

UNIVERSIDAD DE GUADALAJARA

CENTRO UNIVERSITARIO DE CIENCIAS BIOLÓGICAS Y AGROPECUARIAS
DEPARTAMENTO DE CIENCIAS AMBIENTALES
CENTRO DE ESTUDIOS E INVESTIGACIONES EN COMPORTAMIENTO



MANEJO DE VARIOS ESTÍMULOS CONDICIONADOS EN EL MODELO DONAHOE-BURGOS-PALMER

TESIS
QUE PARA OBTENER EL GRADO DE

MAESTRO EN CIENCIAS DEL COMPORTAMIENTO
OPCIÓN ANÁLISIS DE LA CONDUCTA

PRESENTA
ANTONIO PONCE ROJO

DIRECTOR :

DR. JOSÉ E. BURGOS TRIANO

COMITÉ:
DR. FÉLIX HÉCTOR MARTÍNEZ SÁNCHEZ
DR. JULIO AGUSTÍN VARELA BARRAZA
DR. FRANCOIS JACQUES TONNEAU
DR. EMILIO RIBER IÑESTA

GUADALAJARA , JALISCO MEXICO , AGOSTO DEL 2003.

A DOÑA ZENAIDA Y DON ANTONIO

GRACIAS POR TODO

ESTE TRABAJO NO SE HABRÍA PODIDO LOGRAR SIN LA AYUDA Y SOPORTE DE LAS SIGUIENTES PERSONAS E INSTITUCIONES:

CONSEJO NACIONAL DE CIENCIA Y TECNOLOGÍA (CONACYT)

UNIVERSIDAD DE GUADALAJARA

CENTRO UNIVERSITARIO DE CIENCIAS BIOLÓGICAS Y AGROPECUARIAS
DE LA UNIVERSIDAD DE GUADALAJARA

CENTRO UNIVERSITARIO DE LA COSTA
DE LA UNIVERSIDAD DE GUADALAJARA

M. EN C. JEFFRY S. FERNÁNDEZ RODRÍGUEZ,
RECTOR DEL CENTRO UNIVERSITARIO DE LA COSTA DE LA UNIVERSIDAD DE GUADALAJARA

DR. EMILIO RIBES IÑESTA
DIRECTOR DEL CENTRO DE ESTUDIOS E INVESTIGACIONES EN COMPORTAMIENTO DE LA UNIVERSIDAD DE
GUADALAJARA

M EN C. CINTIA ELIZABETH CASTRO SANCHEZ
UNIDAD PARA EL DESARROLLO INSTITUCIONAL DE LA UNIVERSIDAD DE GUADALAJARA

DR. JOSÉ E. BURGOS TRIANO
PROFESOR INVESTIGADOR DEL CENTRO DE ESTUDIOS E INVESTIGACIONES EN COMPORTAMIENTO DE LA
UNIVERSIDAD DE GUADALAJARA

DR. FÉLIX HÉCTOR MARTÍNEZ SÁNCHEZ
PROFESOR INVESTIGADOR Y COORDINADOR DEL PROGRAMA DE POSGRADO EN CIENCIAS DEL
COMPORTAMIENTO, DEL CENTRO DE ESTUDIOS E INVESTIGACIONES EN COMPORTAMIENTO DE LA
UNIVERSIDAD DE GUADALAJARA

DR. FRANCOIS JACQUES TONNEAU
PROFESOR INVESTIGADOR DEL CENTRO DE ESTUDIOS E INVESTIGACIONES EN COMPORTAMIENTO DE LA
UNIVERSIDAD DE GUADALAJARA

DR. JULIO AGUSTÍN VARELA BARRAZA
PROFESOR INVESTIGADOR DEL CENTRO DE ESTUDIOS E INVESTIGACIONES EN COMPORTAMIENTO DE LA
UNIVERSIDAD DE GUADALAJARA

DR. MIGUEL ANGEL GUEVARA PÉREZ
PROFESOR INVESTIGADOR DEL INSTITUTO DE NEUROCIENCIAS DE LA UNIVERSIDAD DE GUADALAJARA

LIC. MARTHA BAÑUELOS HERNÁNDEZ
PROFESOR DOCENTE DEL CENTRO UNIVERSITARIO DE LA COSTA DE LA UNIVERSIDAD DE GUADALAJARA

M EN C. JAVIER VÁZQUEZ NAVARRO
COORDINADOR DE CONTROL ESCOLAR DEL CENTRO UNIVERSITARIO DE CIENCIAS BIOLÓGICAS Y
AGROPECUARIAS DE LA UNIVERSIDAD DE GUADALAJARA

CONTENIDO

FIGURAS.....	5
TABLAS.....	8
ALGORITMOS.....	10
1. INTRODUCCIÓN.....	11
1.2. PREGUNTA DE INVESTIGACIÓN.....	13
1.3. OBJETIVOS GENERALES.....	13
1.3.1. <i>Objetivo general 1</i>	13
1.4. OBJETIVOS ESPECÍFICOS.....	14
1.4.1. <i>Objetivo específico 1</i>	14
1.4.2. <i>Objetivo específico 2</i>	14
1.4.3. <i>Objetivo específico 3</i>	14
2. MARCO CONCEPTUAL.....	15
2.1. ELEMENTOS Y ORGANIZACIÓN DE UNA RED NEURAL.....	23
2.1.1. <i>Unidades de proceso</i>	24
2.1.2. <i>Capas de unidades de proceso</i>	27
2.1.3. <i>Tipos de Arquitectura más comunes</i>	28
2.1.4. <i>Redes neurales de una sola capa y un solo sentido de propagación de señal</i>	28
2.1.5. <i>Redes neurales multicapa de un solo sentido de propagación de señal</i>	29
2.1.6. <i>Redes neurales recurrentes</i>	31
2.1.7. <i>Redes neurales estructuradas en latices</i>	32
2.2. FUNCIONAMIENTO DE UNA RED NEURAL.....	33
2.2.1. <i>Transmisión de los impulsos a través de la estructura de la red</i>	34
2.2.2. <i>Aprendizaje en una red neural</i>	34
2.2.3. <i>Aprendizaje no supervisado</i>	35
2.2.4. <i>Aprendizaje supervisado</i>	40
2.2.5. <i>Aprendizaje por refuerzo</i>	44
2.3. MODELOS BASADOS EN REDES NEURALES PARA LA SIMULACIÓN DE FENÓMENOS CONDUCTUALES.....	46
2.3.1. <i>El Modelo DYSTAL de Alkon, Vogl y Tam</i>	47
2.3.2. <i>El Modelo GMADN de Baxter, et al para simular el aprendizaje asociativo</i>	49
2.3.3. <i>El modelo Schmajuck-Lam-gray (SLG) para simular inhibición latente</i>	52
2.4. EL MODELO DE DONAHOE, BURGOS Y PALMER.....	55
3. MÉTODO.....	59
3.1. EXPERIMENTO 1.....	59
3.1.1. <i>Condición experimental 1</i>	59
3.1.2. <i>Condición experimental 2</i>	62
3.1.3. <i>Condición experimental 3</i>	66
3.1.4. <i>Condición experimental 4</i>	69
3.1.5. <i>Condición experimental 5</i>	72
3.1.6. <i>Condición experimental 6</i>	75
3.2. EXPERIMENTO 2.....	78
3.2.1. <i>Condición experimental 1</i>	78
3.3. EXPERIMENTO 3.....	82
3.3.1. <i>Condición experimental 1</i>	82
3.4. EXPERIMENTO 4.....	84
3.4.1. <i>Condición experimental 1</i>	85
3.5. EXPERIMENTO 5.....	87
3.5.1. <i>Condición experimental 1</i>	88

3.6. EXPERIMENTO 6.....	91
3.6.1. Condición experimental 1.....	91
4. RESULTADOS.....	96
4.1. EXPERIMENTO 1.....	96
4.1.1. Condición experimental 1.....	96
4.1.2. Condición experimental 2.....	98
4.1.3. Condición experimental 3.....	100
4.1.4. Condición experimental 4.....	101
4.1.5. Condición experimental 5.....	103
4.1.6. Condición experimental 6.....	105
4.2. EXPERIMENTO 2.....	107
4.2.1. Condición experimental 1.....	107
4.3. EXPERIMENTO 3.....	110
4.3.1. Condición experimental 1.....	110
4.4. EXPERIMENTO 4.....	111
4.4.1. Condición experimental 1.....	111
4.5. EXPERIMENTO 5.....	112
4.5.1. Condición experimental 1.....	112
4.6. EXPERIMENTO 6.....	114
4.6.1. Condición experimental 1.....	114
5. DISCUSIÓN.....	117
REFERENCIAS	123

FIGURAS

Figura 1. Circuito de ideas cooperantes de Bain (tomado de Olmsted, 1999).....	15
Figura 2. Circuito de ideas cooperantes de Bain con un nodo más resistente ocasionado por una distancia mayor que los del resto del circuito (tomado de Olmsted, 1999).....	16
Figura 3. Circuito de Rashevsky para la operación lógica de la disyunción exclusiva (tomado de Olmsted, 1999).....	17
Figura 4. Tres funciones no lineares de las más usadas en redes neurales.....	25
Figura 5. Unidad de proceso con valores de ejemplo.....	25
Figura 6. Función de salida tipo <i>escalón</i>	27
Figura 7. Capas e interconexiones en una red neural típica.....	28
Figura 8. Red neural de una sola capa.....	29
Figura 9. Red neural totalmente conectada.....	31
Figura 10. Red neural parcialmente conectada.....	31
Figura 11. Red neural recurrente.....	32
Figura 12. Red neural tipo ládice de una dimensión.....	33
Figura 13. Red neural tipo ládice de dos dimensiones.....	33
Figura 14. Arquitectura de un Mapa Auto-organizativo.....	37
Figura 15. Función de sombrero mexicano, mostrando la inhibición lateral.....	39
Figura 16. Tipos de rutas de pesos en el modelo DYSTAL.....	48
Figura 17. Aprendizaje en GMADN.....	50
Figura 18. Memoria en GMADN.....	50
Figura 19. Arquitectura para la simulación del condicionamiento clásico con el GMADN.....	51
Figura 20. Arquitectura para la simulación del condicionamiento operante con el GMADN.....	52
Figura 21. Arquitectura del Modelo SGL de Schmajuck.....	54
Figura 22. Arquitectura del modelo de Donahoe, Burgos y Palmer (1993).....	55
Figura 23. Elementos en una red bajo el modelo de Donahoe, Burgos y Palmer (1993).....	56
Figura 24. Arquitectura usada en las redes neurales de la condición experimental número 1 del experimento 1.....	60
Figura 25. Arquitectura usada en las redes neurales de la condición experimental número 2 del experimento 1.....	63
Figura 26. Arquitectura usada en las redes neurales de la condición experimental número 3 del experimento 1.....	66
Figura 27. Arquitectura usada en las redes neurales de la condición experimental número 4 del experimento 1.....	69
Figura 28. Arquitectura usada en las redes neurales de la condición experimental número 5 del experimento 1.....	72
Figura 29. Arquitectura usada en las redes neurales de la condición experimental número 6 del experimento 1.....	75
Figura 30. Arquitectura usada en las redes neurales de la condición experimental número 1 del experimento 2.....	79
Figura 31. Arquitectura usada en las redes neurales de la condición experimental número 1 del experimento 3.....	82
Figura 32. Arquitectura usada en las redes neurales de la condición experimental número 1 del experimento 3.....	85
Figura 33. Arquitectura usada en las redes neurales de la condición experimental número 1 del experimento 1.....	88
Figura 34. Arquitectura usada en las redes neurales de la condición experimental número 1 del experimento 1.....	92
Figura 35. Nivel de activación de la unidad CR/UR en la primera fase de la condición experimental 1 del experimento 1.....	96
Figura 36. Nivel de activación de la unidad CR/UR en la segunda fase de la condición experimental 1 del experimento 1.....	97

Figura 37. Nivel promedio de activación de la unidad CR/UR durante los 600 ensayos que implicó la condición experimental 1 del experimento 1.....	97
Figura 38. Nivel de activación de la unidad CR/UR en la primera fase de la condición experimental 2 del experimento 1.....	98
Figura 39. Nivel de activación de la unidad CR/UR en la segunda fase de la condición experimental 2 del experimento 1.....	99
Figura 40. Nivel promedio de activación de la unidad CR/UR durante los 600 ensayos que implicó la condición experimental 2 del experimento 1.....	99
Figura 41. Nivel de activación de la unidad CR/UR en la primera fase de la condición experimental 3 del experimento 1.....	100
Figura 42. Nivel de activación de la unidad CR/UR en la segunda fase de la condición experimental 3 del experimento 1.....	100
Figura 43. Nivel promedio de activación de la unidad CR/UR durante los 600 ensayos que implicó la condición experimental 1 del experimento 1.....	101
Figura 44. Nivel de activación de la unidad CR/UR en la primera fase de la condición experimental 4 del experimento 1.....	102
Figura 45. Nivel de activación de la unidad CR/UR en la segunda fase de la condición experimental 4 del experimento 1.....	102
Figura 46. Nivel promedio de activación de la unidad CR/UR durante los 600 ensayos que implicó la condición experimental 1 del experimento 1.....	103
Figura 47. Nivel de activación de la unidad CR/UR en la primera fase de la condición experimental 5 del experimento 1.....	104
Figura 48. Nivel de activación de la unidad CR/UR en la segunda fase de la condición experimental 5 del experimento 1.....	104
Figura 49. Nivel promedio de activación de la unidad CR/UR durante los 600 ensayos que implicó la condición experimental 1 del experimento 1.....	105
Figura 50. Nivel de activación de la unidad CR/UR en la primera fase de la condición experimental 6 del experimento 1.....	106
Figura 51. Nivel de activación de la unidad CR/UR en la segunda fase de la condición experimental 1 del experimento 1.....	106
Figura 52. Nivel promedio de activación de la unidad CR/UR durante los 600 ensayos que implicó la condición experimental 1 del experimento 1.....	107
Figura 53. Nivel de activación de la unidad CR/UR en la primera fase de la condición experimental 1 del experimento 2.....	108
Figura 54. Nivel de activación de la unidad CR/UR en la segunda fase de la condición experimental 2 del experimento 2.....	109
Figura 55. Nivel de activación de la unidad CR/UR en la segunda fase de la condición experimental 3 del experimento 2.....	109
Figura 56. Nivel promedio de activación de la unidad CR/UR durante los 900 ensayos que implicó la condición experimental 1 del experimento 2.....	110
Figura 57. Ejemplo del Nivel de activación obtenido en la unidad CR/UR durante la condición experimental 1 del experimento 3.....	111
Figura 58. Nivel de activación obtenido en la unidad CR/UR durante la condición experimental 1 del experimento 4.....	112
Figura 59. Nivel de activación obtenido en la unidad CR/UR durante la primera fase de la condición experimental 1 del experimento 5.....	113
Figura 60. Nivel de activación obtenido en la unidad CR/UR durante la segunda fase de la condición experimental 1 del experimento 5.....	113
Figura 61. Nivel de activación obtenido en la unidad CR/UR durante la tercera fase de la condición experimental 1 del experimento 5.....	114
Figura 62. Nivel de activación obtenido en la unidad CR/UR durante la primera fase de la condición experimental 1 del experimento 6.....	115
Figura 63. Nivel de activación obtenido en la unidad CR/UR durante la segunda fase de la condición experimental 1 del experimento 6.....	115
Figura 64. Nivel de activación obtenido en la unidad CR UR durante la tercera fase de la condición experimental 1 del experimento 6.....	116

Figura 65. Nivel de activación obtenido en la unidad CR UR durante la cuarta fase de la condición experimental 1 del experimento 6.....	116
Figura 66. Búsqueda de una Arquitectura adecuada para que una Red DBP pueda manejar 2 CS's.	119
Figura 67. Transformación de una arquitectura de red, con base en la reconfiguración a partir de la selección de conexiones significativas.	121
Figura 68. Nuevo sujeto conformado a partir de las dos redes con pesos no significativos podados.....	121

