

## El Efecto Baldwin en la Interrelación entre Evolución y Aprendizaje

José Santos Reyes

Grupo de Sistemas Autónomos – Departamento de Computación  
Facultad de Informática  
Universidade da Coruña  
Campus de Elviña s/n 15071 A Coruña  
santos@udc.es

### Resumen

En este artículo se realiza una síntesis y un análisis de los trabajos más relevantes en el estudio del Efecto Baldwin en la interrelación entre evolución y aprendizaje. Se explican las posibilidades en esa interrelación, centrándose en los trabajos de demostración en el campo computacional del Efecto Baldwin o asimilación genotípica en tiempo evolutivo de los caracteres adquiridos en tiempo de vida. La mayor parte de los trabajos en la bibliografía científica del tema se centran en la capacidad de suavización de la superficie de búsqueda al añadir el aprendizaje en cada individuo de la población genética, por lo que en este artículo se hace énfasis en el aspecto de asimilación genética que aparece siempre que el aprendizaje conlleve un coste en calidad asociado. Se comenta además el uso del Efecto Baldwin en la evolución de controladores en robótica autónoma, usando aprendizaje conexionista y por refuerzo en controladores neuronales y proponiendo medidas cuantitativas de la asimilación genotípica producida al actuar el efecto Baldwin.

**Palabras clave:** Algoritmos Evolutivos, Efecto Baldwin, Estrategia de Lamarck, Robótica Evolutiva, Aprendizaje Conexionista.

### 1. Introducción

La evolución y el aprendizaje son las dos principales fuerzas que promueven la adaptación de los individuos a su entorno. La evolución, operando al nivel de la población, con la selección como su motor, incluye todos los mecanismos de cambios genéticos que ocurren en los organismos a lo largo de las generaciones. El aprendizaje opera en una escala temporal diferente y provee a cada individuo de la habilidad de modificar su fenotipo durante su vida para incrementar su adaptación al entorno y, así, su probabilidad de ser seleccionado en la reproducción. La primera permite la adaptabilidad a

los cambios lentos que ocurren a lo largo de las generaciones del tiempo evolutivo, mientras el aprendizaje permite la adaptación de los individuos a cambios más rápidos en el entorno que los del tiempo evolutivo. En este segundo caso se habla en biología de caracteres adquiridos a lo largo de la vida de un individuo. Lamarck, a principios del siglo XIX, supuso que esos caracteres adquiridos eran heredados en la prole, constituyendo el motor de la evolución. Posteriormente, el establecimiento por Darwin de la selección natural como motor de la evolución y el descubrimiento de las leyes de la herencia desecharon la posibilidad establecida por Lamarck.

En el reflejo de la evolución en el campo computacional los diversos algoritmos evolutivos constituyen un método de búsqueda en paralelo sobre los puntos del espacio representados por cada individuo de la población. Por su parte, la posibilidad de aprendizaje en un individuo, asociada a la adaptabilidad en tiempo de vida en el mundo biológico, se corresponde en computación con una búsqueda local en la vecindad de cada uno de esos individuos considerados en el proceso evolutivo. Al combinar los dos métodos de adaptación, cuando se permite el aprendizaje en cada individuo de la población genética en cada generación del proceso evolutivo, nada impide que lo aprendido computacionalmente, que parte de un genotipo inicial, pase al genotipo de su prole. A esta posibilidad se le denomina, en consecuencia, estrategia de Lamarck.

Aunque no se use esa posibilidad de la estrategia de Lamarck, la utilización del aprendizaje va a tener una influencia lógica en el proceso evolutivo. La razón es que el aprendizaje modifica la calidad de cada individuo y, en consecuencia, la superficie de calidad sobre la que actúa el algoritmo evolutivo. Si la búsqueda local aprende en la dirección correcta, en el sentido de que cada genotipo inicial se acerca a una buena solución, la superficie de búsqueda o paisaje de calidad tenderá a ser más suave, facilitando así el proceso evolutivo.

Pero no termina aquí la interrelación entre evolución y aprendizaje. Se conoce en biología que existe una posibilidad de que los caracteres adquiridos por los individuos terminen apareciendo en los genotipos al cabo de generaciones. Este efecto fue propuesto por el americano James Mark Baldwin en 1896, conociéndose así como "Efecto Baldwin" (Baldwin, 1896). Este Efecto es fácilmente demostrable en el reflejo computacional. Hinton y Nowlan (1987) fueron los primeros en demostrarlo en un sencillo y elegante experimento en el que combinaban un algoritmo genético y una búsqueda local.

Para que el Efecto ocurra, el aprendizaje debe tener un "coste" asociado, de modo que aquellos individuos que necesiten menos aprendizaje a partir de sus genotipos tendrán una calidad mayor con respecto a aquéllos que requieran aprender más para adaptarse a su entorno. De este modo, la selección priorizará y tenderá a elegir a aquellos individuos que requieran aprender menos, con lo que a lo largo de las generaciones aquellos caracteres que en un principio necesitaban aprenderse irán apareciendo paulatinamente en los genotipos, y sin recurrir en ningún momento a Lamarck. Es, por tanto, consecuencia de la selección y del coste inherente al aprendizaje. Se suele decir, tanto en el mundo

biológico como en computación, que lo aprendido revierte o se asimila en el genotipo a lo largo de las generaciones y también que el aprendizaje "guía" la evolución de los genotipos.

En este artículo se comienza describiendo esas diferentes posibilidades en la combinación entre la evolución y el aprendizaje, detallando sus implicaciones en la búsqueda evolutiva. Posteriormente se expone el Efecto Baldwin, con ejemplos en el mundo biológico, para pasar a su reflejo computacional describiendo, primero, el experimento de Hinton y Nowlan. Como gran parte de los trabajos de estudio del efecto Baldwin se han centrado en el primer aspecto de interrelación entre evolución y aprendizaje, el de suavización de la superficie de calidad, mientras pocos estudian el efecto de asimilación genotípica, se realiza un énfasis en este aspecto de asimilación a lo largo de las generaciones en la exposición siguiente de los trabajos más relevantes.

Finalmente se exponen ejemplos en una aplicación práctica y compleja como es en el campo de robótica autónoma la obtención automática de controladores de comportamientos. Tales ejemplos mostrarán los dos aspectos involucrados al estudiar el Efecto Baldwin, es decir, la suavización de la superficie de búsqueda y la asimilación genotípica al actuar el aprendizaje con un coste en calidad asociado. Se finaliza con una breve discusión del trabajo realizado en investigación sobre el Efecto Baldwin y el interés y utilidad de su estudio computacional.

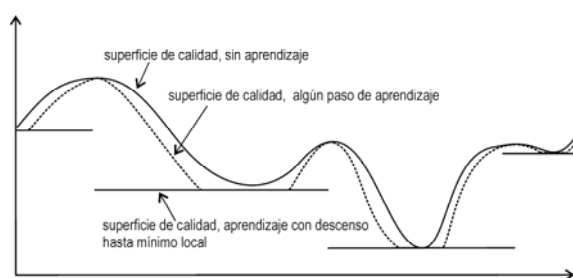
## 2. Posibilidades de interrelación entre evolución y aprendizaje

En la combinación entre ambos métodos cada individuo de la población genética utilizará aprendizaje en tiempo de vida. Tal aprendizaje en cada individuo se puede interpretar como una búsqueda local en cada punto del espacio o genotipo inicial, frente a la búsqueda global y paralela de mejores configuraciones genéticas que realiza la evolución simulada en la escala de tiempo evolutiva.

La búsqueda local que realiza cada individuo moverá sus coordenadas o configuración genotípica hacia otro punto cercano del espacio, cambiando así tal configuración inicial. Caben ahora dos posibilidades: que la nueva configuración, si es elegida por el operador selección, es la que pase su material genético a la hora de generar prole, bien a través de los operadores como cruce y mutación, o bien, en la estrategia elitista, pasando a la siguiente generación si es de alta calidad; la segunda

posibilidad es que la nueva configuración genética no sea la que pase su material genético a la prole, sino que sea la inicial. Sin embargo, aun en esta segunda alternativa, sigue existiendo una influencia clara de la búsqueda local respecto a sólo evolución *darwiniana*, y es que la búsqueda “en tiempo de vida” ha cambiado la calidad de cada genotipo y, por tanto, qué individuos serán seleccionados.

