

---

## Aptitud relativa de los componentes básicos y la hipótesis de los bloques de construcción

---

Stephanie Forrest  
Departamento de Ciencias de la  
Computación Universidad de Nuevo  
México Albuquerque, NM 87131

Melanie Mitchell •  
Laboratorio de IA  
Universidad de Michigan  
Ann Arbor, MI 48109

En D. Whitley (ed.) Fundamentos de los algoritmos genéticos 2,  
Morgan Kaufmann, San Mateo, CA, 1993.

### Resumen

los bloque de construcción La hipótesis establece que el GA funciona bien cuando se recombinan esquemas cortos, de orden bajo y muy ajustados para formar esquemas de orden superior aún más ajustados. Se cree que la capacidad de producir soluciones parciales más y más ajustadas mediante la combinación de bloques de construcción es una fuente principal del poder de búsqueda de GA, pero la comunidad de investigación de GA actualmente carece de descripciones precisas y cuantitativas de cómo el procesamiento de esquemas tiene lugar realmente durante la evolución típica de un Búsqueda GA. Otro problema abierto es caracterizar en detalle los tipos de paisajes de fitness para los que el cruce será un operador eficaz. En este artículo describimos primero una clase de paisajes de fitness (las funciones de "Royal Road") que hemos diseñado para investigar estas cuestiones. A continuación, presentamos algunos resultados experimentales inesperados relacionados con el desempeño del AG en instancias simples de estos paisajes, en los que variamos la fuerza del refuerzo a partir de los "escalones": esquemas de orden intermedio ajustados obtenidos mediante la recombinación de esquemas de orden inferior ajustados. Finalmente, comparamos el desempeño del AG en estas funciones con el de tres esquemas de escalada de uso común,

y descubra que uno de ellos, "escalada de colinas por mutación aleatoria", supera significativamente al GA en estas funciones.

## 1. INTRODUCCIÓN

La investigación sobre los fundamentos de los algoritmos genéticos aspira a responder dos preguntas generales: ¿Cómo funcionan los AG y para qué sirven? Una teoría exitosa de los AG describiría las leyes que gobiernan el comportamiento de los esquemas en los AG y caracterizaría los tipos de paisajes de fitness en los que es probable que el AG funcione bien, especialmente en comparación con otros métodos de búsqueda como la escalada. Esto, por supuesto, requiere una declaración de lo que significa que un GA "se desempeñe bien". Es decir, necesitamos una mejor comprensión de qué es lo que hace bien el GA (por ejemplo, encontrar un óptimo global en lugar de encontrar rápidamente una solución bastante buena).

Nuestra estrategia para responder a estas preguntas consiste en el siguiente enfoque general. Empezamos identificando características paisajes de aptitud que son particularmente relevantes para el desempeño de la AG. Varias de estas características se han discutido en la literatura de AG, incluidas las colinas locales, los "desiertos", el engaño, los bloques de construcción estructurados jerárquicamente, el ruido y la varianza de alta aptitud dentro de los esquemas. Luego diseñamos paisajes simplificados que contienen diferentes configuraciones de tales características, variando, por ejemplo, la distribución, frecuencia y tamaño de diferentes características en el paisaje. Luego estudiamos en detalle los efectos de estas características en el comportamiento del AG. Un objetivo a más largo plazo de esta investigación es desarrollar métodos estadísticos para clasificar cualquier paisaje dado en términos de nuestro espectro de paisajes diseñados a mano, pudiendo así predecir algunos aspectos del desempeño de la AG en el paisaje dado.

Cabe señalar que al plantear este problema en términos del desempeño de la AG en los paisajes del fitness, estamos eludiendo la cuestión de cómo se puede representar mejor un problema particular ante la AG. El éxito de la GA en una función en particular está ciertamente relacionado con cómo se "codifica" la función (Goldberg, 1989b; Liepins & Vose, 1990) (p. Ej., El uso de códigos Gray para parámetros numéricos puede mejorar en gran medida el rendimiento de la GA en algunos problemas), pero como estamos interesados en los sesgos que pertenecen directamente a la AG, simplemente consideraremos el panorama que la AG "ve".

En este artículo describimos algunos resultados iniciales de este programa de investigación a largo plazo. Empezamos centrándonos en el hipótesis de bloques de construcción Holanda, 1975; Goldberg, 1989b), que establece que el AG funciona bien cuando los esquemas cortos, de bajo orden y muy ajustados ("bloques de construcción") se combinan para formar esquemas de orden superior aún más ajustados. En palabras de Goldberg, "... construimos cadenas cada vez mejores a partir de las mejores soluciones parciales de muestreos pasados" (Goldberg, 1989b, p. 41). Se cree que la capacidad de producir soluciones parciales más y más ajustadas mediante la combinación de bloques de construcción es la fuente principal del poder de búsqueda de la AG. Sin embargo, a pesar del supuesto papel central de los bloques de construcción y la recombinación, la comunidad de investigación de AG carece de descripciones precisas y cuantitativas de cómo los esquemas interactúan y se combinan durante

la evolución típica de una búsqueda GA. Por lo tanto, estamos interesados en aislar las características del paisaje implícitas en la hipótesis de los bloques de construcción y estudiar en detalle el comportamiento de la AG (la forma en que se procesan los esquemas y se combinan los bloques de construcción) en paisajes simples que contienen esas características.

Otros investigadores de GA han estudiado estas mismas preguntas utilizando diferentes técnicas. El enfoque más destacado ha sido estudiar los efectos de Engaño de GA sobre el desempeño de la AG (por ejemplo, Goldberg, 1987, 1989a; Liepins & Vose, 1990; Whitley, 1991). Sin embargo, el engaño es solo una de las muchas características de un problema que afecta el desempeño de GA (por ejemplo, ver Liepins & Vose, 1990, y Forrest & Mitchell, 1991). En lugar de estudiar problemas difíciles en los que falla la AG, nuestro enfoque inicial ha sido examinar el comportamiento de la AG en paisajes en los que es probable que funcione bien. Al comprender qué características de esos paisajes conducen a un buen desempeño, esperamos caracterizar mejor la clase de tales paisajes.

Un componente importante de este esfuerzo es definir la clase más simple de paisajes en los que el AG se desempeña "como se esperaba", confirmando así las amplias afirmaciones de la hipótesis de los bloques de construcción. Sin embargo, la tarea de diseñar tales paisajes ha resultado ser sustancialmente más difícil y compleja de lo que anticipamos originalmente. Nuestras elecciones iniciales de paisajes simples han revelado algunos fenómenos sorprendentes e inesperados. La historia de cómo pequeñas variaciones de un paisaje básico pueden hacer que la búsqueda de GA sea mucho menos efectiva revela mucho sobre la complejidad de los GA y señala la necesidad de una teoría más profunda de cómo se descubren los bloques de construcción de orden inferior y se combinan en orden superior. soluciones.

En las siguientes secciones presentamos el Camino real funciones, una clase de funciones no engañosas en las que los bloques de construcción se definen explícitamente. Luego mostramos cómo las variantes simples de estas funciones pueden tener efectos bastante diferentes sobre el desempeño de la AG y discutimos las razones de estas diferencias.

