



**UNIVERSIDAD AUTÓNOMA DEL  
ESTADO DE MORELOS**

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS  
INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS  
CENTRO DE INVESTIGACIÓN EN CIENCIAS

**“EVOLUCIÓN DE FORMAS BÁSICAS DE SEÑALIZACIÓN  
EN AGENTES ARTIFICIALES AUTÓNOMOS”**

TESIS  
QUE PARA OBTENER EL GRADO DE  
MAESTRO EN CIENCIAS  
PRESENTA  
VICTOR HUGO SANTOS PUCHETA

DIRECTOR DE TESIS  
DR. BRUNO LARA GUZMÁN

Cuernavaca, Morelos

Mayo, 2019

---

# Índice general

<b>1. Introducción</b>	<b>10</b>
1.1. Objetivo . . . . .	11
1.2. Preguntas de investigación . . . . .	11
1.3. Antecedentes . . . . .	11
<b>2. Fundamentos teóricos</b>	<b>15</b>
2.1. Problemas de optimización . . . . .	15
2.2. Redes neuronales artificiales . . . . .	17
2.3. Robótica evolutiva . . . . .	20
2.3.1. Algoritmos evolutivos . . . . .	22
2.4. Comunicación: teoría y evolución de las señales . . . . .	26
2.4.1. Señales comunicativas . . . . .	27
2.4.2. Origen evolutivo de las señales . . . . .	29
<b>3. Evolución de la comunicación en contexto fijo</b>	<b>30</b>
3.1. Métodos . . . . .	30
3.1.1. Agente artificial . . . . .	30
3.1.2. Configuración experimental . . . . .	31
3.1.3. Controlador neuronal . . . . .	32
3.1.4. Algoritmo evolutivo . . . . .	35
3.2. Forrajeo . . . . .	36
3.2.1. Prueba de rendimiento . . . . .	37
3.2.2. Prueba de señalización . . . . .	39
3.3. Señalización . . . . .	41

3.3.1. Prueba de rendimiento . . . . .	43
3.3.2. Prueba de señalización . . . . .	43
3.4. Resultados . . . . .	44
3.4.1. Forrajeo . . . . .	44
3.4.2. Señalización . . . . .	46
<b>4. Evolución de la comunicación en contexto aleatorio</b>	<b>50</b>
4.1. Métodos . . . . .	50
4.1.1. Configuración experimental . . . . .	50
4.2. Señalización . . . . .	51
4.2.1. Prueba de rendimiento . . . . .	52
4.2.2. Prueba de señalización . . . . .	52
4.3. Resultados . . . . .	53
4.3.1. Señalización . . . . .	53
4.4. Efectos sobre la señalización de los contextos fijo y aleatorio . . . . .	56
<b>5. Efecto de la arquitectura del sistema de control y parámetros evolutivos</b>	<b>58</b>
5.1. Métodos . . . . .	58
5.1.1. Controlador neuronal . . . . .	59
5.2. Señalización . . . . .	59
5.2.1. Prueba de rendimiento . . . . .	61
5.2.2. Prueba de señalización . . . . .	61
5.3. Resultados . . . . .	62
5.3.1. Señalización . . . . .	62
5.4. Efectos sobre la señalización de la arquitectura del sistema de control y parámetros evolutivos . . . . .	65
<b>6. Análisis del espacio de soluciones usando aprendizaje profundo</b>	<b>67</b>
6.1. Autoencoders . . . . .	68
6.2. Reducción de dimensiones en el espacio genotípico . . . . .	69
6.2.1. Modelo . . . . .	69
6.2.2. Entrenamiento . . . . .	69
6.2.3. Resultados . . . . .	71

6.3. Reducción de dimensiones en el espacio fenotípico . . . . .	76
6.3.1. Modelo . . . . .	76
6.3.2. Entrenamiento . . . . .	76
6.3.3. Resultados . . . . .	77
<b>7. Conclusiones</b>	<b>83</b>
<b>Bibliografía</b>	<b>87</b>

---

# Índice de figuras

1.1. Configuración experimental utilizada por Floreano et al. (2007); Mitri et al. (2009, 2010b). . . . .	12
1.2. Tratamientos utilizados por Floreano et al. (2007); Mitri et al. (2009, 2010b). . . . .	13
2.1. a) Representación gráfica de una red neuronal. b) Modelo no lineal de una neurona artificial $k$ . . . . .	19
2.2. Diferentes arquitecturas de red. . . . .	20
2.3. Diagrama de flujo de un algoritmo evolutivo. . . . .	25
2.4. Ejemplo de un operador de cruce y mutación. . . . .	26
2.5. Diagrama esquemático de un sistema de comunicación general (Shannon, 2001). . . . .	27
3.1. Agente artificial e-Puck. . . . .	31
3.2. Muestra de un entorno simulado en Enki. . . . .	32
3.3. Arena de contexto fijo. . . . .	33
3.4. Función de activación $\tanh(x)$ . . . . .	33
3.5. Arquitectura de red neuronal. . . . .	34
3.6. Proceso de selección artificial en robots. . . . .	36
3.7. Prueba de rendimiento con ausencia de luz. . . . .	38
3.8. Prueba de rendimiento con luz roja. . . . .	39
3.9. Prueba de rendimiento con luz azul. . . . .	40
3.10. Prueba de señalización con luz azul/roja contra roja. . . . .	41
3.11. Prueba de señalización con luz azul/roja contra azul. . . . .	42

3.12. Promedio del desempeño de las 20 simulaciones durante el proceso evolutivo en la familia de forrajeo. . . . .	44
3.13. Tiempo invertido por los 100 mejores individuos en cada área de la arena durante los 1200 ciclos de simulación en la familia de forrajeo. . . . .	45
3.14. Promedio del desempeño de las 20 simulaciones durante el proceso evolutivo en la familia de señalizadores en contexto fijo. . . . .	47
3.15. Tiempo invertido por los 100 mejores individuos en cada área de la arena durante los 1200 ciclos de simulación en la familia de señalizadores en contexto fijo. . . . .	48
4.1. Ilustración de los 32 posibles mapas durante el contexto aleatorio. . . . .	51
4.2. Promedio del desempeño de las 20 simulaciones durante el proceso evolutivo en la familia de señalizadores en contexto aleatorio. . . . .	53
4.3. Comparativa entre la familia de señalizadores de contexto fijo y contexto aleatorio del promedio del desempeño de las 20 simulaciones durante el proceso evolutivo. . . . .	54
4.4. Tiempo invertido por los 100 mejores individuos en cada área de la arena durante los 1200 ciclos de simulación en la familia de señalizadores en contexto aleatorio. . . . .	55
5.1. Arquitectura de red neuronal. . . . .	60
5.2. Promedio del desempeño de las 20 simulaciones durante el proceso evolutivo en la familia de señalizadores en contexto aleatorio (segunda instancia). . . . .	63
5.3. Comparativa entre las diferentes familias de señalizadores del promedio del desempeño de las 20 simulaciones durante las primeras 250 generaciones. . . . .	63
5.4. Tiempo invertido por los 100 mejores individuos en cada área de la arena durante los 3000 ciclos de simulación en la familia de señalizadores en contexto aleatorio (segunda instancia). . . . .	64
6.1. Arquitectura de un autoencoder. . . . .	68
6.2. Arquitectura del autoencoder utilizado para reducir la dimensionalidad de los genotipos. . . . .	70

---

6.3. Funciones de activación sigmoide y ReLU. . . . .	70
6.4. Distribución de los genomas de la familia de forrajeo en el espacio de soluciones a través de las generaciones. . . . .	72
6.5. Distribución de los genomas de la familia de señalizadores de contexto fijo en el espacio de soluciones a través de las generaciones. . . . .	73
6.6. Distribución de los genomas de la familia de señalizadores de contexto aleatorio en el espacio de soluciones a través de las generaciones. . . . .	74
6.7. Distribución de los genomas de las tres familias en el espacio de soluciones a través de las generaciones. . . . .	75
6.8. Arquitectura del autoencoder utilizado para reducir la dimensionalidad de los genotipos. . . . .	77
6.9. Distribución de los fenotipos de la familia de forrajeo en el espacio de soluciones a través de las generaciones. . . . .	79
6.10. Distribución de los fenotipos de la familia de señalizadores de contexto fijo en el espacio de soluciones a través de las generaciones. . . . .	80
6.11. Distribución de los fenotipos de la familia de señalizadores de contexto aleatorio en el espacio de soluciones a través de las generaciones. . . . .	81
6.12. Distribución de los fenotipos de las tres familias en el espacio de soluciones a través de las generaciones. . . . .	82

---

# Índice de cuadros

2.1. Propiedades de los principales canales comunicativos (Alcock, 2009). . 28



## Resumen

El presente proyecto está enmarcado dentro de la robótica evolutiva, un área de investigación que implementa procesos evolutivos inspirados en los principios de la teoría del neodarwinismo para el diseño de robots autónomos.

En el sentido más amplio, la comunicación se considera como cualquier interacción que tiene lugar entre un emisor y un receptor. El emisor transmite información al receptor, quien la utiliza para tomar una decisión. Un aspecto importante para el desarrollo de una estrategia básica de comunicación, es la propiedad emergente de reaccionar ante señales útiles transmitidas mediante algún canal sensorial, sin codificar directamente dicho comportamiento.

En este proyecto, se implementó un sistema de evolución experimental con agentes artificiales que tenían la posibilidad de emitir y percibir luz azul en una arena de forrajeo. Con ello, se analizó cómo un sistema básico de comunicación puede surgir y cómo las diferentes condiciones evolutivas pueden determinar el desarrollo de la misma. Existen enfoques que utilizan tratamientos experimentales durante el proceso evolutivo mediante la implementación de niveles específicos de relación y/o selección (Floreano et al., 2007; Mitri et al., 2009, 2010b). A diferencia de estos enfoques, la contribución de este proyecto se centró en la co-evolución entre emisores y receptores en pequeñas poblaciones genéticamente variables. El enfoque sugerido se implementó a través de la estocástica de los operadores genéticos y cómo estos eventos se propagaron dentro de las poblaciones. Los resultados revelaron que en poblaciones genéticamente variables, el costo de la señalización es un factor clave para desencadenar el desarrollo de una estrategia básica de comunicación. Asimismo, se encontró que el contexto en el cual se desarrollan los agentes, la arquitectura del sistema de control y los parámetros evolutivos, son características importantes para el desempeño de los individuos. Finalmente, en este proyecto fue posible explorar el espacio de soluciones de las secuencias de genotipos y fenotipos mediante el uso de aprendizaje profundo. Lo anterior permitió analizar el efecto de la configuración utilizada en los operadores genéticos.

---

# Capítulo 1

## Introducción

Debido a la dificultad que implica analizar todas las posibles situaciones en las que un grupo de robots podría encontrarse, resulta muy poco práctico pre-programar con anticipación el comportamiento de cada robot, incluida la comunicación. Por lo cual, dentro de la inteligencia artificial se desarrollan técnicas que permiten modelar el comportamiento automático de un robot. Una de estas técnicas es la robótica evolutiva, en la cual se implementa el uso de técnicas evolutivas inspiradas en los procesos biológicos para el desarrollo de controladores en robots autónomos.

Desde una perspectiva biológica la comunicación ha evolucionado millones de veces y puede ocurrir entre diferentes reinos de organismos. Esta característica hace de la comunicación un proceso que puede ser modelado y estudiado desde la perspectiva de la robótica evolutiva. Al aplicar un mecanismo de supervivencia del más apto en un grupo de robots, su rendimiento puede aumentar debido a que solo los individuos con mejor desempeño pasan sus genes a las siguientes generaciones. Tal enfoque también debe ser capaz de desarrollar alguna estrategia básica de comunicación que los robots pueden aplicar para resolver alguna tarea. Esta tesis investiga cómo un sistema evolutivo de este tipo se puede utilizar para desarrollar una estrategia básica de comunicación en agentes artificiales.

## 1.1. Objetivo

Obtener comportamientos de comunicación entre agentes artificiales haciendo uso de algoritmos evolutivos inspirados en el proceso de evolución biológica.

## 1.2. Preguntas de investigación

- ¿Cuáles son las características de las tareas que pueden desencadenar el desarrollo de una estrategia básica de comunicación?
- ¿Cuáles son las características medibles de los sistemas básicos de comunicación evolucionados?
- ¿Cuál es el impacto de los operadores genéticos y la arquitectura del sistema de control dentro del proceso evolutivo?

## 1.3. Antecedentes

Existen enfoques que abordan el tema de la comunicación en robótica evolutiva. En tales enfoques, los mecanismos de presión evolutiva permiten la co-evolución de sistemas de comunicación entre receptores y emisores en una tarea determinada. Las principales investigaciones provienen de los trabajos de [Floreano et al. \(2007\)](#); [Mitri et al. \(2009, 2010b\)](#), que abordan cómo las formas de comunicación, incluyendo la comunicación cooperativa y la señalización engañosa, pueden evolucionar en grupos de robots con redes neuronales simples. Estos estudios se realizaron mediante el uso de un sistema de evolución experimental de forrajeo en una arena que contiene una fuente de alimento y veneno (Ver figura 1.1).

De acuerdo a [Floreano et al. \(2007\)](#), las condiciones evolutivas para el surgimiento de la comunicación en robots dependen de la estructura de parentesco, así como del nivel de selección aplicados durante el proceso evolutivo dentro de las poblaciones. En dicha investigación, se analiza el comportamiento y el desempeño de grandes colonias de robots aplicando dos estructuras de parentesco (relación baja y alta) y dos niveles de selección (régimenes individuales y de nivel de colonia), en donde las

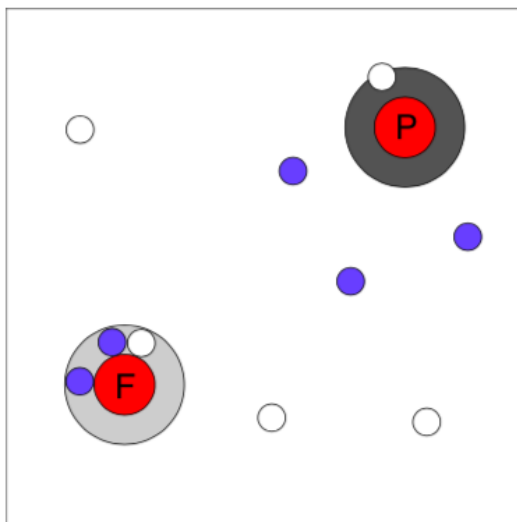


Figura 1.1: Configuración experimental utilizada por Floreano et al. (2007); Mitri et al. (2009, 2010b). En la arena existe una fuente de comida (F) y veneno (P), ambas emitiendo luz roja. Los robots (círculos pequeños) pueden distinguir ambas fuentes al detectar el color de los círculos de papel colocados debajo de cada una ellas mediante sus sensores de color.

colonias con gran relación contenían individuos genéticamente idénticos (Ver figura 1.2). Como resultado, en los experimentos donde los robots estaban muy relacionados y la selección fue aplicada a nivel de colonia se generaron individuos más eficientes que aquellos que evolucionaron mediante otros tipos de selección.

En relación con lo anterior, Mitri et al. (2009, 2010a,b) utilizaron dichas condiciones para estudiar cómo las estrategias de comunicación pueden evolucionar para regular la información provista por señales inadvertidas. En consecuencia, los robots no relacionados produjeron señales poco confiables, mientras que aquellos altamente relacionados señalaron confiablemente la ubicación de la fuente de alimento, obteniendo así un mayor rendimiento. Estos resultados indican que un factor potencialmente importante para el desarrollo de un sistema de comunicación es el nivel de relación genética entre los señalizadores y los receptores.

Estos trabajos se enfocan únicamente en cómo la estructura de parentesco y el nivel de selección influyen conjuntamente en la evolución de la comunicación cooperativa mediante la implementación de colonias de robots. Sin embargo, durante el proceso evolutivo se prescinde de la estocástica del operador genético de selección en

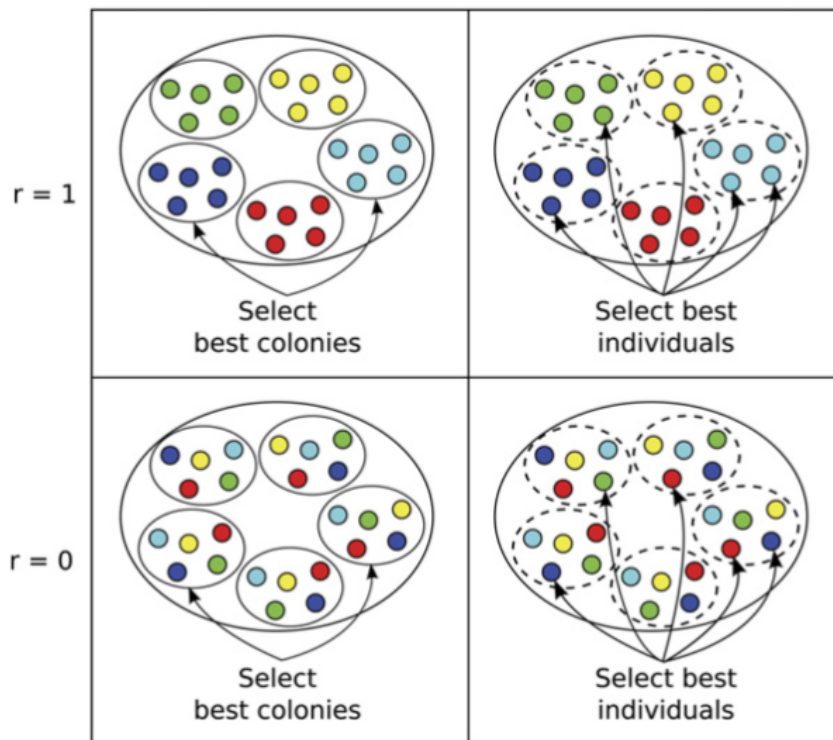


Figura 1.2: Tratamientos utilizados por Floreano et al. (2007); Mitri et al. (2009, 2010b). Ilustración de la composición de la colonia y el régimen de selección en los cuatro tratamientos; alta relación ( $r=1$ ), baja relación ( $r=0$ )

estos estudios.

En la presente investigación se plantea estudiar cómo una estrategia básica de comunicación puede emerger mediante la estocástica de los operadores genéticos en diferentes condiciones experimentales. Esto se logrará sin la necesidad de tratamientos experimentales durante el proceso evolutivo que conducirían a individuos clonados o con niveles específicos de relación y/o selección. En contraste, se introduce el uso de familias para diferenciar entre las distintas condiciones experimentales.

---

# Capítulo 2

## Fundamentos teóricos

### 2.1. Problemas de optimización

Una instancia de un problema de optimización (Rothlauf, 2011) es un par  $(X, f)$ , donde  $X$  es un conjunto de soluciones factibles  $x \in X$  y  $f : X \rightarrow \mathbb{R}$  es una función de evaluación que asigna un valor real a cada elemento  $x$  del espacio de búsqueda. En consecuencia, el problema recae en encontrar una  $x^* \in X$  para la cual:

$$f(x^*) \geq f(x) \quad \text{para toda } x \in X \quad (\text{problema de maximización}) \quad (2.1)$$

$$f(x^*) \leq f(x) \quad \text{para toda } x \in X \quad (\text{problema de minimización}) \quad (2.2)$$

donde  $x^*$  es llamada solución global óptima (o solución óptima) para la instancia del problema en cuestión. Es decir:

$$x^* = \underset{x \in X}{\operatorname{arg\,max}} f(x) \quad (\text{problema de maximización}) \quad (2.3)$$

$$x^* = \underset{x \in X}{\operatorname{arg\,min}} f(x) \quad (\text{problema de minimización}) \quad (2.4)$$

donde *arg max* y *arg min* denotan el elemento del conjunto  $X$  por el cual se alcanza el máximo o mínimo, respectivamente.

Los algoritmos de optimización hacen uso de dos técnicas para encontrar los óptimos globales: a) explorar áreas desconocidas en el espacio de búsqueda, y b) explotar el conocimiento obtenido de puntos previamente evaluados.

El concepto “espacio de búsqueda” se refiere a un conjunto de soluciones candidatas a un problema y a una noción de “distancia” entre las soluciones candidatas (Mitchell, 1998). Dadas dos soluciones  $s$  y  $s'$ , su distancia  $d(s, s')$  se define típicamente como el número mínimo de aplicaciones de una determinada operación básica para transformar la primera solución en la segunda.

Dependiendo de la representación de la solución y de la operación básica involucrada, la distancia puede ser fácilmente calculada. Por ejemplo, si consideramos una representación binaria y una operación básica de cambio de bits, la distancia Hamming puede ser calculada en tiempo lineal en función de la longitud de la cadena de bits (Schiavinotto and Stützle, 2007).

Los problemas de optimización son comunes en muchas disciplinas y varios dominios, como en el caso de las ciencias computacionales o el diseño de controladores en robots autónomos. Por lo cual, se han desarrollado varios métodos para resolver dichos problemas. Estos métodos pueden dividirse en dos categorías: métodos exactos y métodos heurísticos (Rothlauf, 2011).

Los métodos exactos son capaces, teóricamente, de proporcionar una solución óptima, es decir, una solución factible que optimiza (minimiza o maximiza) el valor de la función objetivo. Sin embargo, los métodos exactos no son prácticos de aplicar, a menos que: a) el conjunto de soluciones factibles sea suficientemente pequeño o b) que la estructura y/o el comportamiento de la función objetivo se conozca y se pueda explotar. Un método de optimización exacto es el método de elección si puede resolver un problema de optimización con un esfuerzo que crece polinomialmente con el tamaño del problema. Esta categoría incluye técnicas tales como: enumeración, rama-atadura, método simplex, entre otros (Hooker, 2015).

En contraste, los métodos heurísticos proporcionan una solución factible sin garantía de optimización. En lugar de ello, su objetivo es encontrar buenas soluciones a problemas para los cuales los métodos exactos no son aplicables en la práctica. Por lo general, dichos métodos de optimización son específicos del problema ya que explotan sus propiedades. Los métodos heurísticos incluyen técnicas tales como: inteligencia de enjambre, búsqueda tabú, algoritmos evolutivos, entre otros (Hooker, 2015; Kokash, 2005).



## 2.2. Redes neuronales artificiales

Las redes neuronales artificiales (RNAs) son modelos inspirados en la arquitectura y funcionamiento del sistema nervioso. McCulloch and Pitts (1943) presentaron el primer modelo matemático abstracto de una neurona artificial. En dicho modelo, una neurona es una unidad de cálculo de umbral lineal con entradas múltiples y salida digital única: encendido o apagado.

Posteriormente, Hebb (1949) presentó el primer postulado de aprendizaje para las sinapsis de neuronas biológicas, conocido como regla Hebbiana. Dicha regla fue desarrollada con base a experimentos neurobiológicos y permite a las redes el aprendizaje de patrones. La regla postula que “si las neuronas en ambos lados de una sinapsis se activan sincronizada y repetidamente, la fuerza de la sinapsis es aumentada selectivamente”.

No obstante, fue hasta finales de la década de 1950 y principios de 1960 que la primera RNA fue implementada en una computadora. Rosenblatt (1958, 1962) propuso la mecánica de la neurona artificial única e introdujo el perceptrón como el primer modelo de aprendizaje que permite la corrección del error para la clasificación de problemas linealmente separables. Este modelo considera la suma ponderada de las entradas para generar una señal de salida de tipo lógica que depende de un umbral de activación.

Sin embargo, las limitaciones computacionales de esta primera clase de RNA no fueron superadas antes de la década de 1980. En 1989 se publicaron dos artículos de manera independiente sobre perceptrones multicapa como aproximaciones universales. Funahashi (1989) demuestra que un perceptrón multicapa cuya función de activación sea no constante, acotada y monótona creciente es un buen aproximador universal de funciones. Por otra parte, Hornik et al. (1989) llegan a un resultado similar utilizando funciones de activación sigmoideas, no necesariamente continuas. Los estudios mencionados demostraron la eficacia de este tipo de modelos para resolver problemas que no son linealmente separables.

La neurona artificial es un componente básico de todas las redes neuronales artificiales. Este tipo de modelos permiten procesar entradas ruidosas, como las producidas por sensores de agentes que navegan en un espacio que contiene elementos

aleatorios y/o de incertidumbre. Además, las RNA tienen una capacidad incorporada para adaptar sus pesos sinápticos a los cambios en el entorno circundante. Las RNA son modelos conformados por neuronas que se conectan entre si a partir de pesos numéricos que se ajustan durante el proceso de aprendizaje. En el caso de la robótica evolutiva, los pesos numéricos y/o conexiones de la red son ajustados mediante un proceso de optimización heurístico (algoritmos evolutivos) (Nelson et al., 2009; Stanley et al., 2009; Stanley and Miikkulainen, 2002).