Nada imposibilita utilizar la primera alternativa en el campo computacional. Como se corresponde con la teoría de los caracteres adquiridos de Lamarck, previa a la de selección natural de Darwin, se habla en este caso de estrategias de Lamarck. La segunda alternativa tenderá a suavizar el espacio de búsqueda y, aunque cada individuo no pasa su material genético, modificado por el aprendizaje, a la siguiente generación, existe, sin embargo, una posibilidad de que lo aprendido en tiempo de vida aparezca en el genotipo después de varias generaciones, y sin recurrir a estrategias de Lamarck. En el mundo biológico implica que los caracteres adquiridos en tiempo de vida revierten al cabo de generaciones en el genotipo, lo que se conoce como Efecto Baldwin, en honor a quien lo definió a finales del siglo XIX.



**Figura 1. Ejemplo de superficie de calidad en una búsqueda evolucionista con y sin aprendizaje local.**

La figura 1 ejemplariza el método con la primera alternativa de Lamarck, en la que se muestra una superficie de calidad a lo largo de una dimensión de posibles genotipos. La búsqueda genética es una búsqueda en paralelo sobre múltiples puntos del espacio de calidad, lo que constituye su principal ventaja frente a una búsqueda local a partir de un único punto. Ahora se permite que cada individuo de la población genética realice una búsqueda local en su vecindad a fin de mejorar su calidad, pero la mayor parte de los métodos de aprendizaje tienden a moverse en la dirección de máxima pendiente hacia arriba o hacia abajo, dependiendo de si quiere obtenerse un máximo o un mínimo en el paisaje de calidad. Por tanto, el problema de quedar atrapados en máximos o mínimos locales está presente,

aunque se utilicen técnicas para minimizar esa posibilidad.

En la figura 1, adaptada del trabajo de Whitley y col. (1994), la línea discontinua muestra la calidad de cada cromosoma después de aplicar un algoritmo de aprendizaje local, que lo mueve y acerca hacia el mínimo local más próximo. A medida que se permiten más pasos de búsqueda local, más se aproximará la configuración inicial a la del máximo o mínimo local –líneas continuas horizontales–, hasta llegar al extremo en que prácticamente todos los puntos de la búsqueda genética terminen en una configuración de máximo o mínimo local. Lo problemático aquí es que se ha partido de una población genética cuyos individuos, con el aprendizaje local, se han acercado entre ellos en torno a esos máximos o mínimos locales, por lo que se ha perdido variedad en la población genética y, si se usa una estrategia de Lamarck, esa población con variedad reducida es la que se considera para pasar su material genético a subsiguientes generaciones. Es evidente que reducir la variabilidad de la población genética repercute negativamente en la búsqueda evolutiva, que tiene uno de sus pilares en la diversidad del material genético para buscar configuraciones más aptas.

Boers y Sprinkhuizen (2001) lo resumen muy bien: un algoritmo evolutivo basado en Lamarck usa las codificaciones de los candidatos, una vez mejorados, para generar la siguiente generación. Cuando la superficie de calidad no es demasiado compleja, es probable que el algoritmo evolutivo converja rápido hacia una buena solución. Sin embargo, se reduce la variabilidad genética de la población, dado que las soluciones mejoradas reemplazan a las originales antes de la reproducción, o, en palabras de estos autores, los mínimos locales en la superficie de calidad actuarán como “atractores” para el algoritmo de búsqueda o aprendizaje local, de modo que si la búsqueda se exprime al máximo para los diferentes puntos del espacio, la variabilidad genética se perderá.

Por el contrario, en la segunda posibilidad, lo aprendido en tiempo de vida no revierte en la configuración genética inicial, por lo que se puede pensar que el aprendizaje sólo servirá a cada individuo para mejorar su adaptabilidad al entorno. Pero, como ya se ha remarcado, al depender la calidad final de la búsqueda local, hay una relación directa con la selección y, por tanto, con la evolución. Se puede también cambiar la interpretación y pensar que la evolución, más que buenas configuraciones genotípicas, intentará obtener configuraciones iniciales que tengan una buena predisposición a aprender o, en términos de

Belew y col. (1991), a obtener “semillas” a partir de las cuales comienza a operar el aprendizaje.

Cuanto mejor y más rápido aprenda un individuo, mejor se adaptará a su entorno y más posibilidades tendrá de procrear. Una consecuencia importante del método es que tenderá a suavizar la superficie del espacio de búsqueda, ya que si se parte de un paisaje abrupto, y cada uno de los diferentes individuos de la población genética tiene esa posibilidad de aprender, sus calidades finales tenderán a ser más homogéneas, aunque persistan las diferencias debidas a la configuración genética inicial y al propio método de aprendizaje. De modo que el paisaje tenderá a ser más suave, y, como consecuencia, la búsqueda genética de soluciones óptimas será más fácil, al existir menos máximos o mínimos locales o ser éstos menos pronunciados. Es la misma conclusión a extraer de la figura 1, con la muy importante salvedad de que ahora no se ve reducida la variabilidad genética en la población, al seguir considerando el material genético inicial como fuente para la siguiente generación.

Se ha reseñado que la ventaja de los métodos evolutivos es su búsqueda en paralelo y simultánea en diferentes puntos del espacio de búsqueda, frente a la búsqueda local en la vecindad de cada punto. Pero se debe remarcar que el aprendizaje tiene la ventaja de adaptación a cambios rápidos en el entorno que la evolución, que opera en el orden temporal de los cambios que ocurren a través de las generaciones, puede que no sea capaz de seguir. Nolfi y Parisi (1997) lo demuestran con un ejemplo sencillo en el que someten a un robot a cambios de luminosidad ambiente en su entorno en una misma generación, a los que el robot se adapta mediante aprendizaje en su controlador neuronal, para acercarse en tiempo de vida lo más posible a las paredes usando sus sensores de infrarrojos –cuyos valores de distancia a objetos dependen del nivel de luz ambiente–, mientras la obtención del controlador mediante evolución sólo consigue soluciones de compromiso, no óptimas, que tienen en cuenta las situaciones más desfavorables de luminosidades para poder acercarse a una pared sin chocar. Como concluyen los autores “el aprendizaje complementa a la evolución porque posibilita la adaptación a cambios en el entorno que son demasiado rápidos para que la evolución consiga seguirlos”.

Una última posibilidad en la interrelación es la usada por Floreano y Mondada (1998) al experimentar con la evolución de la propia capacidad de aprendizaje. En este trabajo los autores utilizan redes neuronales artificiales como estructuras de control de comportamientos en un robot autónomo. En su trabajo el genotipo no

codifica los pesos de conexión sino diferentes reglas de aprendizaje *hebbiano* que utilizará un robot para aprender a partir de unos pesos iniciales siempre aleatorios. Al codificar en el genotipo no el controlador de un comportamiento sino la propia capacidad de aprendizaje de un individuo, y si se considera la evolución artificial como método de optimización, se estará optimizando, como apuntan los autores, el propio método de aprendizaje.

Finalmente, en esta relación entre ambos métodos se enmarca también el concepto de “algoritmo memético” (Moscato, 1989; Moscato y Cotta, 2003) que trata de emular la evolución cultural, en la cual los *memes* (término definido por Richard Dawkins en su libro *El Gen Egoísta* en 1976) o unidades mínimas de información y replicación cultural, se ven sometidas también a un proceso de selección y evolución. En estos algoritmos se aplica una búsqueda local, que puede ser cualquier heurística, a fin de mejorar cada solución candidata o agente, como extensión del término individuo usado comúnmente en el contexto de los diferentes algoritmos evolutivos. Los agentes de una población interactúan entre sí mediante cooperación (a través de recombinaciones) y competición (mediante selección y reemplazo), incluyendo un operador mutación que incorpora información externa y en el cual se pueden usar operadores de optimización locales, por lo que en ocasiones se utilizan también los términos de algoritmo evolutivo híbrido o algoritmo evolutivo *lamarckiano*, aunque los autores remarcan que no debe identificarse algoritmo memético como un algoritmo evolutivo más búsqueda local.

