellas la adaptación a las covarianzas (Gallistel), en el presente capítulo ejemplificaremos el primer enfoque.

Hasta el momento hemos supuesto, al igual que la gran mayoría de los teóricos, que los organismos solo detectan sucesos individuales. Existe evidencia de que, en adición, los organismos detectan y se adaptan a una característica de segundo orden: la tasa de ocurrencia de sucesos individuales en su entorno. Por ejemplo seleccionan su tiempo de estancia en diferentes lugares en función de la tasa de ocurrencia de fuentes de reabastecimiento. Finalmente, podemos considerar una propiedad de los entornos que puede considerarse de tercer nivel. La incertidumbre acerca del tiempo, lugar, covarianzas, y tasas puede ser esperada o inesperada. La incertidumbre esperada es la que se estudia en los protocolos tradicionales de condicionamiento. La tarea del organismo es detectar las relaciones estables en su entorno. Su estudio define el campo de las adaptaciones que consisten en encontrar la mejor forma de extraer recursos (explotar) dado cierto conocimiento de la estructura del entorno. La segunda se ha estudiado solo más recientemente y se refiere a cambios frecuentes e inesperados en los tiempos, lugares, covarianzas y tasas que definen la estructura de un entorno (Baum, ; Gallistel. Su estudio define un campo diferente de adaptaciones, que consiste en el estudio de los determinantes de la búsqueda de información (exploración) que indique que el entorno cambió. Bajo estas condiciones no se espera la evolución de un único mecanismo de comportamiento adaptable, por el contrario las diferentes restricciones que definen a los problemas de re abastecimiento y reproducción generan soluciones específicas que pueden lograrse por medio de mecanismos de diversa índole. en este sentido, el estudio del comportamiento adaptable se asemeja más al estudio de la inteligencia artificial y aprendizaje en máquinas.

Una condición necesaria para que el comportamiento se adapte a las propiedades estadísticas de los sucesos biológicamente importantes es que pueda reconocerlos y actuar ante ellos de una forma que pueda incrementar su éxito reproductivo aún cuando su ocurrencia no pueda ser predicha. Desde luego hay un componente perceptual y de categorización cuya discusión esta fuera del dominio del presente capítulo, la tarea a la que nos referimos es determinar entre un mundo poblado de sucesos aquellos que pueden que están ligados al éxito reproductivo. A esos estímulos en la literatura del AEC se les conoce como refuerzos y en el presente capítulo los llamamos sucesos biológicamente importantes SBI, para nuestros propósitos no se limitan a aquellos que son consecuencias de un comportamiento. Por varias décadas del siglo pasado, mucha investigación giró en torno a identificar la característica común a ellos con resultados infructuosos. En un mundo teórico de estímulos y respuestas discretas era comprensible que la atención se centrará en refuerzos como estímulos discretos con propiedades inherentes a ellos y solo modulados por factores motivaciones (comida, privación de ella). Sin embargo, desde una perspectiva evolutiva lo que hace a un suceso un SBI no es su presencia sino lo que un organismo hace en relación a ellos. La comida es un SBI solo si se le come, e igualmente un depredador es un SBI solo si se escapa exitosamente de él. Adicionalmente muchos organismos no cuentan con los receptores a distancia que les permita ubicar SBIs en el espacio. Finalmente no todas las propiedades de un entorno que son SBI se refieren a estímulos discretos, frecuentemente pasar de un lugar a otro lo es también.

En una muy importante contribución Premack() retomó las consideraciones anteriores y propuso que los refuerzos son comportamientos y no estímulos y que el proceso de modificación del comportamiento que se conoce como reforzamiento era la transición entre comportamientos de diferentes valor. Reforzamiento y castigo distinguiéndose por la dirección de la transición en valor. Las transiciones no solo pueden ser entre comportamientos, sino entre situaciones definidas en un tiempo extendido. Lo anterior implica que entre los mecanismos sobre los que puede operar selección natural se incluye no tan solo uno que permita detectar un estímulo discreto como un alimento, sino aun en su forma más simple detectar transiciones entre condiciones que difieren en valor; un mecanismo así le permitiría a los organismos moverse en un espacio de diferencias en valores.