## 2 PIEDRAS PASO A PASO EN EL CRUCE

### PAISAJE

La hipótesis de los bloques de construcción sugiere dos características del paisaje que son particularmente relevantes para la AG: (1) la presencia de esquemas cortos, de bajo orden y muy ajustados; y (2) la presencia de "escalones" intermedios —esquemas de aptitud superior de orden intermedio que resultan de combinaciones de esquemas de orden inferior, y que a su vez pueden combinarse para crear esquemas de aptitud incluso superior. Dos preguntas básicas sobre los trampolines son: ¿Cuánto más alto nivel de condición física deben tener los trampolines intermedios para que la AG funcione bien? ¿Y cómo deben configurarse estos escalones? Para investigar estas preguntas, primero definimos las funciones de Royal Road, que contienen estas características explícitamente.

Para construir una función de Royal Road, seleccionamos una cadena óptima y la dividimos en varios bloques de construcción pequeños, como se ilustra en la Figura 1. Luego asignamos valores a cada esquema de orden inferior y a cada combinación intermedia posible de esquemas de orden inferior, y use esos valores para calcular la aptitud de una cadena de bits  $X$  En términos

Figura 1: Una cadena óptima dividida en ocho bloques de construcción. La función  $R_1(X)$  (donde  $X$  es una cadena de bits) se calcula sumando los coeficientes  $C_s$  correspondiente a cada uno de los esquemas dados de los cuales  $X$  es una instancia. Por ejemplo,  $R_1(1111111100...0) = 8$ , y  $R_1(1111111100...01111111) = 16$ . Aquí  $C_s = \text{orden}(s)$ .

La función  $R_1$ , ilustrado en la Figura 1, se calcula de manera muy simple: una cadena de bits  $X$  obtiene 8 puntos agregados a su idoneidad para cada uno de los esquemas de orden de 8 dados de los que es una instancia. Por ejemplo, si  $X$  contiene exactamente dos de los bloques de construcción de orden 8,  $R_1(x) = 16$ . Asimismo,  $R_1(1111...1) = 64$ . Dicho de manera más general, el valor  $R_1(X)$  es la suma de los coeficientes  $C_s$  correspondiente a cada esquema dado del cual  $X$  es una instancia. Aquí  $C_s$  es igual a pedidos). La contribución al fitness de un escalón intermedio (como la combinación de  $s_1$  y  $s_3$  en la Figura 1) es, por lo tanto, una combinación lineal de la contribución al fitness de los componentes de nivel inferior.

$R_1$  es similar al problema de la “meseta” descrito por Schafer y Eshelman (1991).

Según la hipótesis de los bloques de construcción, R La estructura de bloques de construcción y trampolín de 1 debería trazar un "camino real" para que la AG lo siga hasta el óptimo global. Por el contrario, un algoritmo como el simple ascenso más empinado de la colina, que prueba sistemáticamente mutaciones de un solo bit y solo se mueve en una dirección ascendente, no puede encontrar fácilmente valores altos en dicha función, ya que una gran cantidad de posiciones de un solo bit debe optimizarse simultáneamente para pasar de una instancia de un esquema de orden inferior (p. ej., 11111111 \*\*... \*) a una instancia de un esquema intermedio de orden superior (p. ej., 11111111 \*\*\*\*\* 11111111 \*\*... \*). Si bien es posible que se requiera una búsqueda aleatoria para encontrar los bloques de construcción del nivel más bajo (dependiendo del tamaño de la población inicial y el tamaño de los bloques del nivel más bajo), el aspecto interesante de R 1 está estudiando cómo los bloques de nivel inferior se combinan en los de nivel superior, y este es el aspecto que más nos preocupa. Parte de nuestro propósito al diseñar las funciones de Royal Road es construir una clase de paisajes en forma que distinga el GA de otros métodos de búsqueda como la escalada. En realidad, esto resultó ser más difícil de lo que anticipamos, como se discutirá en la Sección 5.

Esta clase de funciones proporciona un laboratorio ideal para estudiar el comportamiento del GA:

- El paisaje puede variar de varias formas. Por ejemplo, la "altura" de varios escalones intermedios se puede aumentar o disminuir (por ejemplo, la contribución al estado físico puede ser una combinación no lineal de las contribuciones al estado físico de los componentes). Además, el tamaño de los bloques de construcción de menor orden se puede variar, al igual que el grado en que cubren el óptimo. Finalmente, se pueden introducir diferentes grados de engaño al permitir que los esquemas de orden inferior difieran en algunos bits de los trampolines de orden superior, creando efectivamente esquemas de orden inferior que alejen al GA de los buenos esquemas de orden superior. Los efectos de estas variaciones en el comportamiento del AG pueden estudiarse en detalle.
- Dado que el óptimo global y, de hecho, todos los posibles valores de aptitud física, se conocen de antemano, es fácil comparar el desempeño del GA en diferentes variaciones de las funciones de Royal Road.
- Todos los esquemas deseados se conocen de antemano, ya que están explícitamente integrados en la función. Por lo tanto, la dinámica del proceso de búsqueda se puede estudiar en detalle rastreando las ontogenias de los esquemas individuales.

Estamos usando las funciones de Royal Road para estudiar una serie de preguntas sobre los efectos del cruce en varios paisajes, incluyendo las siguientes: Para un paisaje dado, ¿en qué medida el cruce ayuda a la AG a encontrar esquemas altamente ajustados? ¿Cuál es el efecto del cruce sobre los tiempos de espera para que se descubran esquemas deseables? ¿Cuáles son los cuellos de botella en el proceso de descubrimiento? ¿Cómo afecta la configuración de los escalones y el tamaño de los escalones definidos por los escalones el desempeño del AG? Responder a estas preguntas en el contexto de las funciones idealizadas de Royal Road es un primer paso para responderlas en casos más generales.

Primero investigamos el efecto del tamaño del paso de los escalones intermedios en el desempeño del AG. Para hacer esto, comparamos el desempeño de GA en

R 1 con su desempeño en una segunda función R 2, donde las contribuciones al fitness de ciertos escalones intermedios son mucho mayores. R 2 se ilustra en la Figura 2.

R 2 se calcula de la misma forma que R 1: la aptitud de una cadena de bits  $X$  es la suma de los coeficientes correspondientes a cada esquema ( $s_1 - s_{14}$ ) del cual es una instancia.

Por ejemplo,  $R_2(1111111100...011111111) = 16$ , ya que la cadena es una instancia de ambos  $s_1$  y  $s_8$ , pero  $R_2(11111111111111100...0) = 32$ , ya que la cadena es una en vez de  $s_1$ ,  $s_2$ , y  $s_9$ . Por tanto, la aptitud de una cadena depende no sólo del número de esquemas de 8 bits a los que pertenece la cadena, sino también de sus posiciones en la cadena.

La cuerda óptima 11111111...1 tiene aptitud 192, ya que la cadena es una instancia de cada esquema en la lista.

