



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE CIENCIAS

**EMERGENCIA DE PATRONES Y SELECCIÓN
SEXUAL:
UN EXPERIMENTO COMPUTACIONAL**

T E S I S

**QUE PARA OBTENER EL TÍTULO DE:
BIÓLOGO**

P R E S E N T A:

DIEGO TRUJILLO PISANTY



**DIRECTOR DE TESIS:
DR. JOSÉ ANTONIO NEME CASTILLO
2010**



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Agradecimientos

Quiero agradecer por su apoyo en la elaboración de este trabajo a Antonio Neme Castillo por su dirección y por despertar en mí un profundo interés por la computación y la vida artificial. También agradezco a los doctores Arturo Becerra, Pedro Miramontes, Ernesto Vega, Jaime Zuñiga y Carlos Martorell por sus comentarios y el tiempo dedicado a mi trabajo. Por último agradezco a los biólogos Iván Trujillo e Irene Pisanty por su apoyo sobre todo en los campos no académicos.

1. Datos del alumno

Trujillo
Pisanty
Diego
56556986
Universidad Nacional Autónoma de México
Facultad de Ciencias
Biología
303520803

2. Datos del tutor

Dr.
José Antonio
Neme
Castillo

3. Datos del sinodal 1

Dr.
Arturo Carlos II
Becerra
Bracho

4. Datos del sinodal 2

Dr.
Pedro Eduardo
Miramontes
Vidal

5. Datos del sinodal 3

Dr.
Jose Jaime
Zuñiga
Vega

6. Datos del sinodal 4

Dr.
Ernesto Vicente
Vega
Peña

7. Datos del trabajo escrito.

Emergencia de patrones y selección sexual
Un experimento computacional
84 p.
2010

Agradecimientos	i
Tabla de contenido	ii
Resumen	iv
Introducción	1
I. Marco conceptual	3
1.1 Individuos y poblaciones	4
1.2 Genes y alelos	5
1.3 Espacio de soluciones	6
1.4 Selección natural y selección sexual	7
1.5 Reproducción con recombinación	8
1.6 Adecuación	9
1.7 Riqueza y diversidad	11
1.8 Automatas celulares	13
1.9 Espacio de color RGB de 24-bits	14
1.10 Algoritmos genéticos	15
II. Algoritmo espacial de selección sexual	17
2.1 Reglas del autómata	17
2.2 Propiedades de los individuos	18
2.3 Fecundidad y adecuación	19
2.4 Distribución	20
2.5 Selección y reproducción	21
2.6 Suposiciones del algoritmo	22
2.7 El algoritmo en pseudocódigo	23
2.8 Simulaciones realizadas	23
III. Resultados de las simulaciones	30
3.1 Espacio de soluciones regular, completamente explorado y selección por mínima distancia euclidiana	30
3.2 Espacio de soluciones regular, completamente explorado y selección por máxima distancia euclidiana	32
3.3 Cobertura de soluciones inicial reducida a dos puntos distantes y selección por mínima distancia euclidiana	34
3.4 Cobertura de soluciones inicial reducida a dos puntos distantes y selección por máxima distancia euclidiana	36
3.5 Cobertura de soluciones inicial irregular, incompletamente explorada y selección por mínima distancia euclidiana	38
3.6 Cobertura de soluciones inicial irregular, incompletamente explorada y selección por máxima distancia euclidiana	41

IV. Discusión	43
4.1 La influencia del criterio de selección	43
4.2 El efecto de la diversidad inicial	46
4.3 Existencia de patrones	50
4.4 Significancia biológica	50
4.5 Entropía, evolución y trabajo	53
V. Conclusiones	57
5.1 El efecto de la selección sexual en la diversidad	57
5.2 Biología <i>in silico</i>	58
5.3 Computo bioinspirado	59
5.4 Problemas no resueltos	60
Apendice A: Mapas de color	63
Referencias	70

Resumen

Este trabajo es una aproximación práctica al estudio de la selección sexual y su efecto en la diversidad a lo largo del tiempo. En él se discuten conceptos como selección de pareja, reproducción con recombinación, heredabilidad genética y adecuación. El análisis de dichos conceptos se utilizó para crear una serie de reglas algorítmicas que se computan sobre un autómata celular de manera similar a un algoritmo genético. Los individuos del autómata pueden tomar como estado cualquier color del espacio RGB de 24-bits, dicho estado pretende ser análogo a un genotipo. En el algoritmo cada casilla del autómata escoge con cual de sus vecinos se reproduce de acuerdo a un criterio de selección basado en distancia euclidiana de los valores RGB de cada individuo. Después de cada generación se contaron el número de estados existentes y se calculó la entropía de Shannon como un indicador de diversidad de cada generación.

Los cambios en diversidad y riqueza observados durante el tiempo son coherentes con estudios biológicos hechos *in vivo*. Los resultados de este trabajo parecen indicar que los efectos de la selección sexual en la diversidad a lo largo del tiempo son dependientes de la diversidad inicial del sistema y del criterio de selección de los organismos. Dichos resultados también contribuyen al entendimiento de la selección sexual como un proceso de manejo de información que provoca decrementos locales de entropía.

Palabras clave: Selección sexual, evolución, autómata celular, algoritmo genético, modelación, selección dependiente de la frecuencia.

Introducción

Este trabajo pretende ser una aproximación práctica hacia el estudio del efecto de la selección de pareja sobre la variabilidad genética.

La selección de pareja es una 'operación' llevada a cabo por los individuos de una población con reproducción sexual. Dicho fenómeno ocurre cuando los individuos llegan a la categoría de edades en la cual son fértiles y presentan una maduración de los caracteres sexuales tanto primarios como secundarios. Una vez que se encuentran listos para la reproducción los organismos comienzan a buscar a otro individuo con órganos sexuales compatibles y caracteres morfológicos, fisiológicos y/o conductuales atractivos.

Cuando se encuentra una pareja atractiva los organismos se reproducirán de manera sexual con recombinación genética y caracteres heredados. Se puede decir que mientras más descendencia deje un individuo mejor es su adecuación y más copias de sus genes existen en las generaciones subsecuentes.

Los cambios que ocurren en las frecuencias alélicas al repetirse dichos fenómenos durante varias generaciones provocan cambios observables en la riqueza y diversidad tanto fenotípicos como genéticos.

El objetivo central de este trabajo es extraer reglas sencillas del proceso por el cual la selección de pareja crea cambios en la diversidad. Una vez que se crearon reglas abstractas para describir las operaciones mencionadas con anterioridad estas se utilizaron para crear un modelo de computo aplicable a simulaciones. Dicho modelo está basado en un autómata celular cuyos estados son colores.

Las reglas se aplican de manera similar a un algoritmo genético.

El primer capítulo de este documento formaliza los conceptos mencionados en esta breve introducción y da referencia a los trabajos previos de donde fueron tomados. Los conceptos presentados fueron los utilizados tanto para construir la metodología (un programa de computadora) como para discutir los resultados. Además este capítulo intenta entablar un diálogo entre conceptos de biología y computación así como clarificar su relación con esta investigación.

El capítulo dos trata sobre el algoritmo desarrollado para responder las preguntas centrales del trabajo. En este capítulo se explican las reglas del algoritmo así como su origen en conceptos de biología. También se habla sobre las diferentes condiciones en las cuales se corrió y los fenómenos que se pretendieron simular en cada ejecución del programa.

En el tercer capítulo se presentan los resultados de las ejecuciones del programa y se hace una descripción así como una breve explicación del origen de los resultados. Los resultados son discutidos a detalle en el cuarto capítulo donde se habla del efecto de la diversidad inicial del sistema y el criterio de selección de pareja en la diversidad subsecuente. Posteriormente se discute la existencia de patrones espaciales y temporales así como la relación entre evolución, entropía y trabajo.

El quinto y último capítulo concretiza las aportaciones del trabajo al entendimiento del efecto de la selección sexual en la diversidad, el papel de la computación en las ciencias de la vida y las aportaciones de la biología a la computación.

Capítulo 1: Marco conceptual

El estudio del papel evolutivo de la reproducción y la selección sexual ha sido de interés para la ciencia desde antes de Darwin (Barthélemy-Madaule 1979, Sloan 1996). Hoy en día, con los avances en genética de poblaciones, sistemática, computación y modelación es posible simular el papel que juegan diferentes fuerzas en la evolución. Debido al largo tiempo que requiere la evolución natural para generar cambios observables es realmente difícil tener una aproximación experimental para estudiarla. Por esta razón las simulaciones computarizadas comienzan a modificar el método científico y a contribuir en la experimentación (Pagels 1988, Palsson 2000, Kosakovsky 2006).

A su vez, la modelación computadorizada es de utilidad en el desarrollo de software debido a que imitando las estrategias naturales podemos crear algoritmos con aplicaciones industriales. Cada vez son más las publicaciones con palabras como bioinspirado, red neuronal, algoritmo genético, inteligencia de cúmulos, etc. (Cui 2006, Botvinick 2006, Kosakovsky 2006, Garnier 2008).

Este trabajo se trata concretamente de simular la selección sexual con el fin de entender su efecto y comportamiento con respecto a la diversidad genética. Al mismo tiempo se pretende hablar de las herramientas de cómputo que utiliza y que provee el algoritmo desarrollado. El algoritmo es sólo un modelo de la selección sexual construido a partir de supuestos y de la emulación de sólo algunos de los aspectos que se consideran pertinentes. Con el fin de que el modelo sea de utilidad tiene algunos parámetros que son manejables como son la cobertura inicial del espacio de soluciones, el criterio de selección y el número de generaciones que se itera.

1.1 Individuos y poblaciones

La idea biológica de individuo es bastante intuitiva, en casi todos los casos se puede decir que un individuo es una unidad viviente, discreta y autónoma que existe en la naturaleza (Begon 2006). Algunas definiciones toman al individuo como la unidad constructora de las especies ya que una especie es un conjunto de individuos con características observables compartidas (Wiley 1981). A nivel genético los individuos pueden ser difíciles de identificar debido a que en algunos organismos existe la partenogénesis y la propagación vegetativa por lo que dos organismos pueden tener el mismo genotipo (Begon 2006). Los individuos son considerados la principal unidad operacional de este trabajo, todos los demás rasgos y entidades mencionados se consideran emergentes del comportamiento de los individuos.

Los individuos tienen una serie de componentes cuantificables que se utilizan para su estudio. Algunos de estos componentes son la talla, la edad, la longevidad y la fecundidad. En este trabajo no se representa ningún otro más que la fecundidad. La fecundidad se refiere al número de descendientes que tiene un organismo (Krebs 1994). Desde un punto de vista darwinista la fecundidad es un indicador de la adecuación (ver 1.6) de un individuo ya que mientras más 'adaptado' se encuentre éste más descendencia dejará (Arnold 2006). La fecundidad puede verse también como un atributo de las poblaciones, la fecundidad poblacional es un promedio de la fecundidad de todos los individuos (Rockwood 2006, Krebs 1994). Algunos autores (Haupt 2003) toman fertilidad como el número de crías de un individuo y fecundidad como el potencial reproductivo, en este trabajo la fecundidad se refiere al número de descendientes.

Cuando los individuos forman congregaciones se dice que existen en una población. Sin embargo una población es mucho más compleja que sólo un grupo de individuos. Las poblaciones

tienen una serie de propiedades cuantificables que nos pueden ayudar a definir las. El rasgo más evidente de las poblaciones es el tamaño (Rockwood 2006). El tamaño poblacional se refiere al número de individuos que conforman una población por lo que es un valor discreto. Queda claro que el tamaño poblacional varía en el tiempo ya que se ve afectado por la natalidad, la mortalidad, la inmigración y la emigración (Carey 1993). Desde el punto de vista genético la riqueza y la diversidad son otros dos componentes de las poblaciones que se discutirán más adelante. Cabe aclarar que desde un punto de vista ecológico la riqueza y la diversidad son atributos de las comunidades ya que no se refieren a atributos interespecíficos.

1.2 Genes y alelos

En la naturaleza los individuos tienen una serie de caracteres observables que los distinguen entre sí, a estos caracteres nos gusta llamarles fenotipo (Hartl 1998). En una visión reduccionista de la biología es cómodo pensar que el fenotipo es únicamente dependiente de los genes. La forma más sencilla de definir *gen* es decir que es una secuencia de ADN codificante con una posición fija en un mapa cromosómico (Tamarin 2001). Esta definición se puede extender al concepto de locus, el cual se refiere únicamente a la posición que ocupa alguna secuencia de ADN en un cromosoma. El locus no es un objeto tangible sino una referencia al lugar en el cual se encuentra un gen.

Un gen que se encuentra en un locus puede tener varias versiones que codifican el mismo carácter pero lo hacen de manera diferente (Watson 2004). Tomemos por ejemplo la codificación del color de ojos en la mosca *Drosophila melanogaster* en la cual existe un grupo de genes que codifica el color de ojos, sin embargo existe una gran gama de colores de ojo, entonces tenemos que los alelos son las secuencias que codifican los diferentes colores que puede tomar el ojo (Watson 2004). Cabe aclarar y enfatizar que muchos caracteres del fenotipo están codificados por la interacción de más de un gen,

es decir que son epistáticos (Hartl 1998). Se ha propuesto que se deje de utilizar el término gen y que sea completamente sustituido por loci (plural de locus) y alelo, siendo el locus la posición que toma en un cromosoma cualquier secuencia de ADN a la que se le pueda llamar alelo (Gillespie 1998).

1.3 Espacio de soluciones

Dado que existen una gran cantidad de genes en un genoma y que cada uno de estos genes tiene diferentes versiones, es fácil pensar que existen un gran número de combinaciones para construir genomas funcionales basándose tan sólo en los genes y alelos conocidos. Si se modelaran todos los posibles genomas funcionales podríamos decir que tenemos un espacio de soluciones de genomas funcionales.

El concepto de espacio de soluciones se originó de la teoría de probabilidad y de los algoritmos de búsqueda. De manera muy laxa se refiere al conjunto de todas las posibles soluciones que existen para un problema dado, es decir que representa todos los resultados posibles dado un número de variables y sus posibles estados (Coley 1999). En biología, el ejemplo más común de un espacio de soluciones es un cuadro de Punnett en el cual se enlistan todas las posibles combinaciones de alelos resultantes de la unión de gametos y se utilizan para calcular la probabilidad de los fenotipos resultantes.

El término “exploración del espacio de soluciones” aparece varias veces a lo largo de este escrito y se refiere a la manera en la que un algoritmo busca algunas o todas las soluciones que constituyen dicho espacio (Coley 1999). El término “cobertura del espacio de soluciones” se refiere a las soluciones que ya han sido encontradas por un algoritmo.

1.4 Selección natural y selección sexual

Imaginemos que cada genoma de un espacio de soluciones teórico existiera dentro de un organismo. Dada la vasta cantidad de genomas tendríamos organismos con muchas diferencias entre ellos. Algunos de estos organismos serían inviables mientras que otros presentarían caracteres que los hacen más adaptados a ciertas condiciones existentes en la tierra.

Una de las principales características de los seres vivos es la adaptación a su ambiente. La adaptación ha sido uno de los principales argumentos para todas las teorías sobre el origen de la diversidad (Templeton 2006). Darwin estaba consiente de la adaptación bastante antes de utilizar el famoso argumento de los pinzones en defensa de la evolución (Maynard Smith 1999). Resulta que la adaptación es sólo uno de los efectos de la selección natural. La selección natural es un conjunto de fenómenos que resultan en la eliminación de la vida poco adaptada a su ambiente y la supervivencia de los organismos con pequeñas ventajas (Arnold 2006).

Para que exista la selección natural deben existir organismos con variación heredable entre ellos. La principal fuente de variación entre los individuos es la recombinación al momento de la meiosis aunque también la mutación juega un papel importante. El argumento consiste en que aquellos individuos con características favorables dejarán más descendencia (Futuyma 2005). La descendencia no sólo será más numerosa sino que tendrá por lo menos algunas de las características favorables de sus antepasados. La idea de que la herencia es importante para la evolución ya había sido divulgada por otros autores como Lamark y Buffon antes que Darwin pero no fue hasta el redescubrimiento de las leyes de Mendel que se comenzó a entender cómo opera la herencia (Maynard Smith 1999).

Otra gran idea atribuida a Darwin es la selección sexual. La definición moderna de selección sexual difiere un poco de la idea Darwinista original. Hoy en día se considera selección sexual al conjunto de factores fisiológicos, morfológicos y conductuales que limitan o favorecen exclusivamente la reproducción entre individuos. Esto incluye eventos precopulatorios, precigóticos y postcigóticos. Dentro de los eventos precopulatorios se encuentran la incompatibilidad genital, la selección por conducta sexual o cortejo, el uso de caracteres sexuales secundarios, etc (Wiley 1981, Grant 1995). Dentro de los precigóticos los principales mecanismos son la incompatibilidad espermatozoide-ovulo y la agresividad del medio vaginal. El mecanismo postcigótico más discutido es la infertilidad de la cría, la cual es particularmente frecuente en los animales híbridos (Wiley 1981, Futuyma 2005).

A pesar de todos estos mecanismos Darwin definió la selección sexual como el proceso de selección de caracteres sexuales secundarios, es decir aquellos que no participan activamente en la producción de cigotos, por lo que todos aquellos mecanismos ligados al sistema reproductivo y a los genitales no se toman en cuenta. El principal ejemplo que utilizó Darwin fueron los plumajes, cantos y patrones conductuales de las aves para conseguir hembra. Posteriormente se habló de caracteres sexuales en plantas como el color, tamaño, forma y aroma de la flor. A medida que han avanzado los estudios de selección sexual se ha visto que ésta tiene un efecto significativo en la adecuación (Medel 2003, Mayhew 2006). Debido a que la selección sexual propone que los individuos seleccionan algunos caracteres para escoger a su pareja es necesario que exista un criterio de selección en base al cual los organismos definan con quien aparearse.

1.5 Reproducción con recombinación

Entendiendo que la selección natural y la selección sexual operan sobre las pequeñas variaciones entre individuos cabe aclarar que la reproducción sexual es una gran fuente de

variación. De manera evolutiva el rasgo más importante de la reproducción sexual es la recombinación genética (Templeton 2006). La mayoría de los organismos que se reproducen de manera sexual son diploides (es decir que tienen dos copias de cada cromosoma). Uno de los mecanismos que permiten que la descendencia de un organismo sea diploide es recibir la mitad de su material genético de uno de los padres y la otra mitad del otro, por lo tanto las células sexuales (gametos) de los padres deben de contener sólo la mitad del material genético (Watson 2004). El proceso por el cual una célula queda con una sola copia de cada cromosoma se llama meiosis. (Watson 2004).

Durante la meiosis los cromosomas homólogos se acomodan en pares y entrecruzan alguna de sus cromátidas. Después de un enrollamiento las cromátidas se rompen en sitios correspondientes de manera que es posible que al momento de volverse a unir las secciones de una cromátida terminen en la otra creando así nuevas combinaciones cromosómicas. Cabe aclarar que la recombinación ocurre de manera azarosa. Al terminar la meiosis cada uno de estos cromosomas híbridos tiene una secuencia diferente a la del padre a pesar de contener los mismos genes, es decir que se encuentran mezclados los genes de los cromosomas originales en un acomodo diferente (Watson 2004).

La recombinación genética ha sido reconocida como la principal fuente de variación entre los organismos ya que con sólo crear nuevas combinaciones de los alelos existentes se pueden obtener cambios fenotípicos notorios (Gouyon 2002).