Justo un mecanismo de detección de diferencias en el tiempo subyace la forma como la bacteria E-coli resuelve su problema metabólico de re-abastecimiento. Un buen ejemplo de lasa herramientas que pueden resolver este problema lo proporciona E-coli. Sin receptores a distancia, estas bacterias resuelven su problema de re abastecimiento con un sistema que cuenta con a) un mecanismo que les permite detectar un pulso o un cambio en la concentración de alimento presente; b) una memoria de corto plazo que registra el nivel de concentración un instante anterior; c) un mecanismo que permite comparar el nivel de concentración en la memoria de corto plazo con el registrado en la actualidad; d) dos comportamientos, uno de exploración (maromas aleatorias) y otro de exploración (nado directo); e) Una regla de elección que incluye un umbral para la comparación, arriba de él el l comportamiento cambia de exploración (maromas aleatorias) a explotación (nado directo) y finalmente f) un proceso de adaptación que hace que el resultado de la comparación baje del umbral y retorne el comportamiento al de maromas aleatorias. Este es un sencillo modelo de un integrador que permite la adaptación bajo ensayo y error y en el caso de E-coli está bien establecido su sustrato bioquímico (Berg, ), Sin embargo es importante enfatizar que este modelo no surge del conocimiento de la bioquímica, sino de un análisis de los posible mecanismos que podrían conformar una estrategia de ensayo y error, que define un algoritmo de ascenso de colina muy común en la matemática computacional y puede instrumentarse (modelarse) como la sumersión de una cubeta con un orificio en su base en una tina con agua, como un capacitor respondiendo a una carga, como un algoritmo computacional o simplemente dejar sin especificar como se instrumenta. El valor de este modelo radica precisamente en su generalidad para dar cuenta de un sin numero de fenómenos sin tener que especificar los detalles de instrumentación. El integrador puede verse también como un sencillo ejemplo de un modelo de elección compuesto por una variable de decisión (el resultado del proceso de comparación) y una regla de elección, que en este caso incluye un umbral. Estos modelos computacionales de decisión subyacentes una buena parte de la Psicología contemporánea aunque su uso se remonta a la Psicofísica tradicional que inicia en el siglo XIX.

Formalmente un integrador con fuga puede representarse en tiempo discreto con una ecuación muy sencilla

VD(k+1) = (1-a)VD(k) + aR(k)

donde: 0<a<1

VD(k+1) es la carga del integrador después de un pulso R (presentación de comida o cambio en su concentración) que nosotros identificamos como la variable de decisión. La ecuación nos dice que la variable de decisión en un momento k+1 es una suma ponderada de su valor en el momento anterior k y del impacto de un pulso. Alfa es el parámetro que representa el valor de ponderación de los dos factores. En la ausencia de un pulso R, solo observaremos el proceso de adaptación; el valor de decisión decrece como una función del valor de alfa, si este es cercano a cero, (1-a) será muy grande y la adaptación en el tiempo será muy lenta, por el contrario valores de alfa cercanos a uno generarán una rápida adaptación. El valor de alfa también determina la importancia relativa de un pulso. Valores altos de alfa implican simultáneamente un impacto grande del pulso y una rápida adaptación y justo lo inverso para valores pequeños. En ocasiones a (1-a) se le llama beta, pero es importante recordar que aunque alfa y beta pueden interpretarse como dos parámetros diferentes, uno asociado a la adaptación y el otro al impacto del pulso, de hecho no son dos parámetros independientes. El modelo requiere una ecuación adicional que especifica la respuesta que se observara (exploración o explotación) dependiendo de si VD es mayor o menor que el valor de un umbral.

Vale la pena resaltar tres aspectos del modelo del integrador: el primero es que un modelo de integrador asume que el organismo esta siempre actuando y que el impacto de un pulso no es provocar una respuesta en un organismo inerte, sino cambiar de un tipo de comportamiento a otro si el valor de la VD que se genera es mayor o no de cierto umbral. El segundo es que en este modelo los organismos no aprenden acerca de donde, cuando y que estímulos predicen los cambios en las concentraciones; simplemente da cuenta del desplazamiento en el tiempo y en el espacio de un organismo como una función de un gradiente de valor en una variable asociada al éxito reproductivo. Tercero, el sistema de integrador que presentamos es el más simple para modelar comportamiento de orientación. Otras restricciones necesitan de sistemas que combinan más de un integrador con parámetros alfas y umbrales diferentes; estos sistemas son comunes en la neurociencia, en el estudio de la percepción y en la discriminación temporal y la memoria (Dayan y ; McClelland y ; Staddon ).