Algunos de los elementos que se encuentran de manera tradicional en una RNA son (Haykin et al., 2009; Rojas, 2013) (Ver figura 2.1):

- Nodos o neuronas  $k_1, k_2, k_3, \dots, k_m$ .
- Un conjunto de entradas  $x_1, x_2, x_3, \dots, x_m$ .
- Un conjunto de pesos sinápticos  $w_{k1}, w_{k2}, w_{k3}, \dots, w_{km}$  asociados a las entradas.
- Una regla de propagación:

$$v_k = u_k + b_k \quad (2.5)$$

donde:

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (2.6)$$

El término  $b_k$  (*bias*), tiene el efecto de aumentar o disminuir la entrada neta de la función de activación, dependiendo de si es positiva o negativa, respectivamente.

- Una función de activación:

$$y_k = \varphi(u_k + b_k) \quad (2.7)$$

Esta función es utilizada para transformar el nivel de activación de una unidad (neurona) en una señal de salida. Actualmente, existe una serie de funciones de activación comunes en RNA (Karlik and Olgac, 2011).

- La estructura de los nodos.
- La topología de la red.
- El algoritmo de aprendizaje utilizado para encontrar los pesos de la red.

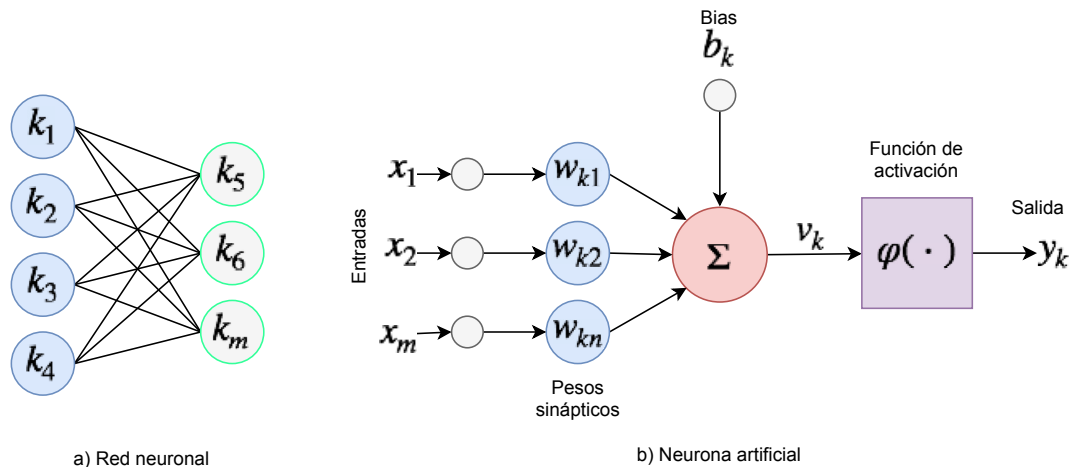


Figura 2.1: a) Representación gráfica de una red neuronal. b) Modelo no lineal de una neurona artificial  $k$ .

En general, podemos identificar tres clases de arquitecturas de red (Haykin et al., 2009):

1. **Feedforward de una sola capa:** Los nodos de la capa de entrada se proyectan directamente sobre una capa de salida de neuronas (nodos de procesamiento), pero no viceversa. La denominación “de una sola capa” se refiere únicamente a la capa de salida, ya que la capa de entrada no realiza ningún cálculo (Ver figura 2.2a).
2. **Feedforward multicapa:** Se distingue por la presencia de una o más capas ocultas, cuyos nodos de procesamiento se denominan neuronas o unidades ocultas. El término “oculto” se refiere a que esta parte de la red neuronal no se ve directamente, ni desde la entrada ni desde la salida. La función de las neuronas ocultas es intervenir entre la entrada externa y la salida de la red de alguna manera útil. Al añadir una o más capas ocultas, la red puede extraer estadísticas de orden superior de su entrada. (Ver figura 2.2b)

3. **Recurrente:** Se distingue de una red *feedforward* en que tiene al menos un circuito de retroalimentación. Además, procesan secuencias de datos a través de lo que se denomina “estado” o “memoria” (Lipton et al., 2015). La presencia de bucles de retroalimentación tiene un profundo impacto en la capacidad de aprendizaje de la red y en su rendimiento (Ver figura 2.2c).

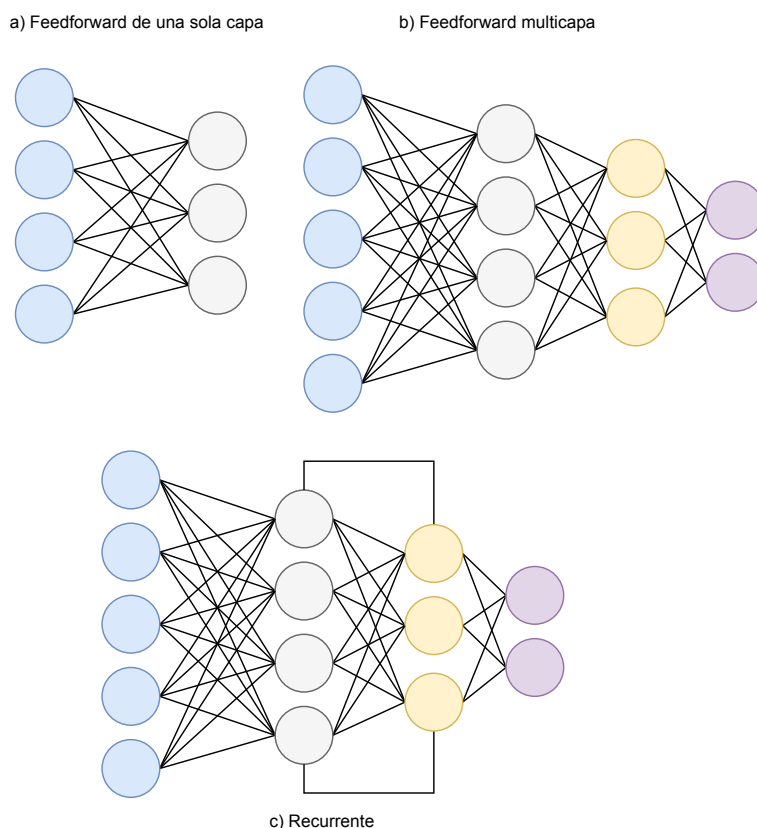


Figura 2.2: Diferentes arquitecturas de red. a) *Feedforward* de una sola capa: incluyendo la capa de entrada, la estructura cuenta con dos capas. Sin embargo, la capa de entrada no se considera debido a que no se realiza ningún cálculo. b) *Feedforward* multicapa: a diferencia de una red de una sola capa, existe (al menos) una capa de “neuronas ocultas” entre la capa de entrada y salida. c) Recurrente: se distingue de una red *feedforward* en que tiene al menos un circuito de retroalimentación.

## 2.3. Robótica evolutiva

El término robótica evolutiva (RE) fue introducido en 1993 (Cliff et al., 1993). Sin embargo, la idea de una inspiración biológica fue abordada tempranamente por Turing (1950), quien habló acerca de la relación entre el proceso de aprendizaje y los procesos evolutivos. No obstante, los orígenes de la RE surgen del trabajo de Friedman (1956), quien propuso aplicar los procesos evolutivos en circuitos de control mediante el uso de “retroalimentación selectiva”. Los circuitos de control que utilizó modelaban las neuronas humanas y tenían la capacidad de ser excitados o inhibidos. Asimismo, era posible formar circuitos más complejos a partir de circuitos simples. Basado en la mutación y selección, Friedman (1956) desarrolló un modelo para generar y evaluar estos circuitos de forma automática.

La RE es una técnica para el diseño de robots autónomos (Nolfi and Floreano, 2000). Este campo de estudio considera a los robots como organismos artificiales autónomos que desarrollan sus propias habilidades en estrecha interacción con el entorno, sin intervención humana. El objetivo de la RE a largo plazo es obtener un proceso automático capaz de diseñar, e incluso construir un robot óptimo dada sólo la especificación de una tarea (Doncieux et al., 2011).

El campo de la RE ha tomado varias direcciones: el relativo a las ciencias cognitivas, la ingeniería y la biología. El primero utiliza la RE para generar agentes artificiales de una manera relativamente libre de predisposiciones que posteriormente pueden ser investigadas y analizadas (Harvey et al., 2005). El segundo enfoque se centra en generar sistemas no necesariamente viables biológicamente, ya que la idea es obtener buenos diseños y comportamientos novedosos (Brambilla et al., 2013; Lipson and Pollack, 2000; Silva et al., 2014). Finalmente, el concerniente a la biología está encaminado al estudio científico sobre los principios de la evolución biológica (Auerbach and Bongard, 2014; Elfwing and Doya, 2014; Kim et al., 2013).

La principal motivación de la RE es generar y optimizar robots capaces de adaptarse a su entorno mediante un proceso análogo a la evolución natural de la supervivencia del más apto. Para ello, la RE se basa en los principios de la teoría del neodarwinismo, también conocida como síntesis evolutiva moderna.

La síntesis evolutiva moderna describe la evolución en términos de variaciones

genéticas en una población que conduce a la formación de una nueva especie (Singh and Singh, 2018). Tomando inspiración de esta teoría, la RE aplica los conceptos de selección, variación y los principios de la herencia para el diseño de robots autónomos.

Durante el diseño evolutivo del robot pueden ser considerados varios aspectos de forma individual o simultánea, como lo son su morfología, sistema de control, arquitectura de control, entre otros (Siciliano and Khatib, 2016). Por ejemplo, si el sistema de control es una RNA, los pesos de conexión se pueden representar a nivel genómico en un vector de valores reales, mientras que los controladores basados en máquinas de estados finitos se pueden describir mediante autómatas (König et al., 2009).

### 2.3.1. Algoritmos evolutivos

El comportamiento que expresa un robot no está ligado específicamente con el controlador, sino que es el resultado emergente de las interacciones del robot con su entorno (Nolfi, 2009). Debido a que los controladores no pueden ser evaluados directamente, la RE utiliza algoritmos evolutivos (AE) para optimizar los controladores candidatos. Asimismo, la RE no está ligada a AE específicos, en lugar de ello se basa en un proceso de auto-organización en el cual la evaluación y la optimización son holísticas (Silva et al., 2016). Los AE operan con una población de individuos  $P(i) = \{x_1^i, \dots, x_n^i\}$ , para la iteración  $i$ , donde cada individuo  $x_n$  representa un punto de búsqueda en el espacio de las soluciones potenciales a un problema dado. De modo que, la población inicial evoluciona sucesivamente hacia mejores regiones del espacio de búsqueda.

Los AE (Yu and Gen, 2010) son algoritmos que realizan tareas de optimización o aprendizaje con la capacidad para evolucionar. Tienen tres características principales:

- **Basado en la población:** Los AE mantienen un grupo de soluciones, llamado población, para optimizar o aprender el problema de forma paralela. La población es un principio básico del proceso evolutivo.
- **Orientado a la aptitud:** Cada solución en una población se llama individuo. Cada individuo tiene su representación genética, llamada código, y una

evaluación del desempeño, llamada valor de aptitud. Los AE prefieren a los individuos más aptos, la cual es la base de la optimización y convergencia de los algoritmos.

- **Basado en la variación:** Los individuos se someterán a una serie de operaciones de variación para imitar los cambios genéticos, proceso fundamental para la búsqueda en el espacio de la solución.

Debido a su naturaleza de prueba y error, los AE requieren de un gran número de evaluaciones durante el proceso de optimización. En consecuencia, la mayoría de los procesos evolutivos son evaluados fuera de línea por medio de simulaciones computacionales basadas en motores físicos que modelan con precisión las propiedades dinámicas de robots reales (Ivaldi et al., 2014). La principal ventaja de evolución fuera de línea es la posibilidad de enfocarse únicamente en el desarrollo del controlador y dejar a un lado los problemas relacionados con los robots físicos.

Por otro lado, algunas investigaciones se han enfocado en la evolución en línea de controladores en hardware real mediante el uso de los recursos computacionales del robot (Bredeche et al., 2009). Sin embargo, debido a que cada evaluación en hardware real implica una cantidad de tiempo considerable este enfoque continúa siendo inviable. Un enfoque alternativo es sintetizar los controladores fuera de línea para posteriormente realizar la transferencia en hardware real (Bongard et al., 2006; O'Dowd et al., 2011). La disponibilidad de motores de física relativamente rápidos permite que la simulaciones fuera de línea puedan utilizarse como un procedimiento de inicialización, reduciendo de esta manera las diferencias entre la simulación y el mundo real (Jakobi, 1997).

Puesto que los AE son técnicas bio-inspiradas, durante el proceso de codificación de las soluciones se deben de considerar algunos términos importantes:

- Las estructuras que codifican las soluciones se llaman genomas.
- La información contenida dentro del genoma se llama genotipo.
- La información que resulta de expresar el genotipo se conoce como fenotipo.
- Cada genoma contiene una serie de valores individuales llamados genes.

En la RE los enfoques tradicionales consisten en optimizar una población de genomas en un espacio de tipo genético. Cada genoma codifica una serie de parámetros del sistema de control de los robots (fenotipo). La metodología básica para el funcionamiento de un AE dentro de la RE es la siguiente (Ver figura 2.3):

- Generar una población aleatoria de genomas.
- Expresar genotipo (Vivir)
- Evaluar la aptitud de cada individuo.
- Seleccionar los individuos para la próxima generación.
- Aplicar mutación y/o cruce para generar la siguiente población.
- Repetir el proceso de traducción, evaluación, selección y variación hasta evaluar una cantidad máxima de generaciones.

En genética evolutiva, la aptitud es la capacidad de los organismos para sobrevivir y reproducirse en el entorno en el que se encuentran (Orr, 2009). En los AE el proceso mediante el cual se obtiene el valor de aptitud de un genoma se llama evaluación. Este proceso utiliza el valor de aptitud para indicar que tan bien un individuo resuelve el problema. En los experimentos de RE estándar, se utiliza una función de aptitud tanto para definir el objetivo como para guiar la búsqueda evolutiva hacia el objetivo (Nelson et al., 2009).

Por otro lado, la selección es un proceso estocástico mediante el cual se eligen los individuos que formarán la nueva generación, de tal forma que incluso los individuos menos aptos tienen cierta posibilidad de sobrevivir. El objetivo de la selección es mantener la diversidad dentro de la población, previniendo una convergencia prematura con soluciones no óptimas mediante algún método de selección (Shukla et al., 2015).

Además, el operador de cruce considera dos o más individuos de una población con el objetivo de generar nuevas soluciones por medio del intercambio genético. Las técnicas de cruce permiten encontrar soluciones óptimas en el menor tiempo posible con un mínimo de generaciones. Para describir la frecuencia con la que se formará el cruce se define una constante de probabilidad. Si no existe cruce, los descendientes



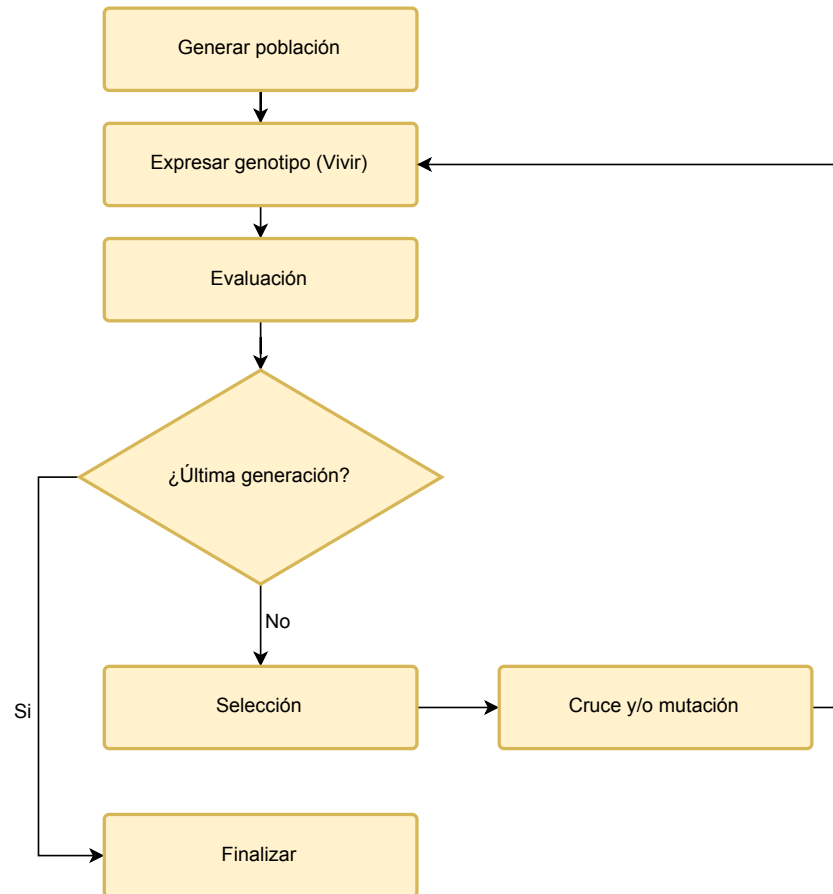


Figura 2.3: Diagrama de flujo de un algoritmo evolutivo. El proceso comienza con una población de genomas de manera aleatoria. Posteriormente, se expresa el comportamiento de los individuos de la generación actual mediante la expresión del genotipo, y se evalúa su aptitud. Con base a los resultados de la evaluación, se realiza la selección de los individuos para la próxima generación y mediante variación (mutación y/o cruza) se genera la siguiente generación. Los procesos de traducción, evaluación, selección y/o variación son ejecutados nuevamente hasta cumplir una cantidad máxima de generaciones.

serán copias exactas de los padres. De lo contrario, los descendientes estarán compuestos de partes del genoma de ambos padres. El impacto de un operador de cruce dentro del espacio de soluciones dependerá de método de cruce utilizado (Umbarkar and Sheth, 2015). (Ver figura 2.4a)

Finalmente, los operadores de mutación son transformaciones estocásticas de un individuo que permite introducir nueva información genética. Estos operadores de mutación permiten diversificar la población existente y explotar áreas del espacio de búsqueda, evitando así caer en extremos locales. La cantidad de genes involucrados durante el proceso de mutación se expresa mediante un valor numérico, llamado porcentaje de mutación. Si no existe ninguna mutación, la descendencia se genera inmediatamente después del cruce sin ningún cambio. De no ser así, se cambian una o más partes de un genoma (Soni and Kumar, 2014). (Ver figura 2.4b)

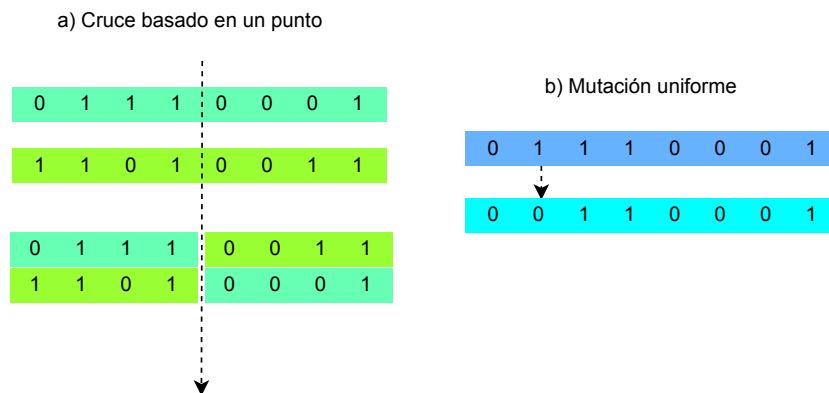


Figura 2.4: Ejemplo de un operador de cruce y mutación. a) Cruce basado en un punto: se selecciona aleatoriamente un punto de corte en los genomas padres y posteriormente las porciones de los cortes se intercambian para formar dos nuevos individuos. b) Mutación uniforme: se altera uno o más genes dentro de un rango específico de acuerdo con una tasa de mutación predefinida.

## 2.4. Comunicación: teoría y evolución de las señales

Varias formas de comunicación han evolucionado entre los diferentes reinos de organismos (Brenner et al., 2008). La comunicación desempeña varias funciones, desde señalización a nivel celular hasta los complejos intercambios lingüísticos humanos. De tal modo que, biológicamente, la comunicación es omnipresente.

En teoría de la comunicación, un sistema de comunicación consiste de cinco partes (Ver figura 2.5) (Shannon, 2001):

1. **Fuente de información:** Produce un mensaje o una secuencia de mensajes para ser comunicados al receptor terminal.
2. **Transmisor:** Opera en el mensaje de alguna manera para producir una señal adecuada para la transmisión a través del canal.
3. **Canal:** Medio utilizado para transmitir la señal del transmisor al receptor.
4. **Receptor:** Realiza la operación inversa de la realizada por el transmisor, reconstruyendo el mensaje de la señal.
5. **Destino:** Persona (o cosa) a quien va dirigido el mensaje.

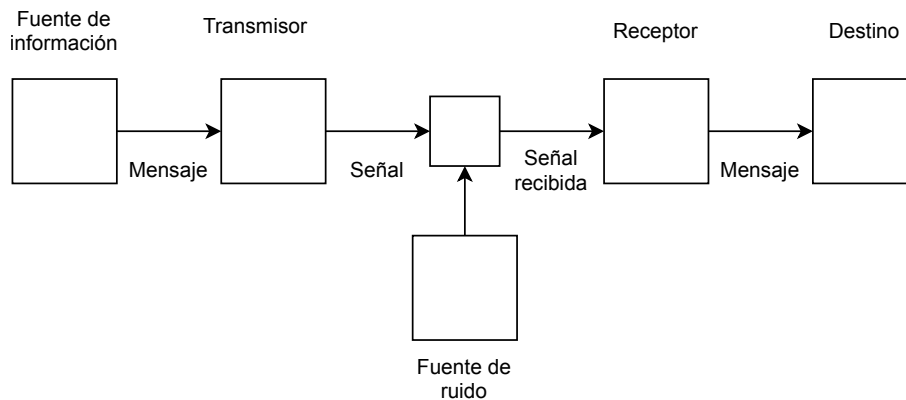


Figura 2.5: Diagrama esquemático de un sistema de comunicación general (Shannon, 2001).

En términos biológicos, la comunicación es el comportamiento o fenotipo por parte de un organismo que altera el patrón de probabilidad de comportamiento en otro organismo de una manera adaptable a uno o ambos participantes (Wilson, 1975). En donde la palabra adaptable implica que la señalización, la respuesta, o ambas, han sido genéticamente programadas en cierta medida por selección natural.

La comunicación implica una interacción entre dos (o más) individuos. En el sentido más amplio, se considera como cualquier interacción que tiene lugar entre un emisor y un receptor. El emisor entrega información al receptor, el cual utiliza esta información para tomar una decisión.

### 2.4.1. Señales comunicativas

Un sistema de comunicación requiere de dos habilidades: la habilidad de producir señales útiles y la capacidad de reaccionar a las señales útiles. Es decir, emitir una señal no es útil a menos que se entienda, y una señal no se entenderá la primera vez que se emite (Smith, 1993). Por lo cual, el termino comunicación implica una serie de conceptos asociados:

- **Señal:** Un comportamiento o fenotipo que altera el comportamiento de otros organismos, que evolucionó a causa de ese efecto, y que es eficaz porque la respuesta del receptor también ha evolucionado (Diggle et al., 2007; Keller and Surette, 2006; Scott-Phillips, 2008; Smith and Harper, 2003). Si se cumplen estas condiciones, entonces la acción es una señal, la reacción es una respuesta, y la interacción general es comunicativa. Si solo la reacción es funcional de esta manera, entonces la acción es un **indicador**; y si solo la acción es funcional de esta manera, entonces es **coercitiva** (Scott-Phillips et al., 2012).
- **Indicador:** Un comportamiento o fenotipo que altera el comportamiento de otro organismo o varios organismos, que es beneficioso para el (los) receptor (es), pero no ha evolucionado debido a ese efecto (Diggle et al., 2007; Keller and Surette, 2006; Scott-Phillips, 2008; Smith and Harper, 2003).
- **Coerción:** Una señal que ha evolucionado debido a su efecto negativo en el (los) receptor (es) en comparación con otros individuos que no reciben la señal (Diggle et al., 2007; Keller and Surette, 2006; Scott-Phillips, 2008; Smith and Harper, 2003).

Las señales comunicativas son transmitidas sobre un amplio rango de canales sensoriales, los cuales incluyen el visual, acústico, táctil y químico (Billen, 2006). Estos canales cuentan con un propósito específico, desde coordinar actividades en grupo hasta influir en la elección de pareja (Grueter and Keller, 2016; Penn and Potts, 1998).

Cada canal sensorial posee propiedades específicas para la transmisión de información; asimismo, cada señal se adecua mejor en un ambiente determinado. Las

Cuadro 2.1: Propiedades de los principales canales comunicativos (Alcock, 2009).

	Canal			
	Químico	Acústico	Visual	Táctil
Rango de alcance	largo	largo	medio	corto
Tasa de transmisión	lenta	rápida	rápida	rápida
¿Salva obstáculos?	sí	sí	no	no
Uso nocturno	sí	sí	no*	sí
Latencia en desaparecer	larga	corta	corta	corta
¿Permite localizar la posición espacial de la fuente?	mal	bien	muy bien	muy bien
Costo energético	bajo	alto	medio	bajo

\* Salvo señales bioluminiscentes.

propiedades de cada canal determinan el costo y beneficio de cada tipo de señal de acuerdo al gasto energético implicado durante su producción (Ver tabla 2.1).