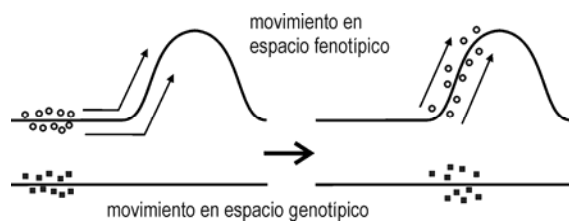
### 3. Efecto Baldwin

En la segunda alternativa de interrelación antes comentada todavía existe una nueva posibilidad, la de que aun sin utilizar una estrategia de Lamarck lo aprendido en tiempo de vida sí revierte en la configuración genotípica, que es lo que se conoce como Efecto Baldwin. Este Efecto fue definido en 1896 por el estadounidense James Mark Baldwin al proponer un “nuevo factor en la evolución” (Baldwin, 1896), mediante el cual los caracteres adquiridos pueden ser heredados indirectamente. Tal efecto también fue descrito de modo independiente por el inglés Conway Lloyd Morgan y el estadounidense Henry Fairfield Osborn en el mismo año.

Baldwin propuso que las habilidades que inicialmente requieren aprendizaje son reemplazadas mediante la evolución por sistemas determinados genéticamente que no requieren el

aprendizaje. Dicho de otro modo, los comportamientos aprendidos pueden convertirse en instintivos en subsecuentes generaciones, sin apelar a la teoría de la herencia de Lamarck. Ese “nuevo factor” se basa en la “plasticidad” del fenotipo: la habilidad que posee un organismo para adaptarse a su entorno durante su vida.

El efecto Baldwin funciona en dos pasos. El primero es la plasticidad del fenotipo que permite a un individuo adaptarse a los cambios en el entorno. Sin embargo, la plasticidad del fenotipo es costosa para un individuo, ya que la adaptabilidad requiere, en el mundo natural, consumo de energía, tiempo o implica peligro. De modo que aparece un segundo paso: con suficiente tiempo, la evolución puede encontrar un mecanismo “rígido” que puede reemplazar al mecanismo de la plasticidad. Esto se debe al coste inherente a la adaptación. La evolución impondrá una presión para encontrar la misma solución, pero que no requiera tal coste, de forma que un comportamiento que ha sido aprendido en el primer paso puede convertirse en instintivo en el segundo paso. El resultado es análogo a la evolución de Lamarck, pero sin existir una alteración directa del genotipo.



**Figura 2. El Efecto Baldwin implica que los movimientos de los fenotipos por aprendizaje conllevan, debido al coste, a un movimiento y asimilación en los genotipos.**

Es evidente que hasta ahora no se ha realizado una diferenciación entre genotipo y fenotipo, aunque los razonamientos son idénticos en este caso, como muestra ahora la figura 2, en la que se esquematiza el movimiento en los correspondientes espacios. Así, los genotipos (pequeños cuadrados) en una determinada generación se trasladan a sus correspondientes fenotipos (pequeños círculos), y la altura de éstos representa la calidad del fenotipo asociado a cada genotipo. Si se pretende maximizar la calidad, el aprendizaje actúa en el espacio de los fenotipos, moviendo éstos hacia zonas del espacio de mayor calidad. Pero, al existir el coste, la evolución tenderá en subsecuentes generaciones a seleccionar aquellos genotipos que estén más cerca de la zona de máxima calidad, “más cerca” en el sentido de que requieran menos aprendizaje, por lo que también existirá el movimiento en el espacio de

los genotipos y que implica la asimilación. La figura denota así claramente que “el aprendizaje guía a la evolución” ya que esta última actúa sobre los genotipos. Cabe reseñar que, en este caso, para poder usar una estrategia de Lamarck, debe existir una traslación inversa del fenotipo modificado hacia el correspondiente genotipo, lo cual, lógicamente es mucho más fácil en el caso de usar traslaciones directas de genotipo a fenotipo, que no impliquen, por ejemplo, procesos de desarrollo del fenotipo.

Un ejemplo clásico en el mundo biológico fue descrito por Conrad Waddington (1953) con la mosca del vinagre. Waddington describió la asimilación genética a lo largo de las generaciones del no desarrollo, frente a choques de calor, de un tipo concreto de venas en sus alas. Partiendo de moscas que no desarrollaban las venas ante el estímulo abrasivo y seleccionando a lo largo de pocas generaciones aquéllas que no desarrollaban las venas en la etapa adulta ante el estímulo, llegó a moscas que no desarrollaban las venas aun sin la presencia del estímulo abrasivo, reconociendo que no se trataba de una evolución de Lamarck y acuñando el término de “asimilación genética”.

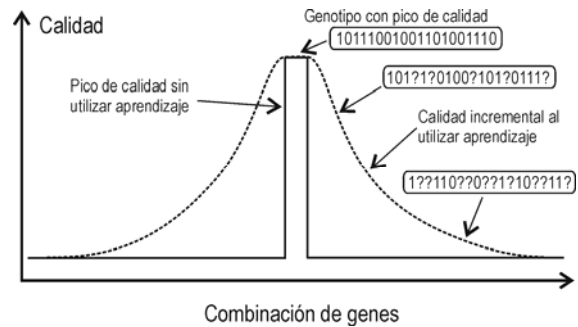
Otro ejemplo en el mundo animal que describe esta asimilación genética es la presencia de callosidades en la piel del avestruz, descrito, entre otros, por Waddington (1942) y Maynard Smith (1993). La piel de la mayoría de vertebrados se vuelve dura y escamosa cuando se expone a algo abrasivo. Este crecimiento de callosidades es una habilidad adaptativa porque es seleccionada por la evolución por su respuesta beneficiosa a estímulos del entorno como colocarse en sitios calientes y duros del suelo. Sin embargo, se observa que estas callosidades aparecen ya en los fetos dentro de los huevos y, por lo tanto, no pueden ser resultado de la exposición a estímulos abrasivos, sino que deben estar genéticamente especificadas. Es decir, la evolución ha ido asimilando en su genotipo la característica del crecimiento de esas callosidades, que les permiten reposar en tales sitios calientes y duros, ya que aquéllos que conseguían desarrollarlos en tiempo de vida tendrían una ventaja evolutiva (cuanto más pronto, menos coste en la adaptación). Por esto, el crecimiento de las callosidades debería aparecer más y más temprano en la vida del avestruz a lo largo de las generaciones, requiriendo cada vez menos estímulos del entorno para producirlas hasta llegar a un punto en el que no se requiriese el estímulo y las callosidades apareciesen en el feto, es decir, se convertirían en genéticamente especificadas mediante la asimilación genética.

### 3.1 Experimento de Hinton y Nowlan

El efecto Baldwin acaparó la atención en computación con el trabajo de Geoffrey Hinton y Steven Nowlan (1987), quienes demostraron el Efecto con un elegante y sencillo experimento. Estos autores utilizaron un algoritmo genético para evolucionar una población de individuos que además emplea una búsqueda local, búsqueda que, como se ha dicho, representa la analogía computacional de la plasticidad del fenotipo en el nivel biológico.

Los autores hacen uso de un ejemplo extremo para demostrar el Efecto: una única combinación de genes –con valores binarios– en un cromosoma, que corresponden a la presencia o ausencia de conexiones entre los nodos de una red neuronal artificial (RNA) asociada al individuo, tiene una calidad determinada, mientras las demás combinaciones son igual de pobres. Es decir, la superficie de calidad es un pico de alta calidad, correspondiente a la única solución óptima, mientras los demás puntos presentan una calidad baja (figura 3, línea continua). En el cromosoma algunos genes son fijos mientras que otros pueden cambiar su valor en el tiempo de búsqueda local (alelos “0”, “1” y “?”). El aprendizaje que utilizan los autores consiste en permitir cambios aleatorios en esos genes sin valor inicial determinado (alelo “?”), que interpretan como interruptores que pueden abrir o cerrar su correspondiente conexión neuronal, de forma que aquellos individuos que estén cerca de la única solución óptima (con más genes fijos correspondientes a esa solución) tendrán mayor posibilidad de encontrarla en un pequeño número de chequeos de combinaciones aleatorias.