### 3 EXPERIMENTOS DE CARRETERA REAL

En un artículo anterior (Mitchell, Forrest y Holland, 1992) informamos algunos resultados iniciales sobre las funciones de Royal Road. Nuestra principal medida de desempeño fue el número

Forrest y Mitchell, 1991): un individuo  $i$  El número esperado de offspring en primavera es  $1 + \frac{F_i - \bar{F}}{\sigma}$ , donde  $F_i$  es el estado físico,  $\bar{F}$  es el estado físico medio de la población, y  $\sigma$  es la desviación estándar. La salida máxima esperada de cualquier cuerda era 1,5, si lo anterior fórmula dio un valor más alto, el valor se restableció a 1,5. Este es un corte estricto, ya que implica que la mayoría de los individuos se reproducirán solo 0, 1 o 2 veces. El efecto de este esquema de selección es ralentizar la convergencia al restringir el efecto que un solo individuo puede tener en la población, sin importar cuánto más en forma sea que el resto de la población. Incluso con esta precaución, observamos algunos

EXPERIMENTO ORIGINAL		
	Evaluaciones de funciones al óptimo	
500 carreras	R 1	R 2
Significar	62099 (error estándar: 1390) 56576	73563 (error estándar: 1794) 66304
Mediana		

Tabla 1: Resumen de los resultados de ejecutar el GA en R 1 y R 2. La tabla da las evaluaciones de función media y mediana tomadas para encontrar el óptimo en 500 corridas en cada función. Los números entre paréntesis son los errores estándar.

interesantes efectos de convergencia prematura (descritos en la siguiente sección). La probabilidad de cruce fue de 0,7 por pareja de padres y la probabilidad de mutación fue 0,005 por bit.

La probabilidad de que una cadena generada aleatoriamente contenga uno de los niveles inferiores

los esquemas de orden 8 son  $8 \cdot \frac{1}{2^8} = \frac{1}{2}$ . Dado que la población inicial tiene 128 al azar individuos generados, hubo en promedio  $\frac{128}{2} = 64$  instancias totales de esquemas de nivel inferior en la población inicial. Es decir, hay una probabilidad de 0,5 de que haya una instancia de cualquier bloque en particular; por lo tanto, dado que hay 8 bloques de nivel más bajo diferentes, habrá un promedio de 4 instancias totales de bloques de nivel más bajo en la población.

### 3.1 EXPERIMENTOS SOBRE R 1 Y R 2

Esperábamos que el GA se desempeñara mejor, es decir, encontrara el óptimo más rápidamente, en R 2 que en R 1 en R 2 hay un camino muy claro a través del cruce de pares de los ocho esquemas iniciales de orden 8 ( $s_1 - s_8$ ) a los cuatro esquemas de orden 16 ( $s_9 - s_{12}$ ), y de ahí a los dos esquemas de orden 32 ( $s_{13}$  y  $s_{14}$ ), y finalmente al óptimo ( $s_{optar}$ ). Creíamos que la presencia de este camino más fuerte aceleraría el descubrimiento del óptimo por parte de la GA, pero nuestros experimentos mostraron lo contrario: la GA se desempeñó significativamente mejor en R 1 que en R 2. Las estadísticas que resumen los resultados de 500 corridas en cada función se dan en la Tabla 1. Esta tabla da la media y la mediana del número de evaluaciones de función tomadas para encontrar el óptimo en 500 corridas cada una. R 1 y R 2.

Si esperamos comprender el desempeño de la AG en general, necesitamos comprender en detalle cuáles son los posibles cuellos de botella para descubrir esquemas deseables. Esto se ha estudiado ampliamente en la literatura sobre el engaño, pero R 2 es una función no engañosa que, no obstante, contiene algunas características que impiden que el AG descubra esquemas deseables tan rápido como en R 1.

¿Qué ralentiza la GA en el caso de R 2? Para investigar esto, tomamos una ejecución típica de GA en R 2 y trazó gráficamente la evolución de cada esquema en el árbol. La figura 3 muestra este rastro para tres conjuntos de esquemas:  $s_1, s_2$ , y  $s_9$ ;  $s_3, s_4$ , y  $s_{10}$ ; y  $s_5, s_6$ , y  $s_{11}$  (ver figura 2). En cada gráfico, la densidad (% de la población) de cada esquema se representa en función del tiempo (generaciones). La densidad se muestrea cada

10 generaciones.

Estos gráficos muestran un fenómeno sorprendente. En el gráfico superior de la Figura 3,  $s_1$  y  $s_2$  aparecen temprano y los ejemplos de ellos se combinan rápidamente para formar  $s_9$ . Una vez que se descubre cada esquema, su densidad en la población aumenta con bastante rapidez a más del 90% de la población. población por generación 60 más o menos. Alrededor de la generación 220 hay una clara caída en las densidades de estos tres esquemas.

La gráfica del medio muestra una evolución muy diferente para  $s_3$ ,  $s_4$  y  $s_{10}$ . Los esquemas  $s_3$  y  $s_4$  ambos están presentes en la población inicial (generada aleatoriamente) (aunque  $s_3$  presencia en la generación 0 no es visible en la trama), pero mientras  $s_4$  sube rápido,  $s_3$  muere en la generación 10, se redescubre fugazmente (junto con  $s_{10}$ ) en la generación 120 (ver señal en el X- eje), y no regresa hasta el final de la carrera, en qué punto lo trae una mutación (junto con  $s_{10}$ ) atrás (ver señal en el X- eje). Esta misma mutación es responsable de crear  $s_{\text{optar}}$  en la generación 535, cuando la carrera termina. El esquema  $s_4$ , después de una rápida subida inicial, entra en una caída pronunciada al mismo tiempo que la caída más leve se puede ver en la gráfica superior de la Figura 3, alrededor de la generación 220.

¿Cuál es la causa de estas caídas y qué previene  $s_3$  de persistir en la población lación? Se puede inferir una respuesta probable del gráfico inferior. Esquema  $s_6$  aparece alrededor de la generación 30, aumenta con bastante rapidez, dando un fuerte repunte alrededor de la generación 220 y llegando a aproximadamente el 95% de la población. Esquema  $s_5$  aparece brevemente alrededor de la generación 20 (parpadeo en el X- eje) y se extingue, pero aparece de nuevo en la generación 220. Su instancia en la población es también una instancia de  $s_{11}$ , e instancias de  $s_{11}$  Levántate muy rápido. Este aumento coincide exactamente con la caída menor en  $s_1$ ,  $s_2$ , y  $s_9$ , y la mayor caída en  $s_4$ . Lo que parece estar sucediendo es lo siguiente: en los primeros casos de  $s_{11}$ , junto con los 16 1 en el quinto y sexto bloques, hay varios ceros en el primer al cuarto bloque. Una instancia de  $s_{11}$  tiene aptitud  $8 + 8 + 16 = 32$ , mientras que una instancia de un esquema de orden 8 como  $s_4$  está en forma 8.