### 2.4.2. Origen evolutivo de las señales

La relación emisor-receptor plantea una problemática que recae en la fuerte dependencia entre ambos actores y los orígenes evolutivos de las señales. Es decir, no está claro por qué un individuo evolucionaría para señalar si ningún receptor ha evolucionado una respuesta a la señal. Para resolver esta problemática, se han planteado dos procesos por los cuales la comunicación podría surgir: la ritualización y la manipulación sensorial (Smith and Harper, 2003).

En la ritualización, las señales evolucionan a partir de comportamientos que originalmente eran solo indicadores. De acuerdo a lo anterior, puede esperarse que los sistemas de comunicación se originen a través de información producida inadvertidamente que evoluciona hasta convertirse en una señal. Por ejemplo, los mamíferos extremadamente asustados a menudo excretan sustancias de desecho. Si los intrusos potenciales comenzaron a utilizar la orina como un indicador territorial, este efecto podría retribuirse a los propietarios por desechar esta sustancia en los límites territoriales (Smith and Harper, 2003).

Por otra parte, en la manipulación sensorial las señales evolucionan debido a que

aprovechan un prejuicio sensorial preexistente y una respuesta conductual asociada en el receptor, que es beneficiosa para el emisor de la señal (Smith and Harper, 2003). Desde este punto de vista, las señales evolucionan a partir de conductas que originalmente solo eran coercitivas. Por ejemplo, los cangrejos violinistas machos (*Uca musica*) algunas veces durante el cortejo construyen pequeños montículos de arena en las entradas de sus madrigueras a las que atraen a las hembras para aparearse (Christy et al., 2003). De manera que, muchas de las exhibiciones durante el apareamiento podrían haber comenzado en escenarios en los que la preferencia por algún objeto o cierto color permitieron la manipulación del receptor por parte del emisor.

En la literatura, los ejemplos de evolución por ritualización son más comunes que los de manipulación sensorial. Asimismo, en la ritualización muchas de las señales probablemente evolucionaron mediante indicadores que otros individuos ya habían utilizado para obtener información. Por lo cual, se ha sugerido que la mayoría de las señales evolucionaron mediante ritualización (Smith and Harper, 2003).

---

## Capítulo 3

# Evolución de la comunicación en contexto fijo

Los experimentos de esta sección fueron divididos en dos condiciones: forrajeo y señalización. La primera de ellas está diseñada para desarrollar en los robots una conducta en donde resultara benéfico en términos de aptitud evitar el área de veneno, localizar el área de comida y permanecer en ella. Por otro lado, la segunda condición está diseñada para desarrollar una conducta en donde emitir luz sobre el área de comida tiene un impacto positivo en términos de aptitud. Posteriormente, se aplicaron dos pruebas (rendimiento y señalización) para comprobar si algún tipo de estrategia básica de comunicación había emergido durante el proceso evolutivo.

### 3.1. Métodos

#### 3.1.1. Agente artificial

Durante los experimentos se utilizó una versión simulada del robot e-Puck (Ver figura 3.1), que consiste en un robot móvil de 7 cm de diámetro y 6 cm de altura. El agente está equipado con dos ruedas con motores diferenciales, un anillo de LED's, una cámara 360° y un sensor de color de piso. La implementación de estas características fue posible debido a que el simulador utilizado es extensible (permite agregar robots personalizados).

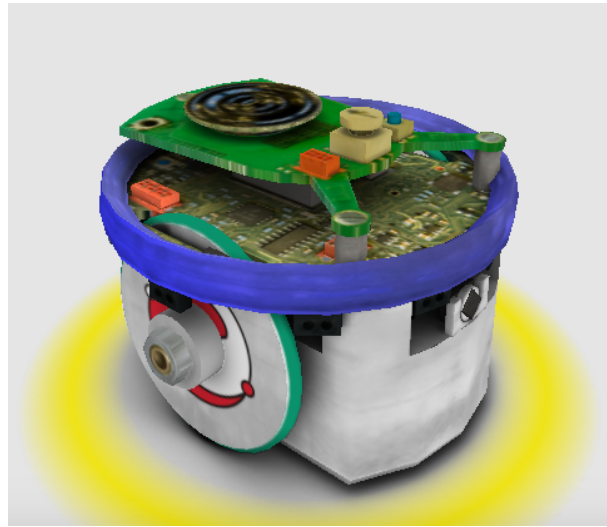


Figura 3.1: Agente artificial e-Puck. Robot simulado durante los experimentos.

El agente y su entorno fueron simulados con Enki (Ver figura 3.2), un simulador de robots de código abierto escrito en C++. Esta plataforma proporciona colisión y soporte físico limitado y es capaz de simular grupos de robots cien veces más rápido que en tiempo real (Magenat et al., 2009).

### 3.1.2. Configuración experimental

Los experimentos se realizaron en una arena cuadrada simulada de  $4m^2$ . La arena contenía una fuente de comida y una de veneno de  $20cm$  de diámetro y  $4cm$  de altura, ambas indicadas de color rojo. Las fuentes fueron colocadas en esquinas opuestas, a  $50cm$  de cada esquina. Se colocó un área circular de color gris con un diámetro de  $25cm$  debajo de la fuente de comida y otra de color negro debajo de la fuente de veneno (Ver figura 3.3). Por lo tanto, un robot podía discriminar entre las dos fuentes usando únicamente sus sensores de piso una vez que pasaba sobre el área de color. Al comienzo de cada simulación los robots se distribuyeron en posiciones aleatorias dentro de la arena (ubicación y ángulo inicial variable). Como única restricción del experimento se determinó que los robots no podían iniciar la simulación dentro de las áreas de comida o veneno. La posibilidad de permanecer cerca de alguna fuente (comida o veneno) hasta el final de la prueba dependía sólo del comportamiento del robot durante el proceso evolutivo.



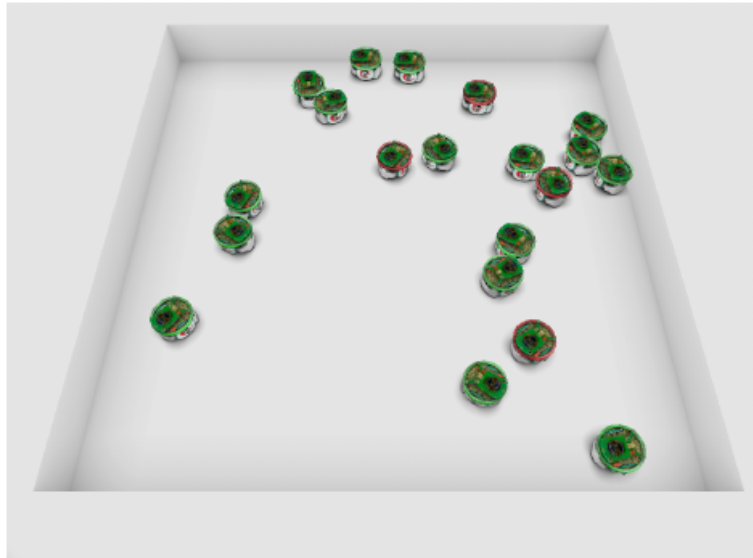


Figura 3.2: Muestra de un entorno simulado en Enki.

Las fuentes de comida y veneno emitían constantemente luz roja; asimismo, el anillo de LED's de los robots podía emitir luz azul. Los robots podían detectar la cantidad de píxeles rojos y azules en toda la arena por medio de su cámara 360°, con una resolución de  $720 * 1$  píxeles. Al mismo tiempo, los robots tenían un ciclo sensorimotor de  $50ms$  durante el cual podían procesar la información visual y la entrada del sensor de color de piso.

### 3.1.3. Controlador neuronal

El sistema de control de cada robot consistió en una red neuronal *feedforward* de una sola capa con 11 neuronas de entrada y 3 de salida, con un total de 33 conexiones (Ver figura 3.5). Las neuronas de la capa de salida recibieron la suma de cada salida de la capa de entrada, multiplicada por el peso de la conexión correspondiente. Se utilizó la función tangente hiperbólica (Ec. 3.1) como función de activación para cada una de las neuronas de salida (es decir, su salida fue entre -1 y 1) (Ver figura 3.4):

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (3.1)$$

Dos de las neuronas de salida se usaron para controlar la velocidad de los motores.

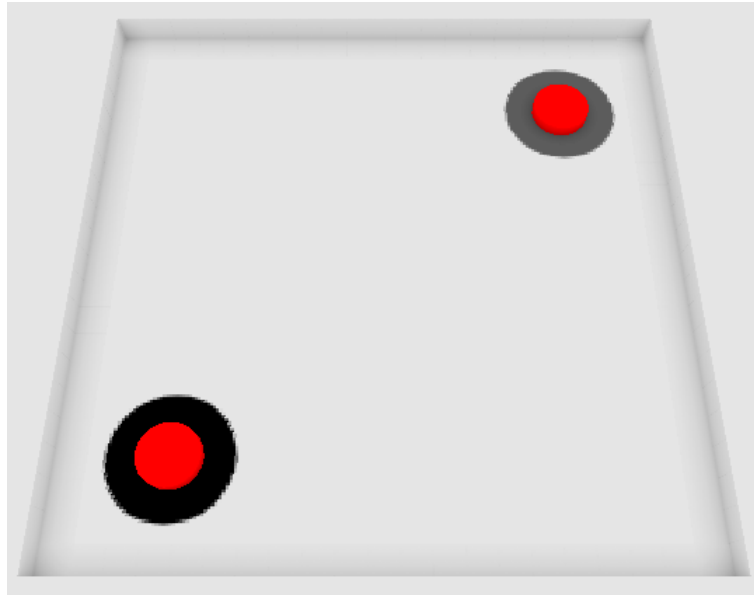


Figura 3.3: Arena de contexto fijo. La arena simulada es una superficie cuadrada con un área de  $4m^2$ , la cual contiene una fuente de comida (círculo negro) y una de veneno (círculo gris), ambas de color rojo.

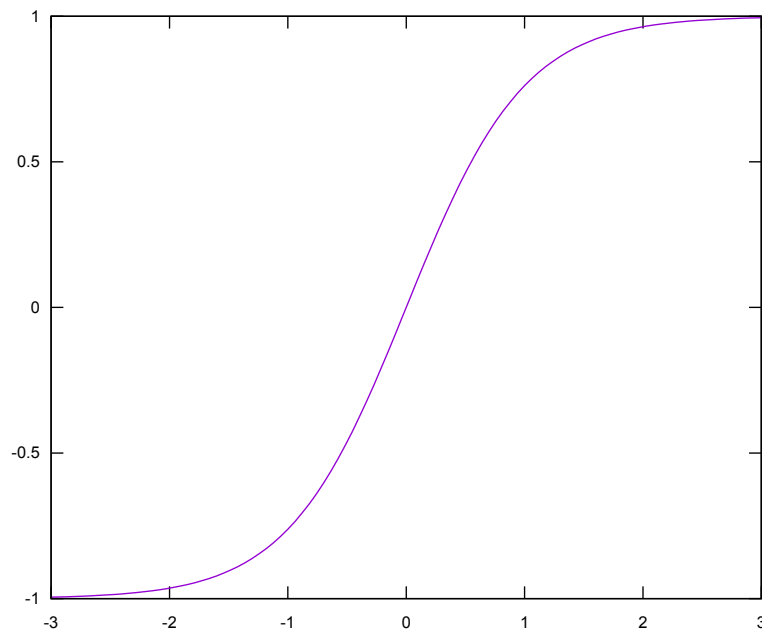


Figura 3.4: Función de activación  $\tanh(x)$ .

La tercera neurona de salida determinó la emisión de luz azul; que era el caso si la salida era mayor que 0.

Los pesos de las conexiones de la red neuronal fueron codificados en una cadena de 264 bits, 8 bits para cada una de las 33 conexiones. Lo que permite tener un rango de 256 valores por conexión en un intervalo de -1 a 1. La codificación de cada uno de los pesos sinápticos a un número real  $r$  en el rango  $[min, max]$  se realizó aplicando la ecuación 3.2.

$$r = min + \left(\frac{i}{n}\right)(max - min) \quad (3.2)$$

donde  $i$  es la cadena de bits decodificada a un número entero y  $n$  es el valor entero máximo representado por la cadena de bits.

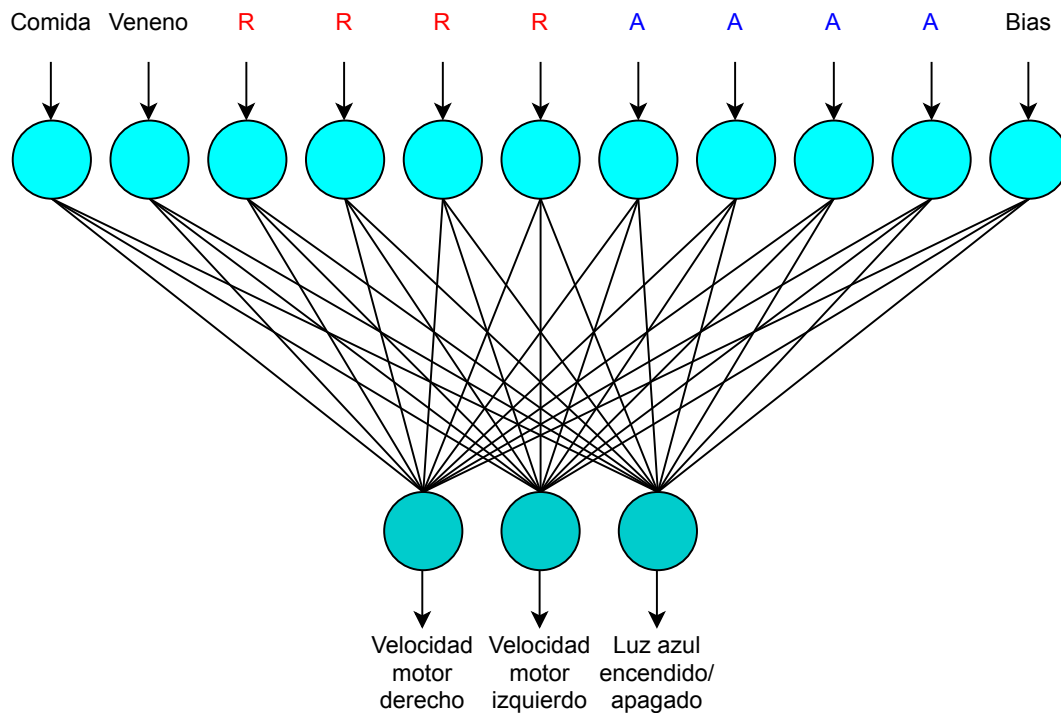


Figura 3.5: Arquitectura de red neuronal. Las primeras dos neuronas de entrada son activadas con el sensor de piso (comida o veneno). La cámara de 360° se dividió en cuatro secciones de 180 píxeles, el promedio de los canales azul y rojo se calculó para cada sección y se normalizó dentro del rango de 0 a 1. Las neuronas de salida con una función continua  $\tanh(x)$  codifican la velocidad de los motores y la emisión de luz azul.

Por otro lado, dos de las neuronas de entrada fueron conectadas al sensor de piso, la primera para detectar la comida y la segunda para detectar el veneno. Si el sensor de piso detectaba alguna de estas fuentes la entrada de la neurona correspondiente era establecida en 1, de lo contrario era establecida en 0. Ocho de las neuronas se utilizaron para codificar la imagen de entrada visual de la cámara de 360°, la cual se dividió en cuatro secciones de 180 píxeles cada una. Para cada sección, el promedio de los canales azul y rojo se calculó y normalizó dentro del rango de 0 y 1, de modo que se usaron cuatro entradas neuronales para el valor rojo y cuatro para el valor azul. Finalmente, la neurona restante se estableció como una neurona “bias” con una entrada de 1 en cualquier condición, esto con la intención de permitir a los robots producir algún comportamiento, incluso si no era percibido dato sensorial alguno.

### 3.1.4. Algoritmo evolutivo

Para todos los experimentos evolutivos se utilizaron poblaciones de 100 genomas, cada generación dividida en grupos de 10 individuos que vivían al mismo tiempo. Después de evaluar el rendimiento de cada individuo, los 100 genomas de la generación se clasificaron linealmente de acuerdo con el valor de aptitud. Posteriormente, se aplicó el algoritmo de selección por ruleta para determinar los genomas candidatos para la reproducción. Por esta razón, la probabilidad  $p_i$  de un individuo de ser seleccionado se calculó utilizando la ecuación 3.3.

$$p_i = \frac{f_i}{\sum_{j=1}^n f_j} \quad (3.3)$$

donde  $f_i$  es el valor de aptitud del individuo  $i$  y  $n$  el número de individuos de la población.

Los genomas seleccionados se emparejaron aleatoriamente para realizar el cruce con un 20% de probabilidad de ser combinados en 2 puntos de cruce aleatorios, generando dos nuevos genomas. Finalmente, se aplicó una mutación genética con una probabilidad del 1% para cada bit. Si la probabilidad de cruce no se alcanzaba, se generaban dos copias de cada genoma emparejado y se aplicaba la misma función de mutación. Todos los genomas resultantes se colocaban en el conjunto genético de

la próxima generación, este proceso era repetido a lo largo de  $n$  generaciones (Ver figura 3.6)

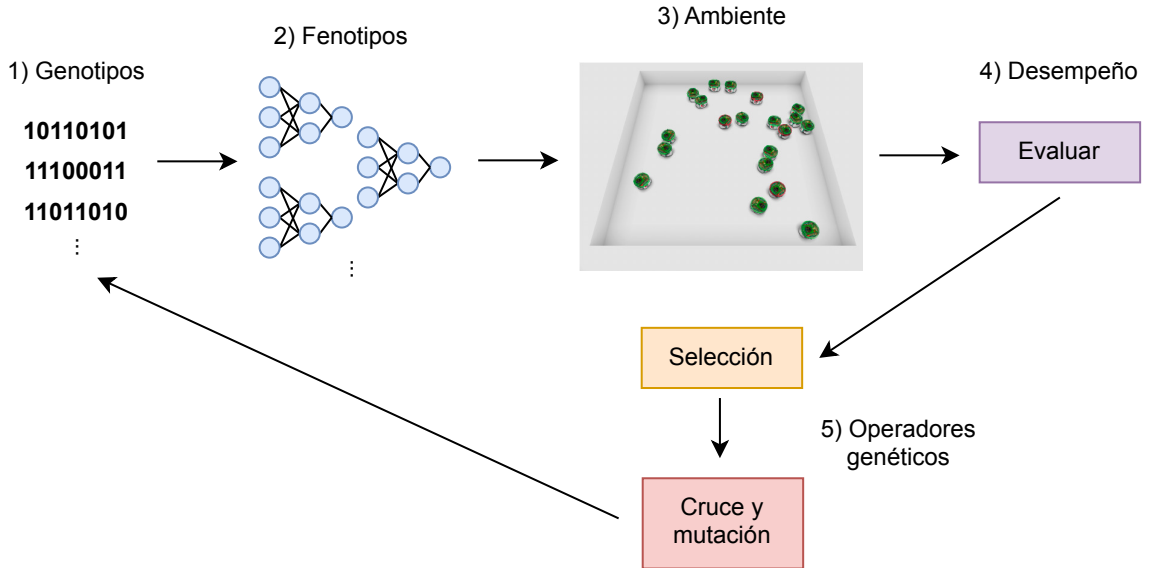


Figura 3.6: Proceso de selección artificial en robots. 1) Se genera una población inicial de genomas que codifican los pesos sinápticos de la RNA. 2) Cada genoma es traducido en su fenotipo y expresado en una RNA. 3) Los robots son colocados en el ambiente y se les permite vivir por un determinado número de ciclos. 4) Al terminar su ciclo de vida el desempeño de cada robot es evaluado. 5) Después de aplicar la evaluación se aplican los operadores genéticos correspondientes (selección, cruce y mutación) y se genera la siguiente población de genomas.

## 3.2. Forrajeo

Esta condición experimental fue diseñada para desarrollar en los robots una conducta en donde resultara benéfico en términos de aptitud evitar el área de veneno, localizar el área de comida y permanecer en ella.

Durante el proceso evolutivo, cada simulación de cada grupo de 10 robots se evaluó durante 1200 ciclos. El rendimiento de cada robot se calculó con la ecuación 3.4:

$$F = \frac{\sum_{i=1}^T t_f(i) - \sum_{i=1}^T t_p(i)}{T}, \quad \text{si } F < 0 \therefore F = 0 \quad (3.4)$$

donde  $t_f$  son los ciclos que el robot permaneció sobre el área de comida,  $t_p$  los ciclos que el robot permaneció sobre el área de veneno y  $T$  el total de ciclos evaluados en cada simulación.

Cabe destacar que el anillo de LED's del robot fue controlado por la RNA incluso cuando no desempeñó ninguna función en esta tarea. Asimismo, la emisión de luz no es penalizada a través de la función de aptitud.

El algoritmo evolutivo se aplicó durante 250 generaciones y el proceso se repitió 20 veces, lo cual resultó en 20 generaciones finales. Por lo tanto, se tomaron los mejores 600 genomas (30 para cada una de las generaciones finales) para formar una familia.

### 3.2.1. Prueba de rendimiento

Se diseñaron tres tipos de escenarios de prueba cada uno con una región de interés (ROI, por sus siglas en inglés *Region of Interest*):

1. Ausencia de luz.
2. Luz roja.
3. Luz azul.

De la muestra total de 600 genomas, 200 de ellos fueron asignados aleatoriamente a una de las tres condiciones experimentales correspondientes a cada área de prueba.

La prueba de rendimiento tenía como objetivo medir el tiempo total que los individuos pasaron en la ROI. Para descartar que los robots utilizaran la ROI como información para realizar las tareas, el sensor de piso fue inhabilitado (las entradas neuronales fueron establecidas a cero) en todas las condiciones. El tiempo total que un individuo pasó en la ROI durante una época (1200 ciclos de vida) fue normalizado a valores entre 0 y 1. Con base en ello, se analizó el tiempo total normalizado que los individuos pasaron en cada ROI.

#### Ausencia de luz

En esta condición se cuantificó el tiempo que los robots pasaron sobre la ROI cuando el ambiente no contenía ninguna información sensorial. El entorno consistió

en una arena cuadrada (Ver sección 3.1.2) con una única ROI delimitada por un área circular de color negro de 20cm de diámetro colocada en el centro de la arena (Ver figura 3.7). Los 200 robots fueron colocados uno a la vez en una esquina de la arena y la simulación fue ejecutada. Es importante tener en cuenta que no existió ninguna pista que indicara la posición de la ROI a los robots.

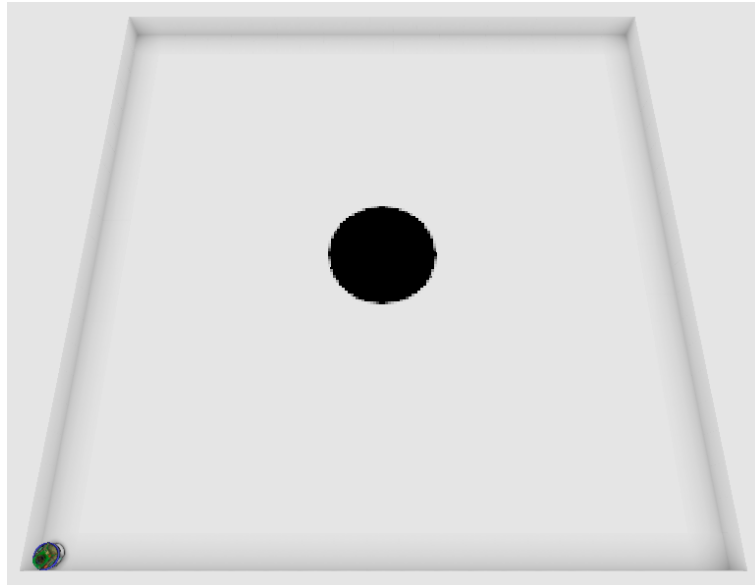


Figura 3.7: Prueba de rendimiento con ausencia de luz. Mapa del entorno para la prueba de rendimiento con ausencia de luz, el círculo negro representa la ROI.

### Luz roja

Al igual que en la condición anterior, el entorno consistió en una arena cuadrada con una única ROI delimitada por un área circular de color negro de 20cm de diámetro colocada en el centro de la arena. Sin embargo, a diferencia de la condición anterior se colocó una fuente de luz roja en el centro de la ROI para indicar la posición de la ROI a los robots (Ver figura 3.8). Los 200 robots fueron colocados uno a la vez en una esquina de la arena y la simulación fue ejecutada.