De este modo, al permitir en cada individuo esa búsqueda local, la superficie de calidad se suaviza en torno a ese punto (línea discontinua), porque los autores rebajan la calidad en función de cuántas combinaciones en los genes no prefijados en el genotipo son necesarias para encontrar el único máximo: mayor calidad cuantas menos pruebas aleatorias necesarias. La figura muestra dos de estos posibles genotipos en función de su distancia al pico de calidad. Así, con este “aprendizaje” o búsqueda local, existe una zona estrecha de calidad incremental en torno a ese punto, llevando a que la evolución pueda descubrir la combinación adecuada de genes. Los autores utilizan la expresiva frase “encontrar la red correcta es como encontrar una aguja en un pajar”, para denotar la situación de búsqueda sin el aprendizaje, en tanto usan la frase “es como si alguien te dice si estás acercándote o no



**Figura 3. Superficie de calidad en el ejemplo de Hinton y Nowlan (1987). Hay un único genotipo con alta calidad, pero la posibilidad de aprendizaje en los genes con alelos “?” suaviza la superficie en torno a ese punto.**

a la aguja que buscas en el pajar” para denotar la situación en la cual actúa el aprendizaje.

La diferencia entre esta búsqueda y una mutación es que ese chequeo o cambio aleatorio no pasa a los hijos. Los cromosomas que representan las RNAs tienen 20 conexiones potenciales representadas en 20 genes, con lo cual existen  $2^{20}$  posibles puntos en el espacio de búsqueda. Con una población de 1000 individuos es difícilmente abordable el encontrar la solución óptima con evolución pura, a no ser por azar. Pero permitiendo a cada individuo de la población aprender (los autores usan 1000 cambios aleatorios), si se encuentra la combinación correcta, este individuo tendrá mayor probabilidad de reproducirse y, aunque no pasa esa óptima combinación a los hijos, la prole tenderá a una combinación de genes más cercana a la correcta. La razón es el coste de la búsqueda, que está implícito en la función de calidad<sup>1</sup>, que determina que un individuo es mejor cuantas menos pruebas aleatorias necesita para encontrar la solución correcta; es decir, cuanto más rápido aprende, menor coste y mejor calidad. Por lo tanto, la evolución irá eligiendo aquellos individuos que requieran cada vez menos coste en la búsqueda, y que son aquellos que están cada vez más cerca de la solución óptima. Si su prole, mediante cruces y mutaciones, produce individuos con un coste todavía menor para encontrar la solución, generará a su vez más prole. Así, la combinación descubierta a través del aprendizaje tenderá a permanecer fija en el genotipo en generaciones sucesivas, es decir, se irá asimilando en el genotipo, y sin recurrir a Lamarck.

<sup>1</sup> La función de calidad la definen los autores como proporcional a  $(1+(19n/1000))^{-1}$ , donde  $n$  es el número de intentos o pruebas que cada individuo ha necesitado para obtener la red correcta. Si no se encuentra el pico de calidad óptimo en los 1000 intentos, la calidad será nula.

### 3.2 Síntesis de trabajos sobre el Efecto Baldwin

Existe una amplia variedad de trabajos en el campo computacional sobre la interrelación entre evolución y aprendizaje y, más específicamente, sobre el Efecto Baldwin, de los que se eligen, a juicio del autor de este artículo, algunos de los más representativos. Muchos tienen como referencia o punto de partida el comentado anteriormente. Por ejemplo, Harvey (1993) completa el anterior experimento con un análisis del efecto de la deriva genética sobre la lenta disminución de los alelos con capacidad de variación. Y French y Messinger (1994) han realizado un trabajo similar con agentes simulados. Al igual que Hinton y Nowlan consideran un único genotipo óptimo en agentes que se deben mover en un entorno reticular en busca de comida. Esa buena codificación de genes conlleva directamente a un buen comportamiento en el nivel fenotípico (los autores usan el término “*good phene*”). Los genotipos no óptimos, sin embargo, poseen un nivel de plasticidad que los puede convertir en la combinación de “*phene*” óptimos en función de la distancia al genotipo óptimo (se usa la distancia de Hamming en las codificaciones binarias). Semejante, por tanto, al caso de Hinton y Nowlan (1987). Los autores muestran cómo la población va adquiriendo paulatinamente individuos con buena codificación genética inicial y fenotípica final a lo largo de las generaciones, en función de ese nivel de plasticidad que representa la probabilidad de que un genotipo, que se encuentra a una determinada distancia genotípica del genotipo óptimo, logre alcanzar el fenotipo óptimo.

Ackley y Littman (1990) realizaron otro trabajo de demostración del Efecto Baldwin. En su experimento un agente debe moverse en un entorno reticular en cuyas posiciones hay comida, predadores y diversas trampas. Cada agente tiene asociado un estado determinado por lo detectado en su rango visual así como un nivel energético función de lo que ha conseguido comer. Cada agente tiene además dos redes neuronales asociadas: una “red de evaluación” que, recibiendo como entrada el estado en el tiempo  $t$ , asocia un valor numérico que representa la bondad de ese estado; y una “red de actuación” que determina para cada estado la acción a tomar en el instante  $t$ . Las únicas acciones posibles son moverse a uno de los cuatro estados adyacentes al actual, que podrán resultar en comer o ser devorado. Los pesos de conexión de ambas redes (perceptrones sin recurrencias) se codifican en el genotipo de cada agente, pero los correspondientes a la red de evaluación son innatos desde el nacimiento y no susceptibles de aprendizaje, al contrario que los pesos de la red de actuación. Esa red de evaluación representa así metas innatas heredadas de los

ancestros como, por ejemplo, “estar cerca de comida es bueno”. Por el contrario, los pesos de la red de actuación se adaptarán para poder incrementar el nivel energético teniendo en cuenta que se seguirán las indicaciones de bondad en los cambios de estado que determina la primera red innata.

Una población de agentes se somete a un proceso evolutivo en el que los agentes deben, además de evitar predadores, mantener su nivel energético por encima de un valor para prevenir su muerte así como poder procrear, mediante clonación con mutaciones, o con el apareamiento con agentes vecinos (con cruce de material genético).

Recalcar, como lo hacen Mitchel y Forrest (1994), que, por tanto, no hay una función de calidad definida a priori, como en el ejemplo de Hinton y Nowlan, y como en la mayoría de los procesos de simple optimización con algoritmos evolutivos, sino que la calidad emerge a partir de muchas acciones e interacciones a lo largo de la vida del agente –hecho que diferencia la utilización de la evolución simulada en los modelos de Vida Artificial respecto a simples aplicaciones de optimización–.

En cada paso cada agente determina su estado y lo evalúa con la primera red innata. Esta evaluación se compara con la del instante  $t-1$ , lo que produce una señal de refuerzo positivo si el agente ha mejorado su evaluación o un refuerzo negativo en caso contrario, señal de refuerzo que se usa en combinación con el aprendizaje supervisado para la modificación de los pesos en la red de actuación (usando el “algoritmo de retropropagación del error por refuerzo complementario”, definido por los mismos autores y que comentamos someramente en el apartado siguiente).

Los autores realizan pruebas de evolución sin aprendizaje, de aprendizaje sólo y pruebas en las que combinan la evolución con el aprendizaje en las redes de actuación. Sus resultados, medidos como el tiempo de extinción del mejor agente, indican que la combinación es la mejor opción seguida del aprendizaje y de la evolución pura. Pero, además, en el primer caso aparece el Efecto Baldwin. Ackley y Littman aprecian un efecto curioso al observar cómo la importancia relativa de la evolución y del aprendizaje cambia a lo largo de la evolución. Concretamente, los autores inspeccionan los genes relacionados con el comportamiento de aproximación a comida en ambas redes. En las primeras generaciones, los genes que codifican ese comportamiento en la red de evaluación (“estar cerca de comida es bueno”) permanecen relativamente constantes en la población, mientras que los genes que codifican los pesos iniciales en la

red de actuación son más variables, lo que indica la importancia de mantener las metas para el aprendizaje y la necesidad de éste para la supervivencia. Pero en generaciones posteriores los genes de la red de evaluación se vuelven más variables a lo largo de la población mientras que los genes que determinan los pesos iniciales en la red de actuación son menos variables, lo que indica ahora que los comportamientos heredados son más significativos que lo aprendido. Ackley y Littman interpretan este hecho como una versión del Efecto Baldwin: inicialmente los agentes se ven forzados a aprender a buscar comida para su supervivencia manteniendo un buen conocimiento en la red de evaluación pero, más tarde en la evolución, el conocimiento genético de que estar cerca de comida es bueno es superado por el comportamiento genéticamente especificado de “aproximarse a comida es bueno”, de forma que la red de evaluación no es necesaria. Es decir, la habilidad inicial de aprender el comportamiento se torna en genéticamente asimilada.