TABLAS

Tabla 1. Patrones de entrada en cada momento temporal presentados a los sujetos en la primera fase de la condición experimental número 1 del experimento 1.....	61
Tabla 2. Patrones de entrada en cada momento temporal presentados en la segunda fase de la condición experimental número 1 del experimento 1.....	62
Tabla 3. Patrones de entrada en cada momento temporal presentados a los sujetos en la primera fase de la condición experimental número 2 del experimento 1.....	65
Tabla 4. Patrones de entrada en cada momento temporal presentados en la segunda fase de la condición experimental número 2 del experimento 1.....	65
Tabla 5. Patrones de entrada en cada momento temporal presentados a los sujetos en la primera fase de la condición experimental número 3 del experimento 1.....	68
Tabla 6. Patrones de entrada en cada momento temporal presentados en la segunda fase de la condición experimental número 3 del experimento 1.....	68
Tabla 7. Patrones de entrada en cada momento temporal presentados a los sujetos en la primera fase de la condición experimental número 4 del experimento 1.....	71
Tabla 8. Patrones de entrada en cada momento temporal presentados en la segunda fase de la condición experimental número 4 del experimento 1.....	71
Tabla 9. Patrones de entrada en cada momento temporal presentados a los sujetos en la primera fase de la condición experimental número 5 del experimento 1.....	74
Tabla 10. Patrones de entrada en cada momento temporal presentados en la segunda fase de la condición experimental número 5 del experimento 1.....	74
Tabla 11. Patrones de entrada en cada momento temporal presentados a los sujetos en la primera fase de la condición experimental número 6 del experimento 1.....	77
Tabla 12. Patrones de entrada en cada momento temporal presentados en la segunda fase de la condición experimental número 6 del experimento 1.....	78
Tabla 13. Patrones de entrada en cada momento temporal presentados a los sujetos en la primera fase de la condición experimental número 1 del experimento 2.....	80
Tabla 14. Patrones de entrada en cada momento temporal presentados en la segunda fase de la condición experimental número 1 del experimento 2.....	81
Tabla 15. Patrones de entrada en cada momento temporal presentados en la tercera fase de la condición experimental número 1 del experimento 2.....	81
Tabla 16. Primer patrón de entradas presentado a los sujetos en la condición experimental número 1 del experimento 3.....	84
Tabla 17. Segundo patrón de entradas presentado a los sujetos en la condición experimental número 1 del experimento 3.....	84
Tabla 18. Primer patrón de entradas presentado a los sujetos en la condición experimental número 1 del experimento 3.....	87
Tabla 19. Segundo patrón de entradas presentado a los sujetos en la condición experimental número 1 del experimento 3.....	87
Tabla 20. Patrones de entrada en cada momento temporal presentados a los sujetos en la primera fase de la condición experimental número 1 del experimento 5.....	90
Tabla 21. Patrones de entrada en cada momento temporal presentados en la segunda fase de la condición experimental número 1 del experimento 5.....	90
Tabla 22. Patrones de entrada en cada momento temporal presentados en la tercera fase de la condición experimental número 1 del experimento 5.....	91
Tabla 23. Patrones de entrada en cada momento temporal presentados a los sujetos en la primera fase de la condición experimental número 1 del experimento 6.....	93

Tabla 24. Patrones de entrada en cada momento temporal presentados en la segunda fase de la condición experimental número 1 del experimento 6.....	94
Tabla 25. Patrones de entrada en cada momento temporal presentados en la tercera fase de la condición experimental número 1 del experimento 6.....	94
Tabla 26. Patrones de entrada en cada momento temporal presentados en la cuarta fase de la condición experimental número 1 del experimento 6.....	95

ALGORITMOS

Algoritmo 1. Aprendizaje en un Mapa Auto-organizativo de Kohonen (procedimiento general).....	40
Algoritmo 2. Aprendizaje por retropropagación del error (procedimiento general).....	42
Algoritmo 3. Propagación de la señal hacia delante en el aprendizaje por Retropropagación del error.....	43
Algoritmo 4. Propagación del error hacia atrás en el aprendizaje por Retropropagación.....	43

1. INTRODUCCIÓN

Conforme la computación se vuelve cada vez más poderosa, los científicos enfrentan con mayor eficiencia el reto de probar sus modelos de ciertos aspectos de la naturaleza, en una forma rápida y eficiente; ante una gran variedad de situaciones, y tomando en consideración el mayor número de posibilidades.

Con base en redes neurales se han establecido modelos del aprendizaje tanto de corte asociacionista como de corte conductual, que han logrado simular gran cantidad de fenómenos conductuales observados en organismos reales. De entre estos modelos, el Modelo Donahoe-Burgos-Palmer (DBP) está cobrando importancia dentro del análisis de la conducta por dos razones principalmente: primero, es uno de los pocos modelos conductuales que con base en redes neurales ha podido simular con éxito un gran número de fenómenos conductuales; segundo, es uno de los pocos modelos conductuales, no solo neurales, sino en general, que asumen que no existen diferencias entre condicionamiento respondiente y condicionamiento operante, que impliquen dos tipos diferentes de aprendizaje.

Sin embargo, a la fecha, existe un problema relacionado con el Modelo DBP que no ha sido estudiado con el suficiente detenimiento: Las implementaciones computacionales del Modelo DBP que han sido desarrolladas a la fecha sugieren que el modelo no puede manejar dos o más estímulos incondicionados al mismo tiempo. Es decir, una vez que una red neural implementada bajo los lineamientos y criterios que establece el Modelo DBP ha sido entrenada para responder de una manera específica ante un estímulo determinado, no es capaz de ajustar su configuración interna para poder responder ante un estímulo distinto.

A la fecha, esta limitación ha impedido que se simulen con el DBP fenómenos conductuales que impliquen la intervención de dos o más estímulos incondicionados, tales como los siguientes: bloqueo, desbloqueo, ensombrecimiento, condicionamiento de segundo orden, adquisición discriminativa, preconditionamiento sensorial, entrenamiento de patrones negativos¹, entrenamiento de patrones positivos², entrenamiento con varios estímulos simultáneos³, entrenamiento con varios estímulos en serie⁴ e inhibición condicionada.

Específicamente, entonces, con este trabajo se intenta abonar al conocimiento del Modelo, tratando de dilucidar si este problema que de facto se da, tiene relación solamente con la arquitectura de las redes usadas, con el tamaño de los parámetros que se usan, o con algunos aspectos relacionados con el propio proceso de entrenamiento.

El interés del presente trabajo fue continuar con la exploración de la plausibilidad conductual del Modelo DBP, como parte integral del trabajo del modelamiento del comportamiento desde una perspectiva bio-conductual.

Finalmente, y de forma más general, este trabajo pretende abonar al campo de las simulaciones computacionales con redes neurales y simulaciones computacionales de fenómenos conductuales. La simulación digital se ha convertido en un recurso cada vez más utilizado en ciencia experimental (Casti, 1997; Emmeche, 1991), incluyendo por supuesto entre ellas, a las ciencias de la conducta.

¹ Negative Patterning, en inglés.

² Positive Patterning, en inglés.

³ Simultaneous Feature-positive, en inglés.

⁴ Serial Feature-positive, en inglés.

Estos *experimentos imaginarios*⁵, como Dawkins (1982), ha llamado a las simulaciones, sufren constantes ataques por los científicos "clásicos", quienes los acusan de poco realismo, como un argumento en su contra. Queda claro que las simulaciones pueden ser realistas en algunos casos, pero la ventaja precisamente es precisamente, que *pueden no serlo*, sin perder por ello su relevancia. Amén de los casos en los que ser *realista*, independiente de la discusión que la definición del término pueda desatar, es imposible: ¡sólo trate de imaginar una simulación realista de la evolución de las especies!

1.2. PREGUNTA DE INVESTIGACIÓN

La pregunta básica que ha guiado el presente trabajo es:

¿Está relacionada la arquitectura o el tamaño de los vectores de entrada, con la imposibilidad que tienen las redes neurales que funcionan bajo del Modelo DBP para aprender a responder ante un estímulo nuevo una vez que han aprendido a responder ante uno diferente?

1.3. OBJETIVOS GENERALES

A continuación se presentan los objetivos que se pretenden lograr con este proyecto de investigación.

1.3.1. OBJETIVO GENERAL 1

Abonar al conocimiento de los alcances y limitaciones del Modelo DBP explorando el papel que la arquitectura de red, el tamaño de los patrones de entrada y el proceso mismo de entrenamiento tienen en el aprendizaje de estímulos nuevos en situaciones de aprendizaje previo.

⁵ Thought Experiments, en inglés.

1.4. OBJETIVOS ESPECÍFICOS

A partir del objetivo general planteado, se establecieron los siguientes objetivos específicos, que se plantean a continuación siguiendo el orden cronológico con el que pretenden ser logrados:

1.4.1. OBJETIVO ESPECÍFICO 1

Conocer de qué manera afecta la arquitectura de red en el aprendizaje de estímulos nuevos en situaciones de aprendizaje previo.

1.4.2. OBJETIVO ESPECÍFICO 2

Conocer de qué manera afecta el tamaño de los patrones de entrada en el aprendizaje de estímulos nuevos en situaciones de aprendizaje previo.

1.4.3. OBJETIVO ESPECÍFICO 3

Conocer de qué manera afecta la configuración previa de la red en el aprendizaje de estímulos nuevos en situaciones de aprendizaje previo.

2. MARCO CONCEPTUAL

Aunque el Conexionismo se puede rastrear hasta la antigua Grecia y un poco más recientemente, en los trabajos de René Descartes o en los de los histologistas de principios del siglo XIX, la primera red neural, fue presentada por Alexander Bain en 1873 en su libro *Mind and Body. The Theories of their relation*, (Olmsted, 1999). Su trabajo fue altamente influenciado por los hallazgos en neuroanatomía, que entre 1858 y 1863 permitieron la identificación microscópica de las fibras nerviosas e incluso de algunos tipos de axones, gracias a las nuevas técnicas de tinción como la basada en el uso del azul de metileno y de la hematoxilina.

Con base en estas ideas, Bain ideó un sistema de umbrales lógicos que servían para conectar ideas, bajo el supuesto de que dada la estructura estática de la anatomía cerebral, un mismo nervio debía participar en la estimulación de varios, en momentos distintos. A este tipo de relación Bain le llamó *cooperación* (Olmsted, 1999)

La Figura 1 muestra un circuito nervioso con base en umbrales, en el que las conexiones representan ideas que se combinan y cooperan para producir una respuesta específica ante un estímulo dado:

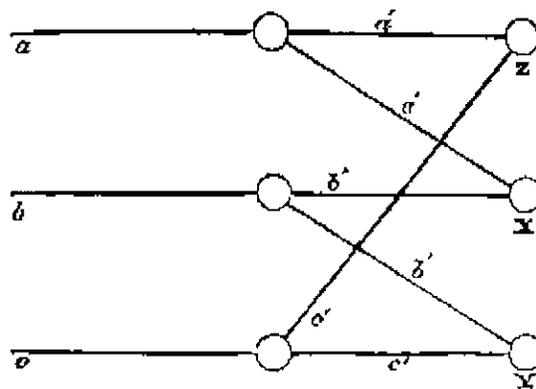


Figura 1. Circuito de ideas cooperantes de Bain (tomado de Olmsted, 1999).

Bain desarrolló más la idea de la resistencia a la estimulación, estableciendo una Teoría de la Adaptación que establecía que la distancia entre nodos en una ruta de estimulación y a la fuerza de la corriente nerviosa eran los dos elementos determinantes del grado de resistencia a la estimulación en partes específicas del circuito (ver Figura 2):

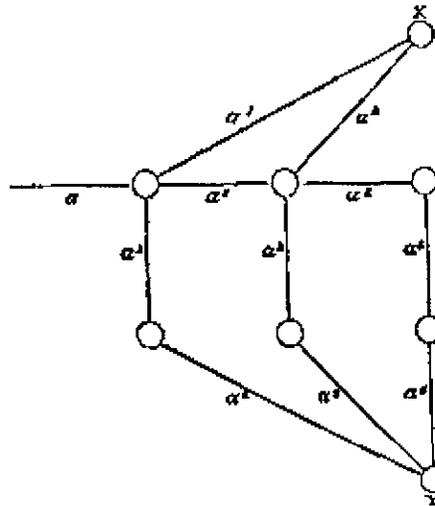


Figura 2. Circuito de ideas cooperantes de Bain con un nodo más resistente ocasionado por una distancia mayor que los del resto del circuito (tomado de Olmsted, 1999).

En 1890 William James rescató la teoría adaptativa de Bain. En esos momentos, el cerebro era concebido como una estructura de redes neurales conectadas al azar, las cuales propagaban o reverberaban corrientes eléctricas en todas las direcciones en una forma análoga a una red de cables metálicos. James propuso que las acciones y los pensamientos se producían como resultado de estas corrientes que fluían en regiones cerebrales (como procesos cerebrales) que tenían un exceso de carga, hacia regiones cerebrales que tenían un déficit de carga eléctrica. La intensidad de los pensamientos y las acciones eran proporcionales a las diferencias de cargas entre regiones. Estas corrientes eléctricas reverberantes fueron llamadas *engramas*. El aprendizaje consistía, según James, en el cambio de las rutas que seguían las corrientes de cargas, hacia nuevas formas, mediante el uso de tres reglas: primero, si los procesos se activaban juntos o en sucesión temporal inmediata, existía la tendencia a la

excitación de uno al otro (James, 1890, p. 566); segundo, las corrientes nerviosas se propagan más fácilmente por las rutas de conducción que más transitadas han sido previamente (James, 1890, p. 563) y, tercero, las rutas que no tienden a usarse producen olvido.

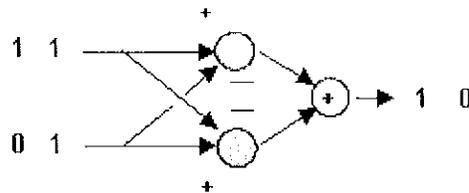


Figura 3. Circuito de Rashevsky para la operación lógica de la disyunción exclusiva (tomado de Olmsted, 1999).

En 1938 Nicolás Rashevsky propuso que el cerebro se organiza con base en operaciones binarias, dado que los potenciales de acción pueden ser vistos como existentes o no existentes en un momento determinado. Rashevsky propuso redes como la que se presenta en la Figura 3, que con base en el uso solamente de operaciones de adición y de sustracción, podía llegar a simular el comportamiento de la operación lógica de la disyunción exclusiva. Uno de los objetivos de Rashevsky era llegar a explicar cómo el comportamiento de los nervios y las redes de nervios podían ser relacionados con procesos psicológicos como el condicionamiento pavloviano (Aizawa, 2002).

En la primera parte de la década de los cuarenta del siglo pasado, Warren McCulloch y Walter Pitts (1943) sentaron las bases para tratar al cerebro como un organismo computacional, basándose en cinco supuestos fundamentales de la neurofisiología de los cuarenta: 1) La transmisión neural es un proceso regido por la Ley del Todo o nada. Es decir, las neuronas son binarias en lo que respecta a su actividad: o transmiten, o no lo hacen. No hay puntos intermedios de transmisión; 2) Para la excitación de una neurona se hace necesario un número fijo de sinapsis en un periodo de adición latente. Las neuronas cuentan con un umbral de excitación fijo, que solamente es rebasado cuando todas las sinapsis que en un tiempo determinado entran a la neurona, producen un estímulo total suficiente

para sobrepasar este umbral; 3) No hay retardos en la actividad del sistema nervioso ajenos al mismo retardo sináptico; 4) Cuando la sinapsis es inhibitoria, la actividad de la neurona implicada se suspende en un periodo de tiempo; y, 5) Las interconexiones entre neuronas no cambian con el tiempo.

Con base en estos cinco supuestos, McCulloch y Pitts idearon un modelo computacional basado en redes de elementos sencillos que podían emular este supuesto funcionamiento del sistema nervioso. Las redes configuradas eran capaces de realizar operaciones lógicas complejas con base en la combinación de módulos sinápticos que realizaban sólo tres operaciones básicas: la conjunción, la negación y la disyunción.

En 1943 Warren McCulloch y Walter Pitts se dieron cuenta de que la consecuencia natural de que el modelo estándar de umbral de neuronas en combinación con los potenciales de acción binarios, producían otro tipo de lógica, llamada lógica de umbrales. Debido a que cada pulso de un potencial de acción es un evento del tipo todo o nada, un valor de umbral de 2 define la operación AND como se muestra en la figura 4. De la misma manera un umbral con valor de 1 define al OR Inclusivo, debido a que tan solo un potencial de acción en cualquiera de las dos líneas es suficiente para producir una salida. Ellos también describieron el uso de la operación de sustracción aunque en sus publicaciones nunca la asociaron explícitamente con la operación condicional.