En el caso de la bacteria E-coli dar maromas era el comportamiento base que aleatoriamente lo podía mover hacia una mejor concentración de comida, comportamiento que metafóricamente llamamos de exploración. En el caso de otras especies, en entornos más ricos se requieren de múltiples y variadas respuestas para moverse en la dirección de un cambio positivo. Killeen () propuso un principio general que genera comportamiento exploratorio y del que puede seleccionarse las respuestas exitosas que mueven al organismo en la dirección de una mejora. En varios experimentos encontró que un SBI induce un incremento en la actividad general de los organismos que al igual que en el caso del integrador decrementa exponencialmente en el tiempo. Esta activación que puede verse también como un elemento de exploración, lleva al organismo a interactuar con su entorno para acercase o alejarse del SBI y le permite detectar si alguna situación o respuesta permite predecirlo.

Sin embargo, el manejo y el exitoso acercamiento o alejamiento a un SBI en un entorno variable requiere que la detección del suceso active un conjunto de respuestas organizadas jerárquicamente y sostenidas por variantes circunstancias del entorno. Hogan () y Timberlake () proponen un sistema justo con estas propiedades y Killeen () argumenta que el refuerzo consiste en el desplazamiento a lo largo de esa organización jerárquica de sistemas de comportamiento que define un gradiente de valor. Estos principios le permiten al organismo navegar situaciones donde los FE se presentan en forma impredecible o cuando ocurren por primera vez en situaciones novedosas.

3. Asignación de Crédito

Sin embargo para distribuir óptimamente el comportamiento entre las diferentes demandas relacionadas a SBI es ventajoso contar con mecanismos de adaptabilidad del comportamiento que permitan la posibilidad de predecirlos y generan un espacio representacional que modela las propiedades estadísticas de la ocurrencia de SBI en el entorno. Son estas propiedades estadísticas las restricciones o filtros que dirigen la evolución de comportamiento y mecanismos de adaptabilidad. La detección y adaptación a las propiedades estadísticas depende a su vez de restricciones genéticas sobre mecanismos previamente seleccionados.

Los organismos no detectan cada posible propiedad física de sus entornos creando una copia fiel de ellos, por el contrario solo detectan aquellas relacionadas con el éxito reproductivo de sus ancestros o con el éxito en su pasado ontogenético. El entorno de los organismos puede verse como una estructura estadística de SBI definida en el tiempo y en el espacio que cambia evolutivamente conforme los organismos se adaptan a ellas. Los procesos evolutivos han operado para seleccionar mecanismos que detectan justo esa estructura estadística de SBI. Diferentes aspectos de esa estructura estadística requiere de soluciones diferentes y por ello no debe sorprender que en lugar de un mecanismo general de aprendizaje observemos soluciones específicas a cada propiedad estadística relevante.

Los procesos evolutivos no tan solo operan en la adquisición de una representación de las propiedades estadísticas del entorno, operan también sobre su uso. Otra forma de decirlo es que son precisamente las demandas de la acción las que determinan las propiedades del conocimiento requerido. Un buen ejemplo son los estudios de evolución experimental de Dunlap y Stephens (2016) con drosophila melanogaster, en los que se observa la evolución del aprendizaje bajo entornos que diferían en su variabilidad (incertidumbre) y confiabilidad del éxito de una acción en un entorno predicho.

En lo que resta del capítulo revisaré principalmente las soluciones al problema de detectar predictores de un SBI discreto. A este problema se le conoce como el de la asignación de crédito y es similar independientemente de que el posible predictor sea un estímulo o una respuesta.

La complejidad computacional de una solución al problema de la asignación de crédito es enorme; incluye el que ante una nueva experiencia un organismo pueda a) identificar al suceso, b) asignarlo a una clase de consecuencias comunes (categorización); c) determinar si la clase de consecuencias es un SBI y d) determinar si existe alguna variable a la que se le pueda atribuir el crédito de su ocurrencia, en otras palabras reducir la incertidumbre acerca de su ocurrencia. El principal problema es que el espacio de posibles candidatos a la asignación de crédito es potencialmente infinito.