Retomando de nuevo el trabajo inicial de Hinton y Nowlan, Nolfi y col. (1994), sin embargo, comentan que tiene varias limitaciones: la modelización del aprendizaje como una búsqueda aleatoria, la no distinción entre las tareas consideradas por el aprendizaje y la evolución, la naturaleza estática del entorno y la correlación entre los espacios de búsqueda considerados en la evolución y en el aprendizaje (estos espacios se consideran “correlacionados” si genotipos cercanos en el espacio de evolución se corresponden a fenotipos cercanos en el espacio fenotípico). Cuando los autores consideran diferentes tareas en el nivel de población (habilidad de un agente para encontrar comida en un entorno reticular) y en el nivel de aprendizaje en cada individuo (predecir la nueva sensorización como consecuencia de las acciones del agente) obtienen que de nuevo la combinación de ambos métodos consigue mejor calidad a lo largo del proceso evolutivo. Los autores concluyen que el aprendizaje fuerza a la evolución a seleccionar individuos que mejoran su calidad con respecto a ambas tareas: las consideradas en la evolución y en el aprendizaje; es decir, a seleccionar individuos que están situados en zonas dinámicamente correlacionadas de ambos espacios (los cambios que resultan en un incremento de calidad por el aprendizaje causan un incremento respecto a la tarea en el espacio de evolución). Harvey (1996, 1997) da incluso una explicación geométrica al explicar cómo el aprendizaje de los pesos del controlador neuronal del agente anterior, aun cuando las tareas son diferentes, la mayor parte de las veces mueve la configuración de pesos hacia una zona correlacionada de ambos espacios, hablando el autor

de un efecto de “reaprendizaje de los pesos perturbados por mutaciones anteriores” y preguntándose si existe “un nuevo factor en la evolución” (Harvey, 1997).

Giles Mayley (1997) apunta que se puede considerar, además, un efecto contrario al Efecto Baldwin, conocido como “*hiding effect*”: como la evolución prioriza a los individuos en función de su calidad en el espacio fenotípico, una vez han aprendido los individuos desde su configuración genética inicial, si el coste de aprendizaje es nulo o muy bajo, el aprendizaje tiende a “esconder” o no discernir las diferencias entre las configuraciones de los genotipos, por cuanto sus fenotipos y calidades finales son similares, lo que constituye una fuerza en sentido contrario al de encontrar las mejores configuraciones genotípicas. Este efecto, añade Mayley, actuará si el coste es bajo y la epístasis es baja, en tanto que a medida que aumenta el nivel de epístasis y coste empezará a dominar el Efecto Baldwin.

Existe otra serie de diversos trabajos que combinan las diferentes posibilidades comentadas, con resultados no concordantes en cuanto a determinar cuál es la mejor estrategia, ya que ésta parece depender en gran medida del problema. Por ejemplo, en (Julstrom, 1999) el autor compara evolución pura, estrategias de Lamarck y de Baldwin en un problema de particionado de nodos en ciclos de mínima longitud de 4. En sus pruebas la estrategia de Baldwin resultó ser la peor, en tanto la estrategia de Lamarck con una búsqueda local en todos los ciclos obtuvo el mejor resultado. Y Ku y Mak (1997) compararon las estrategias de Lamarck y Baldwin en un algoritmo genético celular (la reproducción sólo ocurre entre cromosomas vecinos) en la obtención de redes neuronales recurrentes. La estrategia *lamarckiana* mejoró la convergencia (error de la red) respecto a un algoritmo genético puro, pero no así la estrategia de Baldwin.

Como últimos ejemplos se cita a Whitley y col. (1994), quienes han demostrado que existen funciones para las cuales la utilización de un algoritmo genético sencillo sin aprendizaje, y con la combinación de la estrategia de Lamarck, convergen hacia soluciones óptimas locales, mientras que un algoritmo genético haciendo uso del efecto Baldwin converge a la solución óptima global, aunque esta búsqueda es más lenta que utilizando la estrategia de Lamarck. Y Sasaki y Tokoro (1997) han encontrado que la adaptación mediante una evolución *lamarckiana* (la herencia de caracteres es la que dirige la evolución) es más rápida que la evolución *darwinista* (la mutación y variabilidad genética



dirigen la evolución) cuando los ejemplos a aprender por una red neuronal son los mismos de generación en generación, es decir, en un entorno estático. Pero cuando el entorno es dinámico, cambiando de generación en generación, el mecanismo de Lamarck resultó inestable e inferior a la evolución *darwinista*, manteniendo esta última su adaptación respecto a los cambios dinámicos.

Estos trabajos comentados son algunos de los ejemplos que estudian las diferentes posibilidades, aunque la mayor parte del estudio en el campo computacional se ha centrado en los primeros aspectos de interrelación entre evolución y aprendizaje, mientras la asimilación genética de los nuevos caracteres adquiridos por los mecanismos de plasticidad a consecuencia del coste del aprendizaje ha recibido menos atención en el estudio del Efecto Baldwin, considerada más un aspecto teórico que de importancia en aplicaciones prácticas. Como remarca Turney (1996) “es común pensar en el Efecto Baldwin como la sinergia entre evolución y aprendizaje, pero eso es sólo mitad de la historia, al olvidarse de los costes que involucra el aprendizaje”.

Uno de los pocos trabajos que profundizan en el efecto de la asimilación y su relación con el coste de aprendizaje es el de Mayley (1996), quien extiende el trabajo de Hinton y Nowlan al considerar “espacios NK” de Kauffman con varios máximos locales, cuya rugosidad se controla, en la contribución parcial de calidad de cada uno de los  $N$  genes, con el nivel de epístasis en cada uno de esos  $N$  genes respecto de sus  $K$  genes vecinos. El autor utiliza genotipos y aprendizaje similares a los de Hinton y Nowlan, experimentando con la influencia del coste que implica el aprendizaje en la velocidad de la asimilación genética, y considerando dos tipos de coste en el aprendizaje: uno explícito en el que se resta calidad en función del número de alelos “?” en el genotipo y otro implícito en el que se suma menos calidad cuanto más se tarda en encontrar el mejor máximo durante los 1000 pasos de aprendizaje o búsqueda aleatoria. Sus resultados, imaginables a priori, se resumen en que la evolución pura con un algoritmo genético (AG) con 1000 individuos muestra que a mayor rugosidad (mayor  $K$ ) más dificultad en encontrar un máximo. Con una superficie suave ( $K=0$ ) el aprendizaje perjudica la búsqueda respecto a evolución pura por el precio que se paga por el aprendizaje, pero a medida que aumenta  $K$  la evolución de calidad (del mejor individuo) es más rápida respecto a no usar aprendizaje, y hacia mayor calidad cuanto menor es el coste. El uso de un coste implícito, por su parte, se asemeja a un alto coste explícito. Finalmente, la asimilación, medida como el número de alelos “?”

en la población, sigue igualmente tendencias razonables, como ser más rápida a mayor coste de aprendizaje.

### 3.3 Ejemplos del Efecto Baldwin en robótica evolutiva

Se estudia ahora la aplicación de la interrelación entre ambos métodos en un caso complejo y práctico como es el de obtención automática de comportamientos en robótica autónoma, siguiendo la metodología de robótica evolutiva (Nolfi y Floreano, 2000) en la obtención automática de controladores de comportamientos. Esta disciplina comienza su desarrollo a principios de los 90 (Harvey y col., 1993) al utilizar la evolución simulada para automatizar el desarrollo de arquitecturas de control en la alternativa de robótica basada en comportamientos o de construcción de “abajo hacia arriba” (Arkin, 1998), siguiendo la filosofía de Vida Artificial frente a la clásica de la Inteligencia Artificial simbólica.