La lógica de umbrales se expandió al añadirse más líneas de entrada de manera que la salida del nodo de sumación se convirtió en analógica, debido a que parecía adecuarse más con el modelo estándar de neurona. Aún usando más de dos entradas, las operaciones dejaron de ser lógicas, esto es, dejaron de definir el conjunto básico de clasificadores de patrones usado como bloques de construcción en circuitos más grandes para la clasificación de patrones basados en la distribución de los mis-

mos. En vez de eso, estas operaciones se convirtieron en clasificadores de parámetros que clasificaban los patrones de entrada con base en un parámetro de magnitud de la señal dado por la sumación de las entradas binarias.

La sola existencia de valores análogos por primera vez, permitió que las operaciones se volvieran adaptativas usando factores de multiplicación llamados pesos. La manera exacta en la que esta adaptabilidad fue usada, es lo que define los diferentes tipos de redes neurales de aquí en adelante. Aún, la modificación de un parámetro individual no puede caracterizar completamente la distribución (el patrón) de una entrada de una operación.

El trabajo de McCulloch y Pitts adquirió mayor relevancia cuando a estas redes artificiales de neuronas se les añadió la capacidad de aprender con base en las ideas propuestas por Donald Hebb de la Universidad McGill en Canadá, en 1949 (Hebb, 1949):

...cuando un axón de una neurona A está lo suficientemente próximo para excitar a una neurona B, o cuando ésta toma parte frecuentemente en la excitación para producir su disparo, se da un proceso de crecimiento o cambio de metabolismo en una o ambas neuronas, de modo que la eficiencia de A para producir el disparo de B se ve incrementada" (Hebb, 1949, p. 50).

Este hecho le sugirió a Donald Hebb la posibilidad de haber encontrado el sustrato neurofisiológico del aprendizaje:

"... se debe asumir que este cambio estructural a nivel sináptico sucede en situaciones de aprendizaje [...] ya que varias neuronas que mantienen actividad repetidamente terminan asociadas..." (Hebb, 1949, p. 52).

Aunque en la actualidad se ha descubierto que los supuestos en los que McCulloch y Pitts fundamentaron su trabajo, así como la idea de Hebb acerca del aprendizaje, no son del todo ciertos, sí funcionaron como base de un nuevo modelo computacional de procesamiento de información: las redes neurales.

En las décadas de los cincuenta, sesenta y setenta, los avances en redes neurales se ven enterrados debido a dos factores principalmente: por una parte se debe al éxito comercial que tuvieron las aplicaciones derivadas de la visión simbólica de la mente, como los sistemas expertos y los solucionadores generales de problemas, y por otro lado, debido también al ya famoso artículo de Seymour Papert y Marvin Minsky de 1969 (Minsky y Papert, 1980) en el que se demostraba que los perceptrones de Rosenblatt (Rosenblatt, 1958) eran incapaces de resolver problemas cuyo espacio de posibles soluciones no era separable linealmente.

En 1984 el Conexionismo resurge al mundo de las ciencias cognitivas con una nueva visión, gracias a la comprobación hecha por los mismos Minsky y Papert, de que las redes eran capaces de resolver problemas complejos con algunas modificaciones a las arquitecturas originales descritas y atacadas en su artículo de 1969.

Ahora el Conexionismo es definido, como un enfoque computacional para modelar el cerebro, que se basa en la interconexión de un conjunto de unidades simples para producir comportamientos complejos. Bajo este enfoque, los procesos cognitivos residen y se implementan en las células del sistema nervioso. Actualmente se ve como una de las opciones más fuertes por parte de los filósofos de la mente al ser propuesto como la mejor manera de modelar la cognición humana, gracias a las propiedades que emergen de esta organización de elementos de bajo nivel (Pagels, 1991)

Acorde con McClelland, Rummerlhart y Hinton (1994) los modelos conexionistas pueden ser clasificados por la forma específica en que manejan su representación, en dos categorías: los modelos conexionistas distribuidos y los localistas. Las representaciones distribuidas son vectores en un espacio de estados representacionales que son procesados en muchos nodos en una red. Los modelos localistas por otra parte, usan nodos individuales para representar un concepto completo.

En general, las representaciones distribuidas son en realidad a nivel neurológico más realistas que las localistas, aunque este realismo es sumamente limitado, debido a la abstracción misma que el modelo implica. Muchas de las críticas a los modelos conexionistas que provienen de los neurocientíficos se basan precisamente en la falta de realismo neurológico de las redes conexionistas ya que tienen poca recursión, demasiada inhibición, algoritmos de aprendizaje poco realistas, funciones de transferencia simplistas, y no hay hasta el momento algún modelo que contenga elementos que permitan establecer alguna analogía entre ellos y el gran número de neurotransmisores y de hormonas que afectan a la cognición humana.

El Conexionismo ha originado a la Computación Neural o Neurocomputación (Hecht-Nielsen, 1989) como una alternativa a la computación simbólica (aunque, como ya se ha expuesto en este mismo documento, es históricamente anterior a ésta) como un medio para obtener aplicaciones prácticas y funcionales, que cuenten con las propiedades y ventajas que el procesamiento paralelo y distribuido ofrece (McClelland, Rummelhart y Hinton, 1994). Los sistemas computacionales realizados con base en estos modelos son conocidos como Redes Neuronales Artificiales (RNA)⁶ o, simplemente redes neuronales (RN).

⁶ A menudo se usa también el término *Redes neuronales* (o *Neural networks*, en inglés) para hacer referencia a la Neurocomputación en general.

En las redes neurales el proceso de la información se aleja completamente de las formas tradicionales de proceso en Inteligencia Artificial (la vieja inteligencia artificial y sus derivados), es por ello que hay autores consideran a la computación neural como una disciplina *aparte* de la Inteligencia Artificial (Turban, 1992). Sin embargo, debido a la coincidencia entre las redes neurales y la inteligencia artificial en lo referente a algunos de los fines últimos y muchos de los elementos de trabajo utilizados, se considera a la Neurocomputación como una *disciplina estrechamente asociada* (Haykin, 1994 y Churchland, 1993), o incluso, como una línea de trabajo más en la inteligencia artificial bajo una concepción extendida de la disciplina (Skapura, 1995; Minsky, 1991; Freeman y Skapura, 1991)⁷.

Las características que permiten establecer diferencias entre las redes neurales y otras líneas en inteligencia artificial son las siguientes:

- a) La representación del conocimiento. En las redes neurales el conocimiento es representado por medio de valores numéricos distribuidos en la estructura interna de la red, sin implicar una estructura sintáctica compleja. La configuración completa de estados en un sistema puede representarse por medio de vectores y el trabajo de la red puede verse como una simple relación entre espacios de vectores (uno de entrada y uno de salida). Este tipo de representación ha llevado la reformulación de los problemas como la identificación de patrones o el aprendizaje a niveles matemáticamente tan sencillos como encontrar el vector más adecuado (Minsky, 1991).
- b) El proceso paralelo. El alto grado de interconexión entre las unidades de procesamiento implica un grado elevado de resistencia a errores. Esto se ve reflejado en la robustez del sistema.

⁷ La concepción original de Inteligencia artificial consideraba solamente al proceso simbólico de la información, por lo que, desde la definición misma, las redes neurales y otros enfoques de proceso no simbólico como la computación genética, quedan excluidos. Sin embargo, en los autores más recientes se puede identificar una concepción *ampliada* o *extendida* que incluye ya a otras líneas de trabajo no tradicionales.

Igualmente, la posibilidad de contar con muchas unidades de proceso facilita la realización de tareas complejas que mediante un proceso secuencial necesitarían una mayor inversión de recursos.

- c) El proceso distribuido. Cada unidad de proceso realiza tareas muy sencillas, que no van más allá de la suma de entradas y pesos de conexión, y la redefinición del resultado con base en una función de transferencia. La complejidad del proceso radica más en la arquitectura y funcionamiento globales del sistema, que en la estructura y funcionamiento de sus unidades por separado.

Parece no haber una definición universal acerca de lo que son las redes neurales. Sin embargo, con base en su arquitectura y función, se podría decir que las redes neurales artificiales son dispositivos computacionales compuestos a partir de la interconexión de un cierto número de unidades simples, cuya función principal es la transmisión de información numérica. Los canales de transmisión de la información, las reglas para que esta transmisión se lleve a cabo, e incluso la misma arquitectura específica de la red, dependen de la función para la cual la red ha sido diseñada y entrenada; función relacionada, principalmente, con la identificación de patrones.

2.1. ELEMENTOS Y ORGANIZACIÓN DE UNA RED NEURAL

Existen muchos tipos diferentes de redes neurales, y todos ellos implican la integración de diversos elementos y formas de organización. Sin embargo, todos los tipos contemplan unidades de proceso como la que se describe en el apartado siguiente.

2.1.1. UNIDADES DE PROCESO

Las unidades de proceso en una red neural se llaman neuronas artificiales, nodos, o simplemente neuronas.

Cada unidad de proceso en una red neural se caracteriza por:

- a) Un nivel de actividad. El nivel de actividad de la unidad de proceso es binario y representa el estado de polarización de la neurona biológica,
- b) Un valor de salida. El valor de salida de las unidades de proceso en una red neural representa el disparo de la neurona, que se propagará a otras neuronas a través de las conexiones que ésta mantenga.
- c) Un conjunto de conexiones de entrada. Este conjunto de conexiones representa al conjunto de sinapsis entrantes en una neurona biológica y en las unidades de proceso de las redes neurales está compuesto por el total de conexiones que el nodo recibe de otros nodos. Cada conexión tiene asociado un peso o *fuerza de sinapsis* que determina el efecto de la entrada, estos pesos pueden ser positivos (para una sinapsis excitadora) o negativos (para una sinapsis inhibidora).
- d) Un umbral de disparo, que representa el grado de resistencia al disparo, que las neuronas biológicas presentan y en el caso de las unidades de proceso se expresa como un valor.
- e) Un conjunto de conexiones de salida que representa a las sinapsis salientes de una neurona biológica, y en redes neurales consiste en las conexiones que de una neurona salen hacia otras. Cada conexión de salida tiene asociado también un peso o *fuerza de sinapsis* que determina el efecto de la salida. Al igual que en las conexiones de entrada, estos pesos pueden ser positivos (para una sinapsis excitadora) o negativos (para una sinapsis inhibidora).

- f) Una función de salida o también conocida como función de activación. La interacción de las características anteriormente descritas permite establecer el funcionamiento de la unidad de proceso con base en una función de salida a partir de la cual se calcula el disparo o no que la unidad realizará. Esta función de salida puede ser de varios tipos según el modelo de la red que se esté implementando. En la Figura 4 se muestran las tres funciones no lineales más usadas, en el entendido de que sólo son las más usadas, pero no las únicas (Lipmann, 1987).

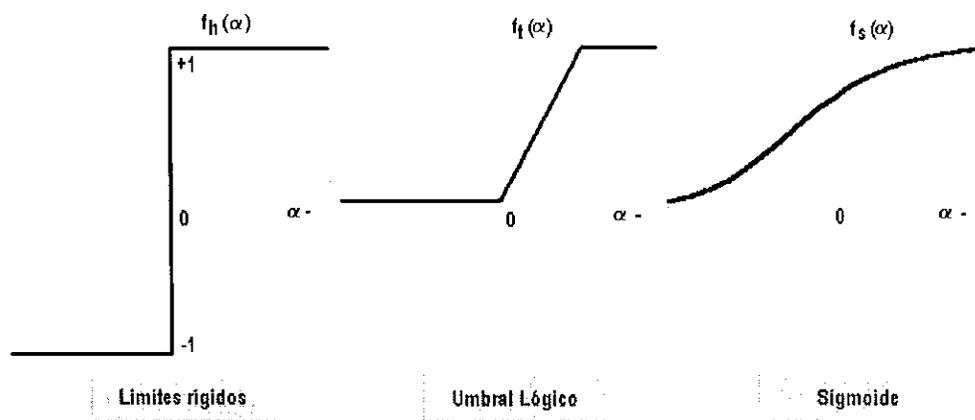


Figura 4. Tres funciones no lineales de las más usadas en redes neurales.

Los elementos que caracterizan a las unidades de procesamiento, listados en la página anterior, se representan matemáticamente por medio de números reales (ver el ejemplo de la Figura 5).

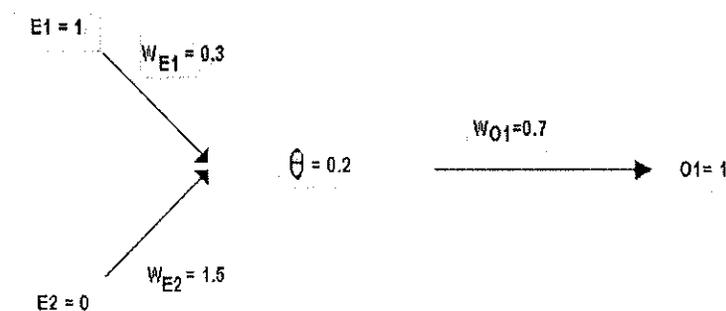
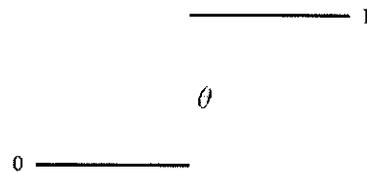


Figura 5. Unidad de proceso con valores de ejemplo.

En la Figura 5, se muestra:

- Una neurona cuyo conjunto de conexiones de entrada esta compuesto por: $E1$ y $E2$
- La fuerza de sinapsis o peso de conexión para $E1$ es $W_{E1} = 0.3$ y para $E2$ es $W_{E2} = 1.5$
- En ambos casos ($E1$ y $E2$) se trata de sinapsis excitadoras.
- El conjunto de conexiones de salida está compuesto por una sola salida que en este caso se ha nombrado $O1$.
- La fuerza de sinapsis o peso de conexión para $O1$ es $W_{O1} = 0.7$
- El umbral de disparo está dado por $\theta = 0.2$
- Dado que la entrada $E1$ es la única entrada activa a la neurona, el efecto total de la entrada es igual a $W_{E1} = 0.3$, que al ser mayor que el umbral implica que la neurona dispare. Es por ello que el nivel de actividad de la neurona de ejemplo es $= 1$.
- Al ser el nivel de actividad de la neurona $= 1$, se da un disparo cuya fuerza es $W_{O1} = 0.7$.
- La función de salida para la neurona de este ejemplo es la que se muestra en la ecuación (1) y en la Figura 6:

$$O1 = \begin{cases} 1 & \text{si } \sum_{i=1}^n E_i \cdot W_{E_i} \geq \theta \\ 0 & \text{de cualquier otra manera} \end{cases} \quad (1)$$



$$f(x) = \left[\sum_{i=1}^n E_i \cdot wE_i \right]$$

Figura 6. Función de salida tipo *escalón*

2.1.2. CAPAS DE UNIDADES DE PROCESO

Las redes neurales típicamente organizan sus unidades de proceso en capas. Estas capas están compuestas de un número de nodos interconectados que usualmente emplean funciones de activación iguales.

Las entradas se presentan a la red neural a través de la primera capa o capa de entrada⁸, la cual comunica con una o más capas ocultas de la red en donde el proceso de la información de entrada se realiza con base en un sistema de pesos de conexión. Las capas ocultas o intermedias, están conectadas a una capa de salida, que es el medio a través del cual la red arroja resultados (ver Figura 7).

⁸ Muchos autores cuentan como capas de procesamiento solamente los nodos en los que se realiza algún tipo de cálculo. Desde este punto de vista, las neuronas de entrada no son consideradas como una capa.

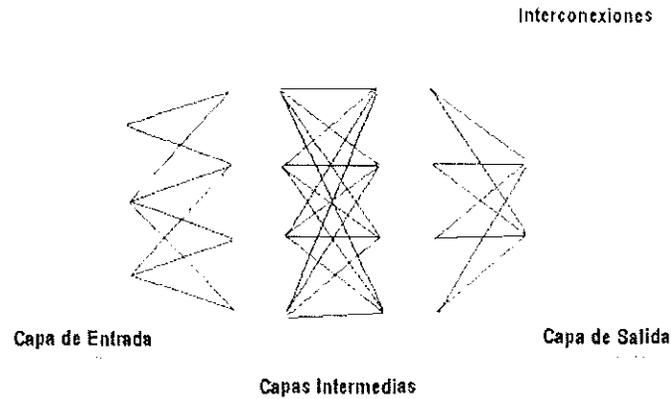


Figura 7. Capas e interconexiones en una red neural típica.