Esta diferencia causa  $s_{11}$  subir muy rápido en comparación con  $s_4$ , e instancias de  $s_{11}$  con algunos ceros en el cuarto bloque tienden a expulsar muchos de los instancias de  $s_4$  en la población. Este fenómeno se ha denominado "autostop", donde los ceros en otras posiciones en la cuerda hacen autostop junto con el muy en forma  $s_{11}$ . Las posiciones más probables para los autostopistas son aquellas cercanas al esquema altamente apto. posiciones definidas, ya que es menos probable que se separen de las posiciones definidas del esquema en el cruce. Estos efectos se observan en la genética de poblaciones reales y han sido discutidos en el contexto de los AG por Schraudolph y Belew (1990) y Das y Whitley (1989), entre otros. Tenga en cuenta que este efecto es pronunciado incluso con la forma relativamente débil de selección utilizada en nuestra AG. (También comparamos el desempeño de GA en R 1 y R 2 utilizando un método de escalamiento de rango lineal (Baker,

1985) en lugar del método de escala sigma descrito anteriormente, y obtuvieron resultados similares a los que se muestran en la Tabla 1.)

Las gráficas que se muestran en la Figura 3 provienen de una sola corrida, pero esta corrida fue típica; el mismo tipo de fenómeno se observó en muchas de las otras carreras en R 2 también. Nuestra hipótesis es que este efecto de autostop es lo que hace que los tiempos relativamente más lentos (en promedio) para que el GA encuentre el óptimo en R 2. El poder del crossover



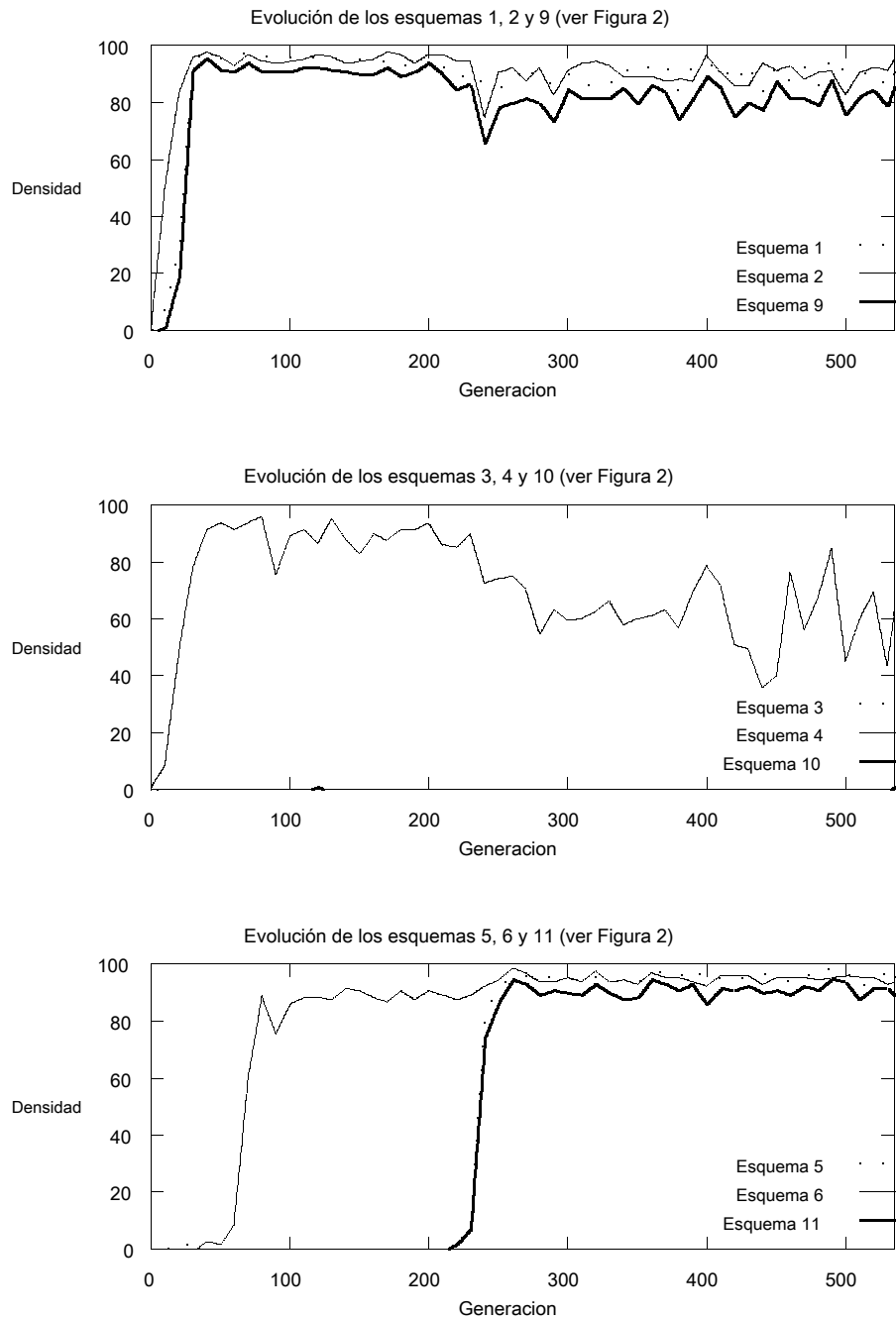


Figura 3: Evolución de tres conjuntos de esquemas en una ejecución típica de GA en R 2. En cada parcela, la densidad de cada esquema (% de la población) se traza frente a la generación. Tenga en cuenta que en la gráfica del medio, los esquemas 3 y 10 son visibles solo como pequeñas protuberancias en el eje x en las generaciones 120 y 535.

TAMAÑO DE LA POBLACIÓN 1024		
Evaluaciones de funciones al óptimo		
200 carreras	R 1	R 2
Significar	37453 (error estándar: 868)	43213 (error estándar: 1275) 36864
Mediana	34816	

Tabla 2: Resumen de los resultados de 200 corridas del GA con un tamaño de población de 1024 en R 1 y R 2.

La combinación de bloques de construcción de nivel inferior se vio obstaculizada, ya que algunos de los bloques de construcción necesarios fueron suprimidos parcial o totalmente por el rápido aumento de bloques de construcción disjuntos. Esto sugiere que la caracterización de un paisaje GA es más que la dirección absoluta de los gradientes de búsqueda. En estas funciones, son las diferencias reales en la adecuación relativa de los diferentes esquemas las que son relevantes.

En R 1, que carece de la aptitud extra que se da a algunos esquemas de nivel intermedio, el problema del autostop no se presenta en un grado tan devastador. La aptitud de un instancia de, digamos,  $s_{11}$  en R 1 tiene solo 16, por lo que su descubrimiento no tiene un efecto tan dramático sobre el descubrimiento y la persistencia de otros esquemas de orden 8 en la función. Contrariamente a nuestras intuiciones iniciales, parece que el refuerzo adicional de algunos escalones de nivel intermedio en realidad daña el AG en estas funciones.