La combinación de ambos métodos permitirá la consecución de los dos aspectos antes expuestos: suavizar la superficie de búsqueda generada en problemas de obtención automática de comportamientos mediante evolución y estudiar la asimilación genotípica al actuar el efecto Baldwin. Además, se exponen ejemplos propios en los que se proponen medidas cuantitativas de cómo va actuando la asimilación en el genotipo de las nuevas características adquiridas por aprendizaje.

En (Silva y Santos, 2004) se detallan ejemplos de la interrelación en robótica evolutiva, utilizando un robot Khepera y el simulador del robot desarrollado por Mitchel (1996), ejemplos de los que se sintetizan aquí los aspectos más interesantes. El robot Khepera es de pequeñas dimensiones (5.5 cm de diámetro) y es el más utilizado en la disciplina de robótica evolutiva. Posee ocho sensores de infrarrojos que pueden actuar bien como detectores de distancias o bien como detectores de intensidad de luz ambiente. Como actuadores posee dos ruedas y se pueden añadir módulos a medida adicionales como una sencilla cámara o una pinza.

Se realizaron pruebas usando como estructuras de control de comportamientos redes neuronales artificiales (RNAs). Las RNAs son las estructuras de control más utilizadas en la disciplina por sus ventajas de tolerancia al ruido –esencial debido al ruido presente en los entornos reales–, y a que en una codificación directa genotipo-fenotipo las primitivas consideradas por el algoritmo evolutivo –los parámetros que definen la red como las conexiones– deben ser, en palabras de Nolfi y

Floreano (2000), “del más bajo nivel posible a fin de evitar selecciones no adecuadas por parte del diseñador humano”. Además, la existencia de algoritmos de aprendizaje bien establecidos en topologías típicas, esencial en este caso de experimentación con aprendizaje en tiempo de vida de los robots, hace más ventajosa su elección.

Se eligió como comportamiento el seguimiento de contornos, suficientemente complejo porque el robot debe seguir un contorno a la mayor velocidad posible y sin chocar, complicado además por el hecho de que el robot debe, partiendo de cualquier posición en el entorno, poder buscar una pared o contorno para seguirlo. Para ello, al igual que Floreano y Mondada (1998), se definió la función de calidad como la suma de:

$$\Delta \text{calidad} = V \cdot (1 - \sqrt{\Delta V})$$

en cada paso o iteración del robot ( $V$  representa la velocidad media, e  $\Delta V$  una medida de la velocidad angular), siempre que el robot detecte un objeto o pared con cualquiera de sus sensores por debajo de un umbral, que define una franja o zona de cercanía a la pared en la que el robot debe mantenerse en su recorrido. La función de calidad intenta así maximizar la velocidad lineal y minimizar los giros bruscos. Además, cada vez que el robot choca, se penaliza su calidad. Ésta, finalmente, es promediada para dos posiciones de partida diferentes del robot.

Remarcar que, aunque no existe un coste de aprendizaje explícito en la función de calidad, como el robot no tiene por qué partir de la zona de cercanía, existe un coste inherente al aprendizaje, puesto que cuanto más tarda un robot en alcanzar esa franja de cercanía o bien la abandone posteriormente, menor calidad obtendrá y, en consecuencia, los individuos o controladores que aprendan más rápido (más cercanos a una solución óptima) tendrán mayor calidad, y al revés. O dicho de otro modo, hay un coste implícito representado por el tiempo que el robot no consigue aproximarse y seguir correctamente una pared.

### Combinación con aprendizaje *hebbiano*

Se utiliza primeramente aprendizaje *hebbiano* en un controlador neuronal para la obtención del comportamiento comentado. En esta línea, Mayley (1996), también ha utilizado aprendizaje *hebbiano* en el comportamiento de seguimiento de paredes. El autor usa una función de calidad similar que implica, como se ha dicho, que cuanto más rápido aprenda un individuo a seguir una pared, menos coste tendrá de aprendizaje y más ventaja evolutiva

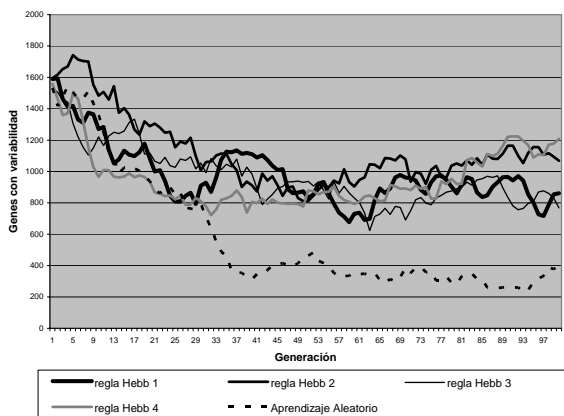
tendrá para pasar su material genético –con unos pesos más cercanos a los óptimos– a las siguientes generaciones. Para el estudio de la asimilación genética el autor asocia en un gen adicional un bit a cada conexión que determina si el peso correspondiente puede modificarse o no en tiempo de vida. El autor comprueba cómo aquellos pesos del controlador neuronal cuyo valor no está determinado genéticamente son cada vez menos necesarios a lo largo de las generaciones, ya que se van asimilando en el genotipo aquellos valores numéricos de los pesos de conexión de la red neuronal que consiguen que ésta realice correctamente el comportamiento.

En (Silva y Santos, 2004) se utilizó un perceptron como controlador, sin necesidad de nodos ocultos (8 nodos de entrada conectados a los sensores y dos nodos de salida sigmoidales de control de velocidad de las dos ruedas), y se usaron las mismas 4 reglas *hebbianas* que Floreano y Mondada (1998), con codificación directa en el genotipo del valor de las conexiones, y que mantienen el peso dentro de un rango  $[-3,3]$  en nuestro caso). Y al igual que Mayley (1996) se asoció a cada gen (peso de conexión) un bit que indica si la conexión es susceptible o no de cambio en el tiempo de aprendizaje.

Los parámetros comunes utilizados en el proceso evolutivo fueron: AG con población de 200 individuos, selección por torneo con ventana de tamaño 20, probabilidad de cruce 0.9, probabilidad de mutación 0.15 y coeficiente de aprendizaje en las reglas de 0.05.

En la evolución de la calidad del mejor individuo a lo largo del proceso evolutivo con las reglas de Hebb, con las que los individuos aprendían durante 1000 iteraciones, se detectó que la mayoría de reglas conseguían obtener mejor calidad respecto al caso sin aprendizaje sólo en las primeras generaciones (Silva y Santos, 2004). Aunque aquí se presenta el posible efecto de asimilación genética, en la pretensión de realizar un mayor énfasis en tal aspecto.

En la figura 4 se muestra ese efecto. Como cada peso de conexión tiene un bit de plasticidad asociado, al existir el coste de aprendizaje, si ocurre la asimilación, ese número de bits activados debería decrementarse en tiempo evolutivo, lo que indica que se van fijando paulatinamente en el genotipo los pesos correspondientes al controlador óptimo. La figura representa así la disminución del número de bits en uno, para las cuatro reglas de Hebb utilizadas, y considerando la totalidad de la población. Se incluyen en esa asimilación los genes



**Figura 4. Asimilación genética utilizando aprendizaje hebbiano y búsqueda aleatoria.**

que, aún manteniendo el bit asociado en 1, no les afecta el aprendizaje por estar continuamente en sus valores extremos.

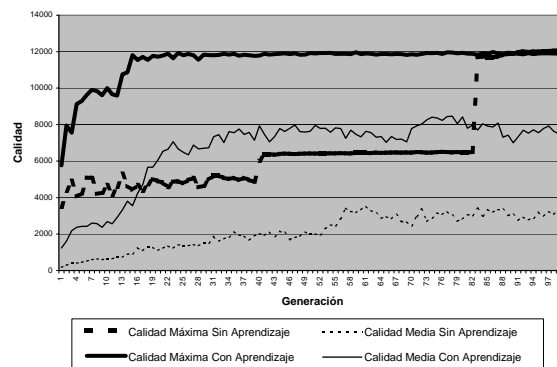
La figura muestra la asimilación, aunque no continua hacia un valor mínimo con las reglas de Hebb, siendo su efecto mejor en el caso de una búsqueda local aleatoria en la vecindad de cada cromosoma, al estilo de la utilizada por Hinton y Nowlan (1987), ya que en este caso los cambios aleatorios lo más probable es que quiten al robot de la trayectoria correcta, por lo que el AG tiende rápidamente a eliminar esos genes con bit asociado en 1 del *pool* genético. En este caso de aprendizaje hebbiano, como también supone Mayley (1996), el algoritmo genético está buscando por genotipos iniciales que combinados con una regla de plasticidad adecuada de Hebb consiguen situar y mantener el robot en la trayectoria requerida, genotipos sobre los que actúa el efecto Baldwin con la consecuente asimilación genética. Es decir, estamos en el caso en que la evolución busca genotipos o semillas acordes a mejorar por las reglas de aprendizaje.