Las redes neurales utilizan una regla de aprendizaje, que modifica los pesos de las conexiones entre las células de acuerdo a los patrones de entrada que le son presentados y sus salidas correspondientes. Así, estos dispositivos aprenden de la manera en que se supone que lo hacen las redes neurales biológicas: reforzando aquellas conexiones que llevan al éxito en el desempeño de una tarea.

2.1.3. TIPOS DE ARQUITECTURA MÁS COMUNES

La manera en la que las neuronas de una red neural se estructuran se relaciona estrechamente con el modelo de aprendizaje a partir del cual la red debe funcionar. En general, se pueden identificar varios tipos de arquitectura de red, de los cuales se presentan a continuación los más comunes.

2.1.4. REDES NEURALES DE UNA SOLA CAPA Y UN SOLO SENTIDO DE PROPAGACIÓN DE SEÑAL⁹

La forma más simple de arquitectura de redes neurales, en cuanto a las capas que las integran, es la red neural de una sola capa (o dos capas, si se considera como capa a las unidades de entrada).

⁹ Single-layer Feedforward Network, en inglés

Este tipo de redes neurales se conforman a partir de una serie de nodos de entrada que están conectados directamente a los nodos de la capa de salida (ver Figura 8).

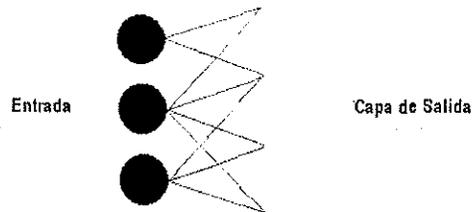


Figura 8. Red neural de una sola capa.

Una de las características principales de este tipo de redes es su imposibilidad para regresar la señal a las unidades de la capa de entrada. Es decir, este tipo de redes propaga la señal en un solo sentido que va desde la entrada hasta éste es de la entrada a la salida (lo que en el ámbito de las redes neurales se conoce como propagación hacia delante o sentido de transmisión delantero).

Las memorias asociativas lineales son un ejemplo típico de esta clase de redes neurales, en las cuales se realiza una asociación entre el patrón de entrada y la información guardada en la estructura de la red. Como en toda red neural, la información se integra en la estructura de la red a partir de las configuraciones de los pesos de interconexión en cada nodo de ésta.

2.1.5. REDES NEURALES MULTICAPA DE UN SOLO SENTIDO DE PROPAGACIÓN DE SEÑAL¹⁰

Este tipo de redes neurales se distingue por la presencia de una o más capas intermedias entre la entrada y la capa de salida. A estas capas intermedias a menudo se les conoce también como capas

¹⁰ Multilayer Feedforward Network, en inglés

escondidas u ocultas¹¹. La función de las unidades de las capas escondidas es dotar a la red de capacidades de cálculo más poderosas, necesarias sobre todo cuando se trata de procesar vectores de entrada sumamente complejos o extensos.

En el caso de las redes neurales multicapa, cada una de las salidas de las unidades de procesamiento en cada capa sirve como entrada para los nodos de la siguiente capa.

Se puede hablar de redes neurales multicapa completamente interconectadas, cuando las neuronas de cada una de las capas están conectadas con todas y cada una de las neuronas de la siguiente capa (ver Figura 9). En cambio, cuando algunas de las neuronas de la siguiente capa no son conectadas a un nodo determinado, entonces se habla de redes neurales parcialmente conectadas (ver Figura 10). Este tipo de redes se usan comúnmente en casos en los que se busca algún tipo de proceso especial en cada uno de los nodos que integran la red, o cuando se ha atribuido un significado a los nodos de las capas intermedias. Un ejemplo claro de este caso es el de los sistemas expertos neurales, en los que los nodos de las capas intermedias están relacionados con padecimientos específicos y los nodos de la capa de salida con tratamientos específicos para estos trastornos (Gallant, 1994).

Tanto en el caso de las redes parcialmente interconectadas, como en el de las redes totalmente interconectadas, se habla de que la propagación de la señal se da en un solo sentido. Y al igual que en las redes de una sola capa, el sentido de propagación es hacia delante.

¹¹ Hidden Layers, en inglés

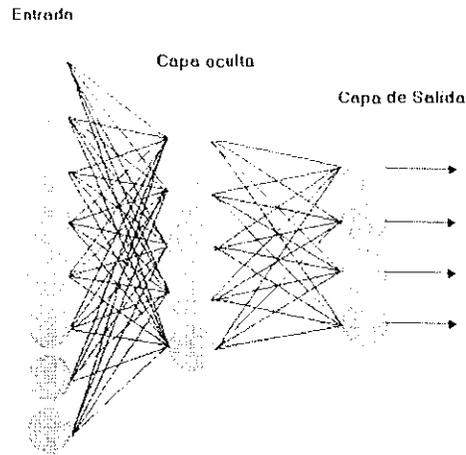


Figura 9. Red neuronal totalmente conectada.

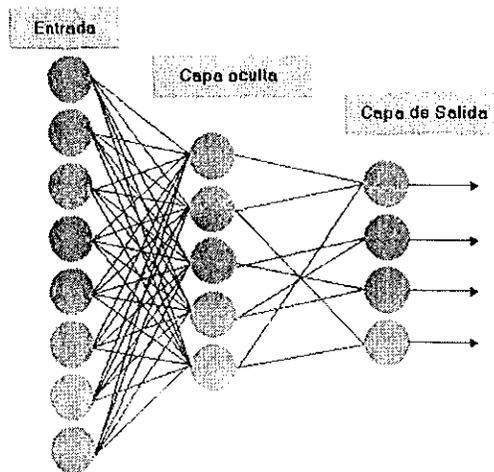


Figura 10. Red neuronal parcialmente conectada.

2.1.6. REDES NEURALES RECURRENTE¹²

Una red neuronal recurrente se distingue por tener por lo menos una conexión de retroalimentación. Es decir, integra un nodo en alguna de las capas intermedias o de salida que tiene conexión, no solamente con alguna o algunas de las neuronas de la capa siguiente, sino que también integra conexiones con neuronas de capas anteriores (ver Figura 11). Este tipo de redes se emplea sobre todo, cuando se requiere lograr algún tipo de retardo en el proceso de las entradas de la red, o

¹² *Recurrent Network*, en inglés

cuando se quiere mantener algún tipo de información latente que propicie un comportamiento de la red no-lineal.

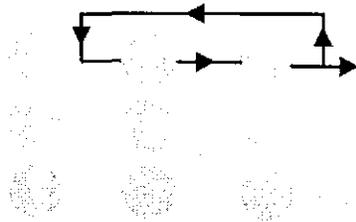


Figura 11. Red neural recurrente.

En la Figura 11 la red neural presentada integra una conexión entre un nodo de la capa de salida y un nodo de la capa intermedia

2.1.7. REDES NEURALES ESTRUCTURADAS EN LÁTICES

Un látice consiste en un arreglo, unidimensional, bidimensional o multidimensional de neuronas con un conjunto correspondiente de nodos de entrada que proporcionan las señales de entrada al arreglo. Las dimensiones del látice hacen referencia al número de dimensiones en el espacio en el que el grafo descansa. La Figura 12 muestra un ejemplo de red neural tipo látice unidimensional, y la Figura 13 muestra uno bidimensional. Nótese que en ambos casos cada nodo de entrada está conectado con cada neurona en el látice. Las estructuras en látices pueden ser vistas como redes neurales de un solo sentido con las neuronas de salida organizadas en filas y columnas (Haykin, 1994)

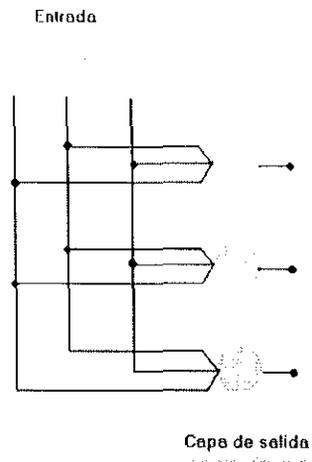


Figura 12. Red neuronal tipo láctice de una dimensión.

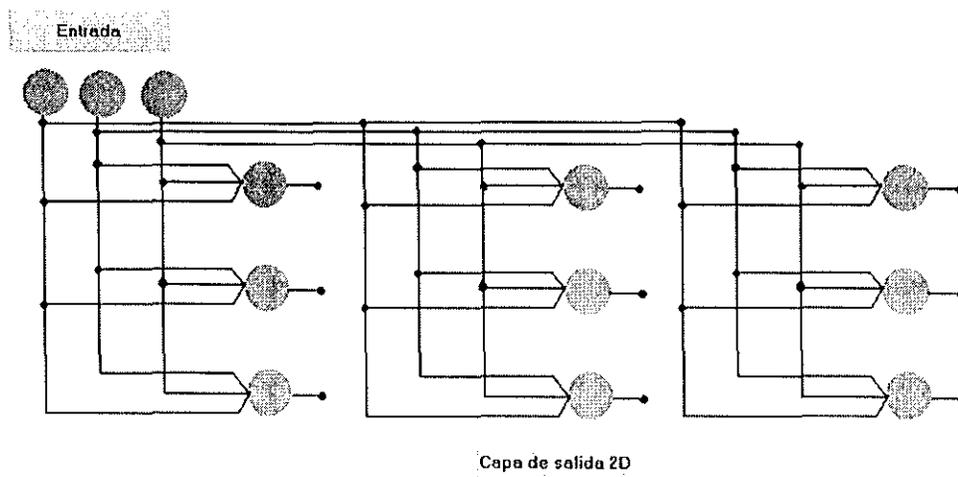


Figura 13. Red neuronal tipo láctice de dos dimensiones.

2.2. FUNCIONAMIENTO DE UNA RED NEURAL

El funcionamiento de una red neuronal se puede dividir en dos grandes procesos: transmisión de los impulsos a través de la estructura interna de la red, y aprendizaje. Estos dos grandes procesos están estrechamente asociados, respectivamente, con las dos funciones básicas atribuidas a una red neuronal: la identificación o clasificación de patrones de entrada (que se realiza mediante la propagación de los impulsos) y la aproximación de funciones, que es la base del aprendizaje de red.

2.2.1. TRANSMISIÓN DE LOS IMPULSOS A TRAVÉS DE LA ESTRUCTURA DE LA RED

Los componentes de una red neural interactúan para transmitir o propagar, datos a través de toda la estructura de la red, de la manera siguiente:

- a) Cada una de las unidades de la capa de entrada se asigna a alguna característica o atributo del problema a tratar. Así, se tendrán tantas unidades en la capa de entrada, como características sobresalientes en el problema a tratar. Las neuronas de esta capa de entrada se alimentan con los valores correspondientes de cada característica del problema en un caso específico.
- b) Las entradas se procesan en cada nodo de la red acorde con las funciones de salida ya definidas y se transmite el valor resultante a cada una de las unidades con las cuales el nodo actual, tiene conexión. Así, los valores resultantes de cada nodo se alimentan al siguiente nodo como entrada para un nuevo procesamiento. La transmisión de las entradas termina cuando los nodos de la capa de salida son alimentados y éstos han ofrecido su resultado al usuario.
- c) La salida de la red representa la solución al problema. En la capa de salida de una red neural se tendrán entonces, tantas unidades como soluciones pueda tener el problema y el nodo activado es el que indica cuál es la solución calculada para el problema procesado.

2.2.2. APRENDIZAJE EN UNA RED NEURAL

Por aprendizaje en una red neural se debe entender al proceso en virtud del cual los pesos de conexión de cada nodo de la red sufren ajustes de manera tal que la capa de salida pueda ofrecer la solución adecuada al problema presentado en la capa de entrada.

Las redes neurales usan formas muy específicas para realizar este ajuste de pesos de conexión, que van de acuerdo al modelo de red bajo el cual están funcionando. Es posible clasificar los méto-

dos de aprendizaje en dos grandes grupos: el aprendizaje no supervisado y el aprendizaje supervisado.

2.2.3. APRENDIZAJE NO SUPERVISADO

El aprendizaje no supervisado es otra modalidad usada para el ajuste de pesos en las redes neurales. En este tipo de aprendizaje a la red no se le proporcionan modelos de respuestas correctas a obtener a partir de ciertas entradas. Simplemente se introducen los ejemplos del conjunto de casos de entrenamiento en el sistema y se espera que la red neural agrupe los ejemplos con base en la cercanía que entre cada uno de ellos existe. El cálculo de cercanía para el agrupamiento posterior se puede realizar de diversas maneras que van desde procedimientos sencillos como los que se basan en la distancia de Hamming¹³, hasta otros más complejos (Skapura, 1995; Maravall Gómez-Allende, 1994)

Para que el sistema realice el procedimiento de agrupamiento de ejemplos se necesita la constante modificación de los pesos de conexión entre las neuronas que integran la red que permita posteriormente, asignar correctamente cada uno de los ejemplos al grupo adecuado. En este caso también se espera que la red neural ya entrenada sea capaz de trabajar eficientemente con casos similares pero no introducidos previamente.

Entre otros, los modelos que se mencionan a continuación son representativos de la línea de investigación basada en el aprendizaje no supervisado.

- Redes de Hopfield (Hopfield, 1982).
- Mapas auto-asociativos de Kohonen (Kohonen, 1982).

- Los Mapas auto-asociativos bidimensionales de Kohonen (Kaski, 1997).

2.2.3.1. Mapas auto-organizativos de Kohonen

Los Mapas Auto-organizativos, son también conocidos como Mapas de Características, SOM's o Mapas de Kohonen, como también se les conoce por su autor, el Finlandés Teuvo Kohonen, quien diseñó este modelo con la finalidad de modelar la manera en que podría explicarse la formación de mapas topológicos en el cerebro por medio del aprendizaje.

Este tipo de redes neurales pueden ser usados como capas de procesamiento para otras redes neurales, o como aplicaciones independientes específicas para clasificación. Ya sea como capa de procesamiento o como aplicación independiente, el tipo de aprendizaje en el cual se basan los Mapas, es el competitivo, que se guía bajo la estrategia de "el ganador se lleva todo".

Existen en la actualidad dos variantes de los Mapas de Kohonen, los mapas para el Aprendizaje con base en la Cuantificación de Vectores y los Mapas con preservación de la Topología (Kohonen, 1982)

La arquitectura de los Mapas Auto-organizativos es sencilla, se integra por una capa de unidades de entrada que se conecta completamente con una capa de unidades de salida de una o dos dimensiones (Ver Figura 14).

¹³ La distancia de Hamming (D_H) puede ser definida como el total de elementos distintos en valor y posición, entre dos vectores. Por ejemplo el vector 1={1,0,1,1,0} se encuentra a una $D_H = 1$, del vector 2={1,1,1,1,0}.

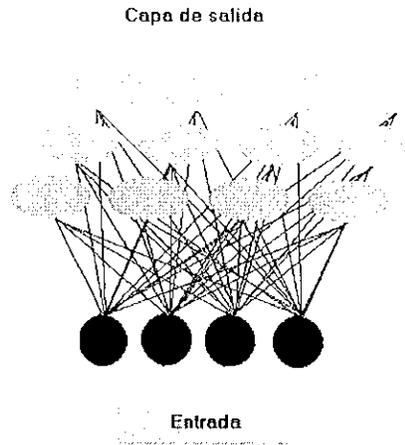


Figura 14. Arquitectura de un Mapa Auto-organizativo.

En este tipo de redes las entradas alimentan a la red directamente a las neuronas de procesamiento dispuestas en una o dos capas de proceso¹⁴, en las que cada salida se calcula de acuerdo a la siguiente ecuación:

$$\text{Salida} = \sum_{i=1}^n E_i \cdot W_{ij} \quad (2)$$

En donde E son las entradas a la red, W_{ij} los pesos que van de la entrada i a la neurona j de la capa de salida, con magnitudes de pesos y de vectores de entrada normalizados.

En este tipo de redes, la neurona con la salida más grande es la neurona ganadora, cuya salida se coloca en 1, y el resto de las salidas de las neuronas de la red se colocan en 0.

El establecimiento de diferencias entre los patrones de entrada se logra con base en las distintas neuronas ganadoras, en donde patrones iguales o similares son relacionados con la misma neurona ganadora.