1.6 Adecuación

Quedando claro que los individuos son todos diferentes entre sí surge el problema de como estimar que diferencias son favorables y como medir que tan favorables son. Se ha discutido que las diferencias favorables son aquellas que le permitan a un individuo dejar más descendencia.

Evidentemente para dejar descendencia los organismos tienen que sobrevivir hasta alcanzar la madurez sexual y luego reproducirse. Por lo tanto la adecuación se puede ver como una medida relativa de tasas de mortalidad y natalidad de un genotipo o fenotipo dado (Stearns 1992).

Desde que se utilizó la frase “supervivencia del más apto” en 1891 se planteó el problema de cómo determinar la adecuación de un individuo. La primera propuesta para cuantificar la adecuación fue desarrollada por Haldane entre 1924 y 1936 y publicada en una serie de artículos titulados *A Mathematical Theory of Natural and Artificial Selection*.

La visión moderna de la adecuación se deriva directamente de la genética de poblaciones y se refiere a la adecuación de un genotipo más que a la de un individuo (Stearns 1992). La adecuación de un genotipo es la contribución promedio de los individuos con ese genotipo a la población a lo largo de una o más generaciones. Debido a que se está midiendo la contribución genética hay dos propiedades importantes que determinan la adecuación de un genotipo. El primero es su fecundidad (el número de descendientes que deja) y el segundo es la supervivencia de los descendientes hasta el siguiente evento reproductivo (Futuyma 2005).

La adecuación de un individuo varía de acuerdo a su fecundidad. Un factor que puede variar la adecuación (y la fecundidad) de un individuo es la frecuencia de individuos fenotípicamente similares a él (Frost 2006). En el caso de la adecuación de un genotipo esto se ve como la adecuación de sí mismo como función de su frecuencia en la población. Cuando la selección natural opera sobre una población en la cual existe adecuación dependiente de la frecuencia se dice que se encuentra bajo selección dependiente de la frecuencia (Mayhew 2006).

El concepto de adecuación en cómputo es un poco más flexible ya que en general se utiliza para comparar el desempeño de un algoritmo para producir un resultado deseado. Como conocemos el resultado que queremos se corren todos los algoritmos en competencia y se observa cuál se acerca más a la solución que deseamos. La idea de que algunos algoritmos (u organismos) tienen adecuación más alta que otros permite organizar el espacio de soluciones en un paisaje de adecuación. Lo que pretende hacer en un paisaje de adecuación es acomodar las soluciones de acuerdo a su adecuación, de manera que hay picos de alta adecuación y valles de baja adecuación lo cual resulta bastante útil para comparar el desempeño de diferentes soluciones (Mitchell 1998).

1.7 Riqueza y diversidad

Al operar la selección natural sobre los individuos en función de su adecuación se producen cambios en las frecuencias alélicas, esto junto con otras fuerzas evolutivas como la mutación y la deriva génica causan la creación de nuevos genomas, cambios de frecuencia en genomas existentes y la desaparición de algunos genomas.

Uno de los problemas que la biología ha tratado de resolver desde sus inicios es contar la cantidad de especies de un lugar. Al número de especies que ocupan ciertos límites espaciales o temporales se le llama riqueza de especies. El concepto de riqueza en realidad sólo quiere decir cuantas cosas diferentes tengo (Futuyma 2005, Begon 2006) por lo que puede ser empleado en genética de poblaciones para crear indicadores como riqueza genotípica (el número de diferentes genomas) y riqueza alélica (el número de diferentes alelos que existen para un gen dado).

Sin embargo la riqueza genotípica es un mal indicador de la información que contiene una población ya que le asigna la misma importancia a cada una de las variables que encuentra. Por esta

razón se crearon los índices de diversidad. Fuera de ser un consenso, se puede decir que la diversidad no sólo se preocupa por el número de cosas diferentes si no también por la abundancia de cada una de ellas por lo cuál se crearon los índices de diversidad. Entre los índices de diversidad se encuentra el índice de Shannon, el cual se utiliza en este trabajo (Begon 2006).

El índice de diversidad de Shannon o entropía de Shannon proviene de las telecomunicaciones y ha sido descrito como 'un indicador de sorpresa'. Lo que la entropía de Shannon hace es describir qué tan homogéneo es un sistema; los sistemas muy homogéneos tienen baja entropía de Shannon (cercana a cero) y los sistemas poco homogéneos presentan alta entropía (cercana a 1). La función de entropía requiere sólo la abundancia de nuestros diferentes estados (ya sean especies o genotipos) y la riqueza de estados. La función de entropía se expresa así (Shannon, 1948):

$$H = - \sum_{i=1}^n p(i) \log_a p(i)$$

Donde H es la entropía, $p(i)$ es la proporción del estado i en el sistema y a es la riqueza de estados. Si definimos que $\log_a 0 = 0$ entonces la curva tiene como mínimo cero, cuando la proporción del estado i es cero o uno (por lo que existe un sólo estado) y como máximo uno, cuando cada uno de los estados se encuentra en igual proporción.

Para que la curva siempre ajuste entre cero y uno es importante que la base del logaritmo sea la riqueza de estados, por esta razón la entropía de Shannon no tiene unidades salvo en base dos (bits) y en base diez (dibits). Sin embargo existen algunos escenarios en los cuales utilizar la riqueza como base del logaritmo eliminaría información relevante, para evitar perder información es importante definir una base en común. En este trabajo se conoce la riqueza máxima a la cual llega cada escenario y se utiliza como base sólo para hacer comparaciones relevantes a ese escenario. Si se quisieran hacer

comparaciones entre pruebas se podría utilizar una riqueza máxima teórica como base del logaritmo.

1.8 Autómatas celulares

Al principio de este trabajo se dijo que uno de los objetivos es la creación de un modelo de cómputo utilizable para responder las preguntas centrales. Después de leer las secciones anteriores debería quedar claro que el modelo que se utilice debe de incluir individuos con interacciones locales y estados diferentes (equivalentes a un genotipo). Además los individuos deben de poderse integrar en una población y los fenómenos poblacionales deben de ser iterables en el tiempo. La herramienta de modelado que mejor cumple con los requerimientos mencionados anteriormente es un autómata celular.

En teoría de la computación se le llama autómata celular a un modelo de cómputo que es dinámico, discreto en el espacio y el tiempo, que opera en una malla regular y que se caracteriza por trabajar en interacciones locales (Schiff 2008). En términos menos técnicos un autómata celular consta de una cuadrícula en donde cada posición (celda) puede estar en un número finito de estados. El estado de cada celda en el tiempo $t+1$ está dado por una regla que toma como variables su propio estado y el de sus vecinos en el tiempo t . En general los autómatas celulares utilizan vecindades de Moore, es decir que cada celda tiene ocho vecinos (figura 1) y topologías toroides pero estos parámetros no son un requisito (Mitchell 2009).

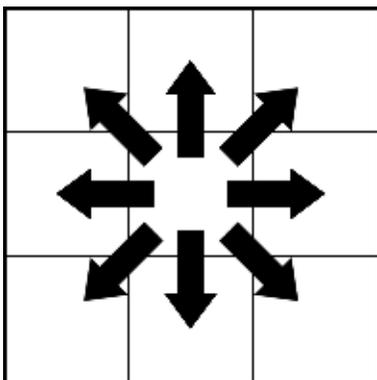


Figura 1: Vecindades de Moore. La casilla central del cuadro mostrado tiene contacto con sus ocho vecinos (Norte, Noreste, Este, Sureste, Sur, Suroeste, Oeste y Noroeste).

Un aspecto interesante de la computación con autómatas celulares es que a partir de reglas sencillas se pueden obtener patrones complejos. Por esta razón los autómatas celulares han sido ligados más de una vez a la biología, sobre todo en campos como embriología, ecología y evolución (Das 1995, Hraber 1997, Akberdin 2006). Incluso hay quienes aseguran que la naturaleza completa es discreta y se puede reducir a computación con autómatas (Wolfram 2002).

Durante los años cuarenta, uno de los inventores de los autómatas celulares, John Von Neuman, demostró que la computación con autómatas podía computar cualquier cosa que una maquina de Turing pudiera computar. Dicha demostración hace que los autómatas celulares cumplan la equivalencia de Turing y por lo tanto sean computacionalmente universales (Mitchell 2009). La integración de la arquitectura de autómatas con la biología y la universalidad computacional han hecho que algunos autores (Wolfram 2002, Kurzweil 2005) consideren que la naturaleza también cumple con la universalidad computacional.

1.9 Espacio de color RGB de 24-bits

Sabiendo que los autómatas celulares requieren de un número finito de estados y que cada estado, para fines de este trabajo, es equivalente a un genotipo es importante definir la estructura informática de los estados. Con el fin de que cada estado sea visualmente único y se pueda ver como análogo a un fenotipo se decidió que los estados de las celdas se codificaran como color de 24-bits.

El espacio de color RGB está constituido por un canal rojo (R), un canal verde (G) y un canal azul (B). El valor de cada canal se codifica en 8-bits por lo cual cada color se codifica en 24-bits. Dado que cada bit puede ser o uno o cero tenemos dos opciones que se pueden acomodar en 24 posiciones. Por lo tanto el espacio completo tiene 2^{24} elementos o 16,777,216 colores diferentes. Cada color esta

codificado por un valor entre 0 y 255 (2^8) por canal (Evening 2002). A lo largo de este texto los valores de RGB se expresan en una lista empezando por el rojo, luego el verde y al final el azul. Por ejemplo el color magenta se codifica como [255,0,255] es decir que tiene 255 rojo, 0 verde y 255 azul.

1.10 Algoritmos genéticos

Un modelo basado en autómatas cumple con casi todos los requisitos del modelo que se plantea construir. Sin embargo las partes de heredabilidad y reproducción diferencial de acuerdo a la adecuación no son elementos de un autómata celular, por esta razón el modelo es también dependiente de un algoritmo genético modificado.

Un algoritmo genético es aquel que emula un proceso evolutivo con el fin de llegar a una o varias soluciones óptimas (Mitchell 1998). En realidad no existe una definición estricta de algoritmo genético y la mayor parte de la literatura no hace distinción entre un algoritmo genético y un algoritmo evolutivo.

Sin embargo hay ciertas características que comparten todos los algoritmos genéticos y evolutivos. Todos ellos tienen una población de cromosomas, una fase de recombinación, una fase de selección de acuerdo a la adecuación y un periodo de mutación (Mitchell 1998).

Los algoritmos genéticos clásicos tienen una fase de selección seguida de una fase de reproducción. Durante la fase de selección se decide cuales organismos del total de la población se reproducirán durante la fase de reproducción. En la siguiente fase de selección se evalúa la adecuación de las nuevas combinaciones creadas y se eliminan aquellas de menor adecuación (Coley 1999). En un algoritmo genético clásico existe un criterio de selección global contra el cual se evalúa toda la

población para determinar la adecuación de cada individuo y en base a este parámetro decidir cuales individuos se reproducen (Coley 1999). Como se verá en el capítulo dos, algunas de estas características se han modificado para adaptar el algoritmo a los objetivos y al procedimiento de este trabajo.

Capítulo 2: Algoritmo Espacial de Selección Sexual

El modelo que se presenta en este capítulo tiene como objetivo simular, a partir de reglas sencillas, el efecto de la selección sexual en la variabilidad genética a lo largo de un tiempo definido. Cada sección de este capítulo explica los aspectos técnicos de cada componente utilizado para generar el modelo; Se explica también la relación que existe entre los elementos del modelo y el proceso biológico que representan.

2.1 Reglas del autómata

Como se ha repetido a lo largo de este texto las reglas del autómata provienen de la abstracción y reducción de fenómenos biológicos relevantes. Antes de profundizar en los detalles técnicos esta sección pretende delimitar las reglas de actualización junto con las propiedades del autómata y correlacionarlas con el fenómeno biológico del cual fueron derivadas.

El primer aspecto relevante es que el espacio simulado tiene una topología plana por lo que hay bordes. El uso de una topología de este tipo en una simulación ecológica tiene sentido ya que las poblaciones terrestres presentan límites geográficos en los cuales el contacto interpoblacional se reduce.

En la naturaleza un individuo que busca aparearse sólo puede tener contacto sexual con aquellos individuos que están en su vecindad o en la vecindad de sus polinizadores. Con el fin de simular un fenómeno espacial de contacto cada casilla del plano sólo tiene acceso a los ocho vecinos que conforman su vecindad de Moore.

Al igual que en la vida cada individuo debe de tener un conjunto de características observables similares a un fenotipo que se encuentre codificado de manera discreta en una secuencia. Por esta razón se decidió que los estados del autómata fueran colores. Además de presentar una característica observable, el color digital se codifica en una cadena de tres variables discretas a las cuales se les asigna un valor numérico. Por lo tanto cada variable se puede ver como un gen y cada valor como un alelo.

Finalmente la regla de actualización del autómata también esta derivada de fenómenos biológicos. En cada generación cada casilla escogerá a uno de sus vecinos para reproducirse, la manera de seleccionar a su pareja se basa en su propio color y en el de sus vecinos. Sin embargo el criterio de selección varía ya que en la naturaleza se ha observado que algunos organismos prefieren parejas similares (Stanton 1989, Medel 2003) y en otros casos hay una preferencia por parejas contrastantes (Marshall 2003). La transferencia de información, al igual que en la reproducción sexual es discreta y tiene un componente aleatorio comparable (más no equivalente) a la recombinación.

El autómata celular propuesto consiste de una cuadrícula de 512 x 512 casillas por lo que el tamaño poblacional es de 262,144 individuos. Esto crea una restricción para explorar la totalidad del espacio RGB ya que este tiene 16,777,216 colores. Por esta razón se considera como espacio de soluciones al espacio de color que contiene sólo los posibles valores RGB que se pueden crear con los valores existentes en la condición inicial.

2.2 Propiedades de los individuos

Cada casilla del autómata se puede considerar análoga a un individuo en una población. Los individuos del autómata tienen un color como principal componente. El color se puede ver como un

carácter codificado por tres genes. Cada gen corresponde a un canal RGB por lo que hay un gen rojo, uno verde y uno azul. Los genes son epistáticos de manera que los tres contribuyen al color final del individuo. Cada gen tiene 256 posibles alelos (cualquier número entero entre 0 y 255) que al combinarse se pueden distinguir de manera fenotípica como un color, a cada combinación de alelos (por lo tanto a cada color) se le denomina genotipo. Esta codificación es particularmente útil porque los colores parecidos tienen genotipos parecidos y se encuentran a poca distancia en el espacio de soluciones.

Los individuos tienen una longevidad fija con el fin de que la adecuación sea únicamente una función de la fecundidad. Todos los organismos de la misma generación se mueren simultáneamente una vez que ha terminado la etapa de reproducción del algoritmo.

2.3 Fecundidad y adecuación

Cómo se dijo al final del capítulo anterior la adecuación en un algoritmo genético clásico se mide de acuerdo a la proximidad de un resultado dado contra un criterio de selección global. En este procedimiento la adecuación de un individuo es únicamente función del criterio de selección. Esto difiere de un sistema biológico en el cual la adecuación es una función de la fecundidad. La fecundidad a su vez esta determinada por un criterio de selección, sin embargo este criterio de selección no es global y es función tanto del estado del propio organismo como del vecino contra el cual se compara. Un ejemplo concreto de esto es el caso de *Raphanus raphanistrum* en el cual las plantas con flores de un color serán fertilizadas por polen de plantas con flores similares por lo cual el estado propio es muy relevante (Stanton 1989, Grant 1995).

Al querer utilizar una definición de adecuación dependiente de la fecundidad es necesario definir por qué algunos individuos tienen más alta fecundidad que otros. En *El origen de las especies* y posteriormente en *El origen del hombre y selección ligada al sexo*, Darwin explica que hay especies (sobre todo animales) en las cuales algunos individuos poseen caracteres (tanto morfológicos como conductuales) que los hacen más atractivos al sexo opuesto. Al ser más atractivos estos organismos son seleccionados más veces para la reproducción por lo que en las siguientes generaciones existen más copias de sus genes. Este fenómeno es conocido como selección de pareja (Futuyma 2002).

En el caso concreto del autómata el color es el único carácter que se selecciona de manera sexual. Los individuos que presentan colores que cumplen de manera satisfactoria con el criterio de selección son escogidos más veces por lo cual sus valores RGB se encuentran en mayor proporción en la siguiente generación. Cabe aclarar que la adecuación no es heredable, dado que puede ocurrir selección dependiente de la frecuencia, la heredabilidad de la adecuación dependerá de si la selección dependiente de la frecuencia es positiva o negativa.

2.4 Distribución

Una de las características de los individuos es que ocupan un lugar geográfico, por esta razón es importante incluir una dimensión espacial en el modelo. La posición en el espacio está determinada por un valor i y un valor j pertenecientes a un plano cartesiano. Una vez que a un individuo se le asigna un valor i,j se representa como un pixel en ese punto. Los individuos en esta versión del algoritmo tienen un coeficiente de dispersión muy alto por lo cual los descendientes de cualesquiera dos individuos pueden 'nacer' en cualquier punto del espacio.

La distribución de individuos se define de manera aleatoria cada generación. Este modo de dispersión se puede cambiar y no es un requisito del algoritmo.

2.5 Selección y reproducción

El instante previo a la reproducción cada individuo selecciona su pareja. El proceso de selección consiste en que cada individuo se compara con sus ocho vecinos para determinar cuál es el que mejor cumple con el criterio de selección.

La elección de un criterio de selección basado en fenómenos biológicos es un tanto problemática ya que en las diferentes especies existen una gran variedad de criterios. Como producto del aumento de probabilidad de expresar alelos recesivos detrimentales al haber autogamia, los individuos han evolucionado una conducta en la cual evitan reproducirse con organismos similares a ellos mismos (Marshall 2003). Sin embargo también existen especies en las cuales la reproducción ocurre frecuentemente entre individuos parecidos. Esto es común en plantas polinizadas por animales en las cuales un polinizador se ve atraído por flores de forma, color, aroma y sabor similar (Grant 1995, Medel 2003). Por esta razón se hicieron simulaciones utilizando tanto el criterio de selección de similitud como el de diferencia.

En el caso particular de este trabajo se mide la distancia euclidiana del valor del estado de una celda en el espacio RGB a los valores de todos sus vecinos. Aquel vecino que mejor cumpla con el criterio de selección de distancia euclidiana será con el que se cruza. En caso de que se encuentre con dos individuos a la misma distancia entonces escogerá al primero con el que se compare. Es importante recordar que colores similares se encuentran a poca distancia en el espacio de soluciones de manera que la distancia euclidiana funciona como una medida de la similitud entre colores.

Una vez que se ha seleccionado la pareja entonces comienza la fase de reproducción. La reproducción debe generar nuevas combinaciones derivadas de los padres. La forma de crear combinaciones consiste en escoger al azar una variable (rojo, verde o azul) del vecino seleccionado. El valor de este gen sustituirá al valor de la misma variable del color central causando que los hijos tengan 2 valores provenientes de uno de los padres y 1 proveniente del otro (Tabla 1).

Padres	Posibles hijos
R1, G1, B1	R2, G1, B1
R2, G2, B2	R1, G2, B1
	R1, G1, B2

Tabla 1: La tabla presentada muestra los posibles combinaciones resultantes de que el color R1,G1,B1 escoja como pareja al color R2,G2,B2

Todos los hijos generados se almacenan en una lista de la cual se escogen al azar para generar una distribución aleatoria en la siguiente generación.