**Combinación con aprendizaje por refuerzo**

Se utiliza ahora aprendizaje por refuerzo como método de búsqueda local en la vecindad de cada individuo del proceso evolutivo. El algoritmo de retropropagación del error por refuerzo complementario, definido por Ackley y Littman (1990), y utilizado en robótica, entre otros autores, por Meeden (1996), combina el aprendizaje supervisado por retropropagación de error con la idea de refuerzo, haciendo así viable el aprendizaje por refuerzo en el caso de controladores neuronales en robótica, en la cual no es posible determinar por el diseñador todas las acciones correctas que debe realizar el robot ante todas las posibles entradas o situaciones, aunque sí indicar de modo natural una

idea de refuerzo positivo o de castigo en la realización del comportamiento requerido.

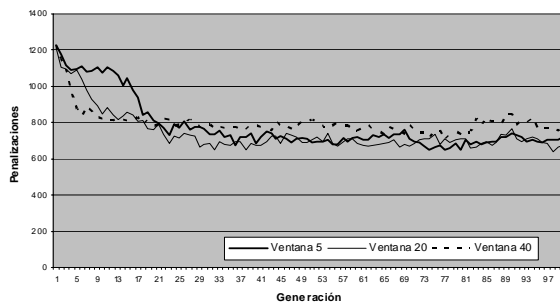
Las salidas del controlador neuronal se asocian ahora a acciones concretas a realizar por el robot. Con el nivel analógico de los nodos de salida se elige la acción que corresponde al de mayor activación o bien se interpretan sus valores de modo probabilístico para determinar unas salidas binarias que determinan la acción o acciones a realizar. La idea general del algoritmo es que, ante refuerzos positivos, el algoritmo modifica los pesos para que la red produzca la misma acción que produjo el refuerzo y en las mismas circunstancias (aproximar la salida hacia la decisión binaria), mientras que si el refuerzo es negativo el algoritmo tratará de variar los pesos de conexión para que la próxima vez se elija el movimiento complementario (por ejemplo, si la acción fue moverse a derecha, el complementario será moverse a izquierda). Es decir, en función del refuerzo, se proporciona la salida deseada para entrenar por retropropagación.



**Figura 5. Calidad del mejor individuo y calidad media con y sin aprendizaje por refuerzo en los individuos.**

Se utiliza ahora un controlador neuronal con 8 nodos de entrada que recogen la información sensorial, 5 ocultos y tres nodos de salida que se corresponden a tres acciones posibles (ir de frente, girar derecha, girar izquierda), de nuevo para el comportamiento de seguimiento de contornos. Que la activación de un nodo de salida corresponda a una acción concreta facilita en gran medida la determinación de la acción complementaria en caso de castigo. Con estas premisas se comprueban ahora los dos aspectos de interés: la suavización del espacio de búsqueda y la asimilación genética.

La figura 5 muestra la evolución de la calidad del mejor individuo y de la calidad media de la población a lo largo de 100 generaciones considerando la posibilidad de aprendizaje –durante 1000 iteraciones o pasos de vida del robot– y



**Figura 6. Efecto de la presión evolutiva en la asimilación genética.**

únicamente evolución, siendo los parámetros del proceso evolutivo comunes al caso anterior. La gráfica denota de modo claro cómo la combinación con el aprendizaje consigue que tanto el mejor individuo, como la media de calidad, lleguen en pocas generaciones a valores mucho mejores respecto al caso de no aprendizaje. La evolución de la calidad máxima indica que el comportamiento óptimo se obtiene en muy pocas generaciones (en torno a la 14), frente a la evolución por separado, que obtiene también el comportamiento óptimo pero tras más de 80 generaciones de proceso evolutivo. Por su parte, la calidad media se mantiene en un nivel mucho más alto en la combinación con aprendizaje a lo largo de todo el proceso evolutivo. Éstas son indicaciones claras de que el aprendizaje ha suavizado en gran medida la superficie de calidad, facilitando así la búsqueda genética. Además, en estos ejemplos en robótica, el aprendizaje no conlleva un tiempo considerable de cómputo, ya que éste se realiza *on-line*, es decir, en cada uno de los 1000 pasos de vida de cada individuo se aplica una o muy pocas iteraciones de entrenamiento con el algoritmo de retropropagación por refuerzo complementario, tiempo que es muy inferior al propio de simulación de los sensores y del movimiento del robot.

Para comprobar si se produce el efecto Baldwin con asimilación genética, se chequean ahora las penalizaciones o número de acciones que resultan en refuerzo negativo, promediadas sobre toda la población (figura 6). Remarcar de nuevo este aspecto del Efecto Baldwin: como a todos los individuos se les permite aprender durante el mismo número de pasos y que lo aprendido no revierte directamente en el genotipo del que parte, la asimilación no tendría por qué tener lugar. Pero, al existir el coste en el aprendizaje, cuanto más cerca del óptimo está el comportamiento del individuo –y, en consecuencia, su genotipo–, mayor calidad, y la evolución tenderá a elegir a aquellos individuos y genotipos que cada vez están más cerca del óptimo y que cada vez cometerán menos penalizaciones, lo

que proporciona esa medida de cómo se va asimilando lo adquirido en tiempo de vida sobre el genotipo. La figura 6 muestra así ese efecto de asimilación para diferentes tamaños de ventana de torneo (5, 20 y 40), mostrando ese número de penalizaciones promediadas sobre todos los individuos de la población. A mayor presión de la selección (mayor tamaño de ventana) más rápida es la asimilación, aunque después de unas 20 generaciones el nivel es similar.

#### 4. Discusión final

Se ha expuesto que la ventaja de los métodos evolutivos es su búsqueda en paralelo en múltiples puntos del espacio de búsqueda, frente a una búsqueda local o aprendizaje que se corresponde con la plasticidad fenotípica en los organismos biológicos. La interrelación entre ambos métodos admite dos posibilidades básicas: aquella en la cual los caracteres adquiridos intervienen en el establecimiento de la nueva población en la generación siguiente (estrategia de Lamarck) y aquella en la que el aprendizaje determina la calidad de cada individuo, aunque no se consideran los nuevos caracteres adquiridos como material genético para la siguiente generación. Esta segunda posibilidad permite suavizar el espacio de búsqueda, acelerando o facilitando la evolución y, además, si el aprendizaje –o plasticidad en general– implica un coste en calidad aparece la asimilación en los genotipos, de modo paulatino a lo largo de las generaciones, de los caracteres adquiridos, que constituye el aspecto más genuino del Efecto Baldwin. Aunque, como se ha reseñado, la mayor parte de trabajos en la bibliografía científica del tema se centran en el aspecto de facilitación de la evolución, mientras pocos inciden en el aspecto de asimilación y su relación con el coste de aprendizaje.

Además de la facilitación de la evolución, ¿qué otras razones hay para usar ambos métodos combinadamente? Una reside en que la principal ventaja de utilizar aprendizaje es que éste permite la adaptación a cambios más rápidos en el entorno que la evolución por sí sola no es capaz de seguir, como se había comentado en el apartado 2. En palabras de Hinton y Nowlan (1987) “el argumento más común a favor del aprendizaje es que algunos de los aspectos del entorno no son predecibles, de modo que es positivamente ventajoso dejar algunas decisiones al aprendizaje mejor que especificarlas genéticamente”. Otra razón, relacionada con la facilitación de la evolución, es la propia eficiencia en la búsqueda. De nuevo, en palabras de los mismos autores: “en búsquedas mediante evolución, que requieren el chequeo de muchas posibilidades

para descubrir una co-adaptación compleja, cada prueba de aprendizaje puede ser de una ayuda casi igual en la búsqueda evolutiva que la producción y evaluación de un nuevo organismo completo". Es decir, se incrementa en gran medida la eficiencia de la evolución porque una prueba de aprendizaje es más rápida y requiere mucho menos gasto que la producción de un organismo completo.