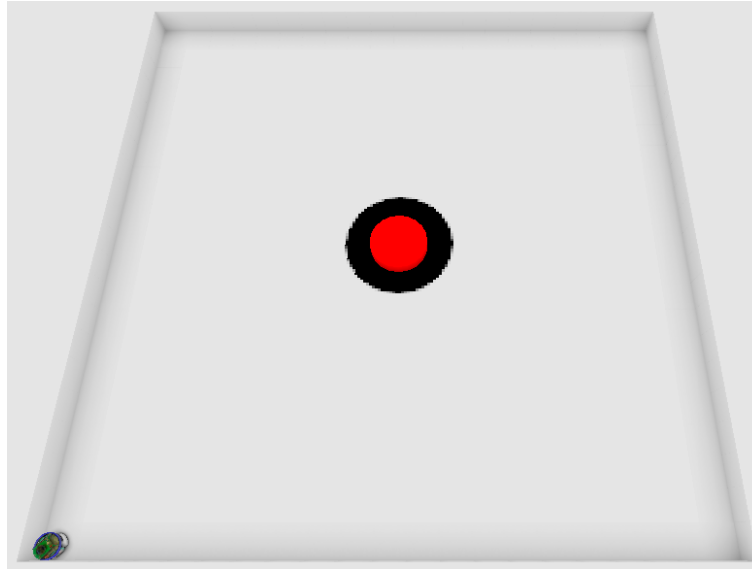


Figura 3.8: Prueba de rendimiento con luz roja. Mapa del entorno para la prueba de rendimiento con luz roja, el círculo negro representa la ROI.

### Luz azul

Para esta condición se utilizó el mismo entorno de la condición con luz roja. No obstante, a diferencia de esta última, la fuente de luz colocada en el centro de la ROI consistía en una luz de color azul. De igual manera que la condición anterior, la fuente de luz azul tenía la función de indicar la posición de la ROI a los robots (Ver figura 3.9). Los 200 robots fueron colocados uno a la vez en una esquina de la arena y la simulación fue ejecutada.

### 3.2.2. Prueba de señalización

Para analizar la preferencia que cada familia desarrolló para cada tipo de luz (azul y roja) durante el proceso evolutivo, se diseñaron dos escenarios de prueba:

1. Luz azul/roja contra luz roja
2. Luz azul/roja contra luz azul

Cada uno de ellos con una ROI por cada fuente de luz. Ambas pruebas fueron diseñadas para analizar la atracción que producen en cada robot cierto estímulo



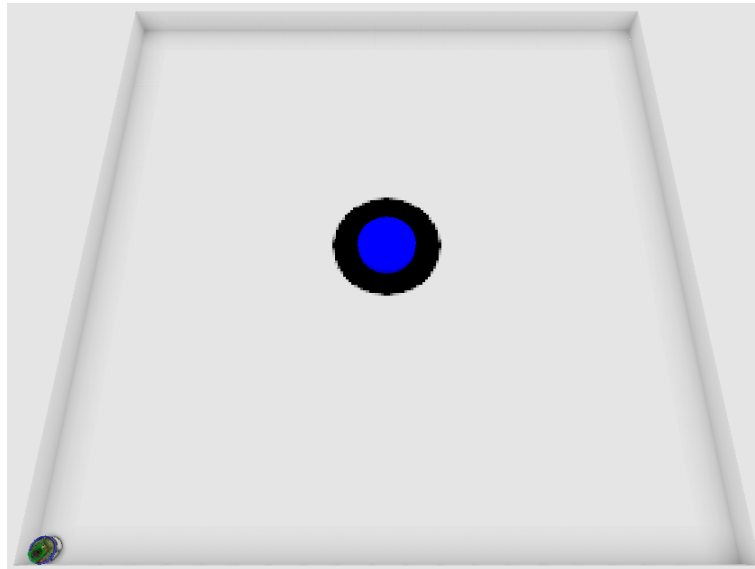


Figura 3.9: Prueba de rendimiento con luz azul. Mapa del entorno para la prueba de rendimiento con luz azul, el círculo negro representa la ROI.

visual y el impacto de ella en la tarea a resolver. La muestra total de 600 genomas fueron evaluados en ambas condiciones.

Al igual que en la prueba de rendimiento, la prueba de señalización tenía como objetivo medir el tiempo total que los individuos pasaron en cada ROI. Asimismo, para descartar que los robots utilizaran la ROI como información para realizar las tareas, el sensor de piso fue inhabilitado (las entradas neuronales fueron establecidas a cero) en todas las condiciones. El tiempo total que un individuo pasó en la ROI durante una época (1200 ciclos de vida) fue normalizado a valores entre 0 y 1. Con base en ello, se analizó el tiempo total normalizado que los individuos pasaron en cada ROI.

### Luz azul/roja contra roja

El entorno consistió en una arena cuadrada (Ver sección 3.1.2), la cual contenía dos fuentes de luz: azul/roja y roja; colocadas a 50cm de una de las dos esquinas opuestas, ambas con su respectiva ROI (Ver figura 3.10). Cada uno de los robots se colocó en la esquina inferior derecha de la arena a la misma distancia de ambas ROI. Los 600 robots fueron colocados uno a la vez en una esquina de la arena y la

simulación fue ejecutada.

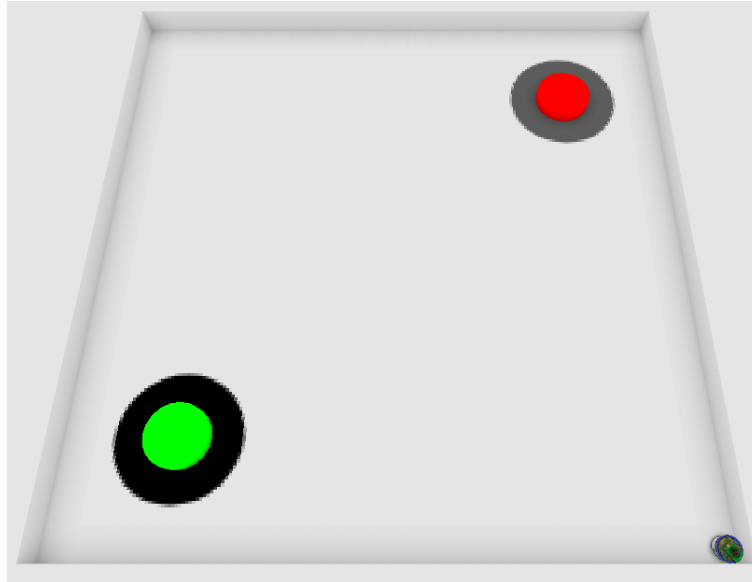


Figura 3.10: Prueba de señalización con luz azul/roja contra roja. Mapa del entorno para la prueba de señalización con luz azul/roja contra roja, cada círculo negro representa una ROI. La fuente verde representa la fuente azul/roja.

### Luz azul/roja contra azul

Para esta condición se utilizó el mismo entorno de la condición con luz azul/roja y roja. No obstante, a diferencia de esta última, el entorno contenía dos fuentes de luz: azul/roja y azul. De igual manera, ambas fuentes fueron colocadas a 50cm de una de las dos esquinas opuestas, ambas con su respectiva ROI (Ver figura 3.11). Cada uno de los robots se colocó en la esquina inferior derecha de la arena a la misma distancia de ambas ROI. Los 600 robots fueron colocados uno a la vez en una esquina de la arena y la simulación fue ejecutada.

## 3.3. Señalización

Esta condición experimental fue diseñada con la intención de desarrollar en los robots una conducta en donde les fuera benéfico en términos de aptitud, emitir luz azul sobre el área de comida. Es decir, los robots debían encender su anillo de LED's

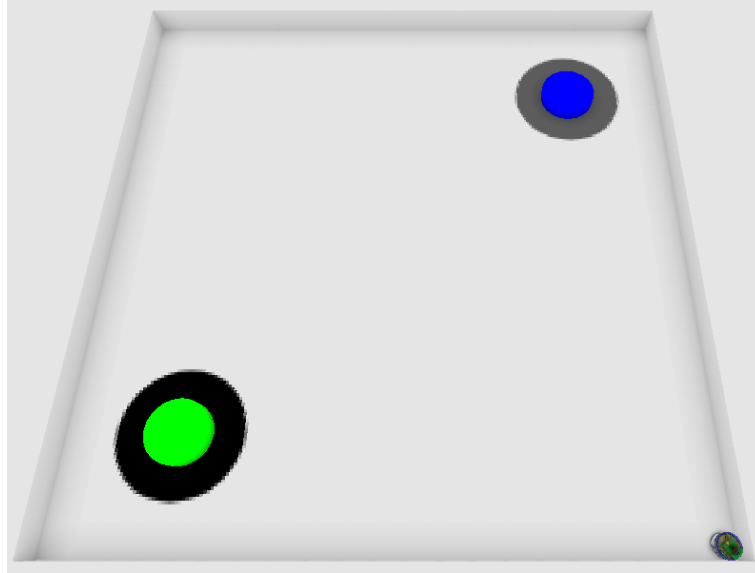


Figura 3.11: Prueba de señalización con luz azul/roja contra azul. Mapa del entorno para la prueba de señalización con luz azul/roja contra azul, cada círculo negro representa una ROI. La fuente verde representa la fuente azul/roja.

cuando se posicionaban sobre el área de comida y evitar esta conducta en cualquier otro lugar de la arena. Durante el proceso evolutivo cada simulación de cada grupo de robots se evaluó durante 1200 ciclos.

Toda señal se transmite por un canal sensorial determinado, lo cual implica un costo (Laidre and Johnstone, 2013; Számadó, 2011). Asimismo, cada canal sensorial posee propiedades específicas de acuerdo al canal empleado para su transmisión (Ver tabla 2.1). En consecuencia, el rendimiento de cada robot se calculó utilizando la ecuación 3.5:

$$F = \frac{\sum_{i=1}^T l_f(i) - \sum_{i=0}^T l_p(i) - \alpha \sum_{i=0}^T l_s(i)}{T}, \quad \text{si } F < 0 \therefore F = 0 \quad (3.5)$$

donde  $l_f$  son los ciclos que el robot emitió luz azul sobre el área de comida,  $l_p$  los ciclos que el robot emitió luz azul sobre el área de veneno,  $l_s$  los ciclos que el robot emitió luz azul,  $\alpha$  es una constante igual a 0.04 que representa el costo de emisión de la señal y  $T$  el total de ciclos evaluados en cada simulación.

A diferencia de Mitri et al. (2010b), la función de desempeño no considera los ciclos que el robot permanece sobre alguna área de la arena sin emitir luz azul. De

manera que, permanecer sobre la fuente de comida sin emitir luz azul no implicaba un incremento en el desempeño de los robots.

Al igual que la condición de forrajeo, el algoritmo evolutivo se aplicó durante 250 generaciones y el proceso se repitió 20 veces, lo cual resultó en 20 generaciones finales. Por lo tanto, se tomaron los mejores 600 genomas (30 para cada una de las generaciones finales) para formar una familia. Asimismo, se aplicaron exactamente las mismas pruebas de rendimiento y señalización que en la condición de forrajeo.

### 3.3.1. Prueba de rendimiento

De la muestra total de 600 genomas, 200 de ellos fueron asignados aleatoriamente a una de las tres condiciones experimentales correspondientes a cada área de prueba:

1. Ausencia de luz.
2. Luz roja.
3. Luz azul.

Se utilizaron exactamente los mismos escenarios de prueba y los mismos parámetros de las pruebas de rendimiento anteriores (Ver sección 3.2.1). Con base en ello, se analizó el tiempo total normalizado que los individuos pasaron en cada ROI.

### 3.3.2. Prueba de señalización

Para analizar la preferencia que cada individuo desarrolló para cada tipo de luz (azul y roja) durante el proceso evolutivo, la muestra total de 600 genomas fueron evaluados en dos condiciones experimentales:

1. Luz azul/roja contra luz roja
2. Luz azul/roja contra luz azul

Se utilizaron exactamente los mismos escenarios de prueba y los mismos parámetros de las pruebas de rendimiento anteriores (Ver sección 3.2.2). Con base en ello, se analizó el tiempo total normalizado que los individuos pasaron en cada ROI.

## 3.4. Resultados

### 3.4.1. Forrajeo

Como resultado del proceso evolutivo, el desempeño de la tarea aumentó considerablemente durante las 250 generaciones (Ver figura 3.12). Para evaluar el desempeño real de los individuos finales se evaluaron los 100 mejores de la última generación de las 20 simulaciones. Como resultado (Ver figura 3.13), se observó que durante los 1200 ciclos evaluados la mayoría de los individuos lograron resolver la tarea de forrajeo. Con respecto a la señalización, dado que se implementó una función de desempeño sin costo ni beneficio (Ver ecuación 3.4), la acción de señalar o no señalar no afectó directamente el desempeño de los robots. Por lo cual, durante los ciclos de simulación no se observó tendencia a señalar el área de comida por parte de los 100 individuos evaluados, ya que la señalización se realizó en las diferentes áreas de la arena (comida y/o veneno y/o fuera de la región de interés).

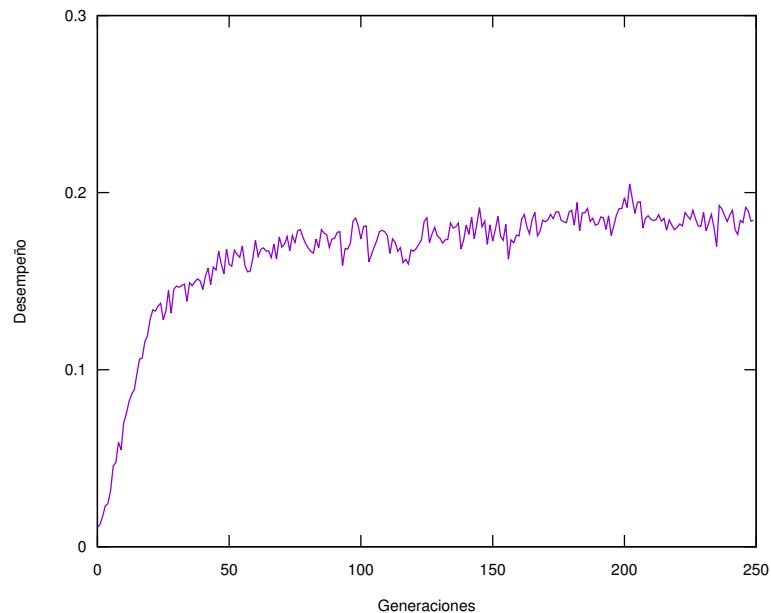


Figura 3.12: Promedio del desempeño de las 20 simulaciones durante el proceso evolutivo en la familia de forrajeo.

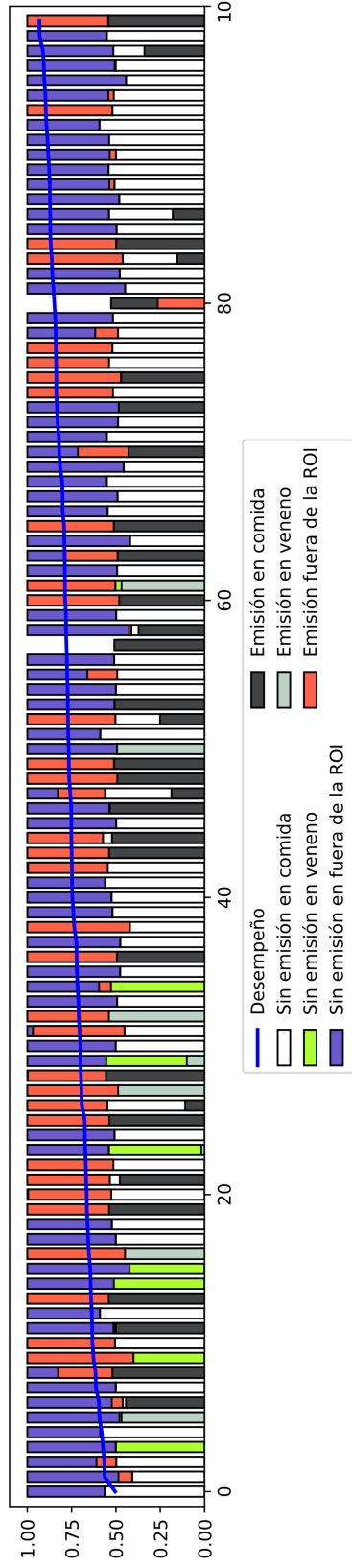


Figura 3.13: Tiempo invertido por los 100 mejores individuos en cada área de la arena durante los 1200 ciclos de simulación en la familia de forrajeo.

### Análisis de la prueba de rendimiento

Hubo diferencias estadísticamente significativas entre las medias de los grupos según lo determinado por el ANOVA de Welch de un factor ( $F(2, 279.247) = 224.397, p < 0.001$ ). Una prueba post hoc de Games-Howell reveló que el tiempo total invertido en el ROI fue significativamente mayor en la condición de luz roja ( $0.379 \pm SEM 0.017, p < 0.001$ ) y en la condición de luz azul ( $0.033 \pm SEM 0.007, p = 0.002$ ) en comparación con la condición de ausencia de luz ( $0.005 \pm SEM 0.001$ ). Además, el tiempo total invertido en la ROI fue significativamente mayor en la condición de luz roja ( $p < 0.001$ ) en comparación con la condición de luz azul.

### Análisis de la prueba de señalización

Debido a que había dos ROI en cada área, se obtuvieron medidas repetidas del tiempo total que un individuo pasó en cada ROI. Una prueba  $t$  de dos muestras emparejadas reveló que, en promedio, los individuos de la condición de luz azul/roja contra luz roja pasaron significativamente más tiempo en el ROI de la fuente roja ( $0.205 \pm SEM 0.009$ ) que en la ROI de la fuente azul/roja ( $0.109 \pm SEM 0.008$ ),  $t(599) = -6.239, p < 0.001, r = 0.247$ . En la condición de luz azul/roja contra la luz azul, los individuos pasaron significativamente más tiempo en la ROI de la fuente azul/roja ( $0.204 \pm SEM 0.009$ ) que en la ROI de la fuente azul ( $0.008 \pm SEM 0.002$ ), determinado por una prueba  $t$  de dos muestras emparejadas,  $t(599) = 18.954, p < 0.001, r = 0.612$ .

Debido a que la acción de señalar o no señalar no afectó directamente el desempeño de los individuos, la emisión de luz azul no tuvo efecto alguno en el proceso evolutivo. Por lo cual, de acuerdo a los resultados el impacto de la respuesta a la luz azul fue menor que la respuesta a la luz roja emitida por los comederos.

#### 3.4.2. Señalización

Como resultado del proceso evolutivo, el desempeño de la tarea aumentó considerablemente durante las 250 generaciones (Ver figura 3.14). Para evaluar el desempeño real de los individuos finales se evaluaron los 100 mejores de la última generación de las 20 simulaciones. Como resultado (Ver figura 3.15), se observó que durante los

1200 ciclos de simulación la mayoría de los individuos lograron localizar la fuente de comida al igual que los individuos de la familia de forrajeo (Ver sección 3.4.1). Dado que la señalización tuvo un impacto en la aptitud de los individuos e implicó un costo (Ver ecuación 3.5), los robots con mejor desempeño fueron capaces de señalar el área de comida y evitar este comportamiento en las demás áreas. Sin embargo, se observó que algunos individuos evolucionaron la estrategia de señalar el área de veneno y/o el área fuera de las regiones de interés (comida y veneno).

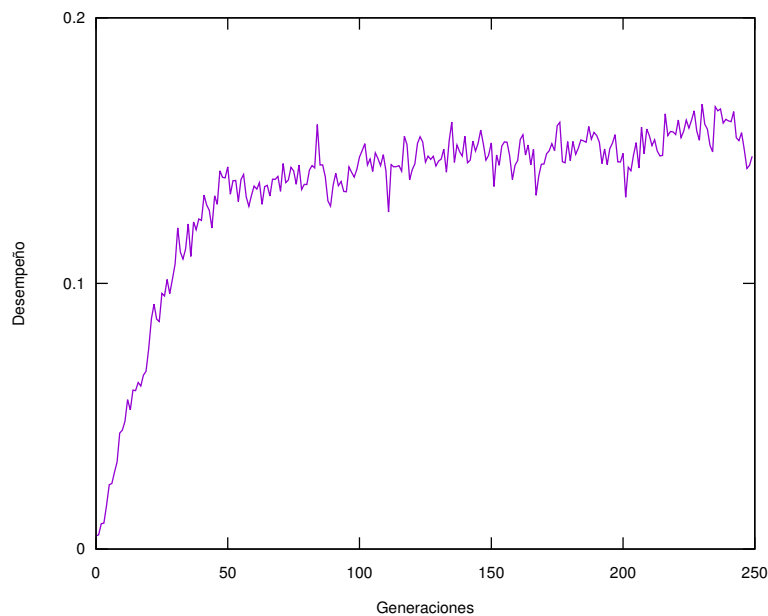


Figura 3.14: Promedio del desempeño de las 20 simulaciones durante el proceso evolutivo en la familia de señalizadores en contexto fijo.

### Análisis de la prueba de rendimiento

Hubo diferencias estadísticamente significativas entre las medias de los grupos según lo determinado por el ANOVA de Welch de un factor ( $F(2, 274.416) = 115.220, p < 0.001$ ). Una prueba post hoc de Games-Howell reveló que el tiempo total invertido en la ROI fue significativamente mayor en la condición de luz roja ( $0.275 \pm SEM 0.018, p < 0.001$ ) y en la condición de luz azul ( $0.046 \pm SEM 0.01, p < 0.001$ ) en comparación con la condición de ausencia de luz ( $0.006 \pm SEM 0.001$ ). Además, el tiempo total invertido en la ROI fue significativamente mayor en la condición de



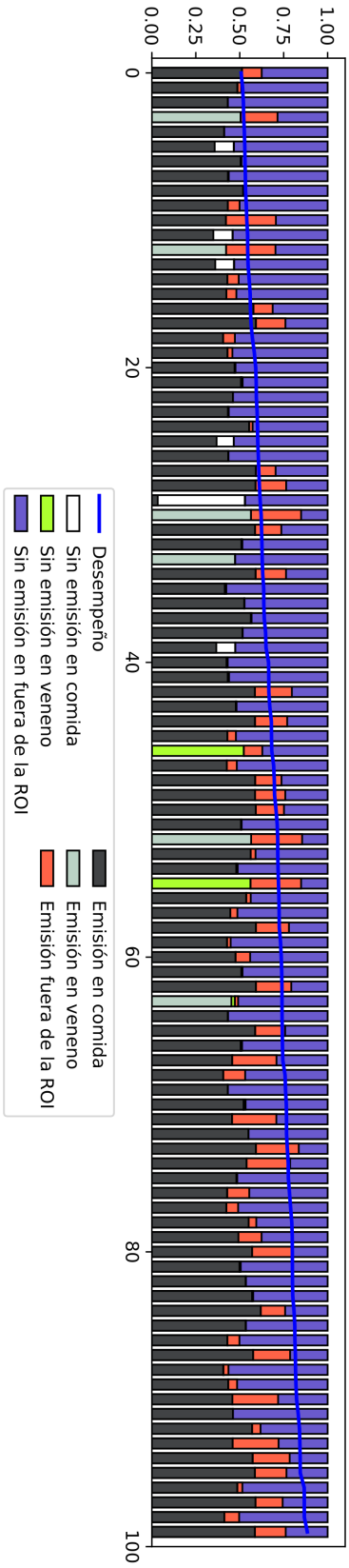


Figura 3.15: Tiempo invertido por los 100 mejores individuos en cada área de la arena durante los 1200 ciclos de simulación en la familia de señalizadores en contexto fijo.

luz roja ( $p < 0.001$ ) en comparación con la condición de luz azul.

### **Análisis de la prueba de señalización**

Debido a que había dos ROI en cada área, se obtuvieron medidas repetidas del tiempo total que un individuo pasó en cada ROI. Una prueba  $t$  de dos muestras emparejadas reveló que, en promedio, los individuos de la condición de luz azul/roja contra luz roja, pasaron significativamente más tiempo en la ROI de la fuente azul/roja ( $0.134 \pm SEM 0.008$ ) que en la ROI de la fuente roja ( $0.089 \pm SEM 0.007$ ),  $t(599) = 3.586$ ,  $p < 0.001$ ,  $r = 0.144$ . En la condición de luz azul/roja contra luz azul, los individuos pasaron significativamente más tiempo en la ROI de la fuente azul/roja ( $0.183 \pm SEM 0.009$ ) que en la ROI de la fuente azul ( $0.007 \pm SEM 0.002$ ), determinado por una prueba  $t$  de dos muestras emparejadas,  $t(599) = 18.204$ ,  $p < 0.001$ ,  $r = 0.596$ .

Al evaluar la respuesta de los individuos ante un único estímulo visual, los resultados mostraron que los individuos desarrollaron una preferencia por la luz roja en comparación con la luz azul. Sin embargo, una muestra significativa desarrolló una atracción hacia la luz azul/roja en comparación con la luz roja al evaluar más de un estímulo visual al mismo tiempo. Lo anterior debido a que la emisión de luz azul tuvo un efecto durante el proceso evolutivo.

---

## Capítulo 4

# Evolución de la comunicación en contexto aleatorio

De acuerdo a los resultados obtenidos en el experimento de contexto fijo (Ver sección 3.4), el experimento de esta sección sólo abordó la condición experimental de señalización. Para evitar que los robots aprendieran la posición directa de los comederos, se implementó el uso de un contexto aleatorio. Posteriormente, se aplicó la prueba de rendimiento y de señalización para comprobar si algún tipo de estrategia básica de comunicación había emergido durante el proceso evolutivo.