2.6 Suposiciones del algoritmo

Como toda abstracción de la naturaleza este algoritmo posee algunas suposiciones:

- El color (fenotipo) de los organismos esta codificado sólo por 3 variables (genes).
- El color de los organismos es el único atributo que se selecciona sexualmente.
- No se encuentran modeladas otras fuerzas evolutivas además de selección sexual y reproducción con recombinación aunque es posible que emerjan.
- Las generaciones son discretas. Por lo tanto los padres mueren al momento de la reproducción.
- Los hijos de los organismos de una generación pueden tomar cualquier punto al azar del espacio en el que vivían sus padres. Es decir que se dispersan aleatoriamente por la totalidad del espacio.

2.7 El algoritmo en pseudocódigo

Con el fin de ilustrar el algoritmo descrito anteriormente con una estructura similar a un lenguaje de programación a continuación se presenta en pseudocódigo:

```
Crear un espacio de  $n*m$  casillas
Individuos = [Crear colores iniciales en formato [R,G,B]]

for casillas in espacio: //Asignar color a las casillas
    Color de la Casilla = random (Individuos)

for GeneraciónActual in Número_De_Generaciones:

    for casillas in espacio:
        Obtener colores de los vecinos de  $i-1$  a  $i+1$  y  $j-1$  a  $j+1$ 

        for Vecinos( $i-1$  a  $i+1$ ,  $j-1$  a  $j+1$ ):
            DistanciaEntreColores =  $\sqrt{(R_0-R_1)^2+(G_0-G_1)^2+(B_0-B_1)^2}$ 

            VecinoSeleccionado = Criterio(DistanciaEntreColores)
            ColorHijo = cambiar un valor RGB con VecinoSeleccionado
            Almacenar ColorHijo en una lista ColoresNuevos

    for casillas in espacio:
        Color de la Casilla = random (ColoresNuevos)
```

2.8 Simulaciones realizadas

Con el objetivo de encontrar patrones espacio-temporales en el Algoritmo Espacial de Selección Sexual (AESS) se crearon 3 condiciones iniciales que se corrieron con dos criterios de selección, por lo cual existen 6 escenarios. El objetivo de las pruebas fue determinar la influencia de la diversidad inicial y el tipo de selección en la diversidad de tiempos subsecuentes.

Las pruebas se realizaron aplicando el algoritmo a la selección entre colores con el fin de tener un resultado visual además de numérico. El espacio durante las pruebas fue de 512 x 512 px por lo cual incluyó 262,144 individuos.

Cada prueba consiste de cien iteraciones del algoritmo para permitir que el sistema llegara a estados cercanos a la estabilidad. Después de cada iteración se calculó la entropía de Shannon como un indicador de diversidad total del sistema. Para poder comparar el cambio en diversidad a lo largo del tiempo se utilizó como base del logaritmo la riqueza máxima registrada durante las cien iteraciones de manera que si la entropía es uno en algún tiempo sabemos que en ese momento existe el mismo número de copias de cada uno de los estados que tomó el sistema en el momento de máxima riqueza. Cabe aclarar que la riqueza y diversidad que se utilizan son genotípicas, es decir que se están contando las diferentes combinaciones de valores RGB como riqueza.

Con el fin de determinar la forma en la que se explora el espacio de soluciones y buscar atractores se registraron en cada generación los valores RGB de la casilla del centro (256,256), de la esquina superior izquierda (0,0) y de un punto al azar de la malla. Esto sería equivalente a coleccionar un individuo de tres localidades durante cien generaciones, secuenciar su genoma y colocar un punto en la posición correspondiente del espacio de soluciones genético. En un escenario biológico un atractor se podría considerar como el genotipo de mayor adecuación.

La distribución inicial y de cada iteración en el espacio fue pseudoaleatoria determinada por el generador de números aleatorios Mersenne Twister implementado en el lenguaje de programación Python (Matsumoto 1998).

Debido a que en la naturaleza existen tanto criterios de selección de similitud como de disimilitud cada condición inicial se corrió utilizando tanto el criterio de selección de mínima distancia euclidiana como el de máxima distancia euclidiana, esto con el fin de poder comparar los efectos del criterio de selección.

ESCENARIO 1:

En la primera condición inicial se intentaron eliminar todas las variables que afectan la adecuación con excepción del criterio de selección de pareja. Para eliminar los fenómenos dependientes de la frecuencia genotípica se creó una condición inicial en la cual todos los genotipos existen sólo una vez. En esta prueba también se pretendió observar el efecto de la selección en una población con riqueza y diversidad máxima.

Para crear las condiciones explicadas anteriormente se construyó un espacio de soluciones en el cual existieran todos los posibles genomas dados los alelos existentes, por esta razón podemos decir que el espacio de soluciones está completamente explorado, como resultado de esto no hay manera de crear combinaciones que no existieran en un tiempo anterior. El espacio de soluciones se construyó generando todas las posibles combinaciones de valores RGB en incrementos de 4 como se muestra en la siguiente lista:

[0,0,0], [0,0,4], [0,0,8], ..., [0,4,0], [0,4,4],[0,4,8], ... ,[252,252,244], [252,252,248], [252,252,252]

La riqueza genética de este espacio es de 252,144 combinaciones (genomas). Si se mapea la posición de cada individuo en el espacio RGB observamos que el espacio es un cubo discontinuo de 252 x 252 x 252 que presenta un punto cada 4 incrementos en cualquier dirección (Figura 2). Esta cobertura inicial se utilizó para las simulaciones uno y dos.

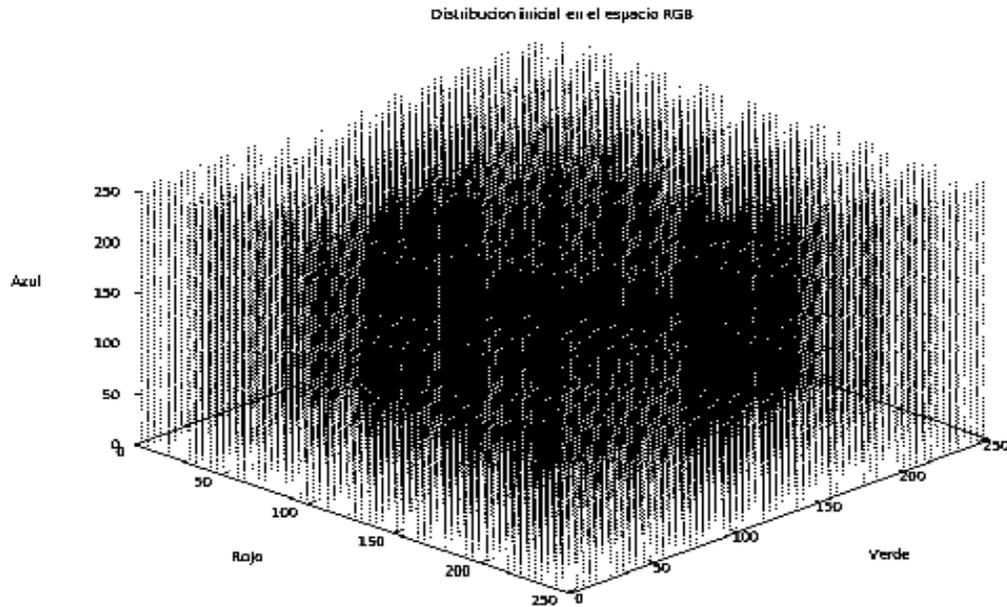


Figura 2: Distribución inicial de los puntos para las pruebas 1 y 2 en el espacio de soluciones RGB. Nótese que todos los posibles valores en incrementos de 4 están ocupados por lo cual el espacio de soluciones esta completamente explorado desde el inicio. La figura muestra una zona más oscura en el centro del cubo, esto no se debe a una mayor densidad de puntos si no a una ilusión óptica.

ESCENARIO 2:

En la condición inicial anterior existen todos los posibles genomas construidos a partir de tres genes con 256 alelos por gen. Dichas condiciones parecerían biológicamente improbables por lo cual en este escenario se reduce el número de posibles alelos a dos por gen, de manera que sólo existen ocho genotipos posibles (2^3). También parecería improbable que existieran todos los posibles genotipos en la primera generación por lo que la condición inicial de esta prueba contiene sólo dos genotipos. Con el fin de mantener neutrales los efectos de la selección dependiente de la frecuencia en la condición inicial los dos genotipos se encuentran en igual proporción

Las características anteriores se cumplieron reduciendo la cobertura inicial del espacio RGB a los puntos $[0,0,0]$ y $[255,255,255]$ (Figura 3) y asignándole aleatoriamente 131,072 píxeles a cada uno de estos colores. Estas condiciones se utilizaron para las simulaciones tres y cuatro.

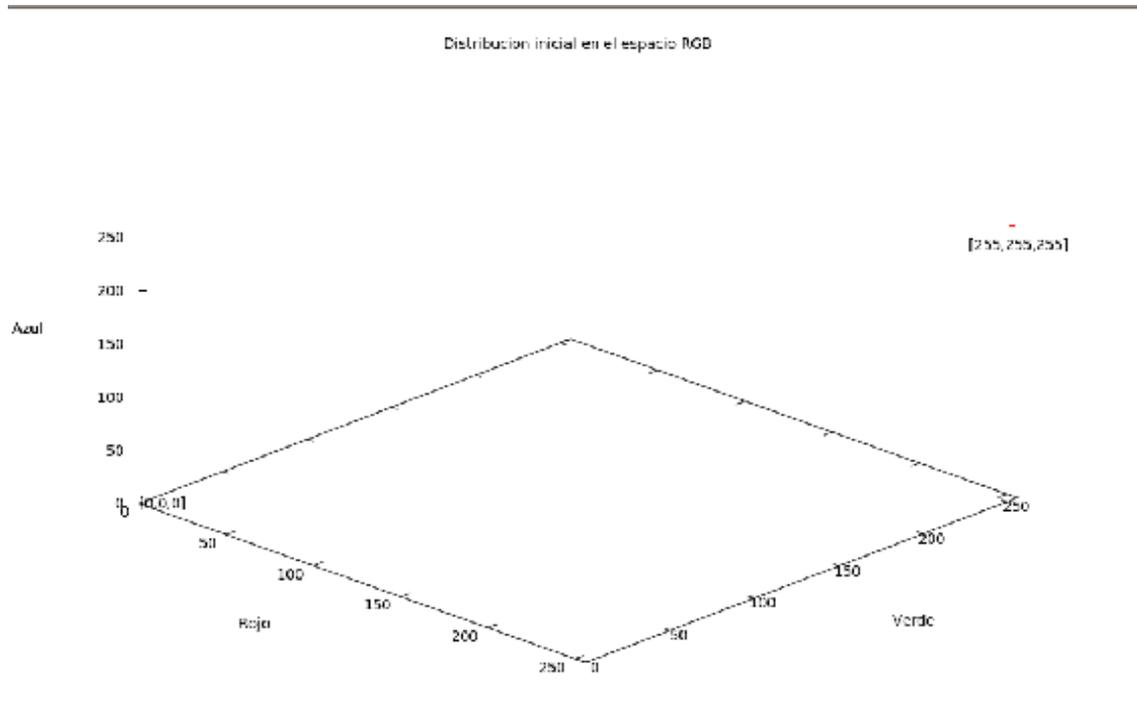


Figura 3: Distribución inicial utilizando sólo dos esquinas ($[0,0,0]$ y $[255,255,255]$) del cubo RGB.

En el caso de selección por mínimo de distancia euclidiana las únicas dos forma de generar nuevas combinaciones es que un cuadro negro $[0,0,0]$ se encuentre rodeado de cuadros blancos $[255,255,255]$ o viceversa (Figura 4).

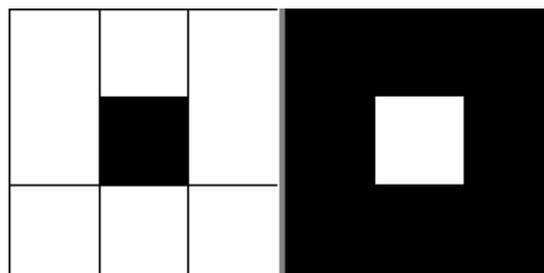


Figura 4: Las únicas dos combinaciones de vecinos que producen nuevos colores bajo la cobertura del espacio RGB en la que sólo se toman las esquinas $[0,0,0]$ y $[252,252,252]$.

ESCENARIO 3:

En los dos escenarios anteriores las soluciones iniciales son simétricas en el espacio RGB. Un espacio de soluciones creado a partir de un genoma real probablemente no lo sería (la discusión sobre como construir un espacio de soluciones a partir de datos de genomas reales no esta dentro del enfoque de este trabajo y probablemente podría ser una tesis por sí sola) por lo que es necesario hacer una simulación con una condición inicial que no cubra el espacio RGB de manera simétrica. En este caso no se ha prestado atención a eliminar de la condición inicial los efectos de la selección dependiente de frecuencia por lo cual algunos genotipos son más comunes que otros.

La construcción de un condición inicial asimétrica en el espacio RGB es complicada por lo cual se ha importado una imagen existente (figura 5), la cual se permutó de manera aleatoria sobre la cuadrícula de 512 x 512 px para crear una condición inicial. El espacio de soluciones de esta imagen se desconoce ya que con los alelos existentes en la condición inicial es posible crear más genomas de los que existen en dicha condición. La cobertura inicial del espacio de soluciones obtenida de esta imagen fue la que se uso en las pruebas cinco y seis.

Capítulo 3 Resultados de las simulaciones

Los resultados de este trabajo se presentan como gráficas para facilitar su lectura (para ver los mapas de color consulte el apéndice A). Para cada una de las pruebas se graficó la entropía de Shannon con respecto al tiempo. También se graficaron los valores RGB que tomó la esquina superior derecha (0,0), el centro del espacio (256,256) y un punto al azar del plano en cada generación. De esta manera se obtuvieron las trayectorias de exploración de cada uno de estas tres casillas.

3.1 Espacio de soluciones regular, completamente explorado y selección por mínima distancia euclidiana

En este caso vemos que el valor inicial de la entropía es máximo, esto se debe a que en la primera generación (T_0) cada individuo es diferente y existe sólo una copia de él, por lo tanto la incertidumbre en el sistema es máxima ($H_0=1.0$). Entre la primera y la segunda generación hay un decremento abrupto de los valores ($H_1\approx 0.95365$), la causa de dicha disminución es que muchos individuos, al haber recombinación, producen combinaciones que ya existían, de esta manera comienza a haber más copias de algunos individuos que de otros. La entropía final de esta prueba (en la generación 99) fue de $H_{99}\approx 0.09509$ lo cual indica un sistema casi homogéneo; Es decir que hay pocas combinaciones y muchos individuos iguales (Figura 6).

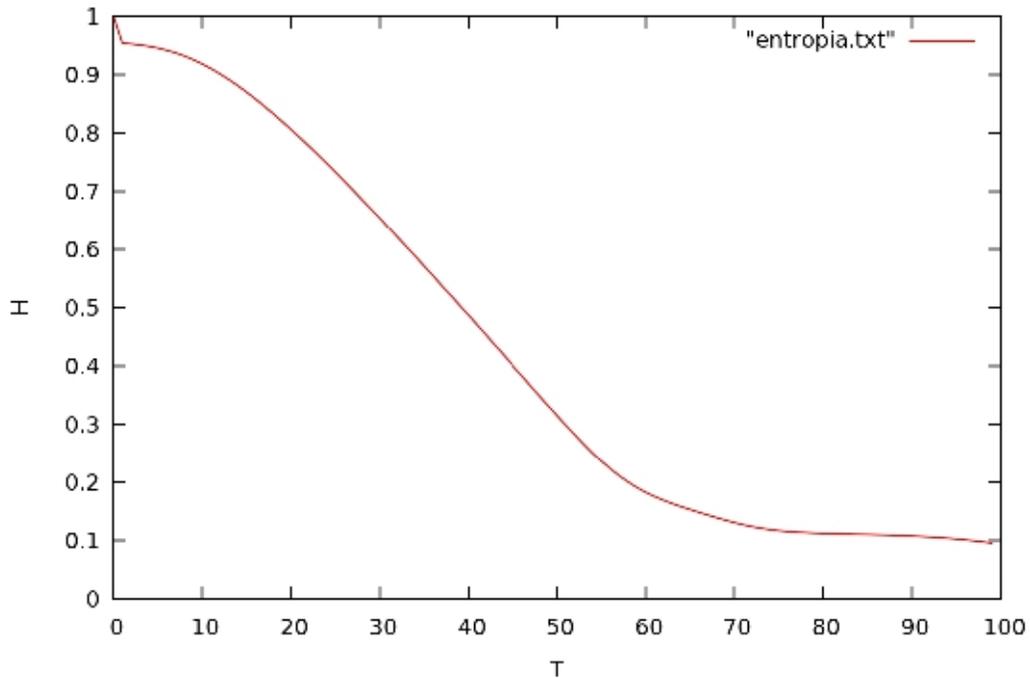


Figura 6: El decaimiento de entropía (H) a lo largo del tiempo. Cada valor de H corresponde a la entropía total de Shannon en la iteración T del programa. El decaimiento abrupto ente las generaciones 0 y 1 se debe a que al crear la población de la generación 1 se eliminan algunas combinaciones y otras se repiten.

Si comparamos las trayectorias de exploración (figura 7) de cada una de las coordenadas muestreadas ((0,0), (256,256), (30,200)) los valores que toman durante las primeras generaciones no parecen tener relación entre sí; Sin embargo conforme avanzamos hacia las últimas generaciones los valores cercanos a [128,128,128] se vuelven más frecuentes. Al parecer con este modo de selección y esta cobertura inicial hay un atractor fijo en el punto [128,128,128], es decir que en las últimas generaciones dicho genotipo es de más alta adecuación.

Exploracion del espacio por tres casillas

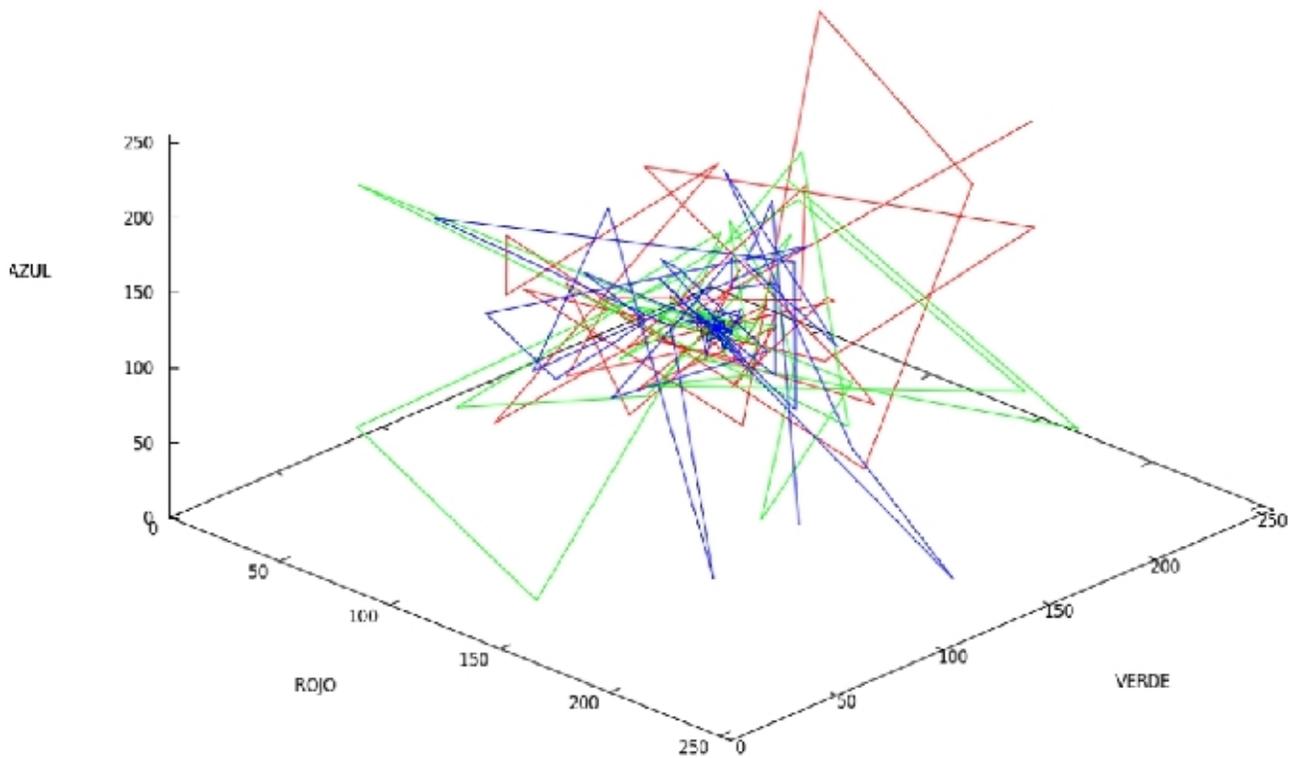


Figura 7: El gráfico anterior muestra los valores que tomaron las casillas 0,0 (línea roja), 256,256 (línea verde) y 39, 200 (línea azul) durante cada generación.