Finalmente, está el interés de interrelación entre diferentes disciplinas. Parisi, Nolfi y Cecconi (1991) comentaban hace más de 10 años diferentes razones por las que los biólogos evolutivos tendían a no considerar el Efecto Baldwin: la ortodoxia reduccionista de éstos, que implica que las causas y mecanismos básicos de la evolución se encuentran sólo en el nivel de la genética; y, como consecuencia, el aprendizaje y el comportamiento han sido en gran medida ignorados en el intento de comprender los mecanismos de la evolución. Otra razón de la falta de atención es que los biólogos evolutivos entienden al aprendizaje y al comportamiento como propios del ámbito de la psicología y la etología, pero no de la biología. Y finalmente, los pocos estudios del Efecto Baldwin en poblaciones reales o simuladas (Parisi y col., 1991). El Efecto Baldwin, sin embargo, está ahora bien determinado en el campo de computación, siendo un ejemplo claro de la positiva interrelación entre diferentes disciplinas. Prueba de ello es cómo los biólogos teóricos han recurrido a las simulaciones de la evolución para, por ejemplo, determinar las condiciones bajo las cuales se facilita la evolución al incorporar la plasticidad (Ancel, 2000).

## Referencias

- Ackley, D.H. and Littman, M.L. (1990), "Generalization and Scaling in Reinforcement Learning". *Advances in Neural Information Processing Systems*, Vol. 2. D.S. Touretzky (Ed.). Morgan Kaufmann, CA. pp. 550-557.
- Ancel, L.W. (2000), "Undermining the Baldwin Expediting Effect: Does Phenotypic Plasticity Accelerate Evolution? *Theoretical Population Biology*, No. 58, pp. 207-319.
- Arkin, R.C. (1998), *Behavior Based Robotics*. MIT Press, Cambridge, MA.
- Baldwin, J.M. (1896), "A New Factor in Evolution". *American Naturalist*, Vol. 30, pp. 441-451.
- Belew, R.K., McInerney, J. and Schraudolph, N.N. (1991), "Evolving Networks: Using the Genetic Algorithm with Connectionist Learning". *Proceedings of the Second Artificial Life Conference*, Addison-Wesley, pp. 511-547.
- Boers, E. and Sprinkhuizen-Kuyper, I. (2001), "Combined Biological Metaphors". *Advances in the Evolutionary Synthesis of Intelligent Agents*. M. Patel, V. Honaver and K. Balakrishnan (Eds.), MIT Press, pp. 153-183.
- Floreano, D. and Mondada, F. (1998), "Evolutionary Neurocontrollers for Autonomous Mobile Robots". *Neural Networks*, Vol. 11, pp. 1461-1478.
- French, R. and Messinger, A. (1994), "Genes, Phenotypes and the Baldwin Effect". In R. Brooks and P. Maes (Eds.), *Artificial Life IV*. Cambridge MA: MIT Press, pp. 277-282.
- Harvey, I. (1993), "The Puzzle of the Persistent Question Marks: A Case Study of Genetic Drift". In S. Forrest (Ed.) *Proceedings of the Fifth International Conference on Genetic Algorithms, ICGA-93*, California: Morgan Kaufmann, pp. 15-22.
- Harvey, I. (1996), "Relearning and Evolution in Neural Networks". *Adaptive Behavior*, Vol. 4, pp. 81-84.
- Harvey, I. (1997), "Is There Another New Factor in Evolution?". *Evolutionary Computation*, Vol. 4 No. 3, pp. 313-329.
- Harvey, I., Husbands, P. and Cliff, D. (1993), "Issues in Evolutionary Robotics". *From Animals to Animats 2, Proceedings of the Second International Conference on Simulation of Adaptive Behavior*, MIT Press, MA, pp. 364-373.
- Hinton, G.E. and Nowlan, S.J. (1987), "How Learning can Guide Evolution". *Complex Systems*, Vol. 1, pp. 495-502.
- Julstrom, B.A. (1999), "Comparing Darwinian, Baldwinian and Lamarckian Search in a Genetic Algorithm for the 4-Cycle Problem". *Late Breaking Papers at the 1999 Genetic and Evolutionary Computation Conference*, pp. 134-138, Orlando-USA.
- Ku, K. and Mak, M.W. (1997), "Exploring the Effects of Lamarckian and Baldwinian Learning in Evolving Recurrent Neural Networks", *Int. Conf. on Evolutionary Computation (ICEC'97)*, pp. 617-620.
- Mayley, G. (1996), "The Evolutionary Cost of Learning". *From Animals to Animats 4. Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, MIT Press, pp. 458-467.

- Mayley, G. (1997), "Guiding or Hiding: Explorations into the Effects of Learning on the Rate of Evolution". *Proceedings of the Fourth European Conference on Artificial Life (ECAL '97)*, P. Husbands and I. Harvey (Eds).
- Maynard Smith, J. (1993), *The Theory of Evolution* (3rd edition). Cambridge University Press.
- Meeden, L. (1996), "An Incremental Approach to Developing Intelligent Neural Network Controllers for Robots". *IEEE Transactions on Systems, Man and Cybernetics Part. B: Cybernetics*, Vol. 26, No. 3, pp. 474-485.
- Mitchel, O. (1996), *Khepera Simulator Package version 2.0*, <http://wwwi3s.unice.fr/om7khep-sim.html>. University of Nice Sophia-Antipolis, France.
- Mitchell, M. and Forrest, S. (1994) "Genetic Algorithms and Artificial Life". *Artificial Life*, Vol. 1, No. 3, pp. 267-289.
- Moscato, P. (1989), "On Evolution, Search, Optimization, Genetic Algorithms and Martial Arts: Towards Memetic Algorithms". *Caltech Concurrent Computation Program, C3P Report 826*.
- Moscato, P. y Cotta, C. (2003), "Una Introducción a los Algoritmos Meméticos". *Revista Iberoamericana de Inteligencia Artificial*, No. 19, pp. 131-147.
- Nolfi, S., Elman, J.L. and Parisi, D. (1994), "Learning and Evolution in Neural Networks". *Adaptive Behavior*, Vol. 3, pp. 5-28.
- Nolfi, S. and Floreano, D. (2000), *Evolutionary Robotics*. MIT Press.
- Nolfi, S. and Parisi, D. (1997), "Learning to Adapt to Changing Environments in Evolving Neural Networks". *Adaptive Behavior*, Vol. 5, No. 1, pp. 75-98.
- Parisi, D., Nolfi, S. and Cecconi, F. (1991), "Learning, Behavior, and Evolution". In F.J. Varela and P. Bourguine (Eds.). *Towards a Practice of Autonomous Systems, Proc. of the First European Conference on Artificial Life (ECAL)*, pp. 207-216. MIT Press.
- Sasaki, T. and Tokoro, M. (1997), "Adaptation toward Changing Environments: Why Darwinian in Nature". *Proceedings of 4th European Conference on Artificial Life (ECAL-97)*, pp.145-153.
- Silva, F. y Santos, J. (2004), "Estudio del Efecto Baldwin en la Obtención de Controladores en Robótica Autónoma". *Actas del Tercer Congreso Español de Metaheurísticas, Algoritmos Evolutivos y Bioinspirados*, pp. 270-277, Córdoba.
- Turney, P. (1996), "Myths and Legends of the Baldwin Effect". *Proceedings of the Workshop on Evolutionary Computing and Machine Learning at the 13th International Conference on Machine Learning (ICML-96)*, Bari, Italy, pp. 135-142.
- Waddington, C.H. (1942), "Canalization of Development and the Inheritance of Acquired Characters". *Nature*, pp. 563-565.
- Waddington, C.H. (1953). "Genetic Assimilation of an Acquired Character". *Evolution* 7, pp. 118-126.
- Whitley, D., Gordon, V.S. and Mathias, K. (1994), "Lamarckian Evolution, the Baldwin Effect and Function Optimization". *Lecture Notes in Computer Science*, Vol. 866. Springer-Verlag. pp. 6-15.