### 4.1. Métodos

El agente artificial, los controladores neuronales y el algoritmo evolutivo utilizado en esta sección, son idénticos a los descritos en la sección 3.1.

#### 4.1.1. Configuración experimental

Parte de la configuración experimental que se utilizó en este experimento se describe en la sección 3.1.2. La única diferencia implicaba un contexto aleatorio, es decir, el entorno cambiaba constantemente a lo largo del proceso evolutivo. Para ello, la arena se seleccionó aleatoriamente de un conjunto de 32 mapas predefinidos. Para cada mapa las fuentes de comida y veneno se colocaron en diferentes posiciones, con

una distancia entre ambas fuentes en un rango de  $1.4m$  a  $1.8m$  (Ver figura 4.1)

Es importante señalar que se implementó un contexto aleatorio para evitar que los robots no solo aprendieran la posición directa de los comederos. Es decir, si bien la tarea era la misma que en un entorno fijo, durante el proceso evolutivo los robots tenían la posibilidad de explorar un mayor número de soluciones para la misma tarea.

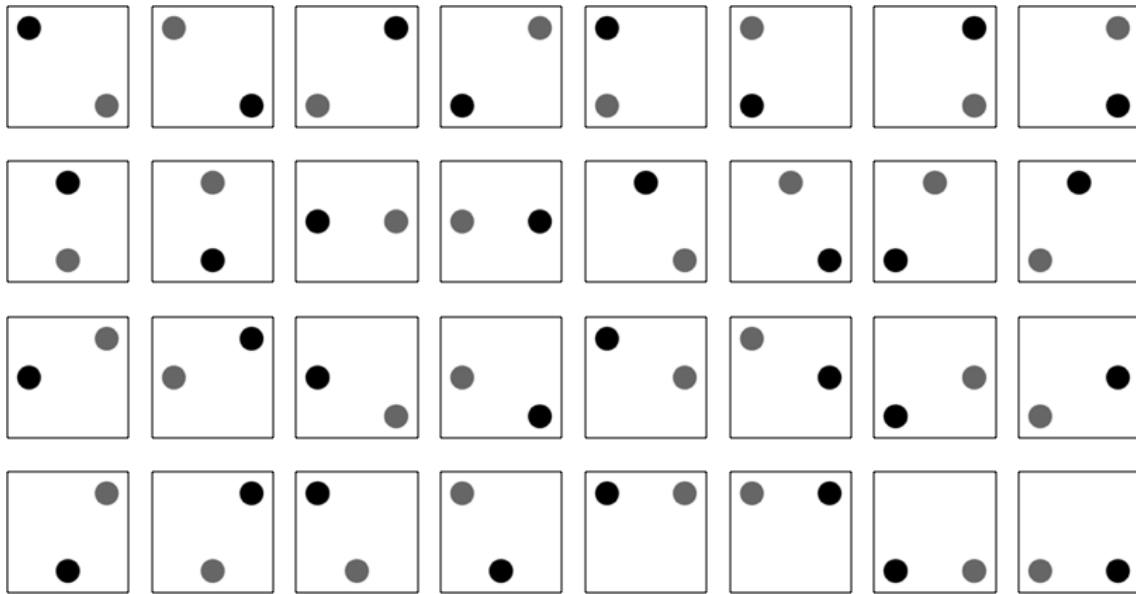


Figura 4.1: Ilustración de los 32 posibles mapas durante el contexto aleatorio. Cada círculo de color representa una ROI: negro para comida y gris para veneno.

## 4.2. Señalización

De igual modo que la sección 3.3, esta condición experimental fue diseñada con la intención de desarrollar en los robots una conducta en donde les fuera benéfico en términos de aptitud, emitir luz sobre el área de comida. Esto es, los robots debían encender su anillo de LED's cuando se posicionaban sobre el área de comida y evitar esta conducta en cualquier otro lugar de la arena. Por lo cual, se utilizó la ecuación 3.5 para evaluar el desempeño de cada individuo.

Durante cada generación y evaluación de cada grupo de robots, la arena se seleccionó aleatoriamente del conjunto de 32 mapas predefinidos (Ver figura 4.1). El

proceso evolutivo en cada simulación de cada grupo de robots se evaluó durante 1200 ciclos.

El algoritmo evolutivo se aplicó durante 250 generaciones y el proceso se repitió 20 veces, lo cual resultó en 20 generaciones finales. Por lo tanto, se tomaron los mejores 600 genomas (30 para cada una de las generaciones finales) para formar una familia.

### 4.2.1. Prueba de rendimiento

De la muestra total de 600 genomas, 200 de ellos fueron asignados aleatoriamente a una de las tres condiciones experimentales correspondientes a cada área de prueba:

1. Ausencia de luz.
2. Luz roja.
3. Luz azul.

Se utilizaron exactamente los mismos escenarios de prueba y los mismos parámetros de las pruebas de rendimiento anteriores (Ver sección 3.2.1). Con base en ello, se analizó el tiempo total normalizado que los individuos pasaron en cada ROI.

### 4.2.2. Prueba de señalización

Para analizar la preferencia que cada individuo desarrolló para cada tipo de luz (azul y roja) durante el proceso evolutivo, la muestra total de 600 genomas fueron evaluados en dos condiciones experimentales:

1. Luz azul/roja contra luz roja
2. Luz azul/roja contra luz azul

Se utilizaron exactamente los mismos escenarios de prueba y los mismos parámetros de las pruebas de rendimiento anteriores (Ver sección 3.2.2). Con base en ello, se analizó el tiempo total normalizado que los individuos pasaron en cada ROI.

## 4.3. Resultados

### 4.3.1. Señalización

De igual modo que la sección 3.4.2, como resultado del proceso evolutivo el desempeño de la tarea aumentó considerablemente durante las 250 generaciones (Ver figura 4.2). Sin embargo, los individuos de contexto aleatorio presentaron un mejor desempeño durante las 250 generaciones (Ver figura 4.3) en comparación con los individuos de contexto fijo. Para evaluar el desempeño real de los individuos finales se evaluaron los 100 mejores de la última generación de las 20 simulaciones. Como resultado (Ver figura 4.4), la mayoría de los individuos lograron localizar la fuente de comida al igual que los individuos de la familia de forrajeo y contexto fijo (Ver secciones 3.4.1, 3.4.2). Los robots con mejor desempeño fueron capaces de señalar el área de comida y evitar este comportamiento en las demás áreas, ya que la señalización tuvo un impacto en la aptitud de los individuos e implicó un costo (Ver ecuación 3.5). No obstante, se observó que algunos individuos evolucionaron la estrategia de señalar el área de veneno y/o el área fuera de las regiones de interés (comida y veneno).

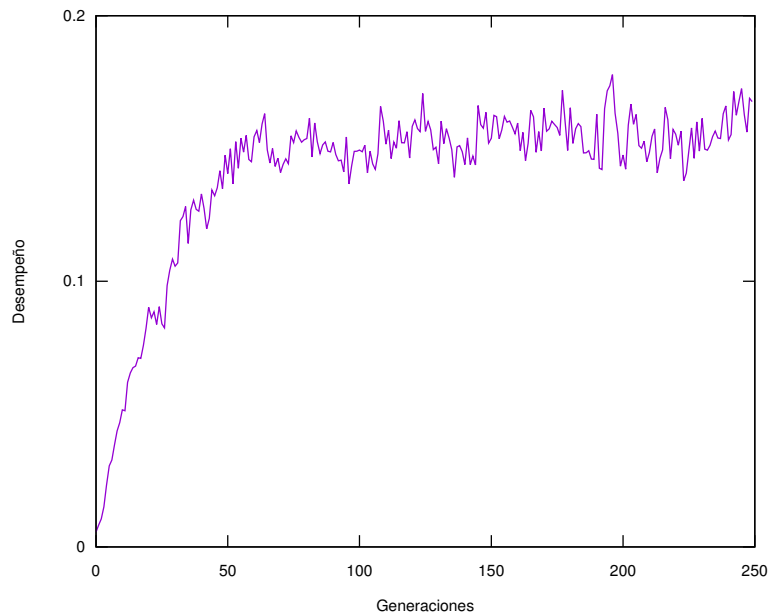


Figura 4.2: Promedio del desempeño de las 20 simulaciones durante el proceso evolutivo en la familia de señalizadores en contexto aleatorio.

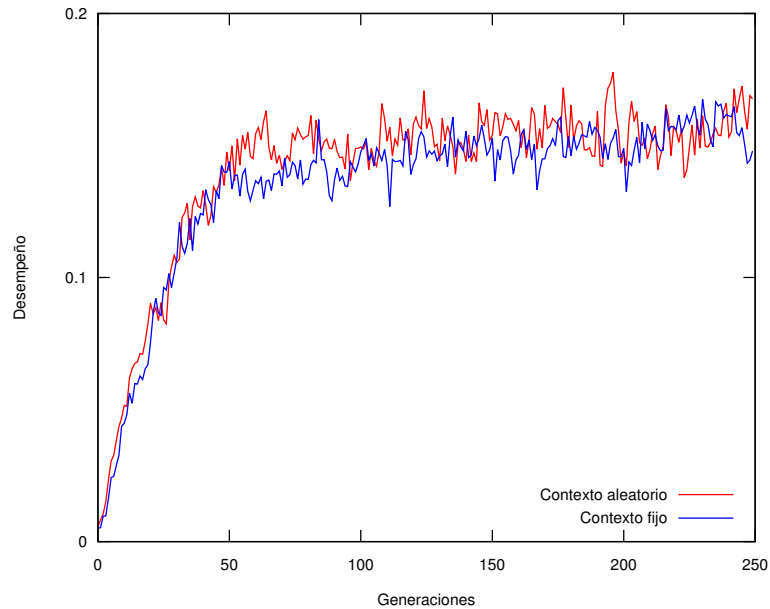


Figura 4.3: Comparativa entre la familia de señalizadores de contexto fijo y contexto aleatorio del promedio del desempeño de las 20 simulaciones durante el proceso evolutivo.

### Análisis de la prueba de rendimiento

Hubo diferencias estadísticamente significativas entre las medias de los grupos según lo determinado por el ANOVA de Welch de un factor ( $F(2, 272.094) = 98.836, p < 0.001$ ). Una prueba post hoc de Games-Howell reveló que el tiempo total invertido en la ROI fue significativamente mayor en la condición de luz roja ( $0.245 \pm SEM 0.018, p < 0.001$ ) y en la condición de luz azul ( $0.073 \pm SEM 0.01, p < 0.001$ ) en comparación con la condición de ausencia de luz ( $0.004 \pm SEM 0.001$ ). Además, el tiempo total invertido en la ROI fue significativamente mayor en la condición de luz roja ( $p < 0.001$ ) en comparación con la condición de luz azul.

### Análisis de la prueba de señalización

Debido a que había dos ROI en cada área, se obtuvieron medidas repetidas del tiempo total que un individuo pasó en cada ROI. Una prueba  $t$  de dos muestras emparejadas reveló que, en promedio, los individuos de la condición de luz azul/roja contra luz roja, los individuos pasaron significativamente más tiempo en la ROI de

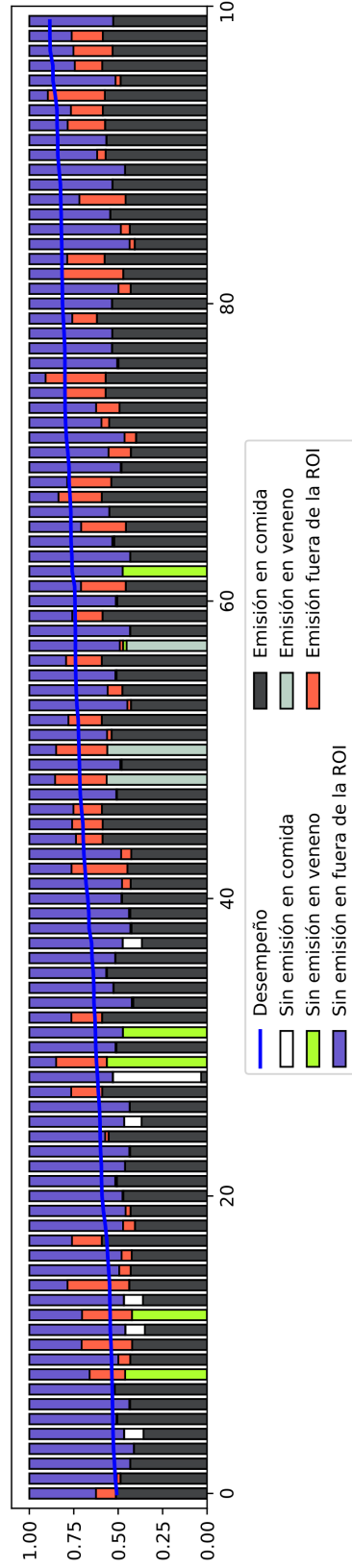


Figura 4.4: Tiempo invertido por los 100 mejores individuos en cada área de la arena durante los 1200 ciclos de simulación en la familia de señalizadores en contexto aleatorio.



la fuente azul/roja ( $0.154 \pm SEM 0.008$ ) que en la ROI de la fuente roja ( $0.112 \pm SEM 0.008$ ),  $t(599) = 2.972$ ,  $p = 0.003$ ,  $r = 0.12$ . En la condición de luz azul/roja contra luz azul, los individuos pasaron significativamente más tiempo en la ROI de la fuente azul/roja ( $0.209 \pm SEM 0.009$ ) que en la ROI de la fuente azul ( $0.011 \pm SEM 0.002$ ), determinado por una prueba  $t$  de dos muestras emparejadas,  $t(599) = 18.941$ ,  $p < 0.001$ ,  $r = 0.612$ .

Al igual que la familia de señalizadores de contexto fijo (Ver secciones 3.4.2) la emisión de luz azul tuvo un efecto durante el proceso evolutivo. La respuesta de los individuos ante un único estímulo visual mostró que desarrollaron una preferencia por la luz roja en comparación con la luz azul. Sin embargo, una muestra significativa desarrolló una atracción hacia la luz azul/roja en comparación con la luz roja al evaluar más de un estímulo visual al mismo tiempo.

## 4.4. Efectos sobre la señalización de los contextos fijo y aleatorio

Para dilucidar los efectos sobre la señalización de los contextos fijo y aleatorio en los cuales los señalizadores evolucionaron, se analizó el tiempo total invertido en la ROI de la fuente azul/roja en la prueba de señalización para ambas condiciones. Se utilizaron los datos de la muestra total de los 600 señalizadores evolucionados en el contexto fijo, y los datos de los 600 señalizadores evolucionados en el contexto aleatorio. En la condición de luz azul/roja contra luz roja, una prueba  $t$  de dos muestras independientes reveló que no hubo diferencias significativas en el tiempo promedio empleado en la ROI de la fuente azul/roja por parte de los individuos evolucionados en el contexto fijo, ( $0.134 \pm SEM 0.008$ ) en comparación con los individuos evolucionados en el contexto aleatorio ( $0.154 \pm SEM 0.008$ ),  $t(1195.915) = -1.637$ ,  $p = 0.102$ . En la condición de luz azul/roja vs luz azul, aunque no hubo diferencias significativas, se observó una ligera tendencia a pasar más tiempo en la ROI de la fuente azul/roja en individuos evolucionados en el contexto aleatorio ( $0.209 \pm SEM 0.009$ ) en comparación con los individuos evolucionados en el contexto fijo ( $0.183 \pm SEM 0.009$ ),  $t(1194.875) = -1.946$ ,  $p = 0.052$ .

Implementar un contexto aleatorio evitó que los robots aprendieran la posición

fija de los comederos, debido a que la arena se seleccionó aleatoriamente de un conjunto de mapas predefinidos. Es decir, aunque la tarea era la misma que en un entorno fijo, durante el proceso evolutivo los robots tenían la posibilidad de explorar un mayor número de soluciones para la misma tarea.

---

## Capítulo 5

# Efecto de la arquitectura del sistema de control y parámetros evolutivos

Con base en los resultados obtenidos en el experimento de contexto aleatorio (Ver sección 4.3), el experimento de esta sección se realizó para explorar cómo la arquitectura del sistema de control y los parámetros evolutivos intervienen en el proceso evolutivo de la comunicación. Por lo cual, después de abordar la condición experimental de señalización se aplicaron las mismas dos pruebas (rendimiento y señalización). Lo anterior, para comprobar si algún tipo de estrategia básica de comunicación había emergido durante el proceso evolutivo con los nuevos parámetros de configuración.

### 5.1. Métodos

El agente artificial y el algoritmo evolutivo utilizado en esta sección son idénticos a los descritos en la sección 3.1. La configuración experimental utilizada es idéntica a la descrita en la sección 4.1.1.

### **5.1.1. Controlador neuronal**

El sistema de control de cada robot consistió en una red neuronal *feedforward* con 41 neuronas de entrada, una capa oculta de 11 neuronas y 3 de salida, con un total de las 443 conexiones (Ver figura 5.1). Se utilizó la función tangente hiperbólica (3.1) como función de activación para cada una de las neuronas de la capa oculta y la capa de salida.

Dos de las neuronas de salida se usaron para controlar la velocidad de los motores. La tercera neurona de salida determinó la emisión de luz azul; que era el caso si la salida era mayor que 0.

Los pesos de las conexiones de la red neuronal fueron codificados en una cadena de 1772 bits, 4 bits para cada una de las 443 conexiones. Lo que permite tener un rango de 16 valores por conexión en un intervalo de -1 a 1. La codificación de cada uno de los pesos sinápticos se realizó aplicando la ecuación 3.2.

A su vez, ocho de las neuronas de entrada fueron conectadas al sensor de piso, las primeras cuatro para detectar la comida y las cuatro restantes para detectar el veneno. Si el sensor de piso detectaba alguna de estas fuentes las entradas de las neuronas correspondientes eran establecidas en 1, de lo contrario era establecidas en 0. Treinta y dos de las neuronas se utilizaron para codificar la imagen de entrada visual de la cámara de 360°, que se dividió en dieciséis secciones de 45 píxeles cada una. Para cada sección el promedio de los canales azul y rojo se calculó y normalizó dentro del rango de 0 y 1, de modo que se usaron dieciséis entradas neuronales para el rojo y dieciséis para el valor azul. Finalmente, tanto en la capa oculta como en la capa de salida, la neurona restante se estableció como una neurona “bias” con una entrada de 1 en cualquier condición. Esto es, con la intención de permitir a los robots producir algún comportamiento, incluso si no era percibido dato sensorial alguno.

## **5.2. Señalización**

Acorde con la sección 4.2, esta condición experimental fue diseñada con la intención de desarrollar en los robots una conducta que les permitiera emitir luz azul al estar posicionados sobre el área de comida y evitar esta conducta en cualquier otro lugar de la arena. Por lo cual, se utilizó la ecuación 3.5 para evaluar el desempeño

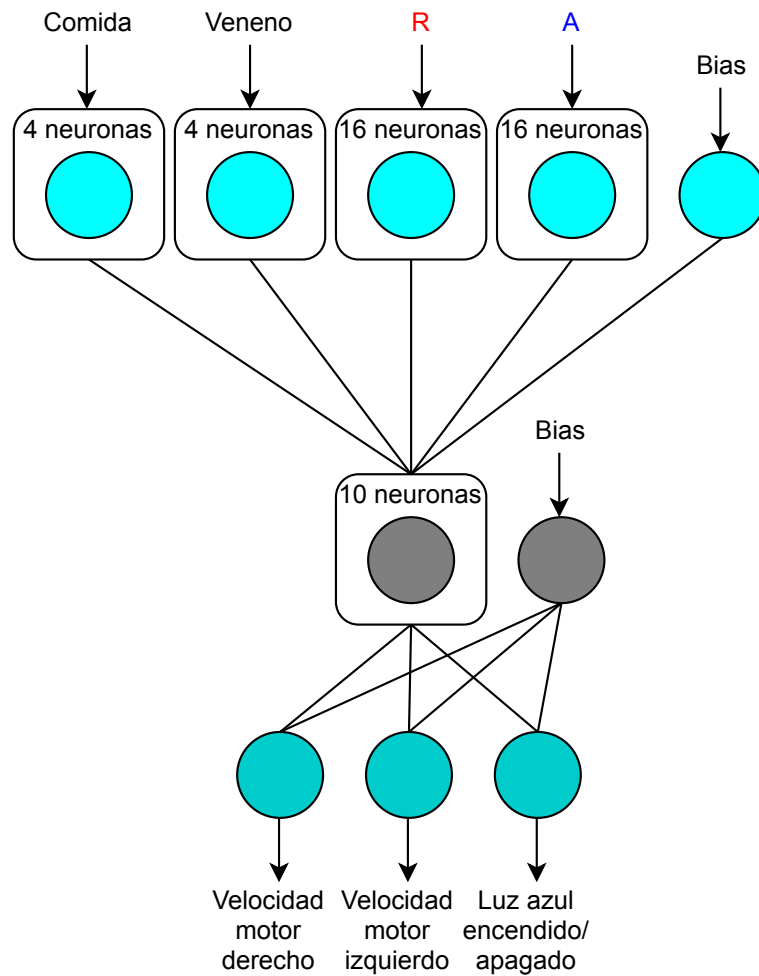


Figura 5.1: Arquitectura de red neuronal. Las primeras ocho neuronas de entrada son activadas con el sensor de piso (comida o veneno). La cámara de 360° se dividió en dieciséis secciones de 45 píxeles, el promedio de los canales azul y rojo se calculó para cada sección y se normalizó dentro del rango de 0 a 1. Las neuronas de salida con una función continua  $\tanh(x)$  codifican la velocidad de los motores y la emisión de luz azul.

de cada individuo.

Igualmente, la arena se seleccionó aleatoriamente del conjunto de 32 mapas pre-definidos (Ver figura 4.1) en cada generación y evaluación de cada grupo de robots. Durante el proceso evolutivo, cada simulación de cada grupo de robots se evaluó durante 3000 ciclos.

El algoritmo evolutivo se aplicó durante 500 generaciones y el proceso se repitió 20 veces, lo cual resultó en 20 generaciones finales. Por lo tanto, se tomaron los mejores 600 genomas (30 para cada una de las generaciones finales) para formar una familia.

### **5.2.1. Prueba de rendimiento**

De la muestra total de 600 genomas, 200 de ellos fueron asignados aleatoriamente a una de las tres condiciones experimentales correspondientes a cada área de prueba:

1. Ausencia de luz.
2. Luz roja.
3. Luz azul.

Del mismo modo, se utilizaron exactamente los mismos escenarios de prueba y los mismos parámetros de las pruebas de rendimiento anteriores (Ver sección 4.2.1). El tiempo total que un individuo pasó en la ROI durante una época (3000 ciclos de vida) fue normalizado a valores entre 0 y 1. Con base en ello, se analizó el tiempo total normalizado que los individuos pasaron en cada ROI.

### **5.2.2. Prueba de señalización**

Para analizar la preferencia que cada individuo desarrolló para cada tipo de luz (azul y roja) durante el proceso evolutivo, la muestra total de 600 genomas fueron evaluados en dos condiciones experimentales:

1. Luz azul/roja contra luz roja
2. Luz azul/roja contra luz azul

Debido a ello, se utilizaron exactamente los mismos escenarios de prueba y los mismos parámetros de las pruebas de rendimiento anteriores (Ver sección 4.2.2). El tiempo total que un individuo pasó en la ROI durante una época (3000 ciclos de vida) fue normalizado a valores entre 0 y 1. Con base en ello, se analizó el tiempo total normalizado que los individuos pasaron en cada ROI.

## 5.3. Resultados

### 5.3.1. Señalización

Debido a las características de este experimento (Ver sección 5), los señalizadores de contexto aleatorio (segunda instancia) de esta condición tenían la posibilidad de explorar el ambiente un mayor tiempo en comparación con las demás familias. Como resultado del proceso evolutivo el desempeño de la tarea aumentó considerablemente durante las 500 generaciones (Ver figura 5.2). En comparación con los individuos de contexto fijo y aleatorio, los individuos de esta condición presentaron un mejor desempeño en las primeras 250 generaciones (Ver figura 5.3), el cual se mantuvo hasta la generación 500. Para evaluar el desempeño real de los individuos finales se evaluaron los 100 mejores de la última generación de las 20 simulaciones. Como resultado (Ver figura 5.4), la mayoría de los individuos lograron localizar la fuente de comida al igual que los individuos de la familia de forrajeo, contexto fijo y aleatorio (Ver secciones 3.4.1, 3.4.2, 4.3.1). Debido a que la señalización tuvo un impacto en la aptitud de los individuos e implicó un costo (Ver ecuación 3.5), los robots con mejor desempeño fueron capaces de señalar el área de comida y evitar este comportamiento en las demás áreas. Sin embargo, se observó que algunos individuos evolucionaron la estrategia de señalar el área de veneno y/o el área fuera de las regiones de interés (comida y veneno).