Se podría pensar que estos resultados se deben en parte a un error de muestreo: dado que los bloques de construcción de orden más bajo tienen una longitud de 8, un GA con una población de 128 no tiene muestras de muchos de los bloques de construcción de orden más bajo en la población inicial. (en promedio, la mitad de los bloques de orden más bajo no estarán representados) y, por lo tanto, debe esperar hasta que los bloques de construcción de orden 8 se creen mediante variación aleatoria. <sup>†</sup> Para probar el efecto de esto en nuestros resultados, realizamos dos experimentos adicionales: (1) ejecutamos el GA en R 1 y R 2 con un tamaño de población de 1024, que da 4 instancias esperadas de cada esquema de orden 8 en la población inicial; y (2) ejecutamos el AG con un tamaño de población 128 en versiones modificadas de R 1 y R 2 en el que los bloques de construcción de orden más bajo tenían una longitud de 4 en lugar de una longitud de 8. En este último caso, hay un promedio de 8 instancias de cada bloque de construcción de orden 4 en la población inicial. Los resultados de estos dos experimentos se dan en las Tablas 2 y 3. Muestran que estas modificaciones, aunque mejoran el rendimiento absoluto del GA (especialmente en el caso de los bloques de construcción de orden 4, lo que hace que la función sea mucho más fácil de optimizar), no cambian la diferencia cualitativa entre el tiempo al óptimo para R 1 y R 2. Esto indica que la diferencia no se debe principalmente a un error de muestreo.

Estos resultados apuntan a un problema generalizado e importante en el desempeño de los AG: el problema de la convergencia prematura. El hecho de que observemos una forma de convergencia prematura incluso en este escenario tan simple sugiere que puede ser un factor en

<sup>†</sup> Nota: este hecho hizo que nuestros experimentos (tiempo para encontrar el óptimo) tuvieran una varianza superior a la normal, razón por la cual realizamos al menos 200 corridas para la mayoría de nuestros experimentos.

ESQUEMAS DE PEDIDO MÁS BAJO LONGITUD 4		
	Evaluaciones de funciones al óptimo	
200 carreras	R 1	R 2
Significar	6568 (error estándar: 198) 5760	11202 (error estándar: 394) 9600
Mediana		

Tabla 3: Resumen de los resultados de 200 corridas de GA en versiones modificadas de R 1 y R 2, en el que los bloques de construcción de menor orden tienen una longitud 4.

VARIANTES DE R2		
	Evaluaciones de funciones al óptimo	
200 carreras	R 2 intrones	R 2 departamento
Significar	75599 (error estándar: 2697) 70400	62692 (error estándar: 2391) 56448
Mediana		

Tabla 4: Resumen de los resultados de 200 corridas de GA en dos variantes de R 2.

cualquier búsqueda de GA en la que la población esté buscando simultáneamente dos o más esquemas de alta aptitud física no superpuestos (p. ej.,  $s_4$  y  $s_{11}$ ), que suele ser el caso. El hecho de que la población pierda esquemas útiles una vez que uno de los buenos esquemas disjuntos Lo que se encuentra sugiere una razón por la que la tasa de paralelismo implícito efectivo de la AG (Holland, 1975; Goldberg, 1989b) puede necesitar ser reconsiderada. (Para otra discusión sobre el paralelismo implícito en los AG, ver Grefenstette & Baker 1989.)

### 3.2 ¿LOS INTRONES SUPRIMEN A LOS HITCHHIKERS?

Para comprender el comportamiento del autostop con mayor precisión, realizamos un experimento que creímos que lo eliminaría hasta cierto punto. Nuestra hipótesis era que el autostop se producía en los lugares que eran espacialmente adyacentes a los esquemas de alta aptitud (p. ej.,  $s_{11}$  sobre). Para reducir este efecto, construimos una nueva función, R 2 intrones, introduciendo bloques de 8 "intrones" —8 \* adicionales: entre cada uno de los bloques de 8 bits de unos. Así en R 2 intrones, Las cadenas tenían una longitud de 128 en lugar de 64. Por ejemplo, en R 2 intrones,  $s_1$  es 11111111 \*\*\*\*\*. . . \*,  $s_2$  es \*\*\*\*\* 11111111 \*\*. . . \*, y su combinación  $s_9$  es 11111111 \*\*\*\*\* 11111111 \*\*. . . \*. El óptimo es la cadena que contiene cada bloque de 8 1, donde los bloques están separados cada uno por ocho loci que pueden contienen 0 o 1. La idea aquí era que un autoestopista potencialmente dañino estaría al menos a 8 bits de distancia del esquema en el que estaba haciendo autostop y, por lo tanto, es probable que se pierda en el cruce. (Levenick, 1991, descubrió que la inserción de intrones en los individuos mejoraba el rendimiento de la AG en un conjunto particular de entornos).

Como se muestra en la columna 1 de la Tabla 4, ejecutar el GA en R 2 intrones arrojó resultados no significativamente diferentes de los de R 2. Esto fue contrario a nuestras expectativas, y las razones de este resultado no están claras, pero una hipótesis es que una vez

instancia de un esquema de orden superior (por ejemplo,  $s_{11}$ ) se descubre, la convergencia es tan rápida que los autostopistas son posibles incluso en lugares relativamente distantes del esquema posiciones definidas.

### 3.3 VARIANDO LOS COEFICIENTES EN $R_2$

Está claro que es necesario algún refuerzo de nivel intermedio para que funcione la AG. Considerar  $R_1$ ; una variante de  $R_1$ , donde  $R_1(x) = 8$  si  $X$  es una instancia de al menos uno de los esquemas de 8 bits, y  $R_1(x) = 64$  si  $X$  es una instancia de todos los esquemas de 8 bits. Aquí, el GA no tendría ninguna razón para preferir una cuerda con un bloque de 16 unos a una cuerda con un bloque de 8 unos y, por lo tanto, no habría presión para aumentar el número de unos. Esquemas intermedios en  $R_1$  proporcionarían lineal refuerzo, ya que la aptitud de una instancia de un esquema de orden intermedio es siempre la suma de la aptitud de las instancias de los esquemas de orden 8 componentes. Algunos esquemas en  $R_2$

proporcionarían fuerte no lineal refuerzo, ya que la idoneidad de una instancia de, digamos,  $s_9$  es mucho mayor que la suma de la aptitud de las instancias del componente orden-8 esquemas  $s_1$  y  $s_2$ . Nuestros resultados indican que el refuerzo no lineal proporcionado por algunos esquemas es demasiado alto: perjudica en lugar de ayudar al desempeño del AG.

¿El refuerzo no lineal ayuda alguna vez a la AG en lugar de obstaculizarla? Para estudiar esto construimos una nueva función,  $R_2$  Departamento, con un esquema de refuerzo no lineal mucho más débil: para esta función,  $C_1 - C_{14}$  se establecen cada uno en el valor plano 1. Aquí el refuerzo lineal es todavía no lineal (una instancia de  $s_9$  tendrá aptitud  $1 + 1 + 1$ , que es mayor que la suma de los dos componentes), pero la cantidad de refuerzo se reduce importantemente.

Los resultados de ejecutar el GA en  $R_2$  Departamento se da en la segunda columna de la Tabla 4. El tiempo promedio hasta el óptimo para esta función es aproximadamente el mismo que para  $R_1$ .

Por lo tanto, la ventaja de menor aptitud en  $R_2$  Departamento no parece perjudicar el rendimiento, aunque no resulta en mejorado rendimiento sobre eso en  $R_1$ .