Los modelos de refuerzo son una propuesta de solución computacional a la asignación de crédito. La solución incluye dos pasos: el primero es la reducción del tamaño inicial del espacio de candidatos incluyendo solo a sucesos que son contiguos, similares, novedosos y o evolutivamente relevantes a los SBI. El segundo es un mecanismo que permita ir reduciendo a través de la experiencia el espacio de candidatos hasta terminar con uno solo. El modelo de refuerzo canónico combina un algoritmo de ascenso de colina con el sesgo de la contiguidad. Bush y Mosteller (1950) formalizaron este clase de modelos que en diferentes variantes han dominado la literatura teórica y experimental en el estudio del aprendizaje a partir de los años 70s del siglo pasado.   
Los modelos de refuerzo son una instancia de modelos de integradores y nuestro propósito en lo que sigue es resaltar como modelos matemáticos similares, pueden usarse para entender fenómenos variados, narrar su evolución y señalar problemas que han llevado a su modificación.

Los organismos enfrentan estímulos que se despliegan en el tiempo, en algunas oportunidades aparecen solos, en otras acompañados (contiguo) de un SBI. En el protocolo estándar de un procedimiento de condicionamiento clásico solo se observa un candidato a la asignación de crédito y cada presentación de él se le conoce como un ensayo. Cada ocurrencia de un ensayo puede estar acompañado o no de un SBI. El modelo de refuerzo tiene una forma similar a la del integrador:

VPx(n+1) = (1-a)VPxn+ aRn (1) donde: 0<a<1

excepto que en este caso el modelo describe el valor predictivo VPn+1 de un estímulo X en el ensayo n+1 como una suma ponderada de dos variables, el VPn del ensayo n y de si el estimulo fue acompañado o no de un SBI (Rn) en el ensayo anterior. R puede tener el valor de 1 o 0 dependiendo de si se presentó o no el refuerzo.

El parámetro alfa determina la importancia de la experiencia acumulada hasta el momento n, relativa a la ocurrencia o no del SBI. Valores cercanos a cero sugieren que la experiencia acumulada es mas importante que una nueva experiencia y valores cercanos a uno sugieren que la presentación de un refuerzo acompañando al predictor X puede minimizar el impacto de la experiencia acumulada hasta ese momento. El parámetro alfa puede interpretarse como una ventana temporal que determina que tan atrás contempla un organismo su experiencia para predecir el futuro.

Una forma alternativa de ver el protocolo experimental es como una serie temporal del estímulo X que tiene el valor de 0 o 1 en cada oportunidad (ensayo) dependiendo de si se presenta o no el SBI. La tarea para el organismo es predecir a partir de la experiencia con la serie temporal el valor que tendrá en la siguiente instancia. Puede comprobarse que la ecuación n representa un filtro que alisa la serie temporal como una media corrida exponencial, donde la importancia de la presentación o no del SBI en una instancia decae exponencialmente como una función de su distancia del presente.

Al igual que para el caso del integrador original, una pequeña manipulación algebraica de la ecuación n genera una forma equivalente del modelo de refuerzo que es el más común hoy en día.

VPn+1 = (1-a)VPn + aRn

VPn+1 = VPn -aVPn + aRn

VPn+1 = VPn + a(Rn -VPn)

Si dejamos que delta = a(Rn -VPn) sea el error de predicción con la expresión hoy mas comúnmente usada en la literatura\_

VPxn+1 = VPxn – aδ

A esta forma equivalente de la ecuación se le conoce como regla delta. En esta forma es más claro que el valor predictivo de un estímulo va cambiando como una función de la magnitud del error de predicción y que el motor del aprendizaje es la reducción de este error, con la solución en equilibrio se obtiene cuando el error el error en predicción es cero.

En ambas interpretaciones el objetivo computacional es buscar la mejor predicción posible, en un caso vista como una serie temporal, en la otra como un algoritmo que reduce errores en predicción. La evidencia de neuronas dopaminergicas que disparan cuando hay un error en predicción le dio un enorme apoyo a estos modelos de refuerzo, sin embargo no puede dar cuenta de los resultados de protocolos experimentales donde potencialmente hay más de un estímulo en el espacio de candidatos, como es el caso en los experimentos de bloque (Kamin).