¹⁴ En el modelo original, la arquitectura de los Mapas de Kohonen emplea una sola capa de proceso, que a la vez es la capa de salida de la red. Sin embargo, modificaciones hechas al modelo entre 1996 y el 2003 han llevado a la inclusión de una capa adicional de salida, para problemas cuya solución implique clasificaciones en dos o más dimensiones.

Para comprender los Mapas auto-organizativos es necesario antes explorar dos procesos que se dan durante el aprendizaje:

a) La Normalización de Vectores, que se da de la siguiente manera:

Si consideramos un vector $A = ax + by + cz$, el vector normalizado A' se obtiene dividiendo cada uno de los componentes de A entre la raíz cuadrada de la suma de cuadrados de todos los componentes. En otras palabras, cada componente es multiplicado por $1/\sqrt{a^2 + b^2 + c^2}$. Tanto el vector de pesos como el vector de entradas son normalizados durante la operación del Mapa Auto-organizativo de Kohonen, debido a que la ley de entrenamiento usa la resta del vector de pesos al vector de entradas. Usando la normalización de los valores en la sustracción, se reducen ambos vectores a estatus de unidades, y con ello la sustracción se hace posible.

b) La Inhibición Lateral, que es un proceso que se deriva de observaciones realizadas a redes neurales biológicas, sobre todo en circuitos neurales relacionados con la visión humana (Kohonen, 1982). Las conexiones laterales en una capa dada, están determinadas por una fuerza que es inversamente proporcional a la distancia entre las unidades. Las conexiones con valores positivos son consideradas conexiones excitatorias y las que tienen conexiones con valores negativos son consideradas conexiones inhibitorias.

La relación entre la fuerza de conexión y la distancia de la neurona ganadora, se representa mediante una función de tipo *sombrero mexicano* (ver Figura 15).

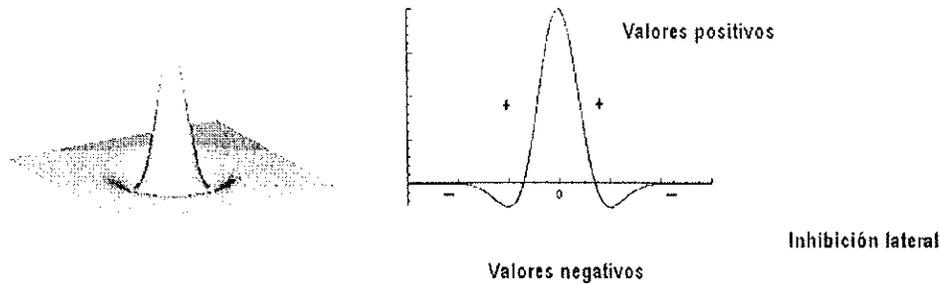


Figura 15. Función de sombrero mexicano, mostrando la inhibición lateral.

El efecto de esta función es la creación de un ambiente competitivo para el aprendizaje (Kohonen, 1982). Solamente las neuronas ganadoras y sus vecinos cercanos participan en el aprendizaje a propósito de un patrón de entrada dado.

La ley de aprendizaje usada en los Mapas Auto-organizativos tiene un solo sentido. Los cambios en los vectores de pesos para una neurona dada sufren ganancias constantes alfa, multiplicados por la diferencia entre el vector de entradas y el vector de pesos en el tiempo $t - 1$:

$$w_i = w_{i-1} + a(X_i - w_{i-1}) \quad (3)$$

Tanto el vector de pesos en $t - 1$ como el vector de entradas son normalizados. Alpha es una constante que se fija al momento de la ejecución del proceso de entrenamiento y sus límites son $[0, 1]$.

En los Mapas Auto-organizativos de Kohonen, existe un parámetro además de Alpha, que se conoce como *tamaño del perímetro de vecindad*¹⁵, que se usa para modelar el efecto de la función de sombrero mexicano, ya descrita. Aquellas neuronas que se encuentran dentro del perímetro de vecindad de la neurona ganadora participan en el entrenamiento en el momento t y en la modificación del vector de pesos. Este parámetro de perímetro de vecindad comúnmente se inicia con un valor mayor, dentro de

¹⁵ *Neighborhood size*, en inglés.

los límites de [0,1] por supuesto, y se va reduciendo conforme avanzan las iteraciones durante el entrenamiento, hasta llegar a señalar solamente a las neuronas ganadoras específicamente.

Actualmente existen muchas versiones del algoritmo para los Mapas Auto-organizativos de Kohonen. A continuación se describe uno de los más usuales y efectivos que implica el uso de una sola capa de unidades de salida, es decir, usa la arquitectura originalmente planteada por Kohonen (Freeman y Skapura, 1991; Kohonen, 1982):

1. Se asume que los nodos de la capa de salida están conectados en un arreglo o array de una o dos dimensiones
2. Se asume que la red construida está completamente interconectada, es decir, todos los nodos o unidades de la capa de entrada se conectan con cada nodo de la capa de salida
3. Todos los pesos $w_{i,j}$ son inicializados con valores aleatorios comprendidos entre [0,1]
4. Se conforma un conjunto de vectores de entrada X
5. Se usa aprendizaje competitivo:
 - 3a. Se elige aleatoriamente un vector de entrada X_i
 - 3b. Se determina la neurona ganadora en la capa de salida, definida como aquel nodo pésimo que cumple con la condición $w_i x \geq w_k x$ solo si los pesos están normalizados
6. Dada la iésima neurona ganadora, se actualizan los pesos de la siguiente manera:
 $w_k t = w_k t - 1 + \mu \mathfrak{N}(i, k)(x - w_k)$, en donde $\mathfrak{N}(i, k)$ es la función de perímetro de vecindad, que tiene un valor de 1 cuando $i=k$ y sale de la distancia comprendida entre $|r_k - r_i|$ entre las unidades i y k del arreglo de unidades de salida. Así, las unidades que se encuentran en vecindad con la neurona ganadora, e incluso la propia neurona ganadora ven modificados sus pesos significativamente. Los pesos asociados con las unidades fuera de la vecindad establecida para la neurona ganadora no cambian significativamente.
7. El proceso se repite hasta que se cumple con un criterio de número de iteraciones alcanzado o hasta que las vecindades entre neuronas no sufren modificaciones
8. La función de establecimiento de perímetro de vecindad más frecuentemente usada es la siguiente:

$$\mathfrak{N}(i, k) = e^{-\frac{|r_k - r_i|^2}{(2\delta^2)}}$$
, en donde δ^2 es un parámetro de amplitud [0,1] que se ve decrementado con el tiempo.
9. El proceso se repite presentando cada patrón del conjunto de vectores, cuidando siempre que la selección del patrón a presentar sea aleatoria. Asimismo, el conjunto de vectores se presentará completo varias veces. El número es determinado por el usuario, pero siempre debe ser un número grande (usualmente se usa un $t \geq 500$).

Algoritmo 1. Aprendizaje en un Mapa Auto-organizativo de Kohonen (procedimiento general).

2.2.4. APRENDIZAJE SUPERVISADO

En el aprendizaje supervisado la adecuación de los pesos de conexión entre cada nodo se realiza evaluando el desempeño general de la red para lograr una salida deseada ante una configuración de patrones de entrada determinados.

En el proceso de aprendizaje supervisado, el entrenamiento de una red consiste en la evaluación de la salida actual que la red neural presenta ante una entrada, y su comparación con la salida que la red *debería presentar* ante esta entrada. La diferencia entre la salida obtenida y la salida deseada se usa para calcular el ajuste que debe darse al interior de la red en cada uno de los pesos de conexión, a fin de que se logre obtener la salida.

A este tipo de aprendizaje se le conoce como aprendizaje supervisado ya que el uso de salidas deseadas ante entradas determinadas permite supervisar el desempeño de la red en cada momento y con ello, la realización de los ajustes de pesos necesarios en la dirección correcta.

Una de las características de las redes neurales asociadas con este tipo de aprendizaje es la capacidad para generalizar los aprendizajes logrados. Se considera que un proceso de aprendizaje supervisado en una red neural ha tenido éxito cuando el ajuste de pesos, realizado a partir de un conjunto de ejemplos, le permite tener un desempeño efectivo ante problemas similares a los ejemplificados, pero nunca antes introducidos en el sistema.

2.2.4.1. La retropropagación del error

Quizá el algoritmo de aprendizaje más famoso en redes neurales sea el que se basa en la retropropagación del error¹⁶. Este algoritmo fue descrito en la compilación de Rumelhart y McClelland (1986) aunque parece haber un reconocimiento generalizado de que los creadores del algoritmo de aprendizaje fue propuesto por Paul Werbos (Werbos, 1974) y por Parker (Parker, 1985).

La arquitectura de una red con aprendizaje basado en la retropropagación el error, tiene un diseño jerárquico consistente en un conjunto de capas de unidades de proceso completamente interconec-

¹⁶ *Backpropagation*, en inglés.

tadas. El proceso de la información bajo la retropropagación tiene por objetivo realizar la aproximación de una función $f : A \subset R^n \rightarrow R^m$ a partir de un subconjunto A de un espacio euclidiano n -dimensional a un subconjunto delimitado $f[A]$ de un espacio euclidiano m -dimensional, por medio de un conjunto de ejemplos de entrenamiento $(x_1, y_1), (x_2, y_2), \dots, (x_k, y_k)$ de la correspondencia, en donde $y_k = f(x_k)$. Se asume que estos ejemplos de la correspondencia f son generados seleccionando de A , x_k vectores al azar de acuerdo con una función fija de probabilidad $p(x)$.

Actualmente existen muchas variables del algoritmo. A continuación se describe uno de los más usuales y efectivos que implica el uso de una sola capa de unidades ocultas (ver: Freeman y Skapura, 1991), en el entendido de que su extensión a más unidades es sencilla.

1. Se aplica un vector de entrada a la red y se calculan los correspondientes valores de salida.
2. Se comparan las salidas obtenidas con las salidas correctas y se determina una medida del error.
3. Se determina en qué dirección (positiva o negativa) debe cambiar cada peso con objeto de reducir el error.
4. Se determina la cantidad en que es preciso cambiar cada peso
5. Se aplican las conexiones a los pesos
6. Se repiten los pasos del 1 al 5 con todos los patrones de entrenamiento hasta que el error para todos ellos se haya reducido hasta una cantidad aceptable

Algoritmo 2. Aprendizaje por retropropagación del error (procedimiento general).

Como podrá observarse en el Algoritmo 1 el trabajo con la retropropagación del error implica la realización de dos fases o momentos en la red neural: una fase de propagación de la señal, que es precisamente la fase de operación o consulta de la red. En esta fase, gracias a la propagación de la señal desde las unidades de entrada, hasta la capa de salida, la red puede establecer una correspondencia entre vectores de entrada y una salida correspondiente. La segunda fase es la de propagación del error hacia atrás o retropropagación. Esta fase que tiene un papel muy importante en el proceso de entrenamiento de la red, ya que gracias a ésta precisamente, es que se pueden ajustar los pesos en la arquitectura interna, para que en la próxima propagación de la señal, la red sea mucho más eficiente. En términos neurocomputacionales, a este ajuste de pesos de le llama también

aproximación de la función. Los pasos que tienen lugar en estas dos fases se presentan en el Algoritmo 3 y el Algoritmo 4.

1. Se localiza la primera unidad de procesamiento de la capa que se encuentre inmediatamente por encima de la capa actual.
2. Se pone a cero el total actual de la entrada.
3. Se calcula el producto del primer peso de conexión de entrada, por la salida de la unidad transmisora.
4. Se añade el producto al total acumulado
5. Se repiten los pasos 3 y 4 para todas las conexiones de entrada
6. Se calcula el valor de salida para esta unidad aplicando la función de salida
7. c en donde x es la entrada total
8. Se repiten los pasos del 2 al 6 para todas las unidades de esta capa
9. Se repiten los pasos del 1 al 7 para todas las capas de la red.

Algoritmo 3. Propagación de la señal hacia delante en el aprendizaje por Retropropagación del error.

Una vez que se ha calculado un valor de salida para todas las unidades de la red, se comparan los valores calculados para las unidades de la capa de salida, con la salida que se ha establecido como deseada, elemento por elemento como se muestra en el Algoritmo 3.

Se calcula un valor de error en cada unidad de salida. Estos términos de error son realimentados en todas las demás unidades de la estructura de la red mediante la siguiente secuencia de pasos:

1. Se localiza la primera unidad de proceso que se encuentre inmediatamente por debajo de la capa de salida.
2. Se pone a cero el error total actual.
3. Se calcula el producto del peso de conexión de la primera salida por el error proporcionado por la unidad de la capa superior.
4. Se añade ese producto al error acumulado
5. Se repiten los pasos 3 y 4 para todas las unidades de salida
6. Se multiplica el error acumulado por $o(1 - o)$ en donde o es el valor de salida de la unidad de la capa oculta que se ha producido durante la operación de la propagación de la señal hacia delante
7. Se repiten los pasos del 2 a 6 para todas las unidades de la capa
8. Se repiten los pasos 1 al 7 para todas las capas
9. Se localiza la primera unidad de proceso que esté en a capa superior a la capa de entrada
10. Se calcula el valor de cambio del peso para la primera conexión de entrada de esta unidad añadiendo una fracción del peso acumulado en esta unidad al valor de entrada de la unidad.
11. Se modifica el término de cambio de peso añadiendo un término de momento, el cual es igual a una fracción del valor del cambio de peso procedente de la iteración anterior.
12. Se guarda el nuevo valor del cambio de peso como valor anterior del cambio de peso para esta conexión.
13. Se modifica el peso de la conexión añadiendo el valor del nuevo cambio de peso de conexión al peso anterior de conexión
14. Se repiten los pasos del 10 al 13 para todas las conexiones de entrada de esta unidad
15. Se repiten los pasos del 10 al 14 para todas las unidades de esta capa
16. Se repiten los pasos del 10 al 15 para todas las capas de la red

Algoritmo 4. Propagación del error hacia atrás en el aprendizaje por Retropropagación

Dado que el algoritmo de retropropagación del error es altamente costoso en términos computacionales, se han realizado numerosos esfuerzos para acelerar su convergencia. Sin embargo, estas

modificaciones aún no han sido lo suficientemente buenas como para desplazar del lugar exitoso que ha ocupado el algoritmo original desde su propuesta (Freeman y Skapura, 1991).

Otros ejemplos de redes neurales que funcionan bajo el esquema del aprendizaje supervisado son, entre muchos, los siguientes:

- Perceptrón (Rosenblatt, 1958).
- Contrapropagación (Freeman y Skapura, 1991).
- Máquinas de Boltzmann (Freeman y Skapura, 1991).

2.2.5. APRENDIZAJE POR REFUERZO

Muchos de los algoritmos para aprendizaje supervisado y no supervisado en redes neurales, requieren de datos de entrenamiento precisos para establecer los pesos de conexión y la conectividad entre los nodos de la red. Sin embargo, en la vida real, es usualmente difícil obtener datos precisos para poder realizar un proceso de entrenamiento y en muchos de los casos es imposible. Es por ello que existe un creciente interés por los algoritmos de aprendizaje con base en refuerzo.

En este tipo de aprendizaje los datos de entrenamiento tienen un papel de evaluadores para el problema y en nada se parece su uso a la retroalimentación "instructiva" en el caso del aprendizaje supervisado. Bajo este paradigma no se tiene un modelo completo del comportamiento de la red deseado; es decir, no se indica durante la fase de entrenamiento exactamente cuál es la salida que se espera de la red ante una entrada determinada. En este caso, la función del supervisor se reduce a indicar mediante una señal de refuerzo si la salida obtenida en la red se ajusta en alguna medida a la deseada y en función de ello se ajustan los pesos con base en un mecanismo de probabilidades. Se podría decir que en este tipo de aprendizaje, la función del supervisor se asemeja más a la de un crítico que opina so-

bre la respuesta de la red, que a la de un maestro que indica la respuesta correcta que se debe generar (Hilera y Martínez, 2000).

El Aprendizaje por refuerzo presenta retos teóricos interesantes. Además de que la señal de refuerzo es solamente aproximada, solamente está disponible una vez en una cadena de eventos o acciones ocurridas. Ello implica que esta señal pudo haber sido causada por una acción realizada muchos pasos atrás o por toda la secuencia completa con múltiples grados de contribución en cada uno de los pasos.