3.2 Espacio de soluciones regular, completamente explorado y selección por máxima distancia euclidiana

Debido a que la cobertura inicial en el espacio de soluciones es igual a la de la prueba 1 (Figura 8) la entropía inicial también es la misma. En esta prueba también se observa un decaimiento abrupto

de la diversidad entre la primera ($H_0=1.0$) y la segunda generación ($H_1=0.95117$). El valor final de la entropía en esta prueba fue de $H_{99}=0.17100$, esto indica un sistema final con poca diversidad pero con diversidad mayor a la prueba anterior. La diferencia de entropía entre esta prueba y la anterior se debe a que en la generación 99 de esta prueba se tiene una población con ocho diferentes combinaciones de valores RGB que difieren mucho entre si.

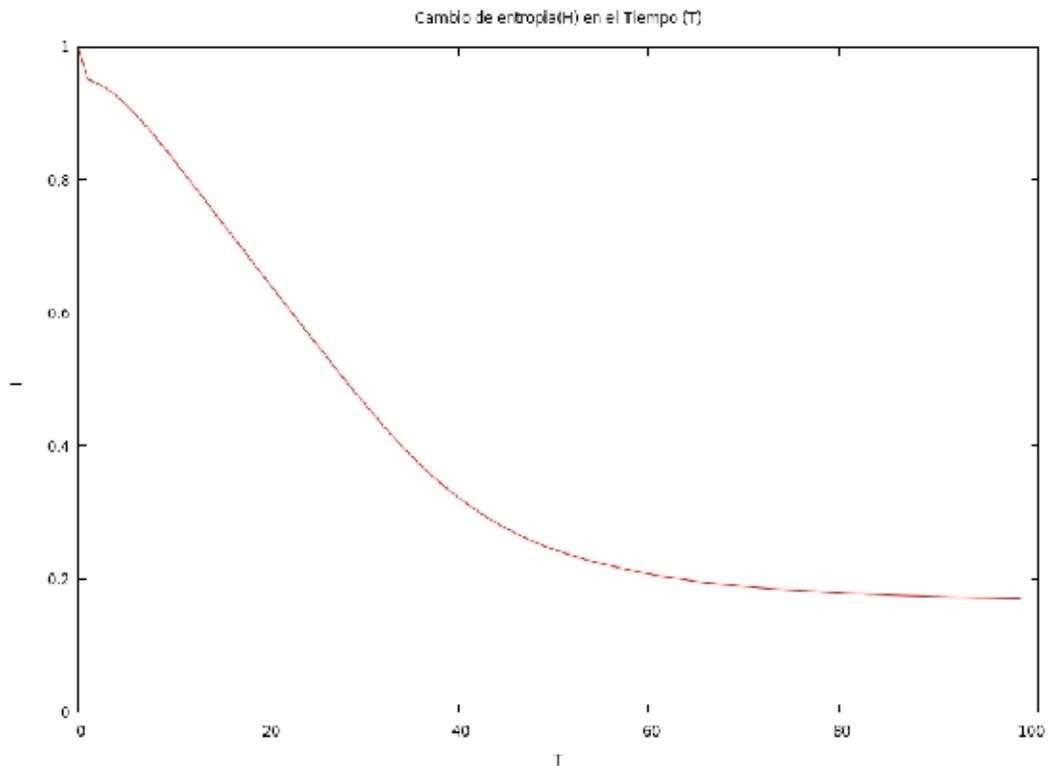


Figura 8: A pesar del cambio en el criterio de selección la curva mantiene una forma igual que en la prueba anterior. Nótese también que el valor de la última generación no es cero como resultado de una riqueza que no tiende a uno.

Al observar los valores que toman los tres píxeles elegidos ((0,0), (256,256) (85,100)) del espacio notamos que existen ocho atractores en esta prueba. La existencia de más de un atractor se debe a que bajo este criterio de selección existen individuos con adecuaciones equivalentes por lo que se mantiene más de una combinación hasta al final (Figura 9).

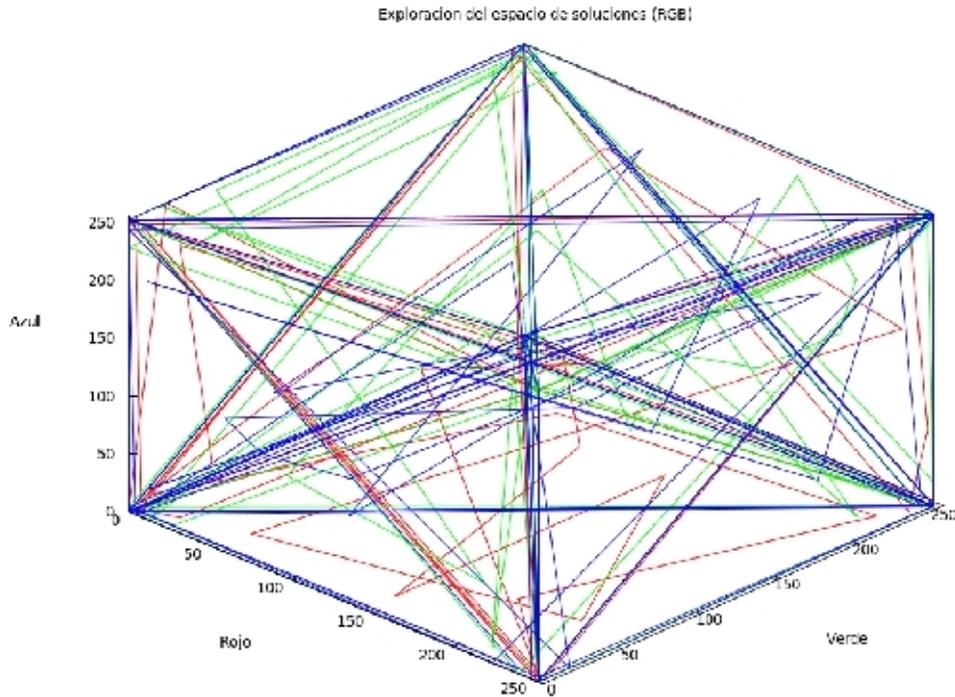


Figura 9: Bajo el criterio de selección de máxima distancia euclidiana existen ocho atractores en un espacio de soluciones regular y completamente explorado. Por esta razón la entropía se mantiene estable en un punto mayor a cero.

3.3 Cobertura de soluciones inicial reducida a dos puntos distantes y selección por mínima distancia euclidiana

La entropía inicial de este escenario es medianamente baja ($H_0=0.33333$), esto se debe a que en este momento la riqueza del sistema no es máxima por lo que se observa una baja diversidad dependiente de una baja riqueza. Debido a que durante esta prueba el criterio de selección fue de mínima distancia euclidiana lo más probable es que cada individuo tome a otro del mismo color como pareja. Dado que los valores RGB de los individuos iniciales son $[0,0,0]$ ó $[255,255,255]$ entonces los hijos de padres del mismo color son iguales a sus padres. Sin embargo cuando un individuo se encuentra rodeado de puros colores diferentes a él entonces se ve obligado a reproducirse con uno de

ellos generando así un nuevo color de 8 posibles. Entendiendo esto queda claro que la entropía no se incrementa más allá de $H_{96}=0.37199$ porque los cambios en diversidad y riqueza no son probables por lo cual el sistema tiende a regresar a tener individuos blancos y negros. Las oscilaciones que se observan en la curva (Figura 10) se deben precisamente a que en algunas generaciones las proporciones de $[0,0,0]$ y $[255,255,255]$ se acercan a 1:1 y en otras se alejan.

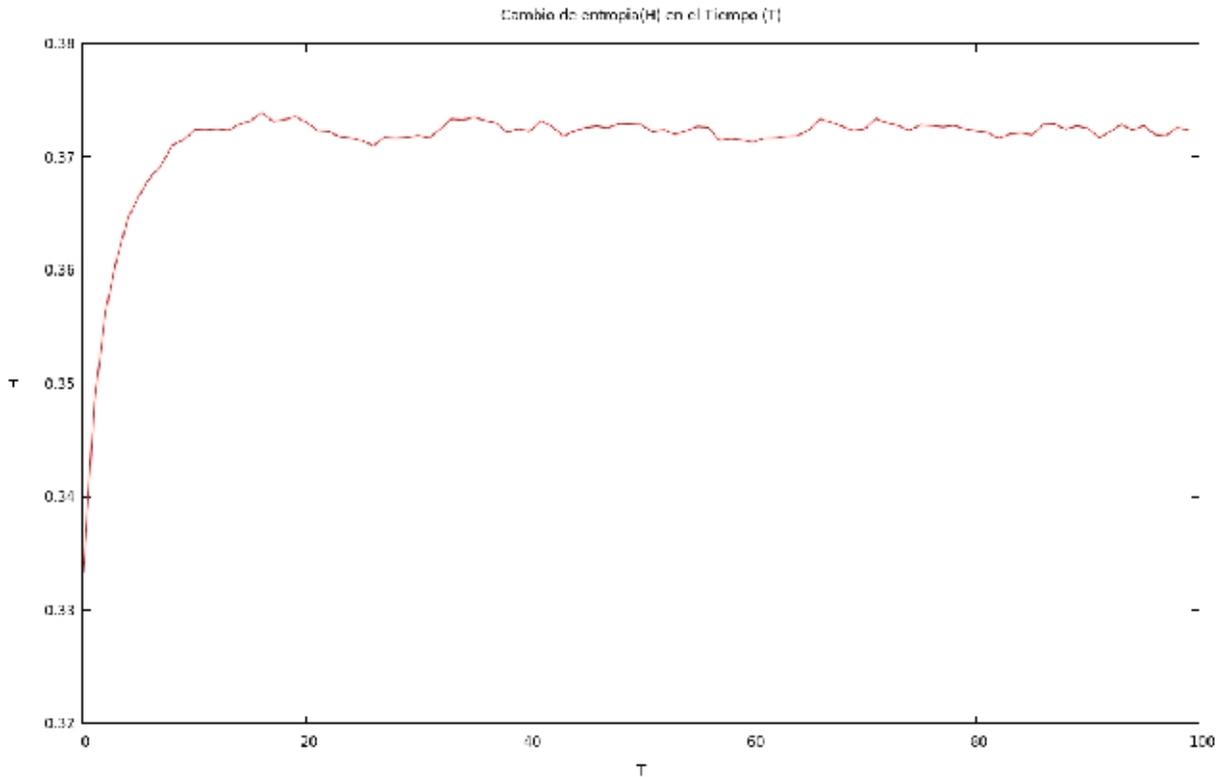


Figura 10: El incremento inicial de entropía durante esta prueba se debe a la creación de 6 combinaciones que no existían en la primera generación. Después de que se han creado todas las posibles combinaciones (8) se observan oscilaciones en la curva debido a los cambios en la proporción de cada color.

La exploración del espacio de soluciones es coherente con las observaciones realizadas sobre la entropía. Al 'muestrear' los píxeles $(0,0)$, $(256,256)$, $(468,362)$ se observa que los valores más frecuentes son el blanco y el negro (Figura 11). Los posibles valores finales que pudieron haber tomado los píxeles en esta ejecución del programa son los mismos que valores que existen al final de la prueba

dos. A diferencia de la prueba dos en este caso vemos que se toman preferentemente los valores blanco y negro. Esto nos lleva a pensar que aunque los valores de las demás esquinas son posibles su adecuación no es equivalente a la de los valores $[0,0,0]$ y $[256,256,256]$

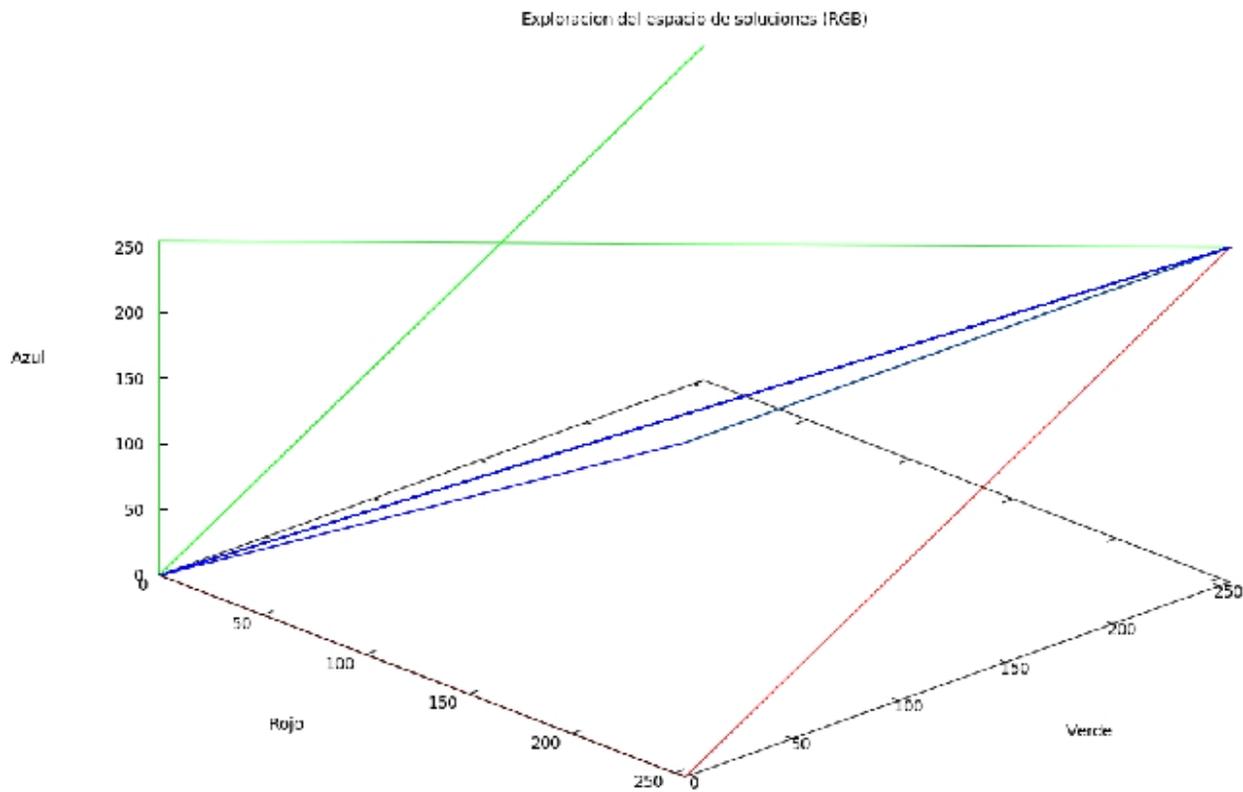


Figura 11: Los valores más tomados por los pixeles $(0,0)$, $(256,256)$ y $(468,362)$ son el blanco y el negro. Rara vez se desvían de estos puntos los valores para tomar el resto de las esquinas del cubo ya que estas son de menor adecuación y por lo tanto menos frecuentes en la población.

3.4 Cobertura de soluciones inicial reducida a dos puntos distantes y selección por máxima distancia euclidiana.

Al repetir la prueba anterior alterando únicamente el criterio de selección vemos que la entropía sube de manera drástica entre la primera generación ($H_0 = 0.33333$) y la tercera generación en la cuál

alcanza un alto valor ($H_2 \approx 0.99228$) cercano al valor final ($H_{99} \approx 1$) (figura 12). Dicho aumento repentino en entropía muestra que en muy pocas generaciones se explora la totalidad del espacio de soluciones (en este caso compuesto de ocho colores) y que cada uno de los genotipos se encuentra en proporciones casi iguales en la población.

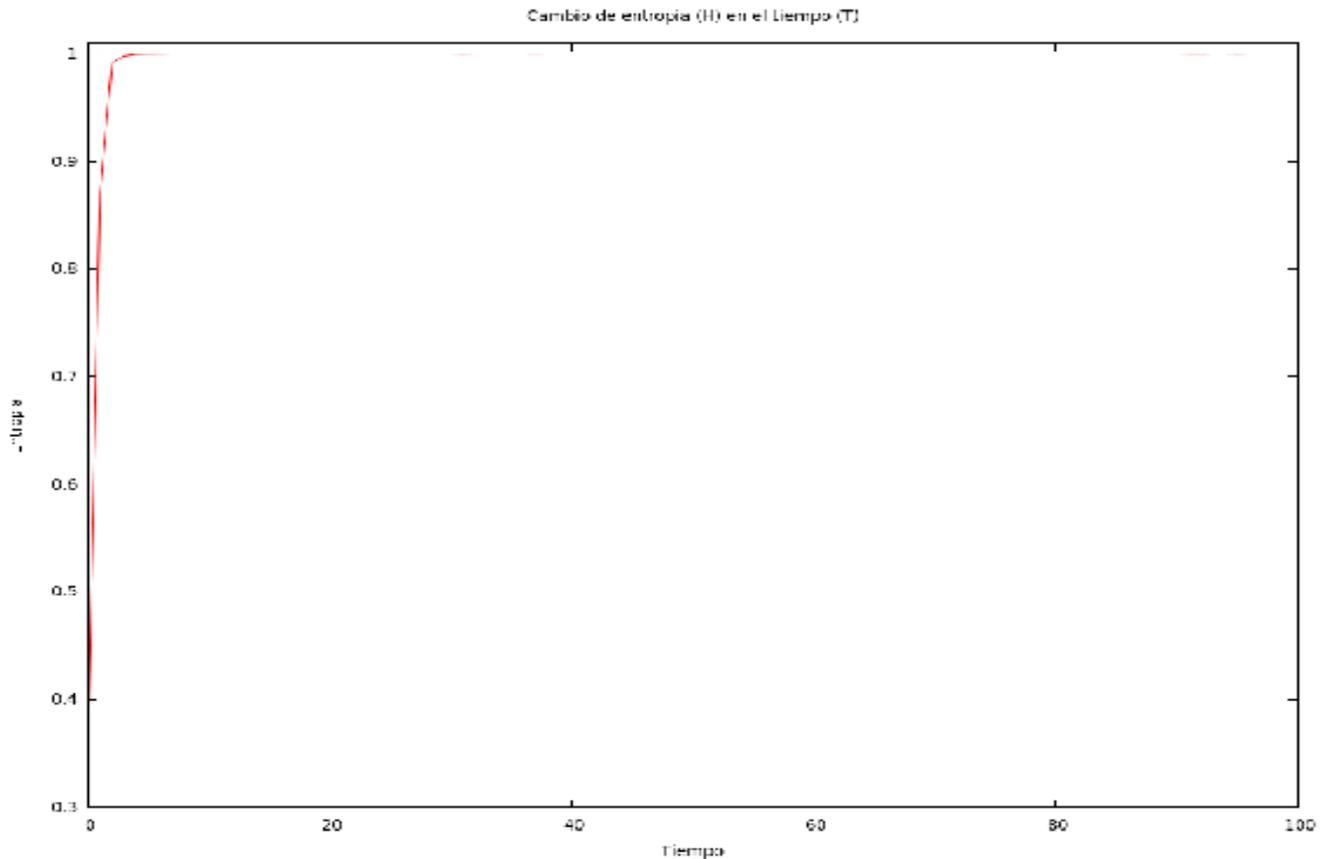


Figura 12: El cambio de diversidad es drástico en este escenario alcanzando valores cercanos a uno en 3 generaciones. Una vez que se ha alcanzado un máximo el sistema se mantiene lo cual sugiere que se conservan proporciones iguales de cada genotipo.