Hasta antes de los años sesenta, los investigadores limitaban sus experimentos a protocolos que presentaban un solo estímulos como posible predictor. Sin embargo, el entorno de los organismos incluye dos o más estímulos presentes en forma simultánea y contigua con un reforzador, y múltiples experiencias con esos estímulos y el reforzador por separado. Rescorla y Wagner () proponen un modelo que da cuenta de la asignación de crédito en este tipo de entornos.

La contribución de Rescorla y Wagner fue modificar el modelo clásico de refuerzo añadiéndole un mecanismo de competencia entre estímulos por la predicción del SBI. Los siguientes supuestos definen al modelo de RyW:

1. Los estímulos en compuesto están conformados por elementos (estímulos) separables.
2. La ecuación computa para cada uno de los elementos de un compuesto, un valor predictivo (VP). Si el compuesto incluye dos estímulos A y B, se computan dos valores: VPA y VPB. El valor de la predicción del compuesto es la suma del valor predictivo de los elementos individuales que integran el estímulo compuesto. Si son dos estímulos, entonces

VPtotal = VPA + VPB.

3. Los dos principios anteriores describen un proceso de competencia entre elementos separables de un compuesto por convertirse en el mejor predictor de un SBI. La ecuación de Rescorla y Wagner captura tanto el principio de competencia entre estímulos como la noción de que el motor del aprendizaje es la reducción de error, excepto que en este caso delta es la discrepancia entre lo que se obtiene (R) y lo que se espera obtener dado la suma los valores de predicción de todos los estímulos presentes simultáneamente (VPtotal ), obteniendo la siguiente ecuación para cada estímulo X

VPxn+1 = VPxn + α(R – VPtotal n)

El modelo de Rescorla y Wagner es muy exitoso para dar cuenta cuenta de la multiplicidad de fenómenos asociados a los problemas de la contiguidad (Miller, ), y su simplicidad y amplia cobertura en relación a modelos alternativos lo hacen un modelo ampliamente aceptado a pesar de las dificultades para dar cuenta de importantes fenómenos, en particular el de la recuperación espontánea después de un protocolo de extinción.

Un segundo problema con los modelos de refuerzo, incluyendo el de Rescorla y Wagner tiene que ver precisamente con su dificultad para modelar la adaptación del comportamiento a cambios abruptos en la presentación de SBI. La rapidez de la adaptación a un cambio en una serie temporal de SBI puede considerarse una función de dos variables, la volatilidad de la serie antes del cambio y la magnitud de este. Hasta hace relativamente pocos años los protocolos experimentales solo consideraban cambios abruptos como en el caso de la extinción y la única volatilidad considerada se le conoce como refuerzo parcial. Esta literatura ha crecido substancialmente e ilustra un problema fundamental con los modelos de refuerzo (Baum,; Gallistel, ; Nevin,: ).

Veamos el problema, en la ecuación () alfa es un parámetro fijo que empíricamente representa la rapidez del aprendizaje. Sin embargo, vimos que teóricamente representa el peso de la experiencia con la serie de SBI relativo al de una nueva experiencia. Vimos que puede verse también como una media corrida exponencial que alisa la serie de SBI experimentada. Considere dos casos extremos, en uno alfa tiene un valor cercano a 1, en este caso un cambio abrupto generaría un error de predicción grande que tendría un efecto importante en el nuevo valor actualizado VPx. Con este valor de alfa los organismos serían sensible a cambios bruscos después de la exposición a una secuencia con poca variabilidad, sin embargo, si la serie es volátil un valor de alfa cercano a 1 haría que los organismos respondieran a ruido como si fueran cambios verdaderos en los valores que generan la serie. En el caso que alfa sea cercana a cero, el pasado es más importante y el organismo es poco sensible a los errores en predicción. Este valor de alfa es muy ventajoso cuando la serie es muy volátil, pues no se ajusta ante ruido pero es desventajosa ante un cambio abrupto en una serie con poca variabilidad en la que le tomará mucho tiempo ajustar su comportamiento al nuevo valor. La evidencia experimental indica que los organismos pueden ajustarse con gran rapidez a cambios abruptos en SBI y que la velocidad del cambio depende de la volatilidad de la presentación de SBI (Gallistel,)

El tema ocupa cada vez un papel más importante en la investigación en AEC y se han presentado dos posibles clases de soluciones a las dificultades de la ecuación de RyW. Una desde la perspectiva de modelos de refuerzo propone que alfa cambia su valor adaptativamente como una función de las propiedades de la serie y lo hace también a través de un mecanismo de reducción de error (Nassar). El segundo enfoque es computacional y asume que la adaptación del comportamiento ante cambios abruptos es el resultado de un algoritmo de detección de cambios común en la literatura de análisis estadísticos de series temporales. Estos temas así como el análisis de las siguientes dificultades del modelo de Rescola y Wagner serán el objeto de otro escrito.