Desde el punto de vista cognitivo y biológico el aprendizaje por refuerzo es más cercano a la forma en la que podemos suponer que se realiza el aprendizaje en organismos superiores, habiendo incluso autores como Cheng-Lian y Chin-Teng Lin (1996) que han inspirado sus algoritmos de aprendizaje por refuerzo en modelos matemáticos del comportamiento que incluso han logrado que se emita respuesta antes de que la señal de refuerzo se proporcione.

El desarrollo del aprendizaje por refuerzo puede ser rastreado y dividido en dos etapas: la primera etapa que comienza en los cincuenta cuando psicólogos y matemáticos desarrollaron modelos computacionales para explicar el aprendizaje en animales y en seres humanos, quienes vieron el aprendizaje como un proceso estocástico (Cheng-Lian y Chin-Teng Lin, 1996)

Casi al mismo tiempo, los cibernetas y los estudiosos del control, estaban realizando esfuerzos independientes en el estudio del aprendizaje estocástico. Su trabajo básicamente se basaba en el uso de autómatas deterministas y después autómatas estocásticos, como un modelo para sistemas de aprendizaje operando en sistemas de ambientes aleatorios no dinámicos (Pedrycz y Waletzky, 1997)

En la primer etapa de los algoritmos para aprendizaje por refuerzo, se hablaba de un proceso "no asociativo" debido a que no había entradas al sistema excepto por la señal de refuerzo (Cheng-Lian y Chin-Teng Lin, 1996)

Ejemplos de algoritmos por refuerzo de fines generales, son: Linear-reward penalti o LRP de Narendra y Tathacher en 1974, (Hilera y Martínez, 2000); Associative Reward-Penalty ARP de Barto y Anandan en 1985 que se aplica en redes de conexiones hacia adelante de dos capas cuyas neuronas de salida presentan una función de activación estocástica, (Hilera y Martínez, 2000), y Adaptative Heuristic Critic presentado por Barton, Sutton y Anderson en 1983, que se utiliza en redes de propagación delantera de tres capas, diseñadas especialmente para que una parte de la red sea capaz de generar un valor interno de refuerzo que es aplicado a las neuronas de salida de la red, (Hilera y Martínez, 2000).

En el siguiente apartado se tratarán algunos modelos que funcionan con base en el aprendizaje por refuerzo, pero que fueron realizados con la finalidad específica de simular fenómenos conductuales.

2.3. MODELOS BASADOS EN REDES NEURALES PARA LA SIMULACIÓN DE FENÓMENOS CONDUCTUALES

Es bajo este modelo de aprendizaje, que podemos ubicar a los modelos específicos para la simulación de fenómenos conductuales. Las redes neurales pueden proporcionarnos elementos para conocer cuáles son los procesos que el aprendizaje implica a nivel cerebral y la manera en la que ocurren (Moore, 1991). En los siguientes párrafos se presentan al lector varios modelos de aprendizaje por refuerzo específicamente dedicados para la simulación y estudio de fenómenos conductuales. Aunque todos ellos tienen un objetivo común, existen diferencias teóricas importantes entre ellos, sobresaliendo principalmente la forma en la que conciben al reforzamiento. Existen modelos que asumen una diferencia entre condicionamiento clásico y condicionamiento operante y con base en ello, se dedican a

simular uno u otro aprendizaje, entre ellos podremos encontrar el Modelo DYSTAL de Alkon, Vogl y Tam (1991), el Modelo GMADN de Baxter *et al* (1991) y por otro lado podemos encontrar a los modelos que asumen que no existe distinción alguna entre los dos condicionamientos, como el Modelo de Donahoe, Burgos y Palmer (1993).

2.3.1 EL MODELO DYSTAL DE ALKON, VOGL Y TAM.

El modelo DYSTAL (Dynamically Stable Associative Learning) es una red neural artificial modificable basada en las características observadas en sistemas neurales biológicos del molusco *Hermissenda* y del hipocampo del conejo.

En el modelo DYSTAL, la red neural, modifica sus pesos con base en a) la convergencia de entradas de "Flujo" colaterales modificables y no modificables b) apareo temporal de esas entradas, y c) actividades pasadas de esos elementos recibiendo las entradas.

La modificación es independiente de la salida del elemento. Como una consecuencia, DYSTAL, muestra a) escalado linear de los esfuerzos computacionales con el tamaño de la red, b) aprendizaje rápido sin un maestro externo y c) habilidad para completar patrones, asociando independientemente ensambles diferentes de entradas y sirviendo como un clasificador de los patrones de entrada.

Este es un modelo realizado con base en el supuesto de la distinción entre condicionamiento clásico y condicionamiento operante, tratándose en este caso de un modelo para el condicionamiento clásico con base en el aprendizaje supervisado.

En las redes neurales típicas los pesos suelen tratarse uniformemente en toda la arquitectura o por lo menos entra las capas que la componen, en este modelo en cambio, se integran rutas de pesos que pueden correr por toda la arquitectura, específicamente de dos tipos (Ver Figura 16):

- a) Rutas de pesos altos, que permiten una rápida y confiable transmisión de señales que produzcan comportamientos estereotípicos o de reflejo, bien definidos, y
- b) Rutas de pesos bajos, que son el resto de las conexiones en la red y son llamadas en este modelo "conexiones colaterales", que inicialmente son de pesos bajos y posteriormente se van incrementando de acuerdo a reglas de aprendizaje bien definidas.

Además, el modelo implica interacciones de pesos no hebbianas y sin retroalimentaciones, independientes de la salida de la red, que son conocidas como "interacciones locales". Estas interacciones producen ajustes de pesos derivados del apareamiento de dos estímulos (que en este modelo se presentan como dos presentaciones sucesivas).

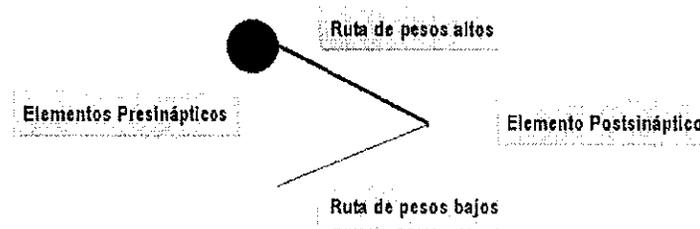


Figura 16. Tipos de rutas de pesos en el modelo DYSTAL.

Entre las ventajas que se reportan para el Modelo DYSTAL se encuentran las siguientes (Alkon, Vogl y Tam, 1991):

1. Escalado lineal de complejidad del algoritmo. En este diseño las salidas no son determinantes de la modificación de los pesos interiores y por ello, su complejidad es de n , con lo que el incremento del número de elementos en la red no implica un esfuerzo computacional como el del algoritmo de retropropagación cuya complejidad es $n \log n$. Ello también facilita su implementación en *hardware*.

2. No se requiere de un maestro externo. Lo que permite a DYSTAL clasificar y reestablecer patrones sin estándares predeterminados de entradas. Así, a diferencia de las redes que dependen de la retropropagación del error como medio de aprendizaje y la minimización del error, DYSTAL no necesita que los patrones aprendidos sean preprogramados en una función de comparación.
3. El modelo es robusto a las variaciones de los parámetros como el radio de incremento de los pesos o el tamaño de cada paso en el decremento de los pesos. Asimismo, el modelo también es robusto ante cambios estructurales, lo que permite conservar las mismas reglas de aprendizaje aún a pesar de añadir elementos colaterales o conexiones de retroalimentación.

2.3.2. EL MODELO GMADN DE BAXTER, ET AL PARA SIMULAR EL APRENDIZAJE ASOCIATIVO.

Baxter *et al*, presentaron en 1987 el Modelo General de Neuromodulación Independiente de la Actividad, o GMADN por sus siglas en inglés (Baxter *et al*, 1991). Ellos asumen, al igual que en el Modelo DYSTAL se hace, que existen diferencias significativas entre condicionamiento clásico y condicionamiento operante, sin embargo, en este caso no se trata de reglas de aprendizaje distinto, sino arquitecturas distintas para el condicionamiento clásico y el operante. Los autores establecen que buscan con este modelo el entendimiento de los eventos que ocurren entre los elementos individuales neurales y en las redes que contribuyen al aprendizaje y la memoria. El enfoque que utilizan para el desarrollo del modelo es analizar las propiedades de las neuronas y de los circuitos neurales que median en formas simples de aprendizaje en el molusco marino *Aplysia*, y desarrollar modelos matemáticos de estos modelos neurales y redes. Desde su perspectiva, un elemento individual adaptativo de tipo neural, que refleje las propiedades bioquímicas y biofísicas de las neuronas sensoriales, simula muchas características del aprendizaje no asociativo y del condicionamiento clásico.

Establecen que aún en redes neurales relativamente simples que incorporen una "regla de aprendizaje" asociativa, empíricamente derivada, se exhibirán algunas características de orden superior del condicionamiento clásico y algunas características elementales del condicionamiento operante.

El aprendizaje en este modelo se da con base en la actividad temporal contigua entre dos elementos, en donde un estímulo potente reforzante activa un sistema un sistema neural de respuesta y un sistema modulador que regula la eficacia de aferencias difusas al sistema de respuesta (Ver Figura 17).

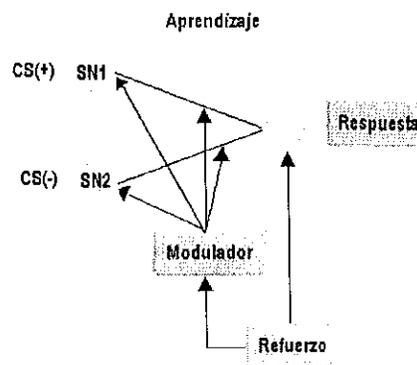


Figura 17. Aprendizaje en GMADN.

En el caso de la memoria, los efectos moduladores amplificados causan un fortalecimiento en la conexión funcional entre la neurona apareada y el sistema de respuesta (Ver Figura 18).

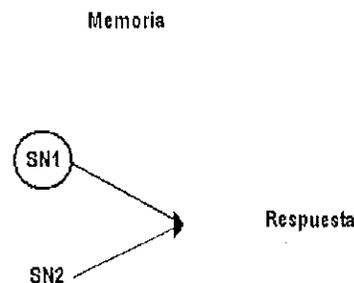


Figura 18. Memoria en GMADN

Para la simulación del Condicionamiento clásico, el GMADN implica el uso de una arquitectura en la que intervienen:

- Neuronas sensoriales idénticas entre sí, que tienen por objetivo la recepción de los estímulos del ambiente. Pueden ser activadas independientemente por estímulos condicionados separados.
- Una neurona facilitadora, que puede ser activada por las neuronas sensoriales si su nivel de entrada excede el umbral preestablecido. Se trata de un elemento no plástico que libera transmisores que inciden sobre la actividad de las neuronas sensitivas.
- Una neurona motora, que es finalmente en donde se observa la respuesta condicionada (Ver Figura 19).

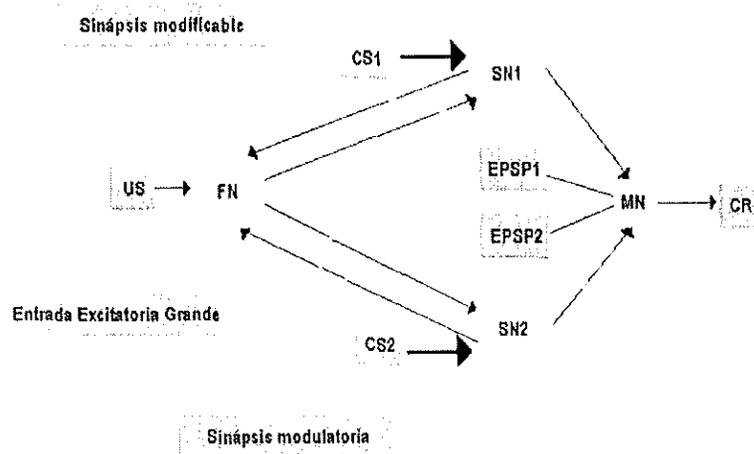


Figura 19. Arquitectura para la simulación del condicionamiento clásico con el GMADN

El condicionamiento operante, por otro lado, requiere que el circuito sea dirigido por dos células que funcionan como generadores de patrones PGA y PGB (en la Figura 20) conectadas por sinapsis inhibitorias. Estas excitan a los elementos asociativos que a su vez estimulan a los elementos

motores. Se necesita además una retroalimentación entre los generadores de patrones a fin reincrementar la duración de su actividad.

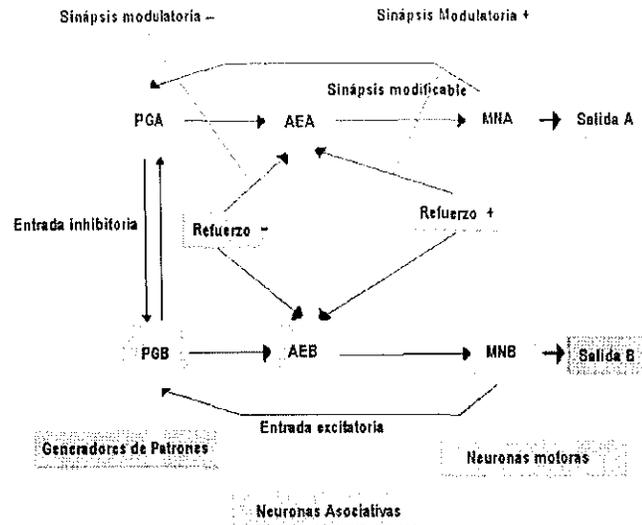


Figura 20. Arquitectura para la simulación del condicionamiento operante con el GMADN

2.3.3. EL MODELO SCHMAJUCK-LAM-GRAY (SLG) PARA SIMULAR INHIBICIÓN LATENTE

Este modelo formaliza algunas de las ideas de Sokolov y de Gray (Schmajuck, Lam y Gray, 1997).

Sokolov establece que los animales construyen una representación interna del ambiente.

Siempre que se detecta alguna novedad (es decir, cuando existe alguna inconsistencia entre lo predicho y los eventos ambientales actuales):

- 1) Tiene lugar una respuesta de orientación, y
- 2) El Modelo interno se modifica

Cuando existe coincidencia entre lo observado y el estímulo predicho, el animal puede responder sin cambio en el modelo neural del mundo que ha construido. En 1971 Gray sugirió que un sistema de

inhibición del comportamiento responde a señales de castigo, señales de no recompensa, estímulos nuevos o miedos innatos, inhibiendo los comportamientos en curso, incrementando su disponibilidad para la acción (nivel de alerta) e incrementando su atención a los estímulos ambientales.

El modelo de red que se describe a continuación provee un mecanismo que retoma las nociones de Sokolov (Schmajuck, Lam y Gray, 1997) acerca de la existencia de un modelo interno del mundo y un sistema de comparación, así como también considera el concepto de Gray acerca de un sistema inhibitorio. Conforme el modelo del mundo genera predicciones de futuros eventos, el sistema de comparación observa lo previsto y lo observado para computar novedades. La novedad es entonces usada para:

- 1) Controlar la atención a los eventos ambientales
- 2) Modificar el modelo del ambiente
- 3) Inhibir los comportamientos en curso

En la Figura 21 se muestra un diagrama con la arquitectura del modelo SGL:

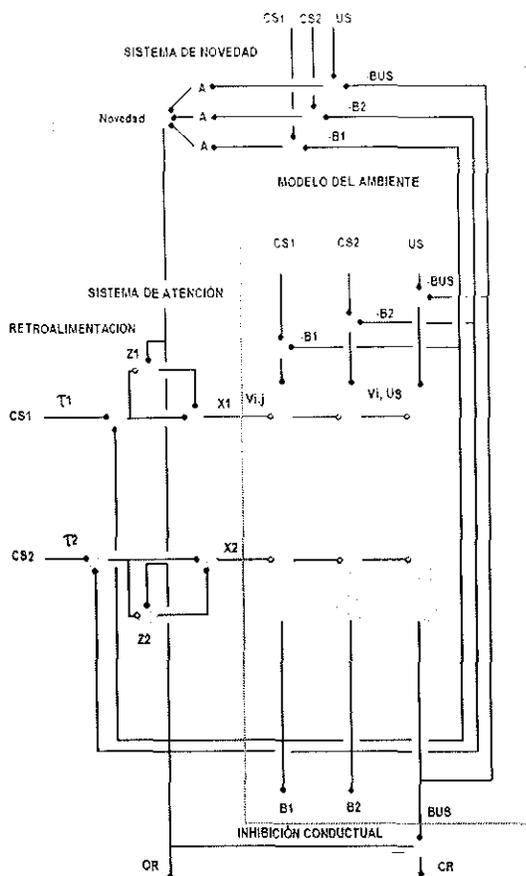


Figura 21. Arquitectura del Modelo SGL de Schmajuck¹⁷

El modelo SGL asume que un Estímulo Condicionado (EC) activa las representaciones internas. El sistema atencional mejora las representaciones internas de los EC activos al mismo tiempo cuando la novedad total del ambiente es grande (incrementando la atención), y decrementa las representaciones internas de aquellos EC activos al mismo tiempo cuando la novedad total es pequeña (decrementando la atención o quitando la atención por completo). La magnitud de las representaciones internas controla el almacenamiento de información al interior del modelo del ambiente (Asociabilidad) y la recuperación de información del modelo (recuperabilidad). Por lo tanto, eventos con grandes representaciones internas muestran una capacidad incrementada de formar asociaciones nuevas con otros eventos ambien-

¹⁷ Us, Estímulo incondicionado, Cs, Estímulo condicionado, t Huella, X Representación interna, ViUS novedad de la asociación, B, Predicción agregada del evento, OR, Respuesta de Orientación, Los puntos negros representan sinápsis fijas y los blancos representan sinápsis variables.

tales, así como una eficacia aumentada para recuperar asociaciones viejas. Eventos con representaciones internas pequeñas, muestran una capacidad decrementada para formar nuevas asociaciones con otros eventos ambientales, así como una eficacia decrementada para recuperar viejas asociaciones.