Las trayectorias de exploración de las casillas (0,0), (256,256) y (468,362) (figura 13) muestran la presencia de ocho atractores correspondientes con el blanco, el negro, los tres colores primarios de la luz y los tres colores secundarios. Es decir que los atractores se encuentran en los puntos del espacio de soluciones de mayor distancia entre ellos. Cabe notar que las líneas oscilan entre los ocho atractores sin mostrar preferencia por ninguno de ellos lo cual sugiere atractores de adecuación equivalente.

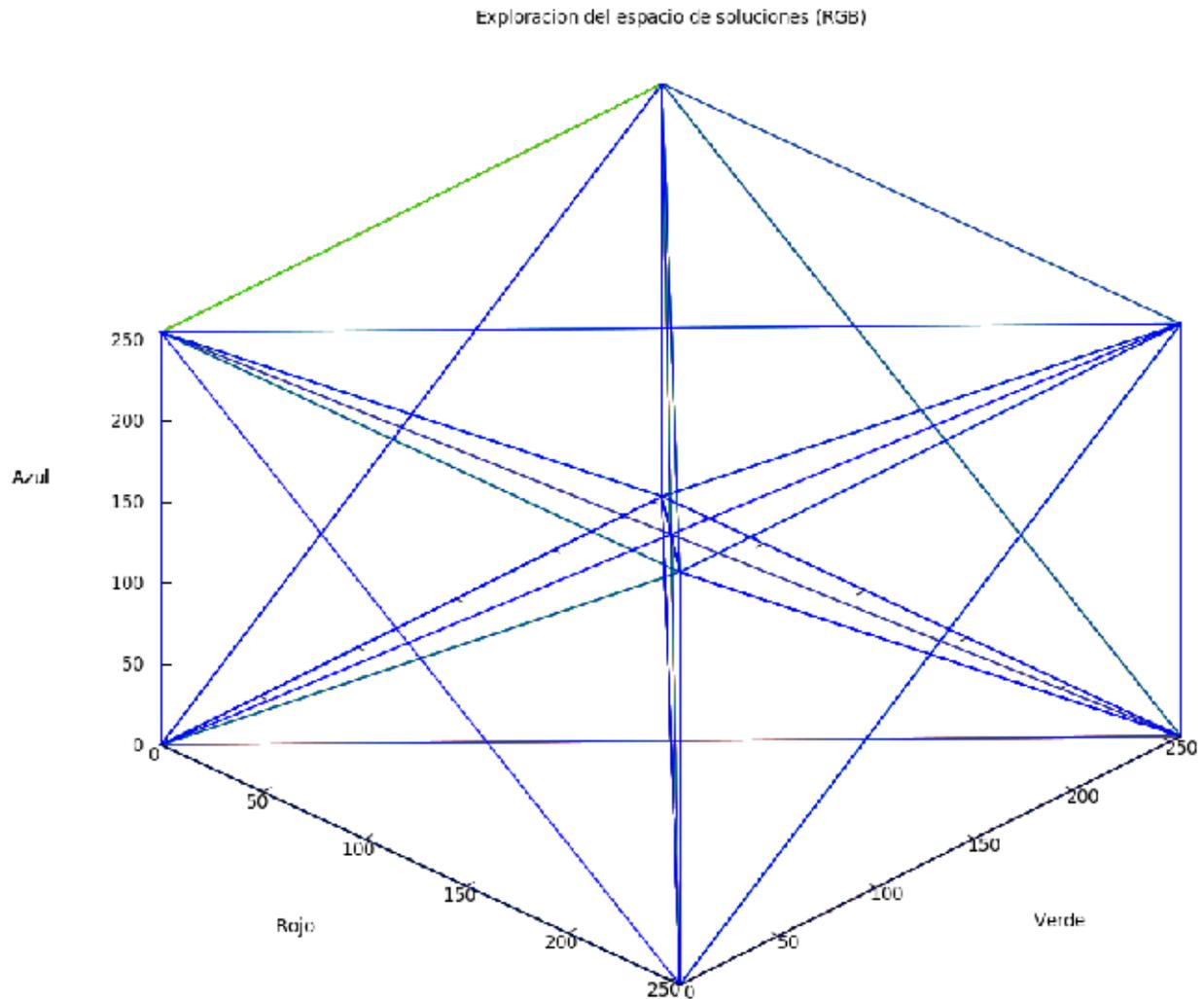


Figura 13: La trayectoria de exploración de las casillas $(0,0)$, $(256,256)$, $(468, 362)$ tiende a ir hacia todas las esquinas del cubo a diferencia de la prueba anterior en la cual había preferencia por dos esquinas. El hecho de que se en el gráfico sólo se observe la línea azul es un artefacto del software de graficado que se debe a que todas las líneas tienen trayectorias similares.

3.5 Cobertura de soluciones inicial irregular, incompletamente explorada y selección por mínima distancia euclidiana

Cuando la cobertura inicial del espacio de exploración es incompleta e irregular observamos

que la entropía inicial ($H \approx 0.50161$) sube durante un periodo de tiempo hasta llegar a un máximo (Figura 14). Este incremento en entropía se debe a que se forman nuevas combinaciones mediante la reproducción y tiene como consecuencia la exploración de partes del espacio que no se encontraban cubiertas al inicio. Dicha cobertura llega a un máximo de entropía ($H \approx 0.97305$) después del cual decae hasta llegar a un valor final de $H \approx 0.00059$ lo cual indica riqueza muy baja y dominancia de un genotipo en la población.

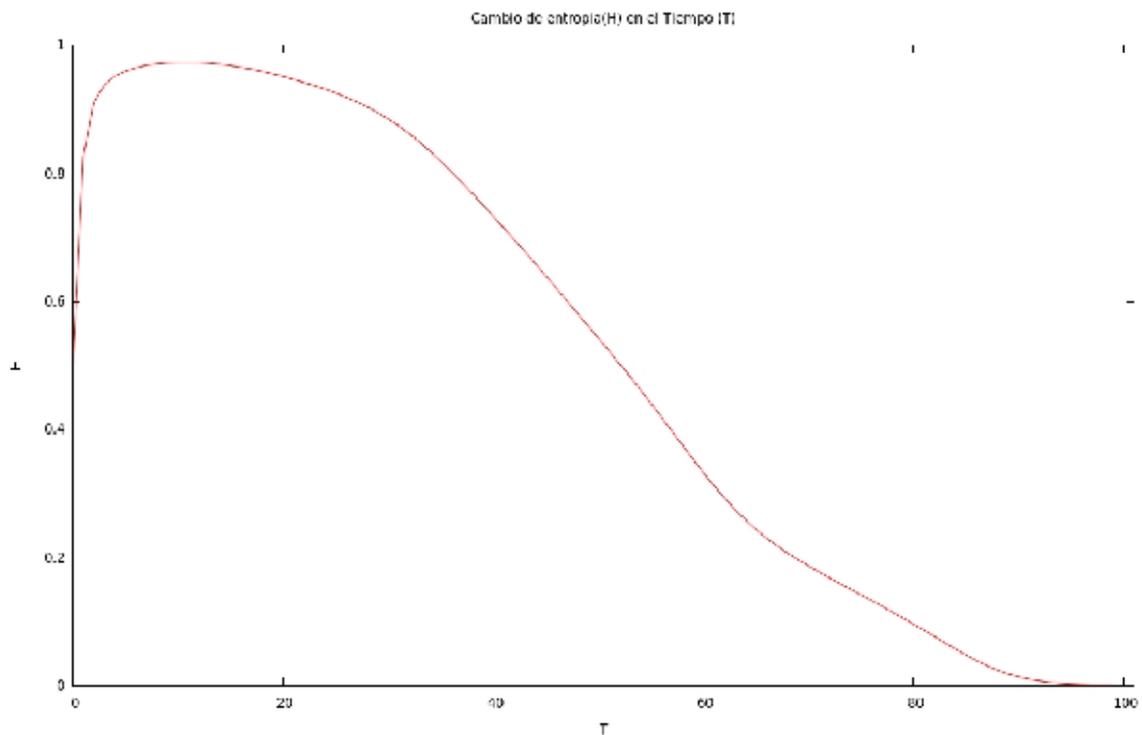


Figura 14: En las primeras generaciones hay un aumento de entropía como resultado de la creación de combinaciones que no existían en la generación anterior. Una vez que se llega a un máximo la entropía decae como en los demás casos.

En esta prueba observamos que los valores tomados por las coordenadas seleccionadas abarcan puntos del espacio que no se encontraban en la cobertura original. Después de incrementar la cobertura los valores finalmente tienden a un atractor fijo cercano a las coordenadas [197,189,174] (Figura 15). Si se comparan la cobertura inicial, la cobertura máxima y la cobertura final se observa que hay sólo 19

puntos en la cobertura final, mientras que había 2,210 al principio y 177,204 en la cobertura máxima. De los 19 colores finales la moda es el color [197,189,174] que corresponde a la posición del atractor. Es decir que dicho genotipo junto con su color asociado es el que permanece en más alta frecuencia al final de la simulación por lo cuál podemos suponer que es el de más alta adecuación.

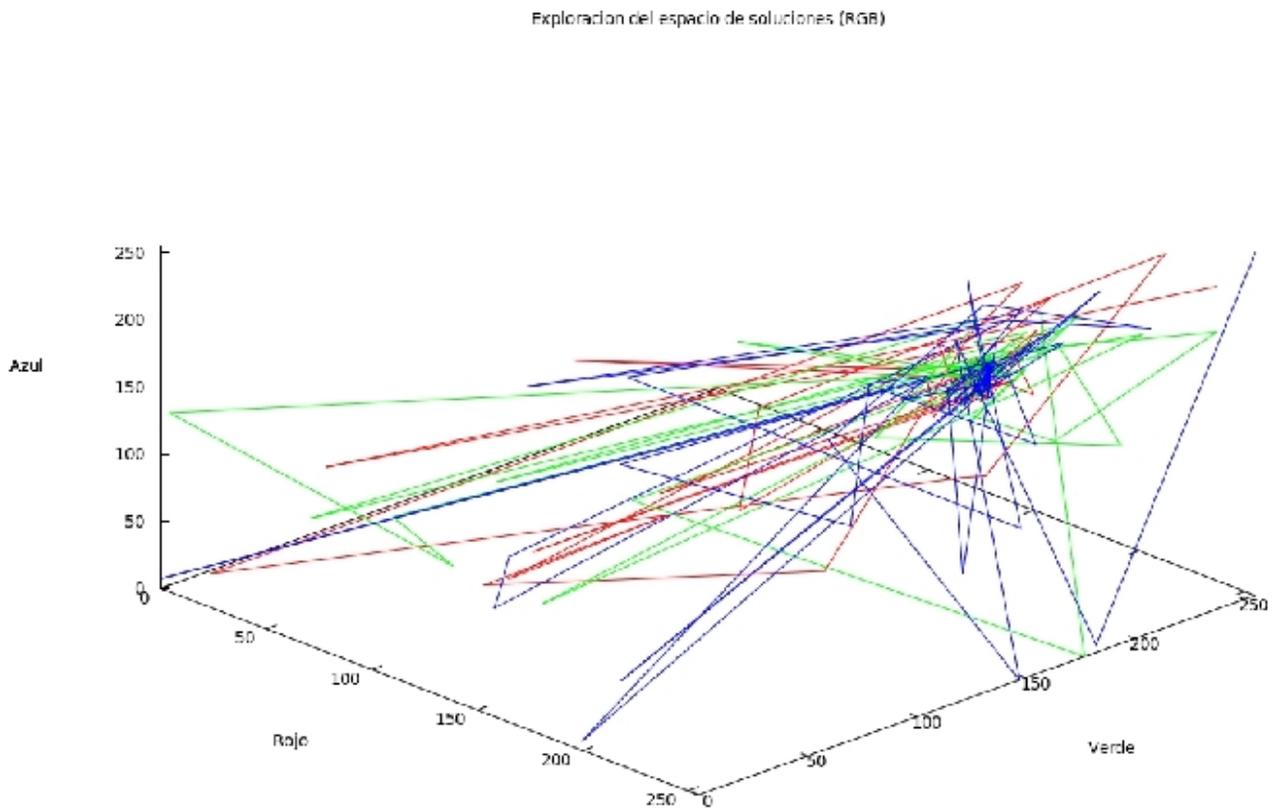


Figura 15: Las trayectorias de exploración para los puntos (0,0), (256,256) y (476,265) ocupan lugares que no se encontraban en la cobertura inicial del espacio de soluciones. En las últimas generaciones los puntos tienden hacia el valor [97,189,174].

3.6 Cobertura de soluciones inicial irregular, incompletamente explorada y selección por máxima distancia euclidiana.

En este caso se presenta la misma condición inicial y por lo tanto la misma entropía inicial ($H \approx 0.50161$) que en la prueba cinco. En este caso también vemos que la entropía (figura 16) aumenta hasta llegar a un máximo ($H \approx 0.88625$) y luego decae hasta establecerse en un mínimo ($H \approx 0.18000$). Dichos cambios de entropía se pueden ver en relación a los cambios de riqueza. La riqueza inicial es de 2,210 genotipos, la riqueza máxima es de 113,454 genotipos y la riqueza final contiene 185 genotipos diferentes.

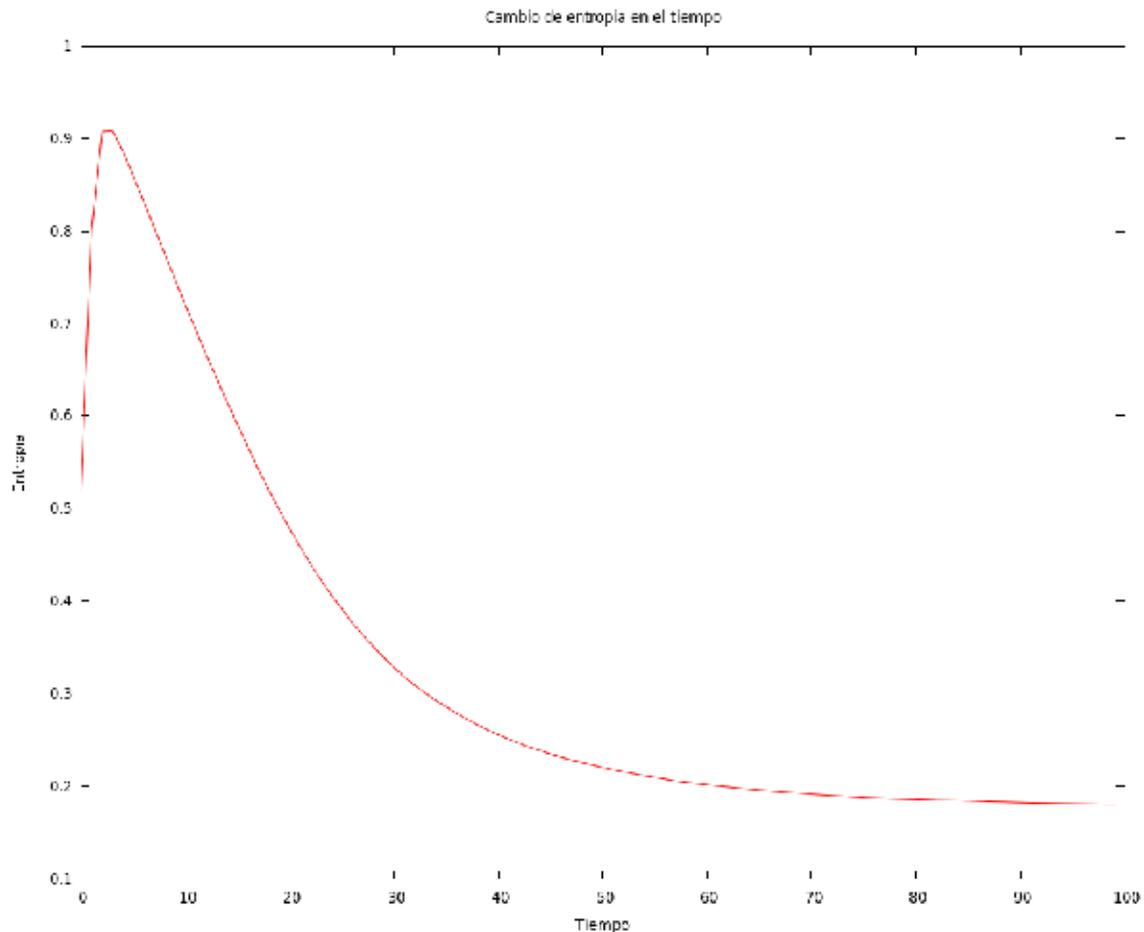


Figura 16: Al igual que en el escenario anterior se observa un aumento de diversidad como resultado de un incremento en la riqueza producto de la selección sexual. Después de llegar a un máximo la entropía decae como resultado de la selección.

Durante la exploración del espacio de soluciones (figura 17) observamos que los valores tomados por las coordenadas seleccionadas toman valores que no se encontraban en la condición inicial. Después de incrementar la diversidad los valores se acomodan en forma de una caja para finalmente tomar valores cercanos a ocho atractores correspondientes a las esquinas. El hecho de que exista más de un atractor es coherente con que la entropía no tienda a cero.