Estos fenómenos pueden estar relacionados con los resultados de Feldman y sus colegas sobre los efectos de las funciones de aptitud súper y sub-multiplicativas sobre la viabilidad evolutiva del cruce (Lieberman y Feldman, 1986; Bergman y Feldman, 1990). Sin embargo, existen varios problemas con la aplicación directa de los teoremas de Feldman. Un problema es que Feldman estudia la viabilidad evolutiva del cruce en lugar del grado en que el cruce ayuda a descubrir individuos de alta aptitud. Nuestro trabajo se concentra en este último. Actualmente estamos investigando cómo se relacionan estas dos preocupaciones. (Esto también fue estudiado por Schaör y Eshelman, 1991.)

## 4. DISCUSIÓN

Los resultados descritos en las dos secciones anteriores muestran que la capacidad de la AG para procesar bloques de construcción depende efectivamente no solo de su presencia, sino también de su aptitud relativa. Si algunos escalones intermedios son mucho más ajustados que los componentes primitivos, entonces la convergencia prematura ralentiza el descubrimiento de

algunos esquemas necesarios. Los intrones simples y un esquema de selección muy suave no parecen aliviar los problemas de convergencia prematura y autostop.

Nuestros resultados señalan la importancia de hacer de la hipótesis de los bloques de construcción una descripción más precisa y útil del procesamiento de los bloques de construcción. Si bien los efectos disruptivos que observamos (autostop, convergencia prematura, etc.) ya se conocen en la literatura de GA, todavía no existe un teorema que los asocie con la estructura de bloques de construcción de un problema dado.

En nuestros experimentos hemos observado que el papel del cruce varía considerablemente a lo largo del curso de la búsqueda de GA. En particular, se pueden identificar tres etapas de la búsqueda: (1) el tiempo que tarda el GA en descubrir los esquemas de orden más bajo, (2) el tiempo que tarda el cruce para combinar esquemas de orden inferior en uno de orden superior. esquema, y (3) el tiempo que tarda el esquema de orden superior en apoderarse de la población. En funciones de varios niveles, como las funciones de Royal Road, estas fases de la búsqueda se superponen considerablemente, y es esencial comprender el papel del cruce y los detalles del procesamiento del esquema en cada etapa (este tema también ha sido investigado por Davis, 1989, y por Schaffer & Eshelman, 1991, entre otros). En trabajos anteriores,

## 5 EXPERIMENTOS CON ESCALADA

Como se mencionó anteriormente, parte de nuestro propósito al diseñar las funciones de Royal Road es construir la clase más simple de paisajes de aptitud física en los que el GA no solo funcionará bien, sino que también superará a otros métodos de búsqueda como la escalada. Además de nuestros experimentos que comparan el desempeño de GA en R1 y R2, comparamos el rendimiento del GA con el de tres esquemas de escalada iterados comúnmente utilizados: escalada de la colina con el ascenso más empinado, escalada de la colina del siguiente ascenso (Mühlenbein, 1991), y un esquema que llamamos " ", Que fue sugerido por Richard Palmer (comunicación personal). Nuestra implementación de estos diversos esquemas de escalada es la siguiente:

- Escalada de pendiente más empinada (SAHC):
  1. Elija una cadena al azar. Llamar a esta cadena actual-colina.
  2. Mute sistemáticamente cada bit de la cadena de izquierda a derecha, registrando la adecuación de las cadenas resultantes.
  3. Si alguna de las cadenas resultantes aumenta la condición física, establezca cima de la colina actual a la cuerda resultante dando el mayor aumento de aptitud.
  4. Si no aumenta el estado físico, guarde cima de la colina actual y vaya al paso 1.  
De lo contrario, vaya al paso 2 con el nuevo actual-colina.
  5. Cuando se haya realizado un número determinado de evaluaciones de funciones, devuelva la cima más alta que se encontró.
- Escalada de pendiente siguiente (NAHC):

ESCALADA EN R2			
Evaluaciones de funciones al óptimo			
200 carreras	SAHC	NAHC	RMHC
Significar	> 256.000 (error estándar: 0)	> 256.000 (error estándar: 0)	6551 (error estándar: 212)
Mediana	> 256.000	> 256.000	5925

Tabla 5: Resumen de los resultados de 200 ejecuciones de varios algoritmos de escalada en R 2.

1. Elija una cadena al azar. Llamar a esta cadena actual-colina.
2. Mute los bits individuales de la cadena de izquierda a derecha, registrando la adecuación de las cadenas resultantes. Si se encuentra algún aumento en la condición física, establezca actual colina a esa cadena de mayor aptitud, sin evaluar más mutaciones de un solo bit de la cadena original. Vaya al paso 2 con el nuevo actual colina pero continúe mutando la nueva cadena comenzando después de la posición del bit en la que se encontró el aumento de aptitud anterior.
3. Si no se encontraron aumentos en la condición física, guarde cima de la colina actual y vaya al paso 1.
4. Cuando se haya realizado un número determinado de evaluaciones de funciones, devuelva la cima más alta que se encontró.

Observe que este método es similar al esquema de "escalada de bits" de Davis (Davis, 1991). En su esquema, los bits se mutan en un orden aleatorio, y actual colina se restablece a cualquier cadena que tenga igual o mejor estado físico que la mejor evaluación anterior.

- Escalada de colinas por mutación aleatoria (RMHC):
  1. Elija una cadena al azar. Llamar a esta cadena mejor evaluado.
  2. Elija un locus al azar para mutar. Si la mutación conduce a una aptitud física igual o superior, establezca mejor evaluado a la cadena resultante.
  3. Vaya al paso 2.
  4. Cuando se ha realizado un número determinado de evaluaciones de funciones, devuelve el valor actual de mejor evaluado.

La Tabla 5 da los resultados de ejecutar estos tres esquemas de escalada en R 2. En cada ejecución, se permitió que el algoritmo de escalada continuara hasta que se descubrió la cadena óptima o hasta que se hubieran realizado 256.000 evaluaciones de funciones y se registró el número total de evaluaciones de funciones realizadas. Como se puede ver, la escalada más empinada y la escalada siguiente nunca encontraron el óptimo durante el tiempo asignado, pero la escalada por mutación aleatoria encontró el óptimo en promedio más de diez veces más rápido que el GA con un tamaño de población de 128, y más de seis veces más rápido que el GA con un tamaño de población de 1024. Tenga en cuenta que la escalada de colinas con mutación aleatoria, como la hemos descrito, difiere del método de escalada de bits utilizado por Davis (1991) en que no muta sistemáticamente los bits, y nunca se rinde y comienza con una nueva cadena aleatoria, sino que continúa deambulando por

mesetas indefinidamente. Larry Eshelman (comunicación personal) ha señalado que el escalador de montañas de mutación aleatoria es ideal para las funciones de Royal Road, de hecho mucho mejor que el escalador de bits de Davis, pero tendrá problemas con cualquier función con mínimos locales. (Eshelman descubrió que el escalador de brocas de Davis funciona muy mal en

R 1, sin encontrar nunca el óptimo en 50 ejecuciones de 50.000 evaluaciones de funciones cada una).