El modelo de Rescorla y Wagner en su forma más completa incluye para cada estímulo predicar un parámetro adicional beta que multiplica también el error de predicción y que afecta consecuentemente la rapidez del aprendizaje. Este parámetro le permite a un organismo reducir el espacio de candidatos a asignación de crédito a lo largo del tiempo, considerando primero aquellos que son más sobresalientes o que atraen más atención. Sin embargo fijar el valor de beta para cada estímulo genera para el modelo un problema similar al que le representa fijar el valor del parámetro alfa, imposibilita cambios adaptativos en los valores del parámetro como una función de su experiencia con los estímulos como predicadores. Como una consecuencia no podría haber una reducción en el espacio de candidatos lo que haría el problema de la asignación de crédito un problema poco tratable. El ejemplo más ilustrativo se encuentra en los experimentos de inhibición latente, en los que se encuentra que la experiencia con un estímulo en forma aislada retrasa su capacidad para adquirir valor predictivo. Hall y Pearce () proponen que el valor de beta disminuye conforme el estimulo se convierte en un mejor predicar de un SBI y consecuentemente, su valor en el espacio de candidatos incrementa si es seguido por un SIB inesperado. Computacionalmente los resultados pueden modelarse también como una computación Bayesiana, donde siempre se inicia con priors que hacen manejable el espacio de candidatos que cuyos valores cambian como una función de sus consecuencias (Daw).

Finalmente el modelo de Rescorla y Wagner no incluye entre sus argumentos las propiedades temporales del entorno (duración de los estímulos, tiempos entre ocurrencias) que se ha encontrado son factores importantes en la asignación de crédito (Gallistel). En particular es un modelo que asume un entorno de sucesos discretos y no un flujo continuo de sucesos en el tiempo, que le presentan al organismo la tarea de predecir no un SBI después de un ensayo, sino su flujo de momento a momento. En el área de la robótica se desarrolló un algoritmo similar al modelo delta pero diseñado para resolver los problemas de modelar entornos en tiempo real en donde la función busca predecir el curso de acción que maximiza una función de valor agregada sobre SBI que ocurren en diferentes tiempos en el futuro. Al algoritmo se le conoce como “diferencias temporales” (Daw) y es una de las áreas de investigación más activas en la actualidad ().

Mucho ha cambiado desde que Reynolds resumió los principios del AEC. El pluralismo metodológico y conceptual que nos acercó a la Biología Evolutiva, a la neurociencia, al aprendizaje en máquinas y a la microeconomía ha resultado en un esquema conceptual que gira alrededor de la evolución del comportamiento adaptable, de explicaciones que describen soluciones a problemas de adaptación y de mecanismos que conforman una caja de herramientas que implementan soluciones óptimas. El capítulo ilustró las ventajas de especificar con claridad el problema de la asignación de crédito y su solución con un sencillo algoritmo basado en un integrador.

Un análisis cuidadoso de los problemas de adaptación ilustra las dificultades que enfrentan modelos como los de un integrador y abre varias avenidas de investigación empírica y teórica. Sin duda la siguiente edición del libro de Reynolds demostraría el enorme avance de nuestra disciplina y su promisorio futuro.

Referencias

Anderson, J. R. (1990). *The adaptive character of thought*. Psychology Press.

Balsam, P., Drew, M., & Gallistel, C. (2010). Time and associative learning. *Comparative Cognition & Behavior Reviews*, *5*, 1-22. doi:10.3819/ccbr.2010.5000

Batterman, R. W., & Rice, C. C. (2014). Minimal model explanations. *Philosophy of Science*, *81*(3), 349-376.