### Análisis de la prueba de rendimiento

Hubo diferencias estadísticamente significativas entre las medias de los grupos según lo determinado por el ANOVA de Welch de un factor ( $F(2, 266.778) = 118.052, p < 0.001$ ). Una prueba post hoc de Games-Howell reveló que el tiempo total inver-

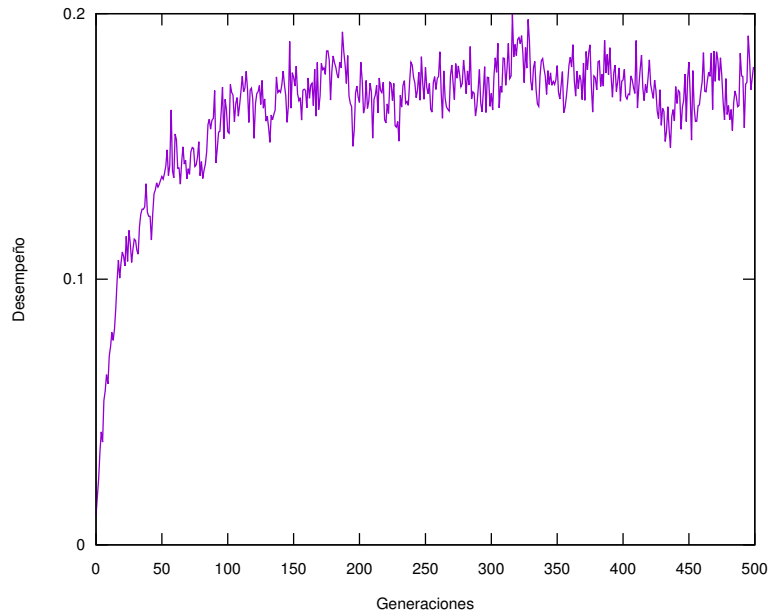


Figura 5.2: Promedio del desempeño de las 20 simulaciones durante el proceso evolutivo en la familia de señalizadores en contexto aleatorio (segunda instancia).

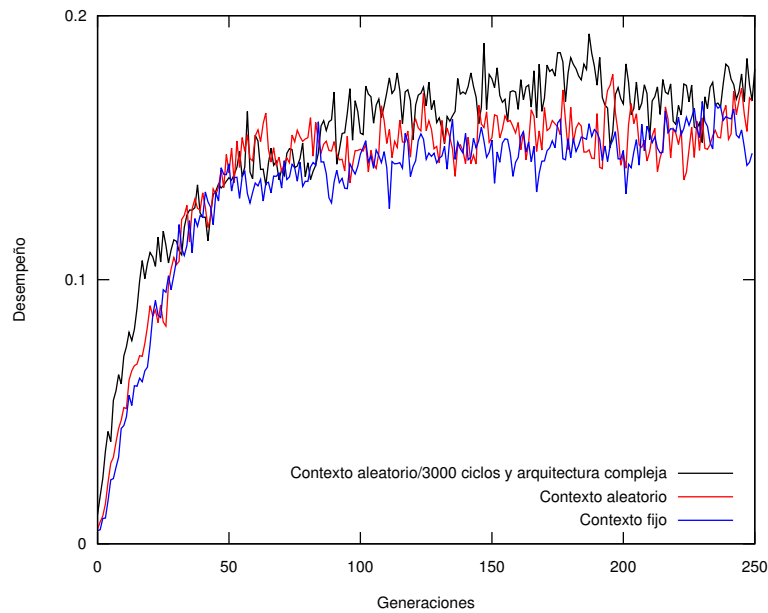


Figura 5.3: Comparativa entre las diferentes familias de señalizadores del promedio del desempeño de las 20 simulaciones durante las primeras 250 generaciones.



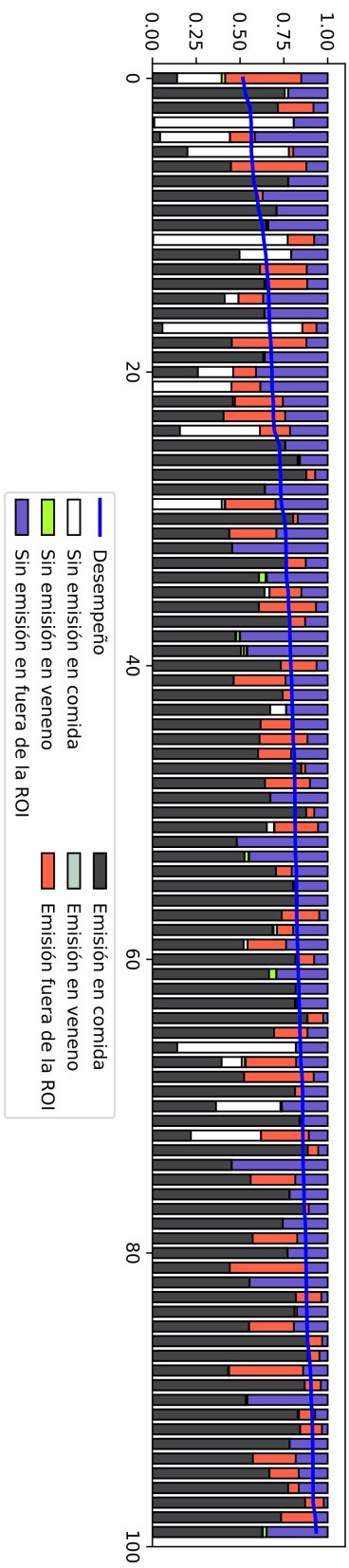


Figura 5.4: Tiempo invertido por los 100 mejores individuos en cada área de la arena durante los 3000 ciclos de simulación en la familia de señalizadores en contexto aleatorio (segunda instancia).

tido en la ROI fue significativamente mayor en la condición de luz roja ( $0.377 \pm SEM 0.025, p < 0.001$ ) y en la condición de luz azul ( $0.091 \pm SEM 0.0164, p < 0.001$ ) en comparación con la condición de ausencia de luz ( $0.003 \pm SEM 0.001$ ). Además, el tiempo total invertido en la ROI fue significativamente mayor en la condición de luz roja ( $p < 0.001$ ) en comparación con la condición de luz azul.

### **Análisis de la prueba de señalización**

Debido a que había dos ROI en cada área, se obtuvieron medidas repetidas del tiempo total que un individuo pasó en cada ROI. Una prueba  $t$  de dos muestras emparejadas reveló que, en promedio, los individuos de la condición de luz azul/roja contra luz roja, los individuos pasaron significativamente más tiempo en la ROI de la fuente azul/roja ( $0.204 \pm SEM 0.013$ ) que en la ROI de la fuente roja ( $0.135 \pm SEM 0.106$ ),  $t(599) = 3.539, p < 0.001, r = -0.329$ . En la condición de luz azul/roja contra luz azul, los individuos pasaron significativamente más tiempo en la ROI de la fuente azul/roja ( $0.268 \pm SEM 0.138$ ) que en la ROI de la fuente azul ( $0.028 \pm SEM 0.005$ ), determinado por una prueba  $t$  de dos muestras emparejadas,  $t(599) = 15.405, p < 0.001, r = -0.169$ .

Dado que la emisión de luz azul tuvo un efecto durante el proceso evolutivo, la respuesta de los individuos ante un único estímulo visual mostró que los individuos desarrollaron una preferencia por la luz roja en comparación con la luz azul. Sin embargo, una muestra significativa de los individuos desarrolló una atracción hacia la luz azul/roja en comparación con la luz roja al evaluar más de un estímulo visual al mismo tiempo. Lo anterior del mismo modo que la familia de señalizadores de arquitectura simple (Ver secciones 4.3.1).

## **5.4. Efectos sobre la señalización de la arquitectura del sistema de control y parámetros evolutivos**

Para dilucidar los efectos sobre la señalización de los señalizadores de contexto aleatorio y contexto aleatorio (segunda instancia) (Ver sección 5), se analizó el tiempo

total invertido en la ROI de la fuente azul/roja de la prueba de señalización para ambas condiciones. Se utilizaron los datos de la muestra total de los 600 señalizadores de contexto aleatorio, y los datos de los 600 señalizadores de contexto aleatorio (segunda instancia). En la condición de luz azul/roja contra luz roja, una prueba  $t$  de dos muestras independientes reveló que los señalizadores de contexto aleatorio (segunda instancia) pasaron significativamente más tiempo en la ROI de la fuente azul/roja ( $0.204 \pm SEM 0.013$ ) en comparación con los señalizadores de contexto aleatorio ( $0.154 \pm SEM 0.008$ ),  $t(1047.287) = 3.173, p = 0.002$ . Asimismo, en la condición de luz azul/roja vs luz azul, reveló los señalizadores de contexto aleatorio (segunda instancia) pasaron significativamente más tiempo en la ROI de la fuente azul/roja ( $0.268 \pm SEM 0.013$ ) en comparación con los señalizadores de contexto aleatorio ( $0.209 \pm SEM 0.009$ ),  $t(1071.952) = 3.473, p = 0.001$ .

Debido a que se realizó un número mayor de ciclos de simulación, los individuos tenían la posibilidad de explorar el ambiente durante mas tiempo en comparación con las demás familias. Por otra parte, el implementar una estructura neuronal de mayor complejidad permitió un mayor número de entradas sensoriales. Lo que permitió incrementar el desempeño y el tiempo de señalización de los individuos durante el proceso evolutivo.

---

## Capítulo 6

# Análisis del espacio de soluciones usando aprendizaje profundo

Para el análisis se utilizaron los resultado de los experimentos realizados en los contextos fijo y aleatorio (Ver secciones 3 y 4). Con base en ello, se obtuvo el material genético y su correspondiente fenotipo de cada una de las tres familias definidas :

- Forrajeo
- Señalizadores de contexto fijo
- Señalizadores de contexto aleatorio

Este material estaba constituido por secuencias de bits en el caso de los genotipos y secuencias de números reales entre -1 y 1 para el caso de los fenotipos. Cada secuencia estaba representada en un espacio de 264 y 33 dimensiones de acuerdo a la longitud de cada secuencia (Ver secciones 3.1.3 y 4.1). Sin embargo, debido a ello resulta complicado el análisis de los datos dentro del espacio de soluciones. Es decir, las  $2^{264}$  posibles soluciones obtenidas al codificar cada peso sináptico en 8 bits, así como su respectivos fenotipos obtenidos al aplicar la ecuación 3.2. Por lo cual, se utilizó aprendizaje profundo para la reducción de dimensionalidad de los datos mediante el uso de *autoencoders*.

## 6.1. Autoencoders

Un *autoencoder* (AE) es una red neuronal entrenada para intentar reconstruir los valores originales de la capa de entrada. Por lo tanto, tiene el mismo número de unidades en la capa de entrada que en la capa de salida. Internamente, un AE tiene una capa oculta  $h$  que describe un código utilizado para representar la entrada. La red consta de dos partes: un codificador representado por la función  $h = f(x)$  y un decodificador que produce una reconstrucción  $r = g(h)$  (Goodfellow et al., 2016) (Ver figura 6.1).

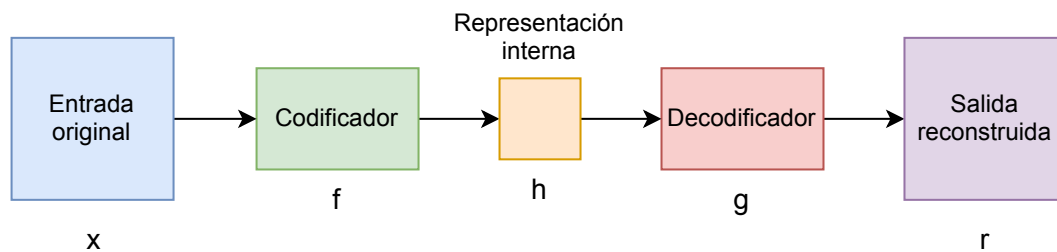


Figura 6.1: Arquitectura de un autoencoder. Un autoencoder consta de dos componentes: el codificador  $f$  (mapeo de  $x$  a  $h$ ) y el decodificador  $g$  (mapeo de  $h$  a  $r$ ), donde  $h$  es la representación interna o código.

De acuerdo a Goodfellow et al. (2016), si un AE tiene éxito simplemente aprendiendo a configurar  $g(f(x)) = x$ , entonces no es especialmente útil. En lugar de ello, los AE están diseñados aprender no solo a copiar perfectamente. Por lo general, están restringidos de tal manera que se les permite copiar sólo de forma aproximada, y copiar sólo los datos de entrada que se asemejan a los datos de entrenamiento.

Los AE pueden aprender representaciones eficientes de datos de baja dimensión en el espacio de dimensiones altas. En otras palabras, la tarea de un AE es, dada una entrada  $x$ , asignar  $x$  a un punto dimensional bajo  $y$  tal que  $x$  pueda ser recuperado de  $y$  (Romero et al., 2017). Por lo cual, la estructura resultante ( $h$ ) se puede elegir para representar los datos en una dimensión más baja.

## 6.2. Reducción de dimensiones en el espacio genotípico

Para realizar el análisis del espacio de soluciones de los genotipos, se utilizó un AE para reducir la dimensionalidad de las secuencias genéticas. Con ello, se redujo de un espacio de 264 dimensiones (Ver sección 3.1.3) a un espacio final de 3 dimensiones.

### 6.2.1. Modelo

El AE consistió en una red neuronal con una capa de entrada con 264 neuronas, las cuales recibían cada uno de los bits que conforman cada una de las secuencias genéticas. Para codificar los datos, se utilizaron 5 capas con 132, 66, 33, 16 y 8 neuronas respectivamente. La representación de los datos en una dimensión más baja, se realizó mediante una capa intermedia con 3 neuronas. Para decodificar los datos, se utilizaron 5 capas con 8, 16, 33, 66 y 132 neuronas respectivamente. Finalmente, una capa de salida de 264 neuronas se utilizó para reconstruir los valores originales de la capa de entrada (Ver figura 6.2). Se utilizó la función sigmoide (Ec. 6.1) como función de activación para la capa intermedia y la capa de salida (Ver figura 6.3). Para las capas restantes, se utilizó la función ReLU (Ec. 6.2) como función de activación (Ver figura 6.3).

$$S(x) = \frac{e^x}{e^x + 1} \quad (6.1)$$

$$f(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } x \geq 0 \end{cases} \quad (6.2)$$

### 6.2.2. Entrenamiento

Para poder llevar a cabo el entrenamiento, se generaron aleatoriamente 5 millones de genomas, de tal forma que la secuencia de ceros y unos quedaran distribuidos de manera uniforme. Asimismo, en conjunto con las secuencias generadas aleatoriamente, se utilizaron las 1,500,000 secuencias resultantes de los 100 individuos evaluados

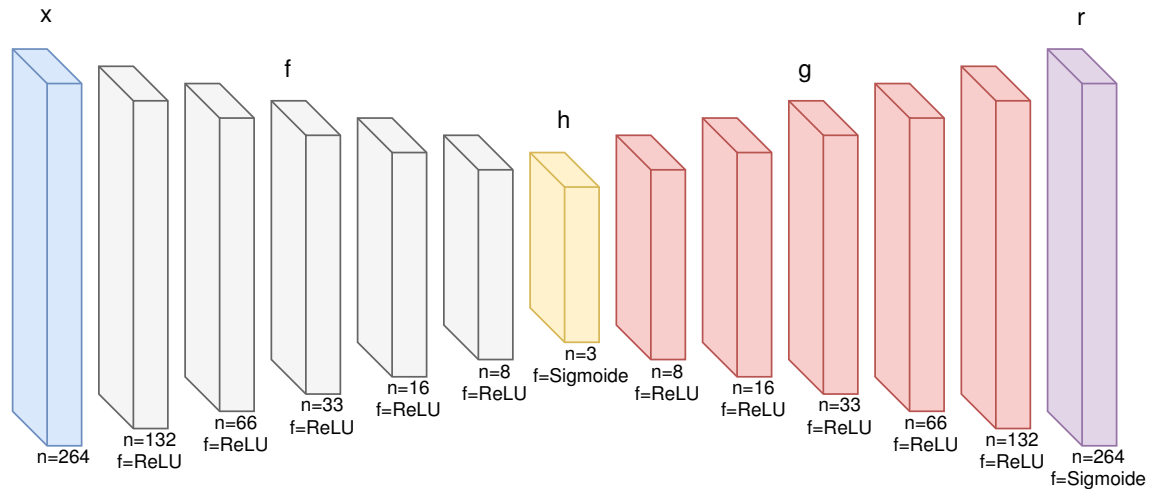


Figura 6.2: Arquitectura del autoencoder utilizado para reducir la dimensionalidad de los genotipos. La red consistió en una capa de entrada  $x$ , un codificador  $f$  de 5 capas, una capa intermedia  $h$  para realizar la representación interna, un decodificador  $g$  de 5 capas y una capa de salida  $r$ . La letra  $n$  representa el número de neuronas de cada capa y  $f$  la función de activación que se utilizó en cada una de las neuronas.

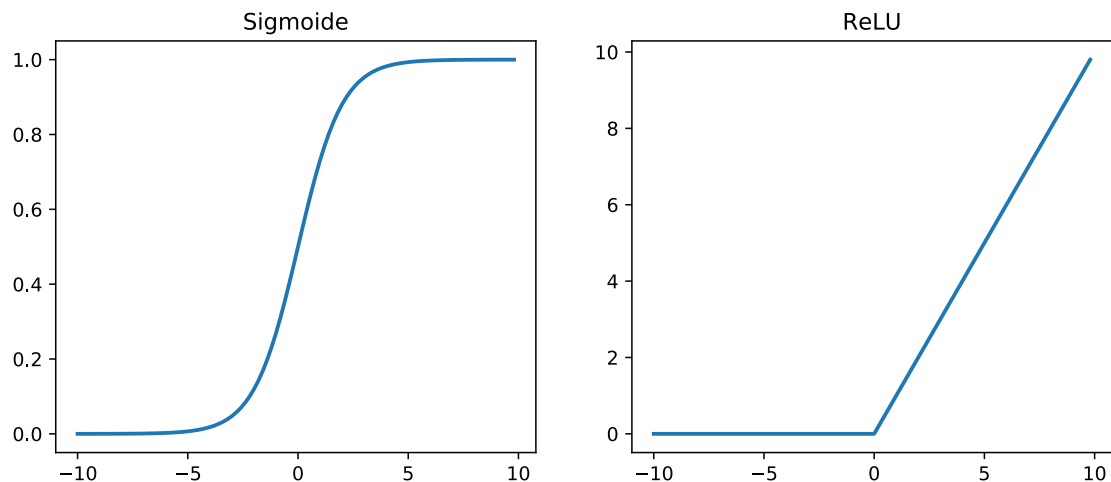


Figura 6.3: Funciones de activación sigmoide y ReLU.

durante 250 generaciones en las 20 simulaciones por cada una de las tres familias. El 80% de estas secuencias se utilizaron como datos de entrenamiento y el 20% como datos de validación. Para optimizar los parámetros del modelo durante el entrenamiento se utilizó el método de optimización estocástico Adam (Kingma and Ba, 2014). Para calcular el desempeño del modelo, el error entre el valor predicho  $\hat{y}$  y el valor de entrada  $y$  se calculó utilizando la función de entropía cruzada binaria (BCE, por sus siglas en inglés *Binary Cross Entropy*) (Ec. 6.3). Mediante la BCE se puede considerar la probabilidad de que un bit sea 1 o 0, ya que naturalmente se aplica a variables aleatorias que pueden estar “activadas” o “desactivadas” (Creswell et al., 2017). Finalmente, el entrenamiento se realizó durante 100 iteraciones sobre el conjunto total de datos. Durante cada iteración los datos fueron procesados en conjuntos de 256 elementos en cada iteración.

$$\mathcal{L}_{BCE} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (6.3)$$

### 6.2.3. Resultados

Al finalizar el proceso de entrenamiento, se utilizó el modelo resultante del AE para realizar el análisis del espacio de soluciones de las secuencias de genomas de las tres familias previamente definidas (Ver sección 6). Como resultado, se visualizó la distribución de los genomas de cada una de las familias a través de las generaciones (Ver figuras 6.4, 6.5, 6.6). Al analizar los genomas de las tres familias, se observó en cada una de ellas una tendencia a converger tempranamente hacia una región del espacio de soluciones en cada una de las simulaciones evaluadas (Ver figura 6.7).

Una explicación de los resultados obtenidos es que las características de los operadores genéticos afectaron la distribución de las secuencias en el espacio de soluciones (Ver sección 3.1.4). Es decir, los porcentajes de cruce y mutación así como el método de selección, permitieron una convergencia temprana durante el proceso evolutivo.



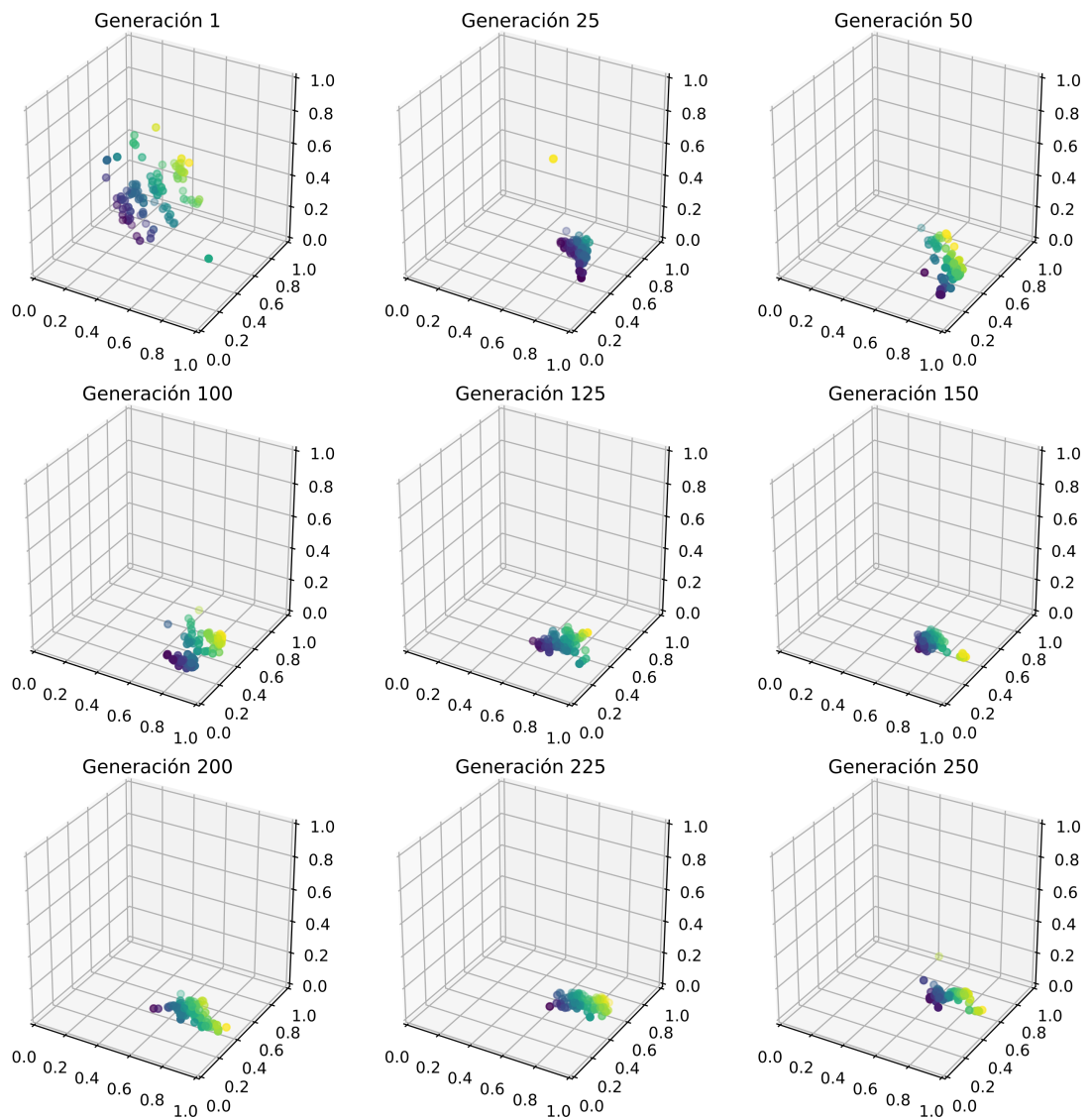


Figura 6.4: Distribución de los genomas de la familia de forrajeo en el espacio de soluciones a través de las generaciones.