Además de basar nuestra comparación en el número de evaluaciones de funciones al óptimo, también comparamos el desempeño promedio en línea (De Jong, 1975) de la GA (tamaños de población 128 y 1024) con escalada de colinas por mutación aleatoria, ambas que se ejecuta en R 2. La Figura 4 traza los resultados de esa comparación. Para una ejecución dada de la GA o de RMHC, el desempeño en línea en un número dado de evaluaciones de función se define como el valor promedio de todas las evaluaciones de funciones realizadas hasta ese momento. Registramos los valores de rendimiento en línea a intervalos de 128 evaluaciones de funciones y repetimos este procedimiento durante 100 ejecuciones. Luego promediamos los valores de desempeño en línea, en cada intervalo de 128 evaluaciones de funciones, en todas las ejecuciones. Por tanto, cada punto de un gráfico de la Figura 4 representa un promedio de los valores de rendimiento en línea durante una serie de ejecuciones. (El número de valores promediados en cada punto varía: dado que GA y RMHC se detienen cuando se encuentra el óptimo, diferentes ejecuciones realizaron diferentes números de evaluaciones de funciones. Solo damos promedios para los cuales había un número significativo de valores para promediar).

Se puede ver en los gráficos que RMHC también supera significativamente a ambas versiones de la AG bajo esta medida.<sup>2</sup>

Estos resultados son una demostración sorprendente de que, al comparar la GA con la escalada en un problema particular o un conjunto de pruebas, es importante cuales se utiliza un tipo de algoritmo de escalada. Davis (1991) también ha señalado este punto.

Las funciones de Royal Road se diseñaron originalmente para servir a dos propósitos bastante diferentes: (1) como un escenario idealizado en el que estudiar el procesamiento de bloques de construcción y el papel del crossover, y (2) como un ejemplo de una función simple que distingue a los GA de Montañismo. Si bien hemos descubierto que ciertas formas de escalar colinas superan al AG en estas funciones (por lo tanto, son inapropiadas exactamente en esta forma para el segundo propósito), cumplen con el primer propósito.

## 6 CONCLUSIONES Y DIRECCIONES FUTURAS

La investigación descrita en este documento es un paso inicial para comprender con mayor precisión cómo se procesan los esquemas en el cruce. Al estudiar el comportamiento del AG en paisajes simples en los que los bloques de construcción deseables se definen explícitamente, hemos descubierto algunos fenómenos imprevistos relacionados con la capacidad del AG para procesar esquemas de manera eficiente, incluso en funciones no engañosas. Las funciones de Royal Road capturan, de una manera idealizada y clara, algunas características del paisaje que son particularmente relevantes para la AG, y creemos que una comprensión profunda de los principios de la AG

---

<sup>2</sup> Es interesante notar que bajo esta medida, la población de 128 GA supera a la de la población de 1024 GA en las etapas iniciales de la carrera.

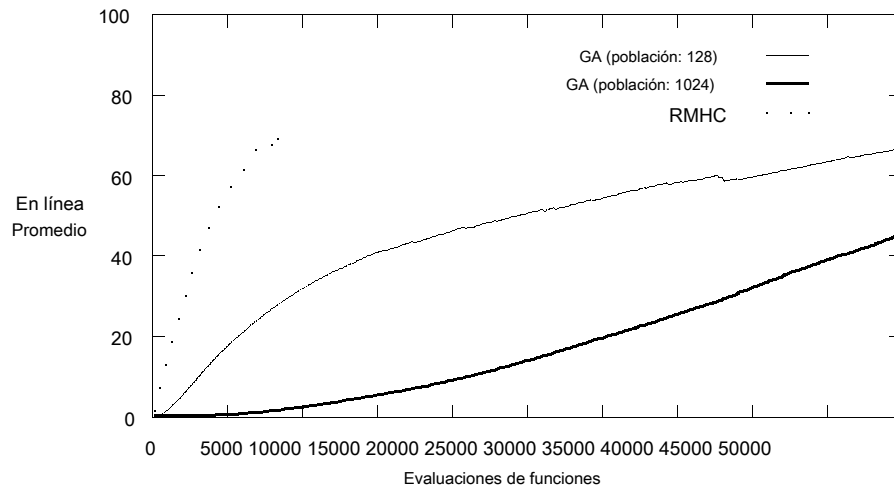


Figura 4: Gráficos del rendimiento promedio en línea de la GA (tamaños de población 128 y 1024) y de escalada de colinas con mutación aleatoria (RMHC), más de 100 carreras. El gráfico de RMHC se detiene en alrededor de 6000 evaluaciones de funciones porque RMHC casi siempre encontró la función óptima en ese momento.



El comportamiento en estos paisajes simples será muy útil para desarrollar teoremas más detallados y útiles sobre el comportamiento de AG.

La investigación informada aquí es un trabajo en progreso y hay varias direcciones para futuras investigaciones. Aquí esbozamos algunos de nuestros planes de investigación a corto y largo plazo.

A corto plazo, planeamos estudiar más detenidamente los cuellos de botella en el descubrimiento de esquemas deseables y cuantificar con mayor precisión la relación entre los valores de aptitud de los diversos componentes básicos y el grado en que se producirán estos cuellos de botella. El autostop es evidentemente un cuello de botella, y debemos comprender mejor de qué manera está ocurriendo y bajo qué circunstancias. Una vez que hemos descrito los fenómenos con más detalle, podemos comenzar a desarrollar un modelo matemático de las competencias de esquemas que observamos (ilustradas en la Figura 3) y cómo se ven afectadas por diferentes esquemas de aptitud de los componentes básicos. Este modelo puede estar relacionado con los modelos propuestos por Vose y Liepins (1991).

Nuestros resultados de autostop necesitan ser analizados y explicados más a fondo, y planeamos un análisis más detallado de los diferentes efectos de varios esquemas de refuerzo no lineales. En particular, se necesitan más detalles en la comparación del desempeño de GA  
mance en R 1 con rendimiento activado R 2 Departamento, y en otros esquemas de coeficientes.

Creemos que existen versiones de paisajes de "caminos reales" que cumplirán nuestro objetivo de encontrar funciones simples que distingan a los AG de la escalada. Por ejemplo, planeamos probar las siguientes variantes: agregar ruido, incluidas todas las combinaciones de esquemas de orden inferior en la lista explícita de esquemas, y permitir que los esquemas se superpongan.

Las funciones de Royal Road exploran solo un tipo de característica del paisaje que es relevante para los AG: la presencia y la adecuación relativa de los bloques de construcción de orden intermedio. Nuestros planes a más largo plazo incluyen la ampliación de los paisajes de clases de fitness que se están investigando para incluir otros tipos de características relevantes; Mitchell, Forrest y Holland (1992) describieron algunas de estas características. También estamos interesados en desarrollar medidas estadísticas que puedan determinar la presencia o ausencia de las características de interés. Estos podrían estar relacionados con el trabajo para determinar la estructura de correlación de los paisajes del fitness (ver Kaufman, 1989; Lipsitch, 1991; y Manderick, de Weger y Spiessens, 1991). Si tales medidas pudieran desarrollarse, podrían usarse para ayudar a predecir la probabilidad de un desempeño exitoso de GA en un paisaje dado.