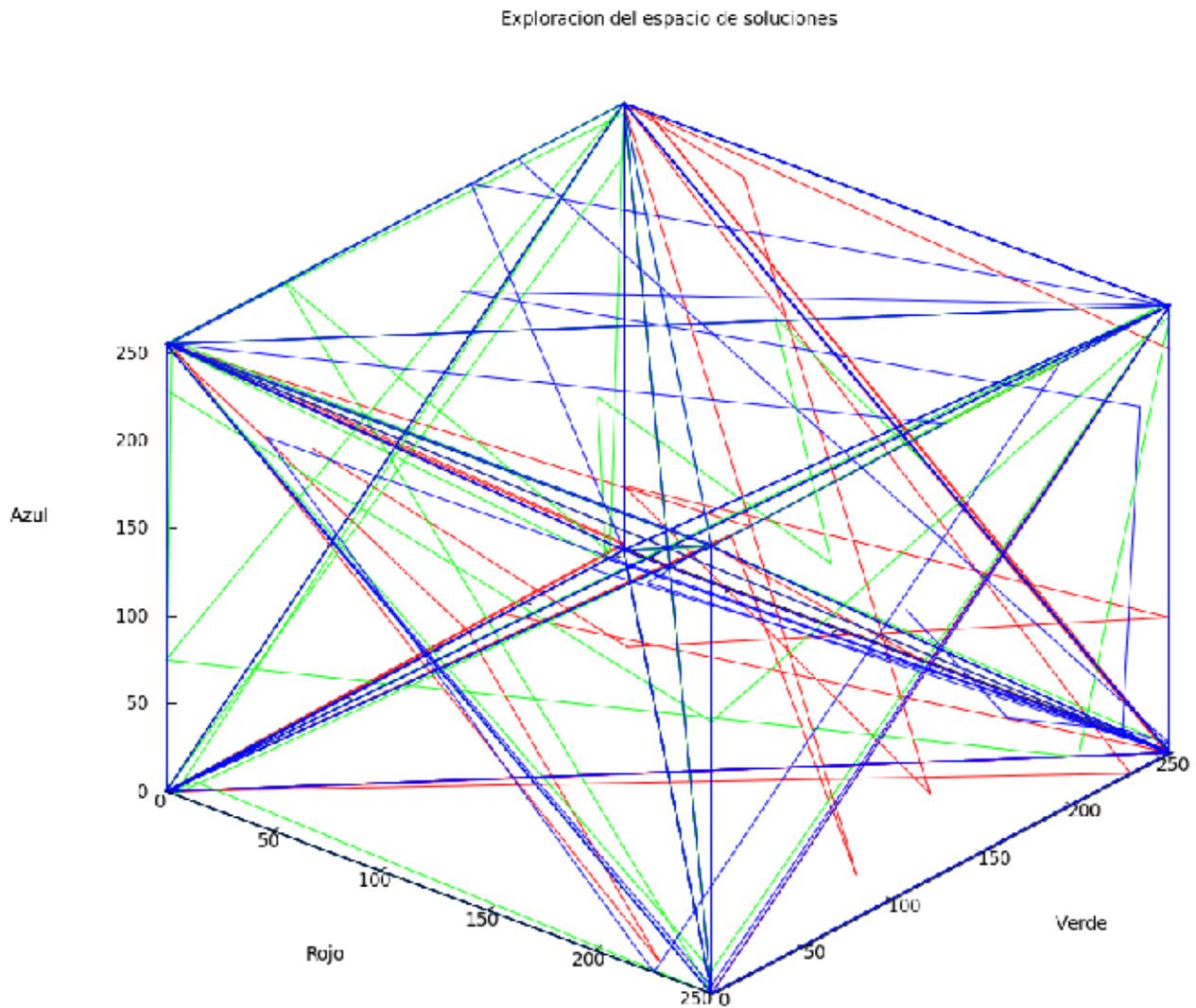


Figura 17: Los valores tomados por las casillas (0,0), (256,256) y (476,265) comienzan dentro de la cobertura inicial del espacio de soluciones para eventualmente crear puntos nuevos que se acomodan en forma de caja . Al final los valores tienden hacia las esquinas indicando la presencia de ocho atractores.

Capítulo 4 Discusión

4.1 La influencia del criterio de selección

El criterio de selección tiene efectos importantes en la diversidad final, en el número de atractores (genotipos de alta adecuación) y en la posición de los mismos en el espacio de soluciones.

Observemos primero los resultados de las simulaciones uno, tres y cinco en las cuales el criterio de selección fue por mínima distancia euclidiana. Si observamos cómo es la exploración del espacio de soluciones vemos que en las pruebas uno y cinco el sistema tiende a un sólo valor. Esto se debe a que bajo el criterio de selección de mínima distancia euclidiana existe sólo un genotipo de alta adecuación en el espacio de soluciones. Debido a que el criterio de selección es por mínima distancia euclidiana la posición del atractor se encuentra en el centroide de las zonas del espacio de soluciones con mayor densidad de puntos, es decir que el genotipo con más alta adecuación se coloca en el punto donde se encuentra la mínima distancia euclidiana promedio.

En la prueba tres encontramos que existe más de un color con adecuación alta. A pesar de que nos encontramos en un sistema que busca la menor distancia euclidiana en este caso encontramos que existen dos puntos con mínima distancia euclidiana promedio por lo que dichos puntos son de adecuación equivalente. En este escenario encontramos también que el sistema a veces se aleja de los puntos de más alta adecuación y toma de manera aleatoria valores lejanos a la máxima adecuación, estas oscilaciones aleatorias del sistema se deben a la reproducción con individuos de menor adecuación como resultado de la distribución aleatoria en el espacio.

Veamos los resultados de las pruebas dos, cuatro y seis en las cuales el criterio de selección fue de máxima distancia euclidiana. Lo primero que resalta en las tres pruebas es la presencia de ocho atractores. Si pensamos que el sistema ahora busca los puntos de máxima distancia euclidiana queda claro que las esquinas de espacios de soluciones cúbicos son el punto de mayor distancia euclidiana promedio y por lo tanto donde se encuentran los genotipos de más alta adecuación. Dado que la distancia euclidiana entre una esquina y su valor opuesto es la misma para todas las esquinas se puede decir que bajo el criterio de máxima distancia euclidiana todas las esquinas son de adecuación equivalente.

El número de genotipos de alta adecuación está directamente relacionado con los efectos que tiene el criterio de selección en la diversidad. En los sistemas en los cuales existe más de un atractor la diversidad final tiende a ser mayor mientras que los sistemas con un sólo atractor tienden a perder diversidad. Esto es un caso claro de pérdida de diversidad como resultado de una pérdida de riqueza. En los casos en los que existe sólo un genotipo al final de la simulación el sistema tiende a eliminar todos los genotipos lejanos al de más alta adecuación por lo cual la riqueza tiende a uno, en la entropía de Shannon un sistema con un único estado tiene una entropía de cero, es decir que la diversidad es mínima. En contraste, los sistemas con más de un genotipo en la última generación tienden a mantenerse con entropía diferente de cero.

Se ha observado en la naturaleza que un juego de alelos bajo selección tenderán a perder diversidad ya que los alelos seleccionados positivamente tenderán a aumentar de frecuencia en la población hasta desplazar a los de menor adecuación (Templeton 2006). Este fenómeno puede dar como resultado selección balanceadora (coexisten varios genotipos en proporciones similares) o selección direccional (domina un genotipo sobre todos los demás). Dichas observaciones son

congruentes con los resultados de este trabajo.

Tomando los resultados de las pruebas uno y cinco vemos que -dado el tiempo suficiente- la diversidad se reduce a individuos muy parecidos entre sí por lo que se generan altas frecuencias de alelos similares, es decir que hay selección direccional. En la prueba tres a pesar de tener un criterio de selección por mínima distancia euclidiana tenemos ocho genotipos de adecuación diferente de los cuales dos representan la máxima adecuación por lo que el sistema tiende a estos dos genotipos. Debido a que existe una tendencia a dos valores podemos pensar que la simulación tres se encuentra bajo selección balanceadora.

En las pruebas con criterio de selección por máxima distancia euclidiana tenemos la existencia de ocho atractores de adecuación equivalente por lo cual los individuos del sistema tienden a tomar los genotipos correspondientes a dichos atractores. Por esta razón podemos afirmar que se encuentran bajo selección balanceadora.

En términos biológicos podemos decir que el criterio de selección tiene efecto sobre la adecuación de los genotipos. Una observación interesante sobre el efecto del criterio de selección en la adecuación es que determina si la selección dependiente de la frecuencia es positiva o negativa. En los casos donde se utilizó el criterio de mínima distancia euclidiana sería favorable para un genotipo que existieran varias copias de él mismo, ya que en la siguiente generación buscará ser seleccionado por individuos similares o iguales a él. En el caso de utilizar máxima distancia euclidiana como criterio de selección vemos que un genotipo se vería afectado por un aumento en frecuencia de genotipos similares a él por lo cual ser poco frecuente en la población aumenta la adecuación de los genotipos. Podemos decir entonces que si el criterio de selección es por mínima distancia euclidiana la selección

dependiente de la frecuencia será positiva y que si el criterio de selección es por máxima distancia euclidiana la selección dependiente de la frecuencia será negativa.

Queda claro que el criterio de selección juega un papel clave en determinar el genotipo hacia el cual tiende el sistema, su posición en el espacio de soluciones y la diversidad final, sin embargo el criterio de selección por si sólo no da una explicación completa sobre el número de atractores, su posición final y la diversidad que se deriva de ellos por lo cual es importante revisar el efecto de la cobertura inicial del espacio de soluciones.

4.2 El efecto de la diversidad inicial

La diversidad inicial determina qué valores puede tomar el sistema. Debido a que no existe la mutación los alelos que existen en la primera generación son todos los que pueden existir en cualquier tiempo. Por esta razón la diversidad inicial determina la forma del espacio de soluciones, es decir que si generamos todas las posibles combinaciones con los alelos presentes en la diversidad inicial obtendríamos un espacio de soluciones con una geometría específica para los alelos existentes. Cualquier cobertura creada después sería un subconjunto de la cobertura creada por todas las posibles soluciones.

La forma que toma el espacio de soluciones dada la diversidad inicial tiene un efecto en el número de atractores que existen y en su posición. Esto queda particularmente claro viendo la prueba uno, en la que tenemos una cobertura de soluciones simétrica y regular. Al escoger al individuo con menor distancia euclidiana de manera local vemos que el sistema tiende hacia los individuos cercanos al centro (el punto de menor distancia euclidiana promedio).

Al observar la prueba cinco vemos que en espacios de soluciones asimétricos e irregulares el punto de adecuación más alta es un punto diferente del centro. La posición del genotipo de más alta adecuación coincide en gran parte con una región de alta densidad de puntos en la diversidad inicial. Al haber más individuos de esta zona lo más probable es que lo más parecido a la mayoría se encuentre en esta región. Es relevante resaltar que los resultados finales en la prueba cinco se derivan de la selección sobre la diversidad máxima y no sobre la diversidad inicial, sin embargo la región de la cobertura inicial con una mayor densidad de puntos se explorará de manera más exhaustiva en la cobertura máxima de manera que las zonas de alta densidad se mantienen más o menos constantes. Al incrementar la densidad de puntos en una región aumenta la probabilidad de que los individuos más parecidos entre sí se encuentren en esta parte del sistema. La interpretación anterior de los resultados es muy similar al fenómeno biológico conocido como selección positiva dependiente de la frecuencia, en este fenómeno la adecuación de un genotipo aumenta mientras más alta sea su frecuencia en la población (Cockerham 1972).

Entendiendo los resultados de la prueba cinco podemos pensar que en la prueba seis existe un fenómeno opuesto, es decir que existe selección negativa dependiente de frecuencia por lo que la adecuación de un genotipo aumenta mientras menor sea su proporción en la adecuación (Gigord 2001). Esto es coherente con el criterio de selección utilizado durante esta prueba. El hecho de que la cobertura inicial del espacio de soluciones fuera irregular no tiene ningún efecto en la posición de los genotipos de alta adecuación ya que en este caso con que existan las esquinas en el momento de más alta riqueza el sistema tenderá a ellas en proporciones iguales como resultado de que la existencia de un color en más altas proporciones reduciría su adecuación.

El número de atractores está en gran parte determinado por la geometría del espacio de

soluciones y por la frecuencia inicial de cada genotipo. Para clarificar la afirmación anterior tomemos como ejemplo la simulación tres. En este escenario vemos que en la condición inicial existen en proporciones iguales dos genotipos (blanco y negro) distantes en el espacio de soluciones. Dado que en el espacio de soluciones no existe un punto que los dos genotipos encuentren igualmente atractivo (por ejemplo un gris) y que los genotipos se encuentran en igual proporción los individuos tienden a reproducirse con los que son idénticos a ellos mismos permitiendo la existencia de dos genotipos de alta adecuación.

Si tomamos ahora como ejemplo las pruebas dos, cuatro y seis podemos ver que los tres escenarios tienen los mismos atractores a pesar de tener espacios de soluciones diferentes. Queda claro que para el criterio de selección de máxima distancia euclidiana los puntos de más alta adecuación son las esquinas del cubo RGB. Por esta razón cualquier espacio de soluciones que incluya las 8 esquinas del cubo RGB y se seleccione bajo máxima distancia euclidiana tendrá ocho atractores.

La diversidad inicial también tiene un efecto en la forma de explorar el espacio de soluciones y por lo tanto en la diversidad subsecuente. Si analizamos las pruebas uno y dos, en las cuales la diversidad inicial corresponde con la máxima diversidad posible, observamos un rápido decaimiento de la diversidad entre las generaciones uno y dos. Este decaimiento se debe a que cualquier combinación creada en el primer evento reproductivo existía en la generación anterior por lo cual empieza a haber varias copias de elementos iguales desde la generación uno, lo cuál resulta en una pérdida de diversidad como resultado de una pérdida de riqueza.

En las pruebas tres, cuatro, cinco y seis existe un fenómeno opuesto. Al tener una diversidad inicial que no corresponde con la máxima diversidad posible dados los alelos existentes, vemos que el

sistema comienza a crear algunas de las combinaciones que le faltan por medio de la reproducción con recombinación. Al llegar a un máximo y como efecto de la selección comienzan a crearse combinaciones repetidas y la diversidad decae de manera similar a las demás pruebas. Por esta razón la posición final del atractor está en función de la diversidad máxima y no de la diversidad inicial.

La prueba tres y cuatro presenta un comportamiento interesante en cuestión de diversidad. La riqueza inicial del sistema es de dos combinaciones mientras que la máxima posible es de ocho, en la segunda generación se crean todas las posibles combinaciones lo cual resulta en un incremento de la diversidad como función de un incremento en la riqueza. En la prueba tres, después de la segunda generación los valores de entropía oscilan de manera cercana a $H_{96}=0.37199$, las oscilaciones aleatorias se deben a los cambios de frecuencia de cada genotipo resultantes de la distribución aleatoria. En el cuarto escenario también se generan todos los posibles genomas en pocas generaciones, sin embargo dado que aquí los ocho atractores son de adecuación equivalente y que el sistema se encuentra bajo selección negativa dependiente de la frecuencia tienden a existir los ocho genotipos en proporciones iguales por lo que la entropía se mantiene en un máximo cercano a uno.

En términos biológicos el espacio de soluciones se puede ver como un *pool* genético ya que contiene todas las posibles combinaciones de genes y alelos existentes en una población (Templeton 2006). Evidentemente un organismo no puede dejar descendencia con alelos no existieran en la población en un tiempo anterior. Para visualizar la relación entre el *pool* genético y los genomas de alta adecuación basta con imaginarnos que todos los genomas del *pool* genético existen en organismos vivos al mismo tiempo (como ocurre en las pruebas 1 y 2), entonces veríamos que no todos los organismos tienen la misma adecuación por lo que en la siguiente generación algunos genotipos se han eliminado y otros han aumentado de frecuencia. Si se repite este fenómeno varias veces veríamos que

los organismos comienzan a presentar sólo algunos de los alelos existentes en la población inicial por lo cual podemos suponer que los genotipos que se presentan en una generación dada son de igual o más alta adecuación que los de la primera generación.

4.3 Existencia de patrones

Al observar los resultados visuales de las pruebas no parece haber formación de patrones espaciales (Apéndice A). Probablemente el espacio no se regionaliza en máximos locales de adecuación como causa de que los individuos creados pueden tomar cualquier punto en el espacio al azar. Sin embargo sí existe un patrón temporal conservado durante las pruebas. Este patrón parece consistir en encontrar un máximo de diversidad después del cual la diversidad decae hasta mantenerse estable en un mínimo. Las pruebas tres y cuatro parece violar ligeramente este patrón, en este caso la entropía aumenta hasta llegar a una estabilidad oscilatoria o se mantiene en un máximo. El hecho de que la entropía llegue a un máximo donde se mantiene sin decaer no viola las premisas dadas para el patrón temporal encontrado, ya que en este caso el máximo de diversidad y el límite mínimo al que la entropía tiende se encuentran cerca del mismo valor.

De manera visual también es discernible un patrón temporal, al observar de manera cronológica las imágenes creadas en cada generación podemos ver que el espacio comienza muy heterogéneo y que conforme el tiempo avanza se va volviendo más homogéneo y tiende a dar colores lisos o combinaciones de pocos colores (ver apéndice A).

4.4 Significancia biológica

Al simular o modelar procesos biológicos de manera computarizada siempre queda la duda de

la fidelidad con la cual se está representando el comportamiento real de un fenómeno. La aproximación más certera para resolver esta duda es hacer experimentos que cumplan las reglas de las simulaciones y comparar el resultado *in vivo* con el resultado *in silico*. En trabajos como este hacer experimentos *in vivo* que permitan resolver las mismas preguntas tomaría mucho tiempo para generar cambios en las frecuencias alélicas por lo cuál se dificulta su realización, aunque sería bueno intentarlo. Como causa de la dificultad técnica de la experimentación esta sección tiene una aproximación más teórica para el problema de la significancia biológica de los resultados del trabajo.

Lo primero que vale la pena discutir es que la computación es un proceso de manejo de información independiente del sustrato en el que se realiza (Bostrom 2003). Esto quiere decir que la computación de ciertos problemas se puede realizar con la misma validez en procesadores de transistores, en bulbos, tarjetas perforadas, engranes, reacciones químicas y bacterias transgénicas (Haynes 2007). Bajo la suposición de la independencia del sustrato podemos pensar que los procesos vivos son procesos de cómputo sin importar el nivel de resolución (desde el acoplamiento de nucleótidos para formar ADN hasta los patrones globales de diversidad) por lo cual la vida artificial puede resolver los mismos problemas computables que la vida natural.

La visión de la vida como procesos de computo ha llevado a la idea de que la vida es un proceso algorítmico regulado por una serie de reglas sencillas aplicadas de manera local durante tiempos muy largos, esto no difiere mucho de un autómatas celular (Mitchel 2009, Wolfram 2002, Dennett 1995). La visión algorítmica de la vida ha recibido ciertas críticas de científicos importantes (Gould 1997). El origen de estas críticas es mezclar la idea de vida algorítmica con la de selección natural algorítmica. Mientras que nadie niega la algoritmidad de la selección natural se ha argumentado -con mucha razón- que la evolución es mucho más que selección natural (Futuyma 2005). Sin embargo esto no

hace que muchos otros fenómenos biológicos no puedan ser descritos como una secuencia de reglas lógicas a manera de un algoritmo.

Otra idea que ha hecho pensar que la vida no es algorítmica es que está llena de contingencias, las contingencias las podemos definir como fenómenos aleatorios y momentos de 'suerte' en la historia de los organismos. Esta también es una concepción errónea ya que es posible estimar la probabilidad de muchos eventos contingentes e integrar la probabilidad de que ocurran en un algoritmo (Denett 1995). Dentro de esta discusión los resultados de este trabajo y cualquier otro trabajo de simulación de procesos biológicos parecen tener una relevancia equiparable al método experimental.

Quedando claro el marco teórico de la validez del cómputo como modelo vale la pena comparar los resultados *in silico* con trabajos anteriores que abordan problemas de cambios en frecuencias alélicas y genéticas en poblaciones bajo selección sexual. Probablemente el trabajo que más se parece al modelo realizado es el de Marshall *et al.* de 2003. En este trabajo se estudia la relación entre la diversidad genética individual y el tamaño del repertorio de canto (un carácter seleccionado sexualmente) en el ave *Arocephalus schoenobaenus*. En dicho artículo se menciona que los repertorios largos se correlacionan de manera positiva con la distancia genética promedio. Los repertorios grandes son preferidos por las hembras ya que garantiza crías con alta variabilidad genética. Este trabajo expone un criterio de selección similar al que se utilizó en las simulaciones en las cuales un carácter fenotípico se utiliza como un indicador de distancia genética entre individuos. En los resultados del trabajo de Marshall *et al* también mencionan que las hembras tienden a escoger a los machos que se encuentran a mayor distancia genética, esto sería análogo a tener un criterio de selección de máxima distancia euclidiana en el modelo de cómputo.