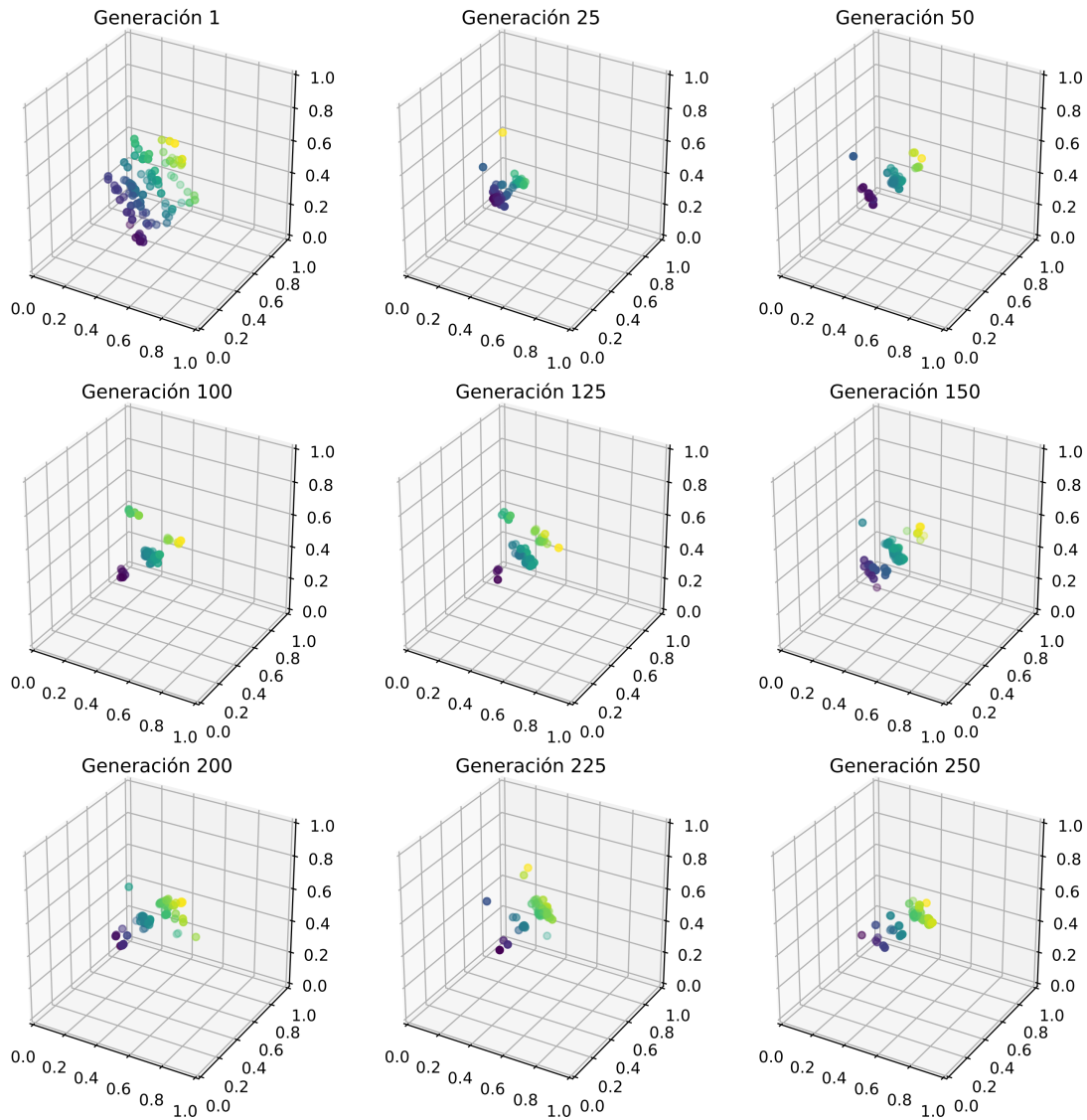


Figura 6.5: Distribución de los genomas de la familia de señalizadores de contexto fijo en el espacio de soluciones a través de las generaciones.

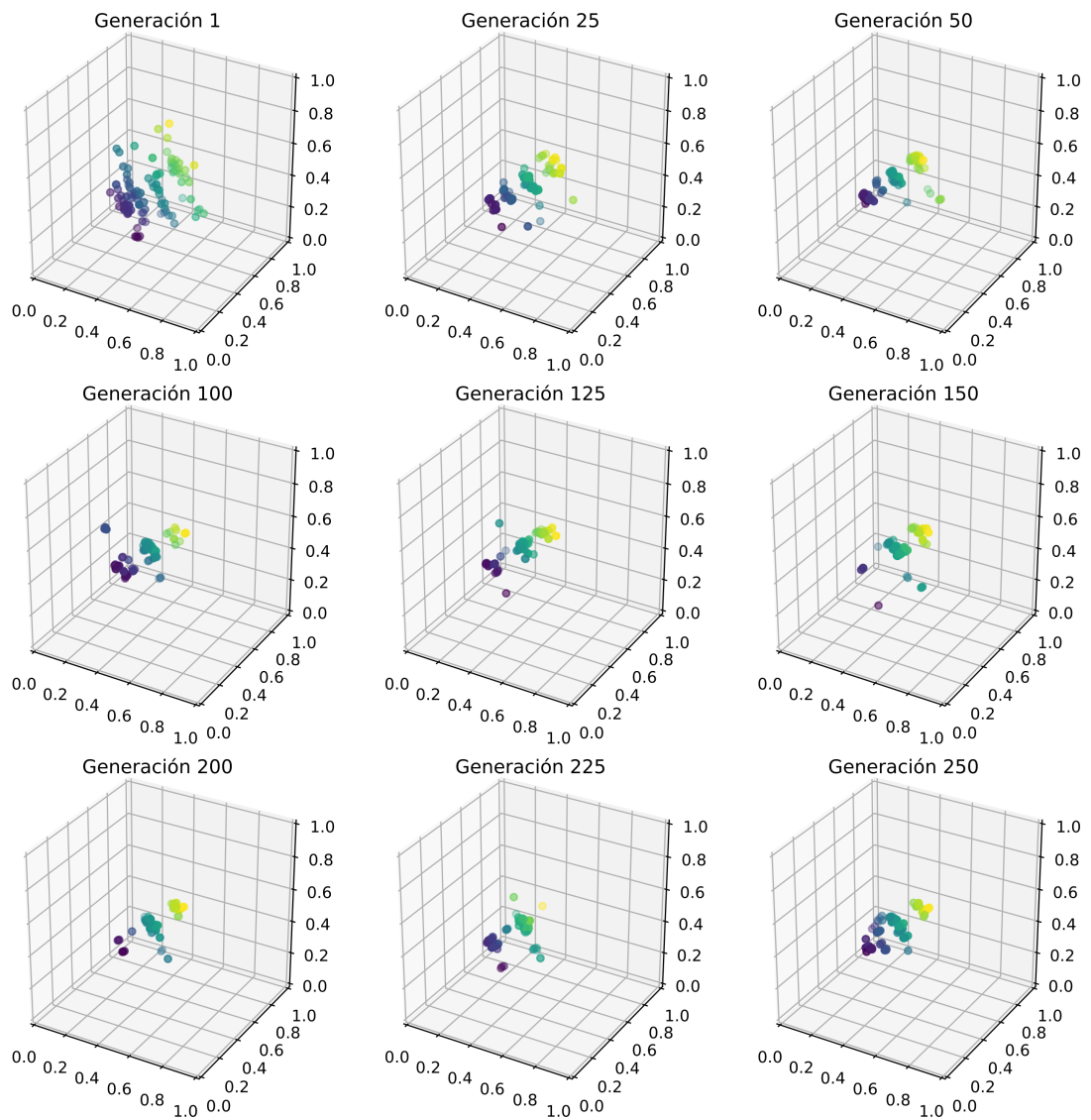


Figura 6.6: Distribución de los genomas de la familia de señalizadores de contexto aleatorio en el espacio de soluciones a través de las generaciones.

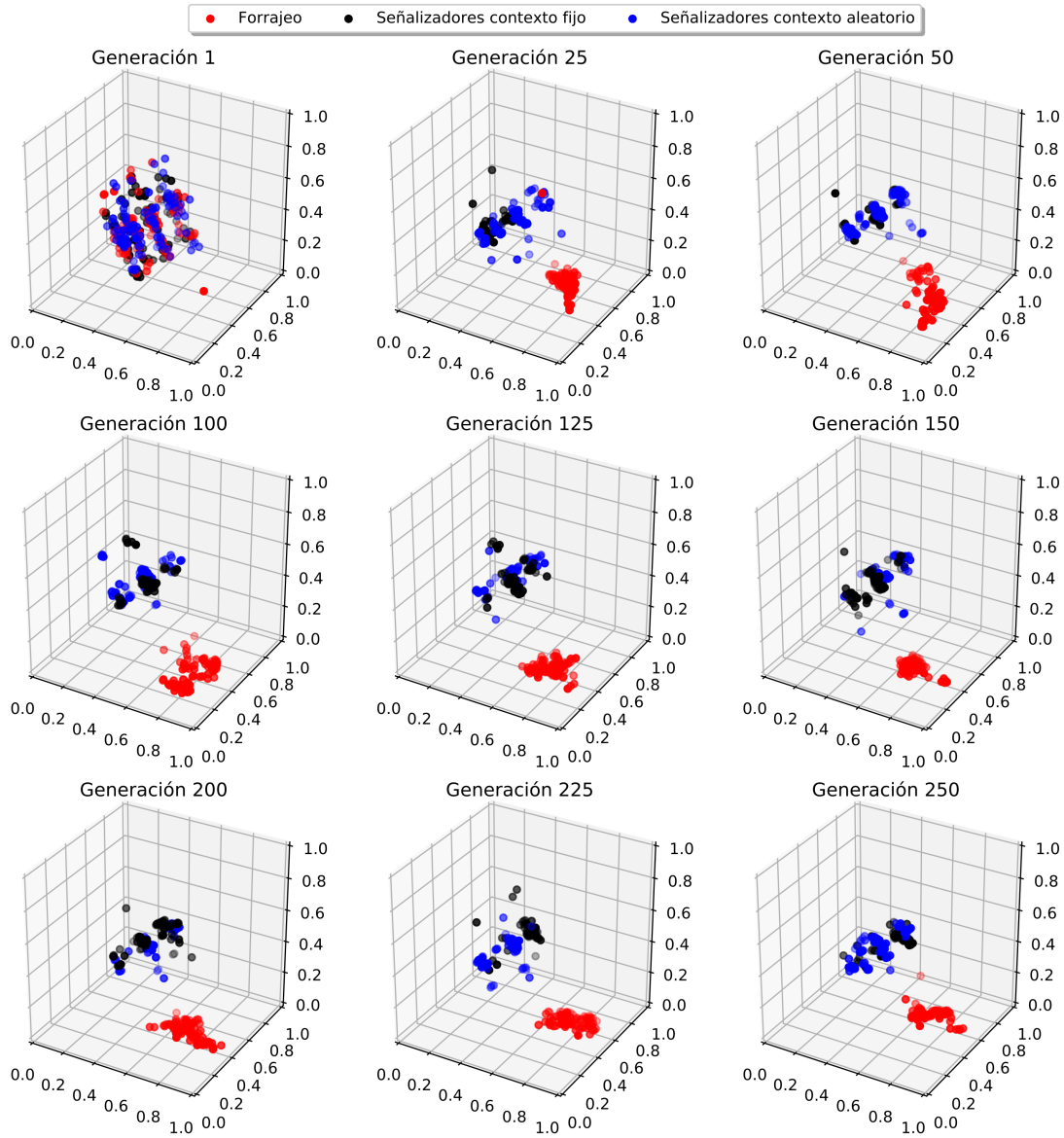


Figura 6.7: Distribución de los genomas de las tres familias en el espacio de soluciones a través de las generaciones.

## 6.3. Reducción de dimensiones en el espacio fenotípico

Para realizar el análisis del espacio de soluciones de los fenotipos, se utilizó un AE para reducir la dimensionalidad de las secuencias de fenotipos. Con ello, se redujo de un espacio de 33 dimensiones (Ver sección 3.1.3) a un espacio final de 3 dimensiones.

### 6.3.1. Modelo

El AE consistió en una red neuronal con una capa de entrada con 33 neuronas, las cuales recibían cada uno de los valores que conforman cada una de las secuencias de fenotipos. Para codificar los datos, se utilizaron 2 capas con 16 y 8 neuronas respectivamente. La representación de los datos en una dimensión más baja, se realizó mediante una capa intermedia con 3 neuronas. Para decodificar los datos, se utilizaron 2 capas con 8, 16 neuronas respectivamente. Finalmente, una capa de salida de 33 neuronas se utilizó para reconstruir los valores originales de la capa de entrada (Ver figura 6.8). Se utilizó la función sigmoide (Ec. 6.1) como función de activación para la capa intermedia y la capa de salida (Ver figura 6.3). Para las capas restantes, se utilizó la función ReLU (Ec. 6.2) como función de activación (Ver figura 6.3).

### 6.3.2. Entrenamiento

Para poder llevar a cabo el entrenamiento, se generaron aleatoriamente 5 millones de fenotipos normalizados a valores entre 0 y 1, de tal forma que la secuencia de números entre 0 y 1 quedaran distribuidos de manera uniforme. Asimismo, en conjunto con las secuencias generadas aleatoriamente, se utilizaron las 1,500,000 secuencias resultantes de los 100 individuos evaluados durante 250 generaciones en las 20 simulaciones por cada una de las tres familias. El 80% de estas secuencias se utilizaron como datos de entrenamiento y el 20% como datos de validación. Para optimizar los parámetros del modelo durante el entrenamiento se utilizó el método de optimización estocástico Adam (Kingma and Ba, 2014). Para calcular el desempeño del modelo, el error entre el valor predicho  $\hat{y}$  y el valor de entrada  $y$  se calculó utilizando la función de error cuadrático medio (MSE, por sus siglas en inglés *Mean*

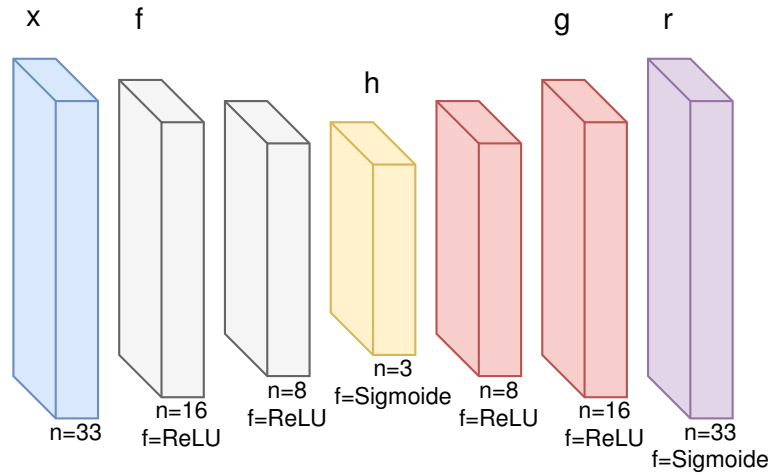


Figura 6.8: Arquitectura del autoencoder utilizado para reducir la dimensionalidad de los genotipos. La red consistió en una capa de entrada  $x$ , un codificador  $f$  de 2 capas, una capa intermedia  $h$  para realizar la representación interna, un decodificador  $g$  de 2 capas y una capa de salida  $r$ . La letra  $n$  representa el número de neuronas de cada capa y  $f$  la función de activación que se utilizó en cada una de las neuronas.

*Squared Error*) (Ec. 6.4). En particular, el MSE es especialmente útil cuando la distribución de los datos consiste en valores entre 0 y 1 (Creswell et al., 2017). Finalmente, el entrenamiento se realizó durante 100 iteraciones sobre el conjunto total de datos. Durante cada iteración los datos fueron procesados en conjuntos de 256 elementos en cada iteración.

$$\mathcal{L}_{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (6.4)$$

### 6.3.3. Resultados

Al finalizar el proceso de entrenamiento, se utilizó el modelo resultante del AE para realizar el análisis del espacio de soluciones de las secuencias de fenotipos de las tres familias previamente definidas (Ver sección 6). Como resultado, se visualizó la distribución de los fenotipos de cada una de las familias a través de las generaciones (Ver figuras 6.4, 6.5, 6.6). Al analizar los fenotipos de las tres familias, se observó en cada una de ellas una tendencia a converger tempranamente hacia una región del espacio de soluciones en cada una de las simulaciones evaluadas (Ver figura 6.7).

Al igual que la sección 6.2.3, una explicación de los resultados obtenidos es que las características de los operadores genéticos afectaron la distribución de las secuencias en el espacio de soluciones (Ver sección 3.1.4). Es decir, los porcentajes de cruce y mutación así como el método de selección, permitieron una convergencia temprana durante el proceso evolutivo.

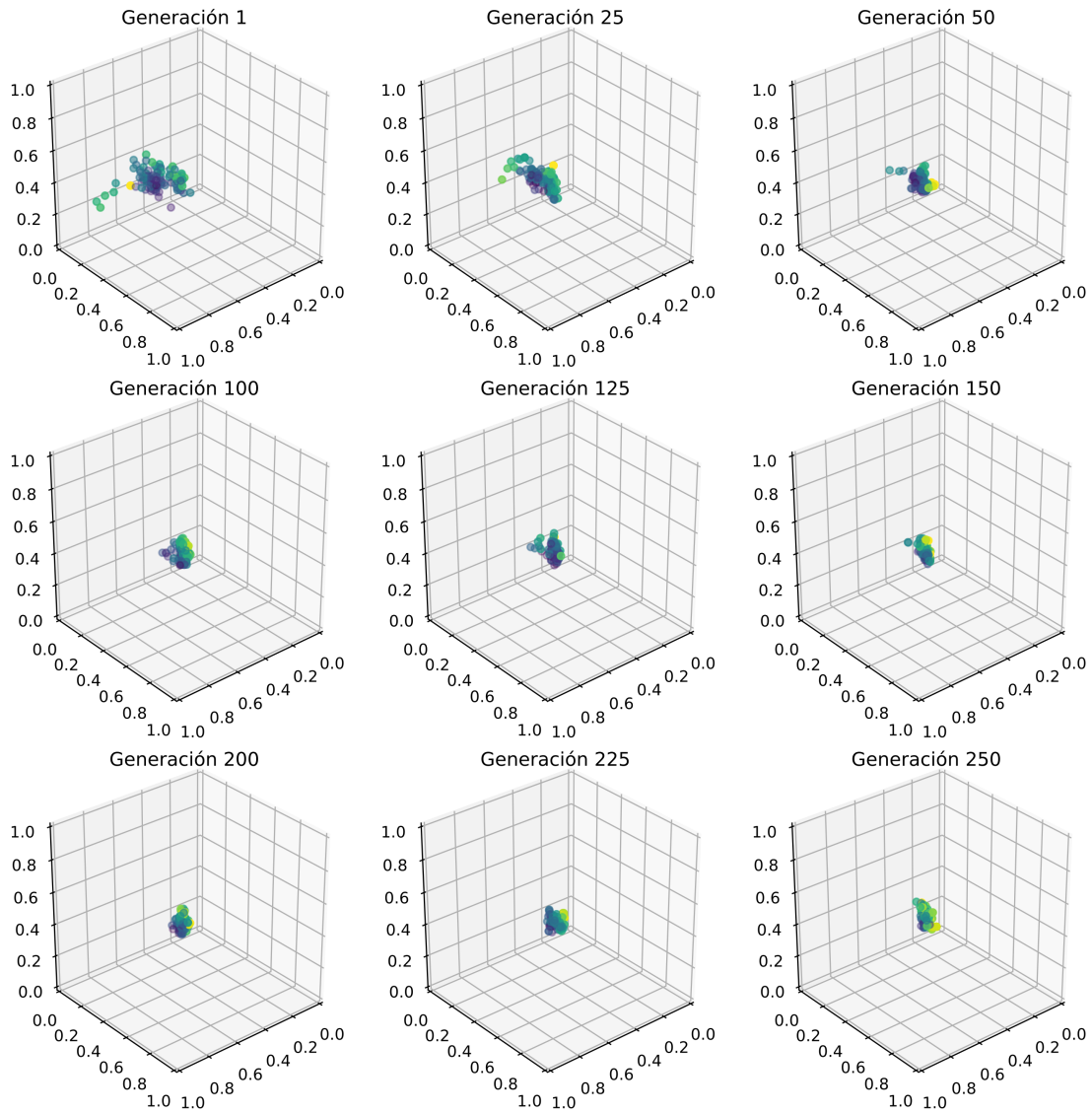


Figura 6.9: Distribución de los fenotipos de la familia de forrajeo en el espacio de soluciones a través de las generaciones.



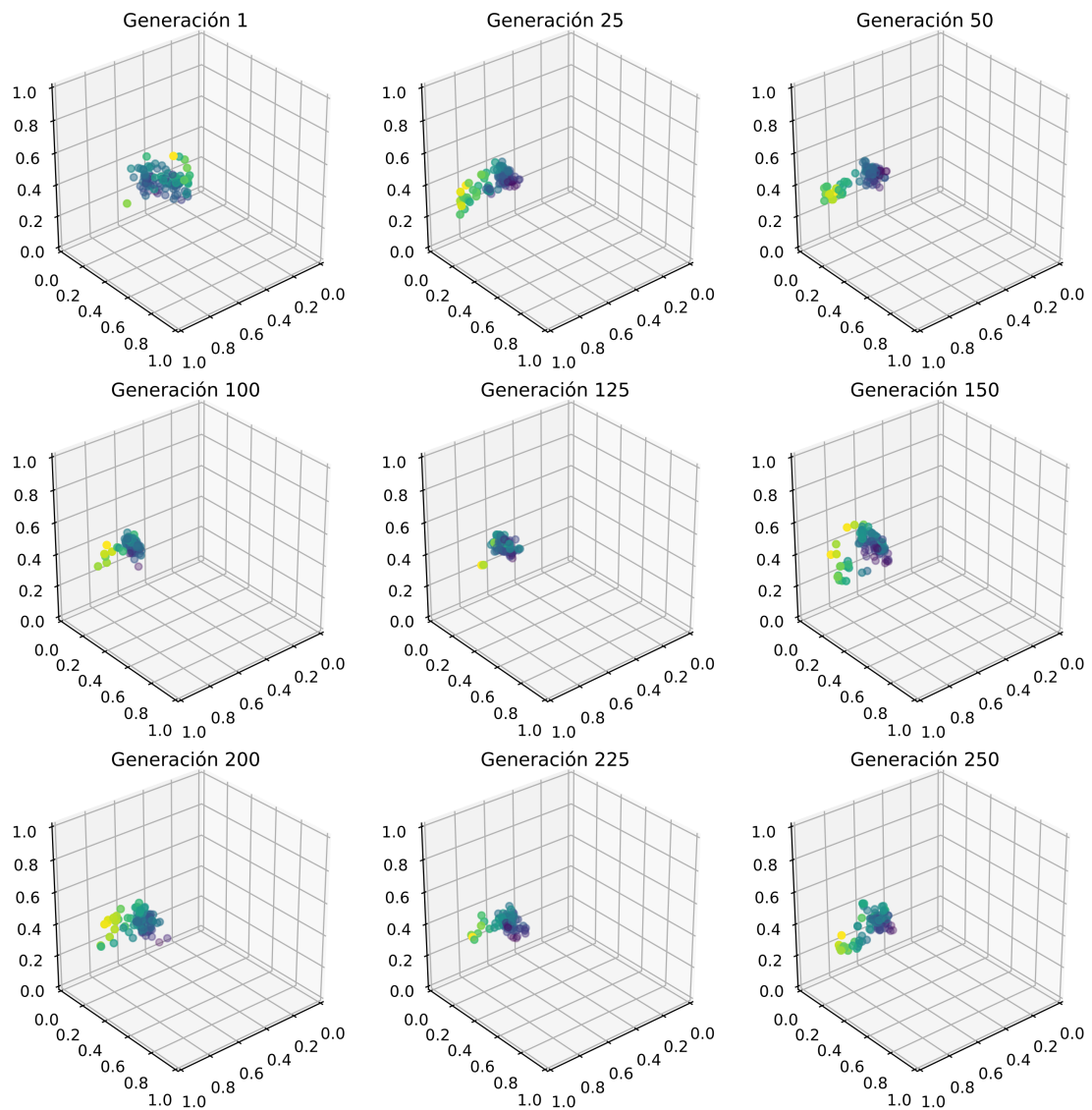


Figura 6.10: Distribución de los fenotipos de la familia de señalizadores de contexto fijo en el espacio de soluciones a través de las generaciones.