2.4. EL MODELO DE DONAHOE, BURGOS Y PALMER.

El modelo de Donahoe, Burgos y Palmer (1993) fue diseñado para la simulación del aprendizaje con base en refuerzos, ya sea instrumental o pavloviano. Es decir, a diferencia de los anteriormente expuestos, parte del supuesto de que no hay diferencia entre condicionamiento clásico y condicionamiento operante. El modelo está compuesto por cinco tipos de unidades de proceso: entradas, unidades de proceso motor, unidades de proceso sensorial, unidades de respuesta y unidades asociativas motoras y sensoriales (Ver la Figura 22 y la Figura 23).

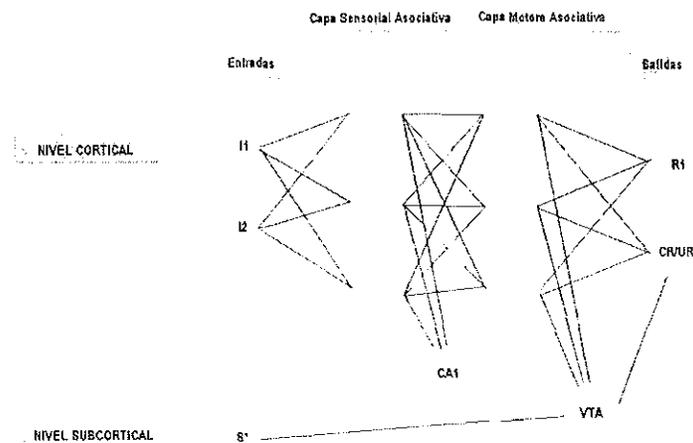


Figura 22. Arquitectura del modelo de Donahoe, Burgos y Palmer (1993)

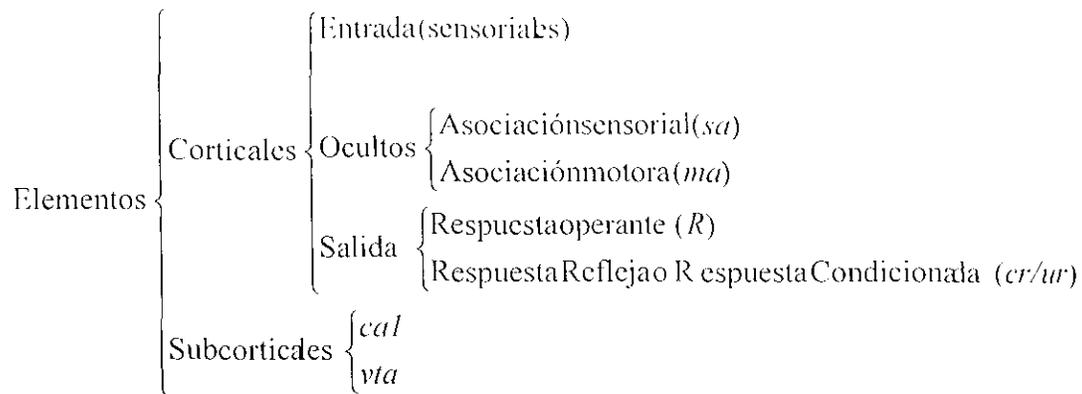


Figura 23. Elementos en una red bajo el modelo de Donahoe, Burgos y Palmer (1993)

Se trata de una red neural interconectada completamente, que basa su desempeño en dos funciones que ocurren durante el paso de un estímulo por toda la arquitectura de la red, hasta la producción de una respuesta: una función de activación, que es la que determina el valor que tendrá el impulso enviado a la siguiente capa de la red, y una función de aprendizaje, que es la que determina el grado de cambio que tendrá cada una de las unidades de la red para adecuarse mejor a la respuesta requerida.

En Donahoe, Burgos y Palmer (1993) La función de activación se define de la siguiente manera:

Siendo $N = \{x \in \mathbb{N} \mid 1 \leq x \leq n\}$ y $P = \{j \in \mathbb{N} \mid m \leq j \leq n\}$ los conjuntos de unidades neurales de procesamiento, en una red neural artificial, siendo N el conjunto de enteros positivos, n el número de unidades en la red y m el número de unidades de salida; y siendo $R = \{x \in \mathbb{R}^+ \mid 0.0 \leq X \leq 1.0\}$ el conjunto de posibles activaciones y valores de los pesos de conexión, donde \mathbb{R}^+ es el conjunto de número reales positivos. $a : P \times T \rightarrow R$ es la función de activación en donde $T \subset N$, representando los elementos de T eventos en el tiempo.

La regla para implementar la función a en las simulaciones de la red neural, se define como sigue:

Siendo $e(j,t)$ el vector de activaciones de las entradas excitatorias al elemento postsináptico j en el tiempo t ; $i(j,t)$ el vector de activaciones de las entradas inhibitorias de j en t y $w(j,t)$ el vector de los pesos excitatorios asociados con j en t , y $w'(j,t)$ el vector de los pesos inhibitorios asociados con j en t , donde $j \in P$ y $t \in T$, asumiendo que $e(j,t), i(j,t), w(j,t)$ y $w'(j,t) \in R^n$, la cantidad de excitación (*exc*) e inhibición (*inh*) producida en j durante t , dada respectivamente por $exc(j,t) = e(j,t) \cdot w(j,t)$ e $inh(j,t) = i(j,t) \cdot w'(j,t)$.

La activación a en t , vista como la probabilidad de disparo, se define como sigue:

$$a_{j,t} = \begin{cases} L(exc_{j,t}) + \tau_j L(exc_{j,t-1})[1 - L(exc_{j,t})] - L(inh_{j,t}) & \text{if } exc_{j,t} > inh_{j,t} \text{ and } exc_{j,t} \geq \theta_{j,t} \\ a_{j,t-1} - \kappa_j a_{j,t-1}(1 - a_{j,t-1}) - L(inh_{j,t-1}) & \text{if } exc_{j,t} > inh_{j,t} \text{ and } exc_{j,t} < \theta_{j,t} \\ 0 & \text{if } exc_{j,t} \leq inh_{j,t} \end{cases} \quad (4)$$

Siendo θ : un umbral excitatorio aleatorio gaussiano y L la función probabilística logística con los parámetros $\gamma = .5$ y $\delta = .1$ ¹⁸:

$$L(x) = \frac{1}{1 + e^{\frac{-(x-\gamma)}{\delta}}} \quad (5)$$

y la función de aprendizaje para el cálculo de los pesos nuevos:

$$\Delta w_{i,j,t} = \begin{cases} \alpha_j a_{j,t} d_t p_{i,t} r_{j,t} & \text{if } d_t > 0 \\ -\beta_j w_{i,j,t-1} a_{i,t} a_{j,t} & \text{if } d_t \leq 0 \end{cases} \quad (6)$$

¹⁸ Valores usados por los autores en las simulaciones reportadas en 1993.

En donde:

$\alpha(j)$ es la tasa de adquisición para las conexiones excitatorias

$\beta(j)$ es la tasa de extinción para las conexiones excitatorias

$d(t)$ es una señal de refuerzo,

$p(i,t) = [a(i,t)w(i,j,t-1)/exc(j,t)]$, y

$r(j,t) = 1 - sum[w(j,t)]$

Este modelo ha demostrado ser efectivo en la simulación de diversos fenómenos conductuales como la superstición (Burgos, 2000) o la generalización (Ponce y Burgos, 2002).

3. MÉTODO

3.1. EXPERIMENTO 1

En el primer experimento se contemplaron seis condiciones experimentales compuestas de dos fases cada uno: una en la que se entrenó a cada uno de los sujetos para responder ante un estímulo determinado y otra en la que se trató de entrenar a los sujetos a que respondieran ante un estímulo distinto al de la primera fase. Para el inicio de la segunda fase los sujetos no sufrieron modificación alguna en los pesos finales resultantes de la primera fase.

Las dos variables independientes en este experimento fueron el tamaño de la red (para el cual se establecieron tres valores: grande, mediano y pequeño) y el tamaño del patrón de entradas (para el cual se establecieron dos valores: grande y pequeño).

3.1.1. CONDICIÓN EXPERIMENTAL 1

3.1.1.1. Sujetos

Se emplearon 20 Redes neurales completamente interconectadas, con: tres entradas para estímulos condicionados y una entrada para estímulo incondicionado; una unidad de salida CR/UR; una sola unidad del tipo Ca1 y una sola unidad del tipo VTA; solamente una capa sensorial asociativa con tres unidades y una capa motora asociativa con tres unidades (ver Figura 24).

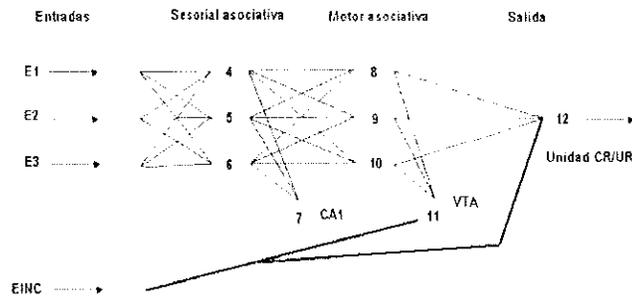


Figura 24. Arquitectura usada en las redes neurales de la condición experimental número 1 del experimento 1. Este grupo fue etiquetado con el nombre de conjunto de "Redes pequeñas para patrones pequeños".

Los parámetros empleados en todas las redes fueron los siguientes: las conexiones entre las unidades fueron inicializados con valores de 0.01, o también conocidos como valores "ingenuos". Conforme el entrenamiento se realizó, los pesos fueron modificados de acuerdo con la función de aprendizaje del modelo, excepto en los casos de las conexiones entre la unidad receptora de los estímulos incondicionados y las neuronas de tipo VTA, así como en la conexión entre esta unidad receptora y la unidad de respuesta CR/UR; en ambos casos, los pesos de conexión son fijos y tienen un valor de 0.9. Los límites tanto de las activaciones, como de los pesos fueron, 0 y 1. En cada una de las capas sensoriales asociativas empleadas en la red, se empleó solamente una unidad del tipo ca1 y para cada capa motora asociativa usada se empleó solo una unidad del tipo VTA. Para la desviación estándar de la función logística, el valor predeterminado fue 0.1; el valor predeterminado para la sumación temporal fue 0.1, igual que el valor de decaimiento. Los valores predeterminados de las tasas de incremento y decremento fueron fijados en 0.5 y 0.1 respectivamente.

3.1.1.2. Aparatos

En todas las simulaciones se empleó una computadora personal portátil con procesador Pentium IV m, con velocidad de 3.5 gigahertz, 512 megabytes de memoria de acceso aleatorio, disco duro de 70 megabytes, unidad de lectura-escritura de discos compactos y video discos en formato DVD, sistema

operativo Microsoft Windows Xp Professional Edition. Equipada con los siguientes paquetes: entorno de programación Microsoft Visual Studio, versión 6.0; procesador de palabras Microsoft Word Xp; hoja de cálculo Microsoft Excel Xp; utilería para dibujo de diagramas Microsoft Visio.NET; el graficador Sigma Plot versión 2001 y el simulador del modelo DBP, creado por el Dr. José E. Burgos Triano, "Selnet".

3.1.1.3. Procedimientos

Durante la primera fase de esta condición experimental se realizó una sesión de 300 ensayos, cada uno de ellos compuesto por 8 momentos temporales. En cada ensayo del experimento se tomó registro del nivel de activación de la unidad de salida (unidad número 12 en la Figura 24), en el momento temporal 7.

Los sujetos fueron entrenados con el conjunto de patrones de entrada que se presenta en la Tabla 1, cuya primer columna contiene el número de la entrada para las redes de este grupo y las siguientes ocho columnas muestran los valores de activación correspondientes en cada uno de los momentos temporales del ensayo. La fila marcada como Entrada S* muestra el valor de la activación en la entrada correspondiente al estímulo incondicionado. En todos los casos, el valor .01 representa un estado mínimo de actividad y el valor 0.9 el estado máximo.

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.9	.9	.9	.9	.9	.9	.9	.9
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.01	.01	.01	.01	.01	.01	.01	.01
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.9

Tabla 1. Patrones de entrada en cada momento temporal presentados a los sujetos en la primera fase de la condición experimental número 1 del experimento I.

Aunque el simulador empleado permite manipular el intervalo entre ensayos (IEEn), no se incluyó en las simulaciones realizadas, asumiendo que se estaba manejando el valor ideal para el tipo de procedimiento empleado.

Los sujetos fueron entrenados durante 300 ensayos, debido a que la experiencia reportada con el trabajo con el modelo, sugiere que 300 ensayos son suficientes para que se llegue a una activación estable en la neurona de salida (Donahoe, Burgos y Palmer, 1993; Donahoe, Palmer y Burgos, 1997)

Durante la segunda fase de esta condición experimental se sometió a los sujetos a una nueva sesión de entrenamiento ante los patrones que se muestran en la Tabla 2. Esta segunda sesión contempló también 300 ensayos

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.9	.9	.9	.9	.9	.9	.9	.9
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.9

Tabla 2. Patrones de entrada en cada momento temporal presentados en la segunda fase de la condición experimental número 1 del experimento 1.

La segunda fase de esta condición experimental se realizó sin reestablecer los pesos de conexión de las redes neurales empleadas. Es decir, esta fase se aplicó a partir de la configuración lograda en la primera fase.

3.1.2. CONDICIÓN EXPERIMENTAL 2

3.1.2.1. Sujetos

Se emplearon 20 Redes neurales completamente interconectadas, con: ocho entradas para estímulos condicionados y una entrada para estímulo incondicionado; una unidad de salida CR/UR; una sola unidad del tipo Ca1 y una sola unidad del tipo VTA; solamente una capa sensorial asociativa con tres unidades y una capa motora asociativa con tres unidades (ver Figura 25).

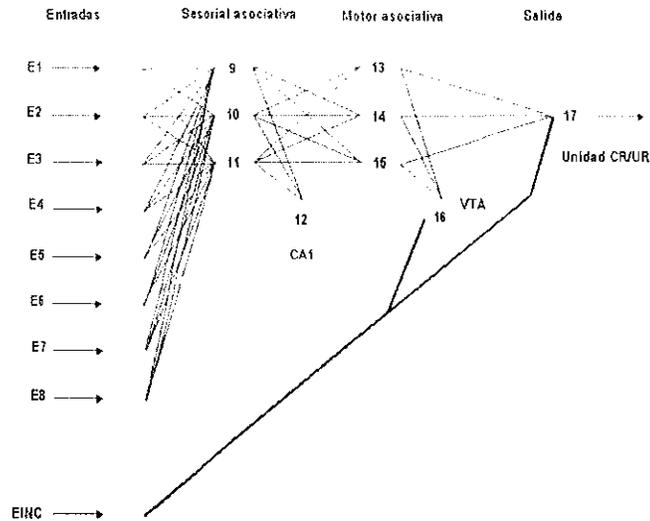


Figura 25. Arquitectura usada en las redes neurales de la condición experimental número 2 del experimento 1. Este grupo fue etiquetado con el nombre de conjunto de "Redes pequeñas para patrones grandes".