El trabajo de 2003 realizado por Medel *et al* también muestra cierta similitud con algunas de las simulaciones realizadas. En este trabajo Medel y su equipo hacen un estudio sobre la preferencia de polinizadores por las diferentes guías de néctar existentes en la flor andina *Mimulus luteus*. Dicha especie se poliniza por cuatro insectos y un colibrí. En los resultados se muestra que los insectos prefieren flores con grandes tamaños de corola y de guía de néctar mientras que los colibríes prefieren corolas y guías un poco más pequeñas. Cabe aclarar que la cantidad de néctar y el contenido sacarino de las flores es independiente al tamaño de corola y a la guía de néctar. En dichos resultados vemos que las plantas con arquitecturas de guía de néctar y corolas pequeñas serán polinizadas preferentemente por colibríes mientras que las grandes se polinizarán por insectos. Podemos pensar entonces que el criterio de selección entre flores con arquitecturas similares es de mínima distancia euclidiana ya que flores similares son visitadas por los mismos polinizadores. Los resultados del trabajo de Medel *et al* son de cierta manera similares a los de la prueba tres, en la cual existen dos fenotipos que se seleccionan fuertemente por mínima distancia euclidiana en un entorno donde existe de manera positiva la selección dependiente de la frecuencia.

4.5 Entropía, evolución y trabajo

Existe la idea errónea de que la evolución está en conflicto con la segunda ley de la termodinámica ya que aparentemente reduce la entropía y aumenta la complejidad de los sistemas. Se han hecho diversas aproximaciones para demostrar que esta idea está equivocada, la mayoría de las publicaciones sobre el tema tienen un enfoque químico en el cual se habla del manejo energético de las reacciones biológicas (Brooks 1988, Avery 2003). En los resultados de la simulación presentada en este trabajo vemos que también hay un decremento de entropía con el tiempo y que el sistema tiende a la homogeneidad y el orden. Sin embargo la aproximación bioquímica no se puede utilizar para explicar el decaimiento de entropía en este trabajo ya que no existen reacciones químicas en el modelo.

Para poder discutir por qué no se está violando la segunda ley de la termodinámica es importante definir la entropía como una medida de la información. Uno de los trabajos clásicos sobre entropía, trabajo e información es el demonio propuesto por Maxwell en 1871; Maxwell proponía una caja dividida en dos secciones con una puerta sin peso conectando las partes, la puerta era operada por un demonio. Cuando el demonio veía venir una partícula rápida de izquierda a derecha la dejaba pasar pero cuando venía una partícula lenta le cerraba la puerta, la regla opuesta aplicaba de derecha a izquierda. Dado el tiempo suficiente el demonio había acomodado partículas rápidas en la derecha y lentas en la izquierda, creando así un desequilibrio térmico sin aparentemente haber hecho trabajo por lo tanto violando la segunda ley de la termodinámica.

Para poder decidir si dejaba pasar o no una partícula el demonio requería información sobre el estado de la partícula -se mueve rápido o lento- por lo que al final del 'experimento' sabríamos con certeza de qué lado hay partículas en cada uno de los estados. Si después de que se han acomodado las partículas se eliminara la división se mezclarían las partículas rápidas y las lentas y ya no tendríamos información certera de dónde encontrar una partícula de cada tipo, por lo cual un incremento en entropía resulta en un decremento de información, el mínimo de información se encuentra cuando el sistema está en equilibrio (Avery 2003).

Para conocer el microestado de las partículas cuando el sistema está en equilibrio habría que preguntarse partícula por partícula si es más lenta o rápida que el promedio, mientras que para conocer el estado de una partícula después de que el demonio ha operado sólo habría que preguntarse si la partícula esta en la izquierda o en la derecha de la caja para conocer su estado energético. Esta relación entre el número de preguntas necesarias para conocer un microestado y el estado general del sistema es lo que une a la entropía termodinámica con la entropía de la información ya que mientras más alta sea

la entropía termodinámica más preguntas binarias se requieren para describir el sistema y por lo tanto más entropía informática (Brillouin 1962).

Establecida la relación entre entropía e información podemos comenzar a analizar el trabajo en el demonio de Maxwell. Según Maxwell si la puerta no tenía peso el demonio no realizaba ningún trabajo. A principios del siglo XX Leo Szilard resolvió el problema del demonio de Maxwell y la contradicción con la segunda ley de la termodinámica diciendo que cuando el demonio decidía qué partículas dejar pasar tenía un requerimiento energético y que su decisión provocaba por lo menos la misma entropía como había eliminado al ordenar las partículas (Mitchell 2009). Esto nos lleva inevitablemente al principio de Landauer en el cual se establece que el trabajo mínimo por bit de información procesado (contestar rápido o lento en este caso) es de $kT \ln(2)$ joules; Donde k es la constante de Boltzmann y T es la temperatura absoluta del sistema (Landauer 1961).

Si pensamos en la vida vemos que todos los sistemas vivos están en desequilibrio termodinámico, por lo tanto se ha argumentado que la vida viola la segunda ley. Al igual que el demonio de Maxwell, la vida hace manejos de información que requieren de energía y trabajo creando una disminución local de entropía pero un incremento global. Bajo estas premisas se simplifica entender por qué el cambio evolutivo no viola la segunda ley, si bien la evolución incrementa el orden y la complejidad de los organismos lo hace a expensas del trabajo que cada individuo hace para alimentarse, reproducirse, evitar la depredación, tener respuestas inmunes y demás actividades biológicas que desempeñe (Avery 2003).

La pérdida de entropía observada en los resultados de las simulaciones se puede explicar de dos

maneras. La primera es una explicación estadística. Si se presenta un sistema muy diverso existe la posibilidad de que se reduzca el número de estados hasta estar en un equilibrio probabilístico en el cual el sistema se acomoda por sí solo en un estado diferente a la entropía máxima, causando un aumento de información. En el demonio de Maxwell esto sería equivalente a no tener el demonio y que las partículas se acomoden de maneras diferentes a la máxima entropía simplemente por una razón estadística. Tomando esta explicación es evidente que la probabilidad de que el sistema tienda a la homogeneidad (como en las pruebas uno y cinco) sin hacer trabajo es casi cero. Dado que observamos que el sistema tiende a la homogeneidad podemos afirmar que esta explicación es incorrecta.

El segundo enfoque es equivalente a la idea de Szilard sobre el demonio. Si analizamos el algoritmo bajo el cual funciona la simulación queda claro que cada casilla requiere cierta información para tomar sus decisiones. El primer manejo de información que hace es asignarse un color, el segundo es comparar su color con el de sus vecinos y determinar cuál es más parecido a ella, posteriormente se reproduce con recombinación y finalmente se le asigna una posición en el espacio a su hijo. Cada una de estas operaciones realiza trabajo y genera por lo menos la misma entropía en los circuitos de la computadora y en el aire circundante como la que elimina dentro de la simulación por lo que la entropía del sistema computadora-aire-simulación debe de permanecer igual o incrementar.

Dadas las explicaciones anteriores y los resultados de este estudio podemos suponer que la selección de pareja y la reproducción sexual son procesos de manejo informático que requieren trabajo y que por lo tanto no violan la segunda ley de la termodinámica. Cabe aclarar que estos dos fenómenos son sólo algunos de los componentes de la evolución por lo cual sólo con los resultados de este trabajo es imposible afirmar que toda la evolución se comporta así.

Capítulo 5 Conclusiones

5.1 El efecto de la selección sexual en la diversidad

De acuerdo con los resultados obtenidos durante este trabajo se puede decir que la selección sexual tiene un efecto reductivo en la diversidad. El hecho de que cualquier tipo de selección reduce la diversidad esta aceptado en la biología (Maynard Smith 1999), la aportación de estas simulaciones es que a partir de reglas reduccionistas y sencillas se llega a un comportamiento similar al esperado en la naturaleza.

Es importante tener claro que lo que observamos como selección sexual en realidad es el resultado de las reglas que determinan la reproducción diferencial de los individuos (a las cuales se les ha llamado criterio de selección a lo largo de este trabajo). Con esto claro se puede comenzar a discutir el aparente efecto orientador que tiene la selección sexual.

Al observar el efecto de la selección en los valores alélicos finales de los individuos pareciera que la selección ha orientado a la diversidad hacia regiones específicas del espacio de soluciones. En realidad el efecto reductor de la selección sexual sobre la diversidad es también un resultado del efecto orientador, esta afirmación se puede hacer porque al restringir a la diversidad a las regiones de alta adecuación del espacio de soluciones se eliminan de la cobertura los individuos que no se encuentren en las regiones de baja adecuación causando una pérdida de riqueza.

Sabiendo que la selección tiene un efecto orientador en la diversidad podemos pensar que la diversidad alélica inicial le marca los límites hacia los cuales puede navegar la diversidad en el tiempo. El criterio de selección simplemente establece en que posición de los límites codificados por la

diversidad inicial se encontrará la diversidad del sistema en el tiempo siguiente. Por esta razón el efecto de la diversidad inicial y de la selección no se pueden separar el uno de otro. Cabe notar que en estas simulaciones una vez que se elimina un alelo de la diversidad se reduce el espacio de soluciones por lo cual no es posible reconstruir el espacio de soluciones inicial ni la diversidad alélica correspondiente. Este fenómeno junto con la mutación ocurre de igual manera en los sistemas vivos por lo cual la evolución es irreversible (Bridgham 2009).

5.2 Biología *in silico*

A lo largo de este trabajo se ha discutido de manera recurrente si es valido reemplazar observaciones de campo y métodos de laboratorio por simulaciones computarizadas y modelos matemáticos. Pareciera que para los fenómenos que son comprobables en campo o en laboratorio las ciencias biológicas no tienen ningún conflicto con la computación y la modelación matemática, algunos ejemplos de esto son los lagos simulados de Van Doorn y el GARP aplicado a predicciones de distribución geográfica de taxones (Van Doorn 1998, Townsend 2005). Sin embargo para problemas en los cuales los periodos de cambio superan el tiempo de vida del humano promedio no queda mucho más que ecuaciones y algoritmos (Mayhew 2006, Templeton 2006).

El escribir programas de computadora puede parecer un proceso tramposo en comparación al trabajo observacional más clásico de la biología. Es muy cierto que las reglas de los programas muchas veces son arbitrarios y que es casi imposible que un investigador programe todas las restricciones físico químicas con las cuales se enfrenta la vida. Sin embargo esto puede resultar en una ventaja ya que dentro de un programa no hay variables extrañas ni errores de muestreo por lo cual todos los resultados son causa sólo de los fenómenos que se están investigando.

Para asegurarnos de que las simulaciones reflejan en verdad procesos biológicos es necesario que los biólogos comiencen a aprender computación de manera que puedan interpretar la significancia biológica y la validez científica de los programas de cómputo desarrollados para resolver problemas de biología. Queda claro que existen problemas en las ciencias de la vida cuya complejidad hace difícil tener una aproximación práctica por lo que el desarrollo de la biología teórica y la vida artificial como modelos de estudio parece inevitable (Palsson 2000).

5.3 Cómputo bioinspirado

La relación entre ciencias de la computación y biología no es sólo benéfica para los biólogos. Las ciencias de la computación han sabido detectar que los procesos biológicos son procesos de cómputo exportables a sustratos no vivientes para resolver problemas.

La integración de la biología a la computación ha disparado el desarrollo de diversas disciplinas como la inteligencia de cúmulos, las redes neuronales artificiales, los algoritmos genéticos, sistemas inmunes artificiales y robótica bioinspirada (Di Caro 2005, Lisboa 2006, Mitchel 1998, Karlin 2008, Spenko 2008). El entendimiento de las reglas básicas de los fenómenos biológicos permite que se expresen como algoritmos que posteriormente se pueden utilizar en alguna de las ramas mencionadas anteriormente o incluso comenzar una nueva línea de investigación en ciencias de la computación.

El algoritmo propuesto en este trabajo se puede aplicar para resolver problemas de optimización y búsqueda. En dichos problemas se tiene una idea de como es el resultado que buscamos pero no se conocen los detalles por lo cual se implementa un algoritmo que dado un criterio de selección explora el espacio de soluciones y arroja los resultados que mejor cumplen el criterio de selección. El ejemplo

más claro de esto es un algoritmo genético clásico (Mitchell 1998). Sin embargo para muchos de estos problemas existe más de una solución por lo que muchas veces se requiere correr el algoritmo genético varias veces para evitar que se 'atore' en máximos locales.

Al incluir una dimensión espacial en el algoritmo genético la búsqueda de soluciones aceptables se lleva a cabo de manera regional por lo que es difícil que la solución entregada por el programa sea un máximo local. En el caso de que exista más de una solución aceptable el algoritmo presentado arrojará varias soluciones (como se vio en las pruebas 2 y 3) por lo que el sesgo hacía máximos locales es improbable.

5.4 Problemas no resueltos

Los objetivos de este trabajo se han cumplido en gran medida. Sin embargo durante la realización surgieron nuevas preguntas y modificaciones que podrían ser interesantes para futuros trabajos.

El primer problema que sería interesante es que no se conoce una aproximación analítica que lleve a los mismos resultados. Si bien para algunas de las pruebas bastaría con medir la distancia euclidiana promedio de cada uno de los colores y seleccionar al que mejor ajuste con el criterio de selección esto seguiría siendo un proceso algorítmico y no analítico. Por otro lado en pruebas como la tres y la cuatro el criterio de selección se computa sobre la máxima cobertura del espacio de soluciones y no sobre la cobertura inicial por lo que se requeriría encontrar también una forma analítica de generar la máxima cobertura. El último factor que dificulta la búsqueda de equivalentes analíticos es que el autómata presentado es no determinístico debido a la cantidad de factores aleatorios que contiene (en la

reproducción y en la distribución). Con estos elementos no es posible demostrar que no existe una solución analítica pero queda claro por qué es difícil de encontrar.

Otro componente en el cual sería interesante profundizar sería la integración de un criterio de dispersión al trabajo. Si se modifica el algoritmo para que los hijos de los organismos se dispersen a una distancia dada de sus padres probablemente veríamos la emergencia de patrones espaciales como resultado de la búsqueda de genotipos de alta adecuación local. Esto permitiría ver el efecto de la selección sexual en la especiación simpátrica.

En esta versión del programa genotipos (secuencias RGB) similares llevan a fenotipos (colores) similares. Sería muy interesante ver qué pasa en el caso de que pequeños cambios genotípicos provoquen fuertes cambios fenotípicos. La única complicación con este tipo de simulación sería que los organismos tendrían que tener un aparato artificial de percepción para poder seleccionar a su pareja a partir de caracteres fenotípicos, dicho sistema artificial sería difícil de desarrollar en software.

Las simulaciones presentadas en este trabajo tienen el problema de que sólo presentan una estrategia reproductiva, es decir que sólo se pueden reproducir sexualmente de manera poliámica. Sería interesante modificar el algoritmo para que permitiera la inclusión de organismos con reproducción asexual, organismos monógamos, sistemas de dominación en los cuales la descendencia proviene de pocos organismos, etc. También sería correcto hacer una versión en la cual los individuos puedan tener más de un evento reproductivo y coexistieran con su descendencia. Dichas modificaciones aumentarían la similitud del sistema de cómputo con algún sistema biológico.

El último problema que sería interesante abordar sería el hacer una versión del algoritmo

utilizando genomas mucho más grandes de manera que el número de combinaciones posibles aumente por lo cual la exploración exhaustiva del espacio de soluciones sea casi imposible. Un espacio de soluciones muy grande probablemente se asemejaría más a un escenario biológico en el cual coexisten varias genomas con adecuaciones equivalentes pero diferentes de la máxima.

Apéndice A: Mapas de color

A continuación se presentan los mapas de color creados durante cada prueba. Debido a que cada mapa mide 512 x 512 px se incluyen en incrementos de cuatro iteraciones, con excepción del primer incremento que es de siete iteraciones. Las imágenes siguen la secuencia 0,7,11,15,....,87,91,95,99. Por razones de espacio cada prueba ha sido dividida en cinco imágenes, cada imagen presentada a continuación incluye 5 generaciones.

Prueba 1: Cobertura de soluciones regular y completamente explorada, selección por mínima distancia euclidiana.



Generaciones 0-7-11-15-19.



Generaciones 23-27-31-35-39.



Generaciones 43-47-51-55-59.

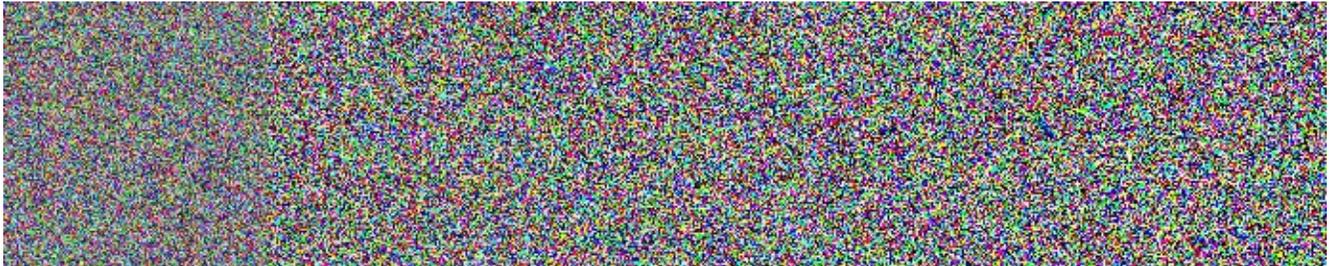


Generaciones 63-67-71-75-79.



Generaciones 83-87-91-95-99.

Prueba 2: Cobertura de soluciones regular y completamente explorada, selección por máxima distancia euclidiana.



Generaciones 0-7-11-15-19.



Generaciones 23-27-31-35-39.



Generaciones 43-47-51-55-59.



Generaciones 63-67-71-75-79.

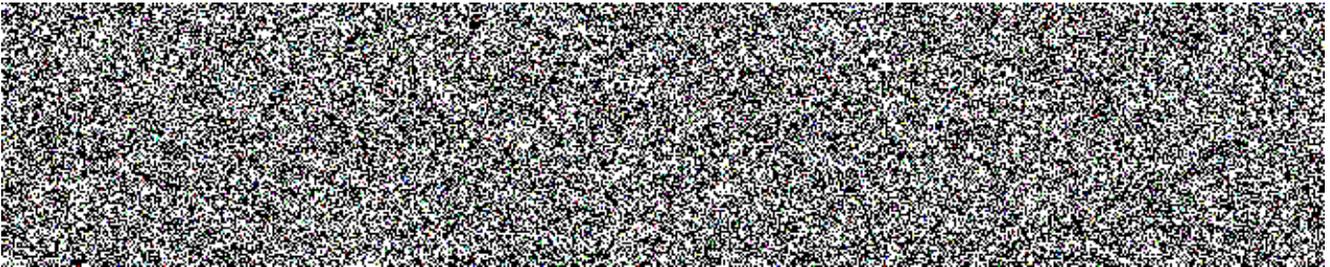


Generaciones 83-87-91-95-99.