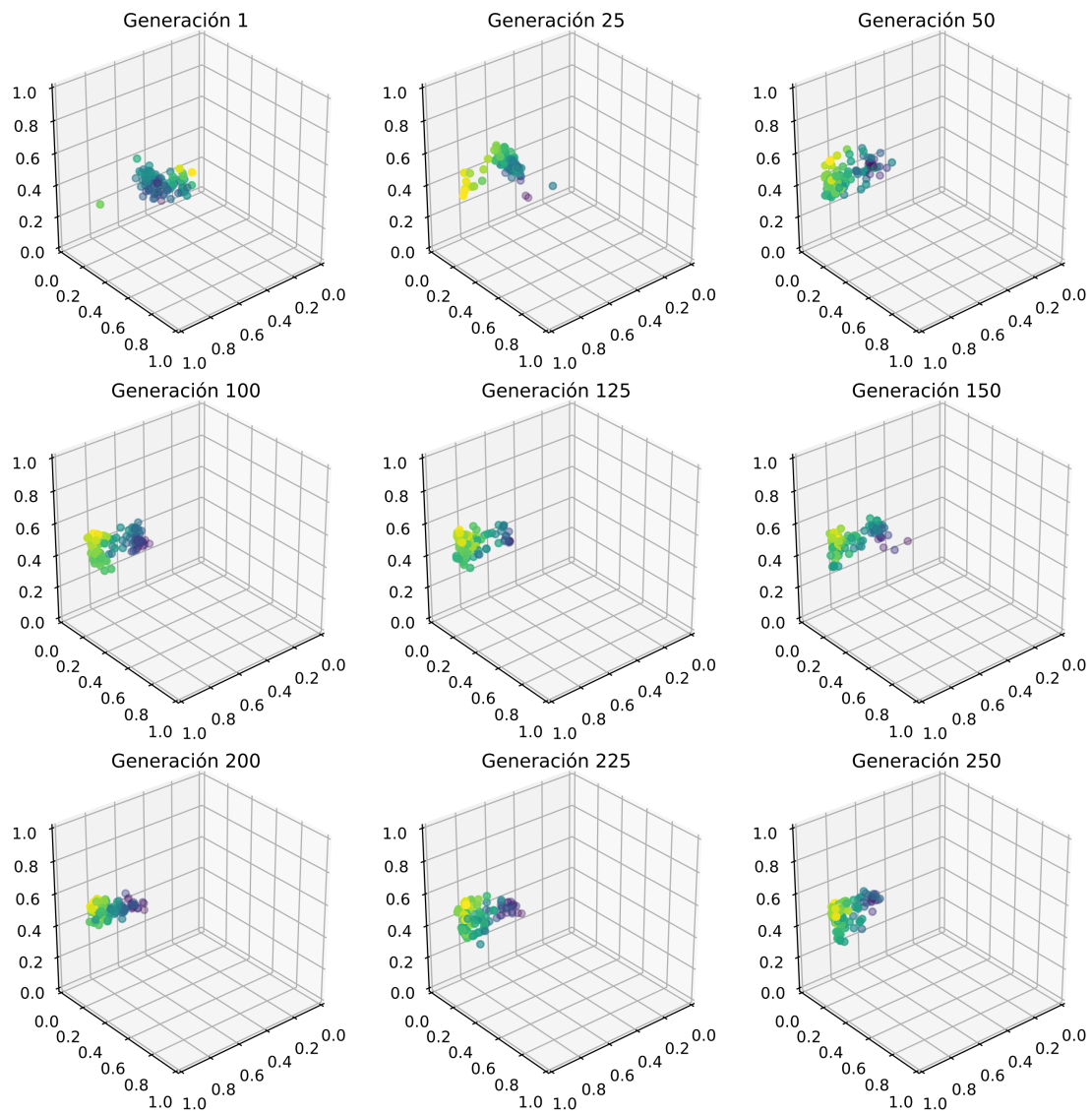


Figura 6.11: Distribución de los fenotipos de la familia de señalizadores de contexto aleatorio en el espacio de soluciones a través de las generaciones.

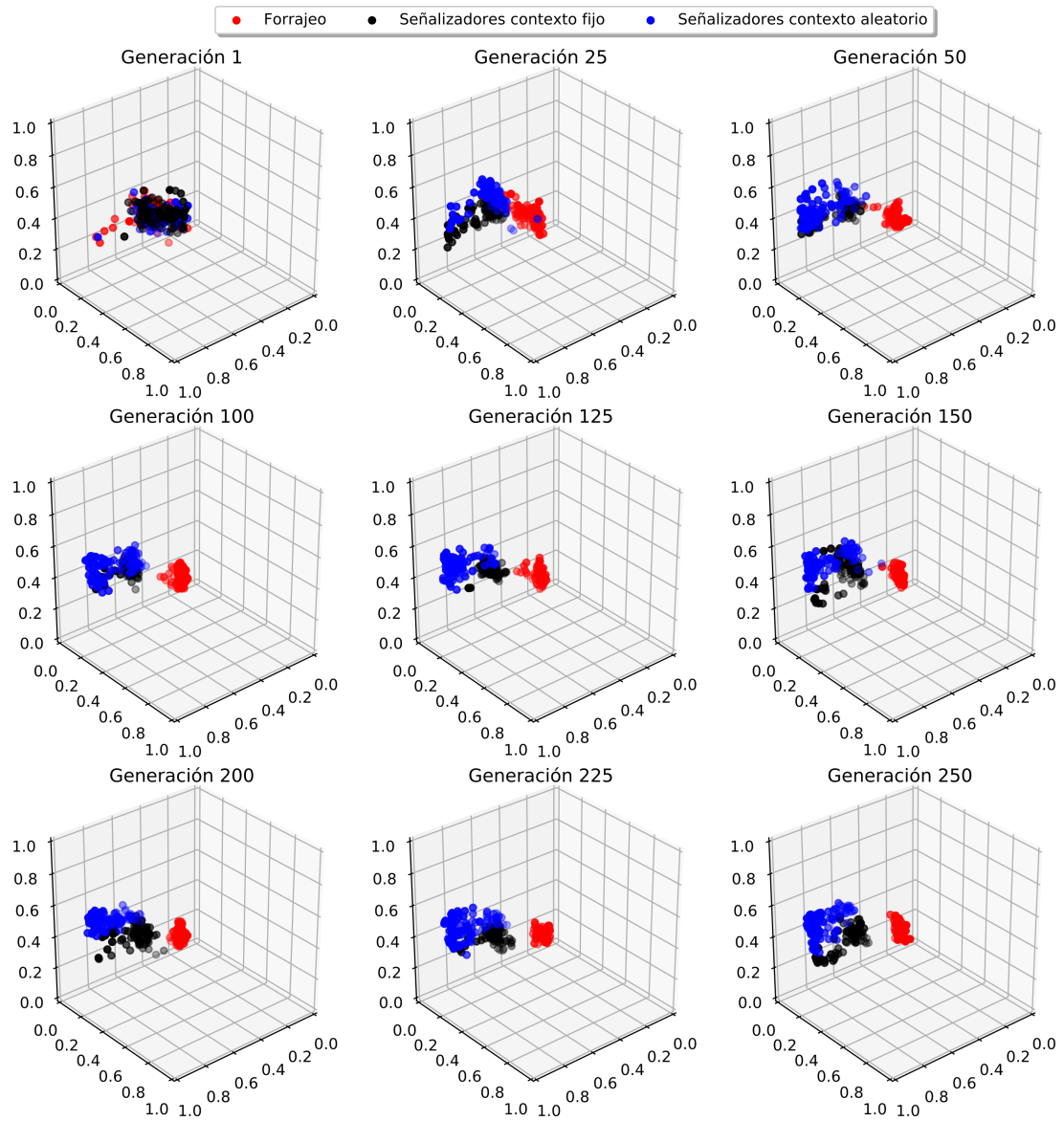


Figura 6.12: Distribución de los fenotipos de las tres familias en el espacio de soluciones a través de las generaciones.

---

# Capítulo 7

## Conclusiones

El trabajo presentado en esta tesis abordó desde un enfoque evolutivo, cómo una estrategia básica de comunicación puede surgir en agentes artificiales a través de la co-evolución entre emisores y receptores. Con este enfoque, se exploró cómo las diferentes condiciones experimentales pueden modificar el proceso evolutivo de las señales.

Las preguntas de investigación del presente proyecto, expuestas en la parte introductoria de la tesis, se responden a continuación con base en los resultados obtenidos en los experimentos:

- ¿Cuáles son las características de las tareas que pueden desencadenar el desarrollo de una estrategia básica de comunicación?

Un sistema de comunicación requiere de dos habilidades: la habilidad de producir señales útiles y la capacidad de reaccionar a las señales útiles. En otras palabras, emitir una señal no es útil a menos que se entienda, y una señal no se entenderá la primera vez que se emite (Smith, 1993).

En la literatura, se ha sugerido que las señales evolucionan a partir de comportamientos que originalmente eran solo indicadores (Scott-Phillips et al., 2012; Smith and Harper, 2003). Es decir, se puede esperar que los sistemas de comunicación se originen a través de información producida inadvertidamente que evoluciona hasta convertirse en una señal. A su vez, en el sentido estricto, cada señal tiene un costo de producción que proviene del componente de eficacia,

incluso si es insignificante en comparación con otros costos y beneficios en el sistema (Laidre and Johnstone, 2013; Számadó, 2011) (Ver tabla 2.1).

Por lo cual, un aspecto importante para el desarrollo de una estrategia básica de comunicación biológicamente plausible, es la propiedad emergente de reaccionar ante señales útiles transmitidas mediante algún canal sensorial sin codificar directamente este comportamiento. Precisamente, durante un proceso evolutivo que involucre una conducta de señalización y el costo de la misma. Es decir, que el acto de señalar afecte directamente la aptitud de cada individuo.

En el caso de la familia de forrajeo, dado que la importancia de la tarea radicaba en localizar la fuente de alimento y evitar la fuente de veneno, la acción de señalar o no señalar no afectó directamente el desempeño de los robots. Por lo tanto, de acuerdo con los resultados obtenidos, el impacto de la respuesta a la luz azul fue menor que la respuesta a la luz roja emitida por los comederos. Esto debido a que la emisión aleatoria de luz azul resultó ser menos significativa en relación con la luz emitida constantemente por los comederos.

Implementar una función de desempeño sin un costo de señalización implica un mayor número de soluciones no óptimas, las cuales existen dentro del espacio de soluciones del problema. Debido a que, una vez que un robot había localizado alguna fuente (comida o veneno), la posibilidad de permanecer en ella hasta el final de la prueba dependía sólo del comportamiento del robot durante su proceso evolutivo. Por lo tanto, no penalizar la libre señalización en la arena implica que la aptitud de los robots no resulte afectada por la señalización en un área diferente a la comida.

En las familias señalizadoras, ya que la señalización tuvo un impacto en la aptitud de los individuos e implicó un costo, los robots con mejor desempeño fueron capaces de señalar el área de comida y evitar este comportamiento en las demás áreas. Los resultados obtenidos en la prueba de rendimiento mostraron que los robots desarrollaron una mayor preferencia por la luz roja que por la luz azul. A pesar de ello, los resultados de la prueba de señalización mostraron que una muestra significativa de los robots desarrolló una atracción hacia la luz azul/roja en comparación con la luz roja. Esto es, los resultados experimentales sugieren que esta atracción surgió a través de la información

inadvertida producida por otros robots al señalar la fuente de alimento, que se convirtió en una señal.

- ¿Cuáles son las características medibles de los sistemas básicos de comunicación evolucionados?

En la comunicación, el emisor entrega información al receptor, el cual utiliza esta información para tomar una decisión. El medir el rendimiento y la confiabilidad de la estrategia de comunicación a través de los ciclos de vida que el robot pasó en la región de interés en diferentes condiciones, permitió analizar el rendimiento de la respuesta en los receptores y la tarea a resolver ante diferentes entradas visuales (luz roja, azul y roja/azul). En contraste con otros enfoques (Floreano et al., 2007; Mitri et al., 2009, 2010b) que registran la posición del robot en movimiento en relación con su posición original (atracción y repulsión), utilizando un único estímulo visual durante ciclos cortos.

Por otra parte, en la familia de señalizadores de contexto aleatorio, los controladores evolucionados demostraron un mejor desempeño. Esto sugiere que la aleatoriedad de la ubicación de los comederos mejoró el rendimiento en la resolución de la tarea. El contar con un contexto diferente en cada generación evaluada y para cada grupo de robots concluyó en controladores más genéricos. Es decir, el contexto aleatorio evitó que los robots aprendieran la posición directa de los comederos. Ya que, aunque la tarea era la misma que en un entorno fijo, durante el proceso evolutivo los robots tenían la posibilidad de explorar un mayor número de soluciones para la misma tarea.

Mediante el uso de aprendizaje profundo (*autoencoders*), fue posible analizar el material genético y su correspondiente fenotipo de las familias que compartían la misma longitud en sus secuencias. Con base en ello, se obtuvieron representaciones de datos de baja dimensión en el espacio de dimensiones altas. En síntesis, al analizar las secuencias de cada familia, se observó en cada una de ellas una tendencia a converger tempranamente hacia una región en cada una de las simulaciones evaluadas. Por lo cual, los resultados sugieren que los porcentajes de cruce y mutación así como el método de selección permitieron una convergencia temprana durante el proceso evolutivo.

- ¿Cuál es el impacto de los operadores genéticos y la arquitectura del sistema de control dentro del proceso evolutivo?

En comparación con otros estudios (Floreano et al., 2007; Mitri et al., 2009, 2010b), la contribución de esta tesis se centró en cómo la co-evolución entre emisores y receptores en pequeñas poblaciones genéticamente variables puede llevar a la evolución de una estrategia de comunicación básica. El enfoque sugerido se implementó a través de la estocástica de los operadores genéticos, dejando de lado el uso de tratamientos experimentales durante el proceso evolutivo que conduciría a individuos clonados o con niveles específicos de relación y/o selección.

En consecuencia, se observó que el impacto evolutivo de la variación en la señalización puede evolucionar sin nexos familiares y sin aplicar ninguna estructura de parentesco o nivel de selección. Las estrategias de señalización evolucionaron a través de los procesos estocásticos en el orden de nuevas mutaciones y/o eventos de cruce y/o cómo se propagan dentro de las poblaciones.

Por otra parte, los señalizadores que implementaron una estructura neuronal de mayor complejidad y un número mayor de ciclos de simulación, fueron capaces de resolver la tarea al igual que los señalizadores con una estructura neuronal simple. Sin embargo, implementar dichas configuraciones resultó en un mayor desempeño de los individuos durante el proceso evolutivo en comparación con las demás familias de señalizadores. Por consiguiente, los resultados de la prueba de señalización mostraron que los señalizadores de contexto aleatorio (segunda instancia) pasaron mayor tiempo en la región de interés de la fuente azul/roja en comparación con los señalizadores de contexto aleatorio.

---

# Bibliografía

Alcock J (2009) *Animal behavior: An evolutionary approach*. Sunderland, Mass.: Sinauer Associates.

Auerbach JE and Bongard JC (2014) Environmental influence on the evolution of morphological complexity in machines. *PLoS computational biology* 10(1): e1003399.

Billen J (2006) Signal variety and communication in social insects. In: *PROCEEDINGS OF THE SECTION EXPERIMENTAL AND APPLIED ENTOMOLOGY-NETHERLANDS ENTOMOLOGICAL SOCIETY*, volume 17. p. 9.

Bongard J, Zykov V and Lipson H (2006) Resilient machines through continuous self-modeling. *Science* 314(5802): 1118–1121.

Brambilla M, Ferrante E, Birattari M and Dorigo M (2013) Swarm robotics: a review from the swarm engineering perspective. *Swarm Intelligence* 7(1): 1–41.

Bredeche N, Haasdijk E and Eiben A (2009) On-line, on-board evolution of robot controllers. In: *International Conference on Artificial Evolution (Evolution Artificielle)*. Springer, pp. 110–121.

Brenner K, You L and Arnold FH (2008) Engineering microbial consortia: a new frontier in synthetic biology. *Trends in biotechnology* 26(9): 483–489.

Christy JH, Baum JK and Backwell PR (2003) Attractiveness of sand hoods built by courting male fiddler crabs, *uca musica*: test of a sensory trap hypothesis. *Animal*



- Behaviour* 66(1): 89 – 94. DOI:<https://doi.org/10.1006/anbe.2003.2196>. URL <http://www.sciencedirect.com/science/article/pii/S0003347203921968>.
- Cliff D, Husbands P and Harvey I (1993) Explorations in evolutionary robotics. *Adaptive behavior* 2(1): 73–110.
- Creswell A, Arulkumaran K and Bharath AA (2017) On denoising autoencoders trained to minimise binary cross-entropy. *arXiv preprint arXiv:1708.08487* .
- Diggle SP, Gardner A, West SA and Griffin AS (2007) Evolutionary theory of bacterial quorum sensing: when is a signal not a signal? *Philosophical Transactions of the Royal Society B: Biological Sciences* 362(1483): 1241–1249.
- Doncieux S, Mouret JB, Bredeche N and Padois V (2011) Evolutionary robotics: Exploring new horizons. In: *New horizons in evolutionary robotics*. Springer, pp. 3–25.
- Elfwing S and Doya K (2014) Emergence of polymorphic mating strategies in robot colonies. *PloS one* 9(4): e93622.
- Floreano D, Mitri S, Magnenat S and Keller L (2007) Evolutionary conditions for the emergence of communication in robots. *Current biology* 17(6): 514–519.
- Friedman GJ (1956) *Selective feedback computers for engineering synthesis and nervous system analogy*. Master’s Thesis, UCLA, Engineering.
- Funahashi KI (1989) On the approximate realization of continuous mappings by neural networks. *Neural networks* 2(3): 183–192.
- Goodfellow I, Bengio Y, Courville A and Bengio Y (2016) *Deep learning*, volume 1. MIT press Cambridge.
- Grueter C and Keller L (2016) Inter-caste communication in social insects. *Current opinion in neurobiology* 38: 6–11.
- Harvey I, Di Paolo E, Wood R, Quinn M and Tuci E (2005) Evolutionary robotics: A new scientific tool for studying cognition. *Artificial life* 11(1-2): 79–98.

- Haykin SS, Haykin SS, Haykin SS and Haykin SS (2009) *Neural networks and learning machines*, volume 3. Pearson Upper Saddle River, NJ, USA:.
- Hebbs D (1949) The organization of behavior. *Wiley and Sons, New York, NY, USA* .
- Hooker JN (2015) Toward unification of exact and heuristic optimization methods. *International Transactions in Operational Research* 22(1): 19–48.
- Hornik K, Stinchcombe M and White H (1989) Multilayer feedforward networks are universal approximators. *Neural networks* 2(5): 359–366.
- Ivaldi S, Padois V and Nori F (2014) Tools for dynamics simulation of robots: a survey based on user feedback. *arXiv preprint arXiv:1402.7050* .
- Jakobi N (1997) Evolutionary robotics and the radical envelope-of-noise hypothesis. *Adaptive behavior* 6(2): 325–368.
- Karlik B and Olgac AV (2011) Performance analysis of various activation functions in generalized mlp architectures of neural networks. *International Journal of Artificial Intelligence and Expert Systems* 1(4): 111–122.
- Keller L and Surette MG (2006) Communication in bacteria: an ecological and evolutionary perspective. *Nature Reviews Microbiology* 4(4): 249.
- Kim S, Laschi C and Trimmer B (2013) Soft robotics: a bioinspired evolution in robotics. *Trends in biotechnology* 31(5): 287–294.
- Kingma DP and Ba J (2014) Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* .
- Kokash N (2005) An introduction to heuristic algorithms. *Department of Informatics and Telecommunications* : 1–8.
- König L, Mostaghim S and Schmeck H (2009) Decentralized evolution of robotic behavior using finite state machines. *International Journal of Intelligent Computing and Cybernetics* 2(4): 695–723.

- Laidre ME and Johnstone RA (2013) Animal signals. *Current Biology* 23(18): R829–R833.
- Lipson H and Pollack JB (2000) Automatic design and manufacture of robotic life-forms. *Nature* 406(6799): 974.
- Lipton ZC, Berkowitz J and Elkan C (2015) A critical review of recurrent neural networks for sequence learning. *arXiv preprint arXiv:1506.00019* .
- Magenat S, Waibel M and Beyeler A (2009) Enki-an open source fast 2d robot simulator.
- McCulloch WS and Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics* 5(4): 115–133.
- Mitchell M (1998) *An introduction to genetic algorithms*. MIT press.
- Mitri S, Floreano D and Keller L (2009) The evolution of information suppression in communicating robots with conflicting interests. *Proceedings of the National Academy of Sciences* 106(37): 15786–15790.
- Mitri S, Floreano D and Keller L (2010a) Evolutionary conditions for the emergence of communication. In: *Evolution of Communication and Language in Embodied Agents*. Springer, pp. 123–134.
- Mitri S, Floreano D and Keller L (2010b) Relatedness influences signal reliability in evolving robots. *Proceedings of the Royal Society of London B: Biological Sciences* : rspb20101407.
- Nelson AL, Barlow GJ and Doitsidis L (2009) Fitness functions in evolutionary robotics: A survey and analysis. *Robotics and Autonomous Systems* 57(4): 345–370.
- Nolfi S (2009) Behavior and cognition as a complex adaptive system: Insights from robotic experiments. *Philosophy of Complex Systems, Handbook on Foundational/Philosophical Issues for Complex Systems in Science*. Elsevier .

- Nolfi S and Floreano D (2000) *Evolutionary robotics: The biology, intelligence, and technology of self-organizing machines*. MIT press.
- O’Dowd PJ, Winfield AF and Studley M (2011) The distributed co-evolution of an embodied simulator and controller for swarm robot behaviours. In: *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*. IEEE, pp. 4995–5000.
- Orr HA (2009) Fitness and its role in evolutionary genetics. *Nature Reviews Genetics* 10(8): 531.
- Penn D and Potts WK (1998) Chemical signals and parasite-mediated sexual selection. *Trends in Ecology & Evolution* 13(10): 391–396.
- Rojas R (2013) *Neural networks: a systematic introduction*. Springer Science & Business Media.
- Romero J, Olson JP and Aspuru-Guzik A (2017) Quantum autoencoders for efficient compression of quantum data. *Quantum Science and Technology* 2(4): 045001.
- Rosenblatt F (1958) The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review* 65(6): 386.
- Rosenblatt F (1962) *Principles of neurodynamics* .
- Rothlauf F (2011) *Design of modern heuristics: principles and application*. Springer Science & Business Media.
- Schiavinotto T and Stützle T (2007) A review of metrics on permutations for search landscape analysis. *Computers & operations research* 34(10): 3143–3153.
- Scott-Phillips TC (2008) Defining biological communication. *Journal of evolutionary biology* 21(2): 387–395.
- Scott-Phillips TC, Blythe RA, Gardner A and West SA (2012) How do communication systems emerge? *Proceedings of the Royal Society of London B: Biological Sciences* 279(1735): 1943–1949.

- Shannon CE (2001) A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review* 5(1): 3–55.
- Shukla A, Pandey HM and Mehrotra D (2015) Comparative review of selection techniques in genetic algorithm. In: *Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE), 2015 International Conference on*. IEEE, pp. 515–519.
- Siciliano B and Khatib O (2016) *Springer handbook of robotics*. Springer.
- Silva F, Duarte M, Correia L, Oliveira SM and Christensen AL (2016) Open issues in evolutionary robotics. *Evolutionary computation* 24(2): 205–236.
- Silva F, Duarte M, Oliveira SM, Correia L and Christensen AL (2014) The case for engineering the evolution of robot controllers. In: *14th International Conference on the Synthesis and Simulation of Living Systems*. pp. 703–710.
- Singh V and Singh K (2018) Modern synthesis .
- Smith JM (1993) *The theory of evolution*. Cambridge University Press.
- Smith JM and Harper D (2003) *Animal signals*. Oxford University Press.
- Soni N and Kumar T (2014) Study of various mutation operators in genetic algorithms. *International Journal of Computer Science and Information Technologies* 5(3): 4519–4521.
- Stanley KO, D’Ámbrosio DB and Gauci J (2009) A hypercube-based encoding for evolving large-scale neural networks. *Artificial life* 15(2): 185–212.
- Stanley KO and Miikkulainen R (2002) Evolving neural networks through augmenting topologies. *Evolutionary computation* 10(2): 99–127.
- Számádó S (2011) The cost of honesty and the fallacy of the handicap principle. *Animal Behaviour* 81(1): 3–10.
- Turing AM (1950) Computing machinery and intelligence. *Mind* 59(236): 433–460. URL <http://www.jstor.org/stable/2251299>.

- Umbarkar A and Sheth P (2015) Crossover operators in genetic algorithms: A review. *ICTACT journal on soft computing* 6(1).
- Wilson EO (1975) Sociobiology, the modern synthesis. *Belknap, Cambridge, MA* .
- Yu X and Gen M (2010) *Introduction to evolutionary algorithms*. Springer Science & Business Media.



UNIVERSIDAD AUTÓNOMA DEL  
ESTADO DE MORELOS



Instituto de  
Investigación en  
Ciencias  
Básicas y  
Aplicadas

**INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS**

Coordinación de Programas Educativos

Posgrado en Ciencias



**DR. VICTOR BARBA LÓPEZ  
COORDINADOR DEL POSGRADO EN CIENCIAS  
PRESENTE**

Atendiendo a la solicitud para emitir DICTAMEN sobre la revisión de la TESIS titulada "Evolución de Formas Básicas de Señalización en Agentes Artificiales Autónomos" que presenta el alumno Victor Hugo Santos Pucheta (5620171101) para obtener el título de **Maestro en Ciencias**.

Nos permitimos informarle que nuestro voto es:

NOMBRE	DICTAMEN	FIRMA
Dr. Edgar Francisco Roman Rangel ITAM	Aprobado	
Dr. Jorge Hermsillo Valadez CInC-UAEM	Aprobado	
Dr. Fernando Martín Montes González Universidad Veracruzana	APROBADO	
Dr. Bruno Lara Guzmán CInC-UAEM	Aprobado	
Dr. Juan Manuel Rendón Mancha CInC-UAEM	Aprobado	