Baum, W. M. (2012). Rethinking reinforcement: Allocation, induction, and contingency. *Journal of the Experimental Analysis of Behavior*, *97*(1), 101-24. doi:10.1901/jeab.2012.97-10

Baum, W. M. (2010). Dynamics of choice: A tutorial. *Journal of the Experimental Analysis of Behavior*, *94*(2), 161-74. doi:10.1901/jeab.2010.94-16

Baum, W. M. (2012). Extinction as discrimination: The molar view. *Behavioural Processes*, *90*(1), 101-10. doi:10.1016/j.beproc.2012.02.01

Baum, W. M. (1981). Optimization and the matching law as accounts of instrumental behavior. *Journal of the experimental analysis of behavior*, *36*(3), 387-403.

Bindra, D. (1974). A motivational view of learning, performance, and behavior modification. *Psychological review*, *81*(3), 199.

Bolles, R. C. (1972). Reinforcement, expectancy, and learning. *Psychological Review*, *79*(5), 394.

Brown, P. L., & Jenkins, H. M. (1968). Auto‐shaping of the pigeon's key‐peck. *Journal of the experimental analysis of behavior*, *11*(1), 1-8.

Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world. *Trends in Cognitive Sciences*, *10*(7), 294-300. doi:10.1016/j.tics.2006.05.00

Craver, Carl and Tabery, James, "Mechanisms in Science", *The Stanford Encyclopedia of Philosophy*(Spring 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2017/entries/science-mechanisms/>

Frigg, Roman and Hartmann, Stephan, "Models in Science", *The Stanford Encyclopedia of Philosophy*(Spring 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2017/entries/models-science/>

Gallistel, C. R., Mark, T. A., King, A. P., & Latham, P. E. (2001). The rat approximates an ideal detector of changes in rates of reward: implications for the law of effect. *Journal of Experimental Psychology: Animal Behavior Processes*, *27*(4), 354.

Gallistel, C. R. (2012). On rationalism and optimality: Responses to the Miller and Nevin Commentaries. *Behavioural processes*, *90*(1), 87-88.

Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science (New York, N.Y.)*, *349*(6245), 273-8. doi:10.1126/science.aac607

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological review*, *84*(3), 279.

Godfrey-Smith, P. (2006), “The Strategy of Model-based Science”, *Biology and Philosophy*, 21: 725–740.

Godfrey-Smith, P. (2009). *Theory and reality: An introduction to the philosophy of science*. University of Chicago Press.

Godfrey-Smith, P. (2017). *Other Minds: The Octopus and the Evolution of Intelligent Life*. HarperCollins UK.

Herrnstein, R. J. (1969). Method and theory in the study of avoidance. *Psychological review*, *76*(1), 49.

Herrnstein, R. J. (1970). On the law of effect. *Journal of the experimental analysis of behavior*, *13*(2), 243-266.

Kamin, L. J. (1969). Selective association and conditioning. *Fundamental issues in associative learning*, 42-64.

Kitcher, Philip (1989). Explanatory unification and the causal structure of the world. In Philip Kitcher & Wesley Salmon (eds.), *Scientific Explanation*. Minneapolis: University of Minnesota Press. pp. 410-505.

Marr, D., & Vision, A. (1982). A computational investigation into the human representation and processing of visual information. *WH San Francisco: Freeman and Company*, *1*(2).

McElreath, Richard. 2015. Statistical Rethinking: A Bayesian Course with Examples in R

and Stan. Boca Raton: Chapman & Hall/CRC.

Niv, Y., & Schoenbaum, G. (2008). Dialogues on prediction errors. *Trends in Cognitive Sciences*, *12*(7), 265-72. doi:10.1016/j.tics.2008.03.00

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological review*, *87*(6), 532.

Pincock, C. (2011). *Mathematics and scientific representation*. Oxford University Press.

Premack, D. (1965). Reinforcement theory. In *Nebraska symposium on motivation* (Vol. 13, pp. 123-180).

Rescorla, R. A. (1972). Informational variables in Pavlovian conditioning. *Psychology of learning and motivation*, *6*, 1-46.

Strevens, M. (2008). *Depth: An account of scientific explanation*. Harvard University Press.

Tinbergen, N. (1963). On aims and methods of ethology. *Ethology*, *20*(4), 410-433.

van Fraassen, Bas C. (2008), *Scientific Representation: Paradoxes of Perspective*, Oxford: Oxford University Press.

Weisberg, M. (2012). *Simulation and similarity: Using models to understand the world*. Oxford University Press.