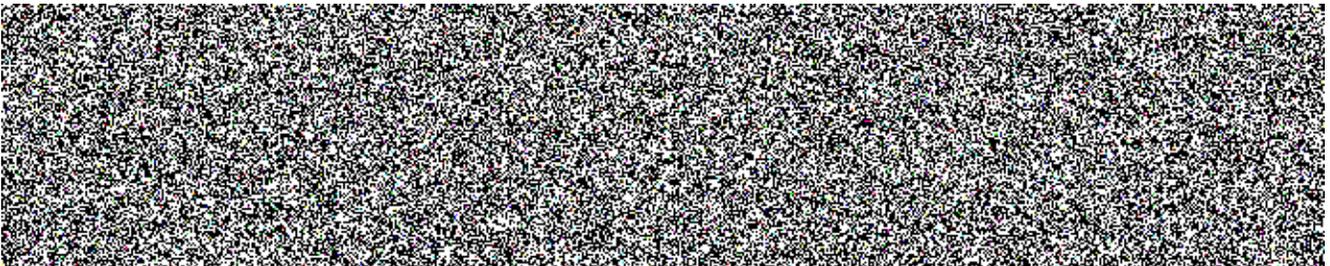
Prueba 3: Cobertura de soluciones reducida a 2 puntos opuestos, selección por mínima distancia euclidiana.



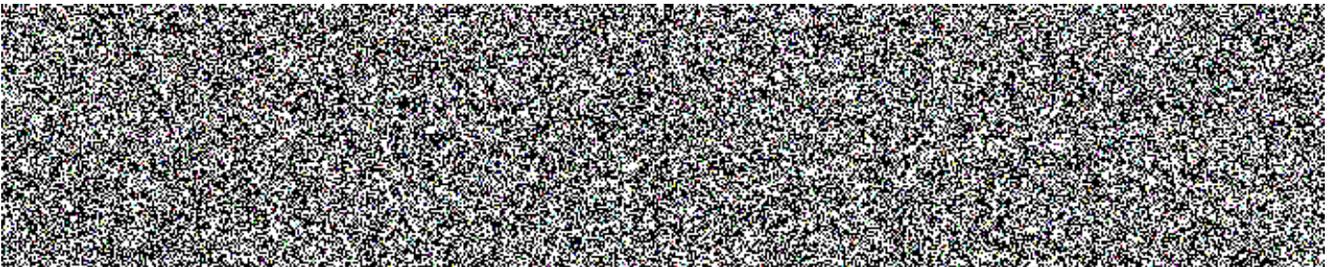
Generaciones 0-7-11-15-19.



Generaciones 23-27-31-35-39.



Generaciones 43-47-51-55-59.



Generaciones 63-67-71-75-79.



Generaciones 83-87-91-95-99.

Prueba 4: Cobertura de soluciones reducida a 2 puntos opuestos, selección por máxima distancia euclidiana.



Generaciones 0-7-11-15-19.



Generaciones 23-27-31-35-39.



Generaciones 43-47-51-55-59.



Generaciones 63-67-71-75-79.



Generaciones 83-87-91-95-99.

Prueba 5: Cobertura de soluciones importada de una imagen. selección por mínima distancia euclidiana.



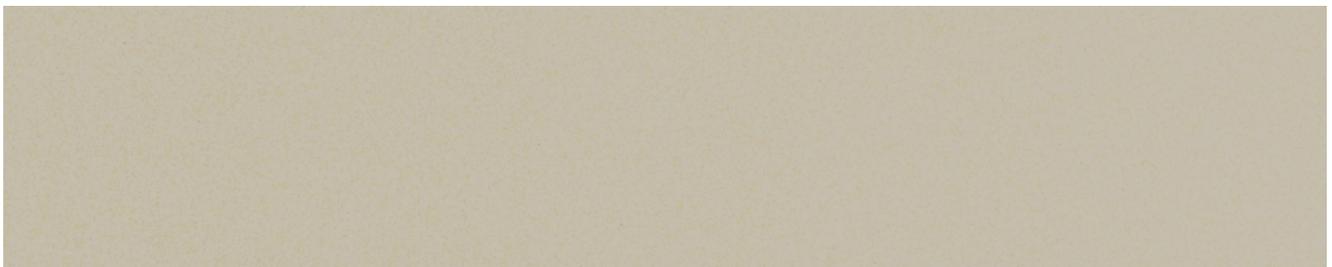
Generaciones 0-7-11-15-19.



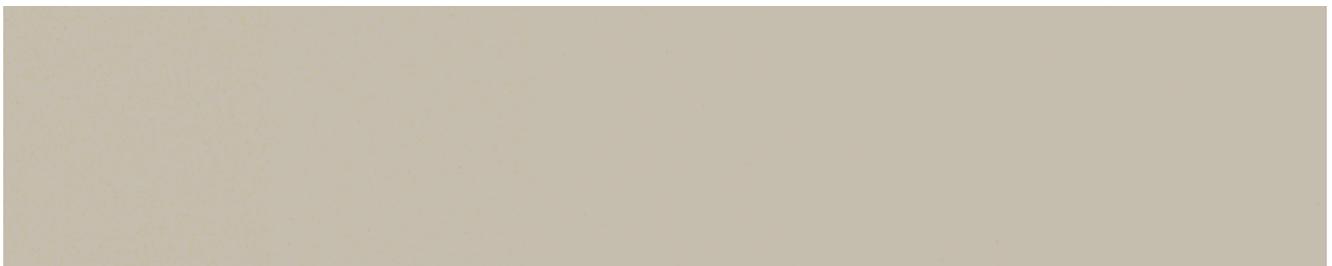
Generaciones 23-27-31-35-39.



Generaciones 43-47-51-55-59.

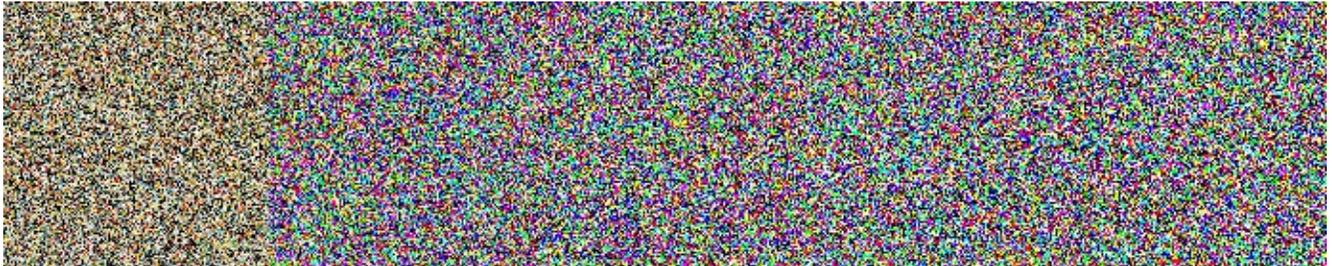


Generaciones 63-67-71-75-79.



Generaciones 83-87-91-95-99.

Prueba 6: Cobertura de soluciones importada de una imagen. selección por máxima distancia euclidiana.



Generaciones 0-7-11-15-19.



Generaciones 23-27-31-35-39.



Generaciones 43-47-51-55-59.



Generaciones 63-67-71-75-79.



Generaciones 83-87-91-95-99.

Referencias

Capítulo 1

Barthélemy-Madaule M. 1979, *Lamarck ou le mythe du précurseur*, París, Éditions du Seuil p.11-62.

En: Barahona, Suárez, Martínez (Comp.) 2001, *Filosofía e Historia de la Biología*, Facultad de Ciencias UNAM. México D.F., p.65-106.

Sloan P.R. 1996, *Natural History*, En: Barahona, Suárez, Martínez (Comp.) 2001, *Filosofía e Historia de la Biología*, Facultad de Ciencias UNAM. México D.F., p.41-64.

S. L. Kosakovsky, D. Posada, M.V Gravenor, C.H. Woelk, S.D. Frost; 2006, *Automated Phylogenetic Detection of Recombination Using a Genetic Algorithm*, Mol. Biol. Evol. Vol.23-10, p.1891-1901

B. Palsson; 2000, *The challenges of in silico biology*, Nature Biotechnology, Vol.18, Nov.

H. Pagels; 1988, *The Dreams of Reason: The Rise of the Sciences of Complexity*, Simon & Schuster, E.U.A.

X. Cui, J. Gao, P.E. Potok; 2006, *A flocking based algorithm for document clustering analysis*, Journal of Systems Architecture, Junio, 2006.

M. M. Botvinick, D.C. Plaut; 2006, *Short-Term Memory for Serial Order: A Recurrent Neural Network Model*, Psychological Review, Vol.113-2, p.201-233.

S. Garnier, C. Jost, J. Gautrais, M. Asadpour, G. Caprari, et al.; 2008 *The Embodiment of Cockroach*

Aggregation Behavior in a Group of Micro-robots , Artificial Life, Vol.14-4, p.387-408.

E. O. Wiley; 1981, *Phylogenetics: The theory and practice of phylogenetic systematics*, John Wiley and Sons inc. Publishing. E.U.A., p. 21.

L. L. Rockwood; 2006, *Introduction to population ecology*, Blackwell Publishing, E.U.A., p.5-7, 86.

M. Begon, C. R. Townsend, J. L. Harper; 2006, *Ecology: From individuals to ecosystems*, 4 ed., Blackwell Publishing. Oxford, Reino Unido, p. 89-95, 471.

C. J. Krebs; 1994, *Ecology: The experimental analysis of distribution and abundance*, 4 ed. Harper Collins. E.U.A., p.160-161.

M. L. Arnold; 2006, *Evolution Through Genetic Exchange* , Oxford University Press, Reino Unido, p. 91-92.

A. Haupt, T. T. Kane; 2003, *Population Handbook*, Population Reference Bureau. Washington, E.U.A., p. 13-15.

J. R. Carey; 1993, *Applied Demography for Biologists with Special Emphasis on Insects* , Oxford University Press, Nueva York, E.U.A., p. 3-9.

D. L. Hartl; 1998, *Genetics: Principles and Analysis*, 4 ed., Jones and Bartlett Publishers, Toronto, Canadá, p. 19-20, 41.

R. H. Tamarin; 2001, *Principles of Genetics*, 7 ed. McGraw-Hill, E.U.A., p. 9.

J. D. Watson, T. A. Baker, S. P. Bell, A. Gann, M. Levine, et al.; 2004, *Molecular Biology of the Gene*, 5 ed., Pearson Education. San Francisco, E.U.A., p.9-11.

J. H. Gillespie; 1998, *Population Genetics: A concise guide*, John Hopkins University Press. Meryland, E.U.A., p. 5-8.

D. A. Coley; 1999, *An intriduction to genetic algorithms for scientists and engineers*, World Scientific Publishing. Singapur, Singapur. p.3-5.

A. R. Templeton; 2006, *Population Genetics and Microevolutionary Theory*, John Wiley & sons inc., E.U.A., p. 341-360, 3-5.

P. H. Gouyon, J.P. Henry, J. Arnould; 2002, *Gene Avatars: The Neo-Darwinian Theory of Evolution*, Kluwer Academic Publishers. Nueva York, E.U.A., p.204-214.

S. C. Stearns; 1992, *The Evlution of Life Histories*, Oxford University Press, Nueva York, E.U.A., p. 10-16.

J. Maynard Smith; 1999, *Evolutinary Genetics*, Oxford University Press. Oxford, Reino Unido, p.3-8.

V. Grant; 1995, *Sexual Selection in Plants: Pros and Cons*, Proceedings of the national academy of science, Vol. 92, p.1247-1250.

- D. J. Futuyma; 2005, *Evolution*, Sinauer Associates Inc. Massachusetts, E.U.A., p.141, 386, 283-287.
- P. Frost; 2006, *European hair and eye color : A case of frequency-dependent sexual selection?* ,
Evolution and Human Behavior, Vol. 27, p. 85-103.
- P. Mayhew; 2006, *Discovering evolutionary ecology: Bringing together ecology and evolution*, Oxford
University Press, Nueva York, E.U.A., p. 140
- R. Medel, C. Botto-Mahan, M. Kalin-Arroyo; *Pollinator Mediated Selection on the Nectar Guide
Phenotype in the Andean Monkey Flower, Mimulus luteus* , Ecology, Vol.84-7, p. 1721-1732.
- M. Mitchell; 1998, *An introduction to genetic algorithms*, MIT Press. Massachusetts, E.U.A. p.5-7.
- C. E. Shannon; 1948, *A Mathematical Theory of Communication*, The Bell Systems Technical Journal,
Vol. 27, p. 379–423, 623–656.
- J. L. Schiff; 2008, *Cellular Automata: A Discrete View of the World*, John Wiley & Sons, Nueva Jersey,
E.U.A, p.39-43.
- M. Mitchell; 2009, *Complexity: A guided tour*, Oxford University Press. Oxford, Reino Unido, p.145-
149.
- I.R. Akberdin, E.A. Ozonov, V.V. Mironova, D.N. Gorpichenko, N.A. Omelyanchuk, et al.; 2006, A

cellular automaton model of morphogenesis in Arabidopsis thaliana, Biophysics, Vol.51-1, p.s91-s94.

P.T. Hraber, T. Jones, S. Forrest; 1997, *The ecology of Echo*, Artificial Life, Vol.3, p.165-190.

R. Das, J.P. Crutchfield, M. Mitchel, J.E. Hanson; 1995, *Evolving Globally Synchronized Cellular Automata*, Proceedings of the Sixth International Conference on Genetic Algorithms.

S. Wolfram; 2002, *A new kind of science*, Wolfram Media Inc. Illinois, E.U.A., p.465-471.

R. Kurzweil; 2005, *The singularity is near*, Penguin. Nueva York, E.U.A., p.85-91.

M. Evening; 2002, *Adobe Photoshop 7.0 for Photographers*, Focal Press. Oxford. Reino Unido, p.2-4.

Capítulo 2

M. L. Stanton, A. A. Snow, S. N. Handel, J. Berezky; 1989, *The impact of a flower-color polymorphism on mating patterns in experimental populations of wild radish (Raphanus raphanistrum)*, Evolution, Vol. 43-2, p.335-346.

R. C. Marshall, K. L. Buchanan, C. K. Catchpole; 2003, *Sexual selection and individual genetic diversity in songbird*, Proceedings of the Royal Society of London, Vol. 270 sup., p. s248-s250.

V. Grant; 1995, *Sexual selection in plants: Pros and cons*, Proceedings of the National Academy of Science, Vol. 92, p. 1247-1250.

R. Medel, C. Botto-Mahan, M. Kalin-Arroyo; 2003, *Pollinator-Mediated Selection on the Nectar Guide Phenotype in the Andean Monkey Flower*, *Ecology*, Vol. 84-7, p. 1721-1732.

D. J. Futuyma; 2005, *Evolution*, Sinauer Associates Inc. Massachusetts, E.U.A., p.332-333.

M. Matsumoto, T. Nishimura; 1998, *Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator*, *ACM Transactions on Modeling and Computer Simulation* Vol. 8-1, pp.3-30. En: <http://docs.python.org/library/random.html>

Capítulo 4

A. R. Templeton; 2006, *Population Genetics and Microevolutionary Theory*, John Wiley & sons inc., E.U.A., p. 341-360, 2-5.

C. C. Cockerham, P. M. Burrows, S. S. Young, T. Prout; 1972, *Frequency Dependant Selection in random mating populations*, *American Naturalist*, Vol. 106, p. 493-515.

L. D. B. Gigord, M. R. Macnair, A. Smithson; 2001, *Negative frequency-dependent selection maintains a dramatic flower color polymorphism in the rewardless orchid *Dactylorhiza sambucina* (L.) Soo*, *Proceedings of the National Academy of Sciences*, Vol. 98-11, p. 6253-6255.

J. H. Gillespie; 1998, *Population Genetics: A concise guide*, John Hopkins University Press. Meryland, E.U.A., p.77-79.

N. Bostrom; 2003, *Are you living in a computer simulation?*, Philosophical Quarterly, Vol.53-211, p.243-255.

K. A. Haynes, M. L. Broderick, A. D. Brown, L. Harden, J. O. Dickson, et al; 2007, *Computing with living hardware*, IET Synthetic Biology, Vol.1-1,2, p.44-47.

M. Mitchell; 2009, *Complexity: A guided tour*, Oxford University Press. Oxford, Reino Unido, p.145-159.

S. Wolfram; 2002, *A new kind of science*, Wolfram Media Inc. Illinois, E.U.A., p.465-471.

D. Dennet; 1995, *Darwin's Dangerous Idea*, Penguin, E.U.A., p.52-60.

S.J. Gould; 1997, *Darwinian Fundamentalism*, The New York review of books, Vol.44-10, p.34-37.

J. Avery; 2003, *Information Theory and Evolution*, World Scientific Publishing, Singapur, p.35-77.

R. C. Marshall, K. L. Buchanan, C. K. Catchpole; 2003, *Sexual selection and individual genetic diversity in songbird*, Proceedings of the Royal Society of London, Vol. 270 sup., p. s248-s250.

R. Medel, C. Botto-Mahan, M. Kalin-Arroyo; 2003, *Pollinator-Mediated Selection on the Nectar Guide Phenotype in the Andean Monkey Flower*, Ecology, Vol. 84-7, p. 1721-1732.

D.R. Brooks, E.O. Wiley; 1988, *Evolution as Entropy: Towards a unified theory of biology*, University of Chicago Press, Chicago, E.U.A., 2°ed. p.

G.S. van Doorn, A.J. Noest, P.Hogeweg; 1998, *Sympatric speciation and extinction driven by environment dependent sexual selection*, Proceedings of the Royal Society of London, Vol.265, p.1915-1919.

A. Townsend Peterson, R. Scachetti-Pereira and W. W. Hargrove; 2004, *Potential Geographic Distribution of Anoplophora glabripennis (Coleoptera: Cerambycidae) in North America*, American Midland Naturalist, Vol.151-1, p.170-178

L. Brillouin; 1962, *Science and information theory*, Academic Press Inc., Nueva York, E.U.A. p.162-182.

R. Landauer; 1961, *Irreversibility and Heat Generation in the Computing Process* , IBM Journal of Research and Development, vol. 5, pp. 183-191.

Capítulo 5

J. H. Gillespie; 1998, *Population Genetics: A concise guide*, John Hopkins University Press. Meryland, E.U.A., p.55-59.

B. Palsson; 2000, *The challenges of in silico biology*, Nature Biotechnology, Vol.18, p.1147-1150.

G. Di Caro, F. Ducatelle, L. M. Gambardella; 2005, *AntHocNet: an adaptive nature-inspired algorithm for routing in mobile ad hoc networks*, European Transactions on Telecommunications (Special Issue on Self-organisation in Mobile Networking), Vol.16, p.443–455.

J. T. Bridgham, E. A. Ortlund, J. W. Thornton; 2009, *An epistatic ratchet constrains the direction of glucocorticoid receptor evolution*, Nature, Vol.461, p. 515-520.

P. J. Lisboa, A. F. G. Taktak; 2006, *The use of artificial neural networks in decision support in cancer: A systematic review*, Neural Networks, Vol.19, p.408–415.

M. J. Spenko, G. C. Haynes, J. A. Saunders, M. R. Cutkosky, A. A. Rizzi; 2008, *Biologically Inspired Climbing with a Hexapedal Robot*, Journal of Field Robotics, vol.25-4,5, p.223–242.

M. Mitchell; 1998, *An introduction to genetic algorithms*, MIT Press. Massachusetts, E.U.A.

J. Karlin, J. Rexford, and S. Forrest; 2008, *Autonomous security for autonomous systems*, Computer Networks, Vol.52, p.2908-2923.