#### Expresiones de gratitud

La investigación informada aquí fue apoyada por el Centro de Estudios No Lineales en el Laboratorio Nacional de Los Alamos, las Universidades Occidentales Asociadas y la subvención de la NSF IRI9157644 (apoyo a S. Forrest); la Sociedad de Becarios de Michigan y el Departamento de EECS de la Universidad de Michigan (apoyo a M. Mitchell); y el Instituto Santa Fe (apoyo a ambos autores). John Holland ha colaborado con nosotros en gran parte de este trabajo y es una fuente constante de asistencia y consejos invaluable. Agradecemos a Richard Palmer por sugerir el algoritmo de "escalada de colinas de mutación aleatoria", y

Larry Eshelman por sus comentarios e informes de experimentos sobre varios esquemas de escalada. También agradecemos a Robert Axelrod, Aviv Bergman, Arthur Burks, Michael Cohen, Marcus Feldman, Rick Riolo y Carl Simon por muchas discusiones útiles sobre estos temas.

#### Referencias

- JE Baker (1985). Métodos de selección adaptativos para algoritmos genéticos. En JJ Grefenstette (Ed.), *Actas de la Primera Conferencia Internacional sobre Algoritmos Genéticos y sus Aplicaciones*. Hillsdale, Nueva Jersey: Lawrence Erlbaum Associates.
- A. Bergman y MW Feldman (1990). Más sobre selección a favor y en contra de la recombinación. *Biología teórica de poblaciones*, 38 (1), 68–92.
- R. Das y D. Whitley (1991). Los únicos problemas desafiantes son engañosos: búsqueda global resolviendo hiperplanos de orden 1. En RK Belew y L. B Booker (Eds.), *Actas de la Cuarta Conferencia Internacional sobre Algoritmos Genéticos*. San Mateo, CA: Morgan Kaufmann.
- LD Davis (1989). Adaptación de probabilidades de operador en algoritmos genéticos. En JD Schaffer (Ed.), *Actas de la Tercera Conferencia Internacional sobre Algoritmos Genéticos*. San Mateo, CA: Morgan Kaufmann.
- LD Davis (1991). Escalada de bits, sesgo de representación y diseño de conjuntos de pruebas. En RK Belew y L. B Booker (Eds.), *Actas de la Cuarta Conferencia Internacional sobre Algoritmos Genéticos*. San Mateo, CA: Morgan Kaufmann.
- KA De Jong (1975). Un análisis del comportamiento de una clase de sistemas adaptativos genéticos. Tesis doctoral inédita, Universidad de Michigan, Ann Arbor, MI.
- S. Forrest y M. Mitchell (1991). La ejecución de algoritmos genéticos en polinomios de Walsh: algunos resultados anómalos y su explicación. En RK Belew y LB Booker (Eds.), *Actas de la Cuarta Conferencia Internacional sobre Algoritmos Genéticos*. San Mateo, CA: Morgan Kaufmann.
- DE Goldberg (1987). Algoritmos genéticos sencillos y el mínimo problema engañoso. En LD Davis (Ed.), *Algoritmos genéticos y recocido simulado (Notas de investigación en inteligencia artificial)*. Los Altos, CA: Morgan Kaufmann.
- DE Goldberg (1989a). Algoritmos genéticos y funciones de Walsh: Parte II, Engaño y su análisis. *Sistemas complejos*, 3: 153-171.
- DE Goldberg (1989b). *Algoritmos genéticos en búsqueda, optimización y aprendizaje automático*. Reading, MA: Addison Wesley.
- JJ Grefenstette y JE Baker (1989). Cómo funcionan los algoritmos genéticos: una mirada crítica al paralelismo implícito. En JD Schaffer (Ed.), *Actas de la Tercera Conferencia Internacional sobre Algoritmos Genéticos*. San Mateo, CA: Morgan Kaufmann.
- JH Holland (1975). *Adaptación en sistemas naturales y artificiales*. Ann Arbor,

MI: Prensa de la Universidad de Michigan.

SA Kauffman (1989). Adaptación a paisajes de fitness accidentados. En D. Stein (Ed.), Conferencias en Ciencias de la Complejidad, 527–618. Reading, MA: Addison-Wesley.

JR Levenick (1991). La inserción de intrones mejora la tasa de éxito del algoritmo genético: siguiendo el ejemplo de la biología. En RK Belew y LB Booker (Eds.), Actas de la Cuarta Conferencia Internacional sobre Algoritmos Genéticos. 123-127. San Mateo, CA: Morgan Kaufmann.

U. Liberman y MW Feldman (1986). Un principio de reducción general para modificadores genéticos de recombinación. Biología teórica de poblaciones, 30 (3), 341–371.

GE Liepins y MD Vose (1990). Problemas de representación en la optimización genética. Revista de Inteligencia Artificial Experimental y Teórica, 2, 101-115.

M. Lipsitch (1991). Adaptación en paisajes accidentados generada por interacciones locales de genes vecinos. En RK Belew y LB Booker (Eds.), Actas de la Cuarta Conferencia Internacional sobre Algoritmos Genéticos. San Mateo, CA: Morgan Kaufmann.

B. Manderick, M. de Weger y P. Spiessens (1991). El algoritmo genético y la estructura del panorama del fitness. En RK Belew y LB Booker (Eds.), Actas de la Cuarta Conferencia Internacional sobre Algoritmos Genéticos. San Mateo, CA: Morgan Kaufmann.

M. Mitchell, S. Forrest y JH Holland (1992). El camino real de los algoritmos genéticos: paisajes de fitness y rendimiento de GA. En Actas de la Primera Conferencia Europea sobre Vida Artificial. Cambridge, MA: MIT Press / Bradford Books.

H. Mühlenbein (1991). Evolución en el tiempo y el espacio: el algoritmo genético paralelo. En GJE Rawlins (Ed.), Fundamentos de los algoritmos genéticos, 316–337. San Mateo, CA: Morgan Kaufmann.

JD Schaör y LJ Eshelman (1991). Sobre el crossover como estrategia evolutivamente viable. En RK Belew y LB Booker (Eds.), Actas de la Cuarta Conferencia Internacional sobre Algoritmos Genéticos, 61–68. San Mateo, CA: Morgan Kaufmann.

NN Schraudolph y RK Belew (1990). Codificación de parámetros dinámicos para algoritmos genéticos. Informe técnico CSE CS 90-175. Departamento de Ingeniería y Ciencias de la Computación, Universidad de California, San Diego.

R. Tanese (1989). Algoritmos genéticos distribuidos para la optimización de funciones. Tesis doctoral inédita, Universidad de Michigan, Ann Arbor, MI.

M. Vose y G. Liepins. (1991). Equilibrios puntuados en la búsqueda genética. Sistemas complejos 5, 31–44.

LD Whitley (1991). Principios fundamentales del engaño en la búsqueda genética. En G. Rawlins (Ed.), Fundamentos de los algoritmos genéticos. San Mateo, CA: Morgan

Kaufmann.