Los parámetros empleados en todas las redes fueron los siguientes: las conexiones entre las unidades fueron inicializados con valores de 0.01, o también conocidos como valores "ingenuos". Conforme el entrenamiento se realizó, los pesos fueron modificados de acuerdo con la función de aprendizaje del modelo, excepto en los casos de las conexiones entre la unidad receptora de los estímulos incondicionados y las neuronas de tipo VTA, así como en la conexión entre esta unidad receptora y la unidad de respuesta CR/UR; en ambos casos, los pesos de conexión son fijos y tienen un valor de 0.9. Los límites tanto de las activaciones, como de los pesos fueron, 0 y 1. En cada una de las capas sensoriales asociativas empleadas en la red, se empleó solamente una unidad del tipo ca1 y para cada capa motor asociativa usada se empleó solo una unidad del tipo VTA. Para la desviación estándar de la función logística, el valor predeterminado fue 0.1; el valor predeterminado para la sumación temporal fue 0.1, igual que el valor de decaimiento. Los valores predeterminados de las tasas de incremento y decremento fueron fijados en 0.5 y 0.1 respectivamente.

3.1.2.2. Aparatos

En todas las simulaciones se empleó una computadora personal portátil con procesador Pentium IV m, con velocidad de 3.5 gigahertz, 512 megabytes de memoria de acceso aleatorio, disco duro de 70 megabytes, unidad de lectura-escritura de discos compactos y video discos en formato DVD, sistema operativo Microsoft Windows Xp Professional Edition. Equipada con los siguientes paquetes: entorno de programación Microsoft Visual Studio, versión 6.0; procesador de palabras Microsoft Word Xp; hoja de cálculo Microsoft Excel Xp; utilería para dibujo de diagramas Microsoft Visio.NET; el graficador Sigma Plot versión 2001 y el simulador del modelo DBP, creado por el Dr. José E. Burgos Triano, "Selnet".

3.1.2.3. Procedimientos

Durante la primera fase de esta condición experimental se realizó una sesión de 300 ensayos, cada uno de ellos compuesto por 8 momentos temporales. En cada ensayo del experimento se tomó registro del nivel de activación de la unidad de salida (unidad número 17 en la Figura 25), en el momento temporal 7.

Los sujetos fueron entrenados con el conjunto de patrones de entrada que se presenta en la Tabla 3, cuya primer columna contiene el número de la entrada para las redes de este grupo y las siguientes ocho columnas muestran los valores de activación correspondientes en cada uno de los momentos temporales del ensayo. La fila marcada como Entrada S* muestra el valor de la activación en la entrada correspondiente al estímulo incondicionado. En todos los casos, el valor .01 representa un estado mínimo de actividad y el valor 0.9 el estado máximo.

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.9	.9	.9	.9	.9	.9	.9	.9
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 4	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 5	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 6	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 7	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 8	.01	.01	.01	.01	.01	.01	.01	.01
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.9

Tabla 3. Patrones de entrada en cada momento temporal presentados a los sujetos en la primera fase de la condición experimental número 2 del experimento 1.

Aunque el simulador empleado permite manipular el intervalo entre ensayos (IEEn), no se incluyó en las simulaciones realizadas, asumiendo que se estaba manejando el valor ideal para el tipo de procedimiento empleado.

Los sujetos fueron entrenados durante 300 ensayos, debido a que la experiencia reportada con el trabajo con el modelo, sugiere que 300 ensayos son suficientes para que se llegue a una activación estable en la neurona de salida (Donahoe, Burgos y Palmer, 1993; Donahoe, Palmer y Burgos, 1997)

Durante la segunda fase de esta condición experimental se sometió a los sujetos a una nueva sesión de entrenamiento ante los patrones que se muestran en la Tabla 4. Esta segunda sesión contempló también 300 ensayos.

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.9	.9	.9	.9	.9	.9	.9	.9
Entrada 4	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 5	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 6	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 7	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 8	.01	.01	.01	.01	.01	.01	.01	.01
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.9

Tabla 4. Patrones de entrada en cada momento temporal presentados en la segunda fase de la condición experimental número 2 del experimento 1.

La segunda fase de esta condición experimental se realizó sin reestablecer los pesos de conexión de las redes neurales empleadas. Es decir, esta fase se aplicó a partir de la configuración lograda en la primera fase.

3.1.3. CONDICIÓN EXPERIMENTAL 3

3.1.3.1. Sujetos

Se emplearon 20 Redes neurales completamente interconectadas, con: tres entradas para estímulos condicionados y una entrada para estímulo incondicionado; una unidad de salida CR/UR; dos unidades del tipo Ca1 y dos unidades del tipo VTA; dos capas sensoriales asociativas con tres unidades cada una y dos capas motoras asociativas con tres unidades cada una (ver Figura 26).

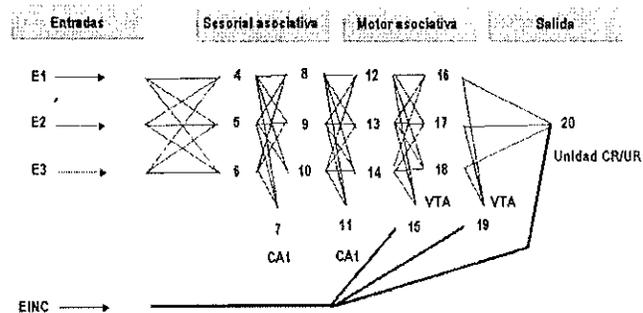


Figura 26. Arquitectura usada en las redes neurales de la condición experimental número 3 del experimento 1 Este grupo fue etiquetado con el nombre de conjunto de "Redes medianas para patrones pequeños".

Los parámetros empleados en todas las redes fueron los siguientes: las conexiones entre las unidades fueron inicializados con valores de 0.01, o también conocidos como valores "ingenuos". Conforme el entrenamiento se realizó, los pesos fueron modificados de acuerdo con la función de aprendizaje del modelo, excepto en los casos de las conexiones entre la unidad receptora de los estímulos incondicionados y las neuronas de tipo VTA, así como en la conexión entre esta unidad receptora y la unidad de respuesta CR/UR; en ambos casos, los pesos de conexión son fijos y tienen un valor de 0.9. Los lími-

tes tanto de las activaciones, como de los pesos fueron, 0 y 1. En cada una de las capas sensoriales asociativas empleadas en la red, se empleó solamente una unidad del tipo ca1 y para cada capa motor asociativa usada se empleó solo una unidad del tipo VTA. Para la desviación estándar de la función logística, el valor predeterminado fue 0.1; el valor predeterminado para la sumación temporal fue 0.1, igual que el valor de decaimiento. Los valores predeterminados de las tasas de incremento y decremento fueron fijados en 0.5 y 0.1 respectivamente.

3.1.3.2. Aparatos

En todas las simulaciones se empleó una computadora personal portátil con procesador Pentium IV m, con velocidad de 3.5 gigahertz, 512 megabytes de memoria de acceso aleatorio, disco duro de 70 megabytes, unidad de lectura-escritura de discos compactos y video discos en formato DVD, sistema operativo Microsoft Windows Xp Professional Edition. Equipada con los siguientes paquetes: entorno de programación Microsoft Visual Studio, versión 6.0; procesador de palabras Microsoft Word Xp; hoja de cálculo Microsoft Excel Xp; utilería para dibujo de diagramas Microsoft Visio.NET; el graficador Sigma Plot versión 2001 y el simulador del modelo DBP, creado por el Dr. José E. Burgos Triano, "Selnet".

3.1.3.3. Procedimientos

Durante la primera fase de esta condición experimental se realizó una sesión de 300 ensayos, cada uno de ellos compuesto por 8 momentos temporales. En cada ensayo del experimento se tomó registro del nivel de activación de la unidad de salida (unidad número 20 en la Figura 26), en el momento temporal 7.

Los sujetos fueron entrenados con el conjunto de patrones de entrada que se presenta en la Tabla 5, cuya primer columna contiene el número de la entrada para las redes de este grupo y las siguientes ocho columnas muestran los valores de activación correspondientes en cada uno de los momentos

temporales del ensayo. La fila marcada como Entrada S* muestra el valor de la activación en la entrada correspondiente al estímulo incondicionado. En todos los casos, el valor .01 representa un estado mínimo de actividad y el valor 0.9 el estado máximo.

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.9	.9	.9	.9	.9	.9	.9	.9
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.01	.01	.01	.01	.01	.01	.01	.01
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.9

Tabla 5. Patrones de entrada en cada momento temporal presentados a los sujetos en la primera fase de la condición experimental número 3 del experimento 1.

Aunque el simulador empleado permite manipular el intervalo entre ensayos (IEEn), no se incluyó en las simulaciones realizadas, asumiendo que se estaba manejando el valor ideal para el tipo de procedimiento empleado.

Los sujetos fueron entrenados durante 300 ensayos, debido a que la experiencia reportada con el trabajo con el modelo, sugiere que 300 ensayos son suficientes para que se llegue a una activación estable en la neurona de salida (Donahoe, Burgos y Palmer, 1993; Donahoe, Palmer y Burgos, 1997)

Durante la segunda fase de esta condición experimental se sometió a los sujetos a una nueva sesión de entrenamiento ante los patrones que se muestran en la Tabla 6. Esta segunda sesión contempló también 300 ensayos

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.9	.9	.9	.9	.9	.9	.9	.9
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.9

Tabla 6. Patrones de entrada en cada momento temporal presentados en la segunda fase de la condición experimental número 3 del experimento 1.

La segunda fase de esta condición experimental se realizó sin reestablecer los pesos de conexión de las redes neurales empleadas. Es decir, esta fase se aplicó a partir de la configuración lograda en la primera fase.

3.1.4. CONDICIÓN EXPERIMENTAL 4

3.1.4.1. Sujetos

Se emplearon 20 Redes neurales completamente interconectadas, con: ocho entradas para estímulos condicionados y una entrada para estímulo incondicionado; una unidad de salida CR/UR; dos unidades del tipo CA1 y dos unidades del tipo VTA; dos capas sensoriales asociativas con tres unidades cada una y dos capas motoras asociativas con tres unidades cada una (ver Figura 27).

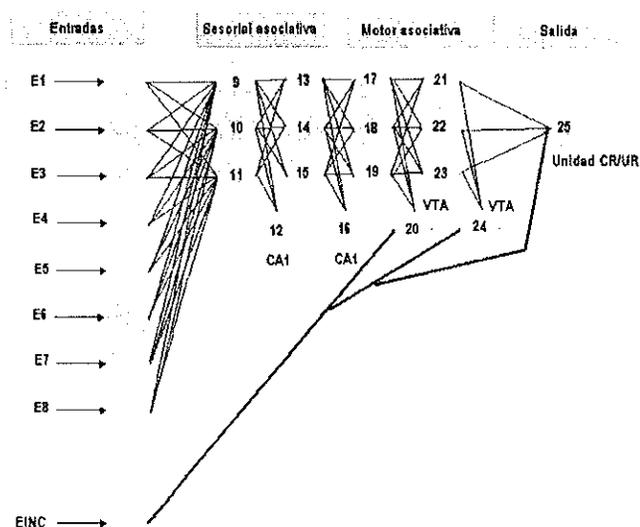


Figura 27. Arquitectura usada en las redes neurales de la condición experimental número 4 del experimento 1. Este grupo fue etiquetado con el nombre de conjunto de "Redes medianas para patrones grandes".

Los parámetros empleados en todas las redes fueron los siguientes: las conexiones entre las unidades fueron inicializados con valores de 0.01, o también conocidos como valores "ingenuos". Conforme el entrenamiento se realizó, los pesos fueron modificados de acuerdo con la función de aprendizaje del modelo, excepto en los casos de las conexiones entre la unidad receptora de los estímulos incondicionados y las neuronas de tipo VTA, así como en la conexión entre esta unidad receptora y la unidad de respuesta CR/UR; en ambos casos, los pesos de conexión son fijos y tienen un valor de 0.9. Los límites tanto de las activaciones, como de los pesos fueron, 0 y 1. En cada una de las capas sensoriales

asociativas empleadas en la red, se empleó solamente una unidad del tipo ca1 y para cada capa motor asociativa usada se empleó solo una unidad del tipo VTA. Para la desviación estándar de la función logística, el valor predeterminado fue 0.1; el valor predeterminado para la sumación temporal fue 0.1, igual que el valor de decaimiento. Los valores predeterminados de las tasas de incremento y decremento fueron fijados en 0.5 y 0.1 respectivamente.

3.1.4.2. Aparatos

En todas las simulaciones se empleó una computadora personal portátil con procesador Pentium IV m, con velocidad de 3.5 gigahertz, 512 megabytes de memoria de acceso aleatorio, disco duro de 70 megabytes, unidad de lectura-escritura de discos compactos y video discos en formato DVD, sistema operativo Microsoft Windows Xp Professional Edition. Equipada con los siguientes paquetes: entorno de programación Microsoft Visual Studio, versión 6.0; procesador de palabras Microsoft Word Xp; hoja de cálculo Microsoft Excel Xp; utilería para dibujo de diagramas Microsoft Visio.NET; el graficador Sigma Plot versión 2001 y el simulador del modelo DBP, creado por el Dr. José E. Burgos Triano, "Selnet".

3.1.4.3. Procedimientos

Durante la primera fase de esta condición experimental se realizó una sesión de 300 ensayos, cada uno de ellos compuesto por 8 momentos temporales. En cada ensayo del experimento se tomó registro del nivel de activación de la unidad de salida (unidad número 25 en la Figura 27), en el momento temporal 7.

Los sujetos fueron entrenados con el conjunto de patrones de entrada que se presenta en la Tabla 7, cuya primer columna contiene el número de la entrada para las redes de este grupo y las siguientes ocho columnas muestran los valores de activación correspondientes en cada uno de los momentos temporales del ensayo. La fila marcada como Entrada S* muestra el valor de la activación en la entra-

da correspondiente al estímulo incondicionado. En todos los casos, el valor .01 representa un estado mínimo de actividad y el valor 0.9 el estado máximo.

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.9	.9	.9	.9	.9	.9	.9	.9
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 4	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 5	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 6	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 7	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 8	.01	.01	.01	.01	.01	.01	.01	.01
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.9

Tabla 7. Patrones de entrada en cada momento temporal presentados a los sujetos en la primera fase de la condición experimental número 4 del experimento 1.

Aunque el simulador empleado permite manipular el intervalo entre ensayos (IEEn), no se incluyó en las simulaciones realizadas, asumiendo que se estaba manejando el valor ideal para el tipo de procedimiento empleado.

Los sujetos fueron entrenados durante 300 ensayos, debido a que la experiencia reportada con el trabajo con el modelo, sugiere que 300 ensayos son suficientes para que se llegue a una activación estable en la neurona de salida (Donahoe, Burgos y Palmer, 1993; Donahoe, Palmer y Burgos, 1997)

Durante la segunda fase de esta condición experimental se sometió a los sujetos a una nueva sesión de entrenamiento ante los patrones que se muestran en la Tabla 8. Esta segunda sesión contempló también 300 ensayos

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.9	.9	.9	.9	.9	.9	.9	.9
Entrada 4	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 5	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 6	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 7	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 8	.01	.01	.01	.01	.01	.01	.01	.01
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.9

Tabla 8. Patrones de entrada en cada momento temporal presentados en la segunda fase de la condición experimental número 4 del experimento 1.

La segunda fase de esta condición experimental se realizó sin reestablecer los pesos de conexión de las redes neurales empleadas. Es decir, la sesión se realizó a partir de la configuración lograda en la primera fase.

3.1.5. CONDICIÓN EXPERIMENTAL 5

3.1.5.1. Sujetos

Se emplearon 20 Redes neurales completamente interconectadas, con: tres entradas para estímulos condicionados y una entrada para estímulo incondicionado; una unidad de salida CR/UR; tres unidades del tipo Ca1 y tres unidades del tipo VTA; tres capas sensoriales asociativas con tres unidades cada una y tres capas motoras asociativas con tres unidades cada una (ver Figura 28).

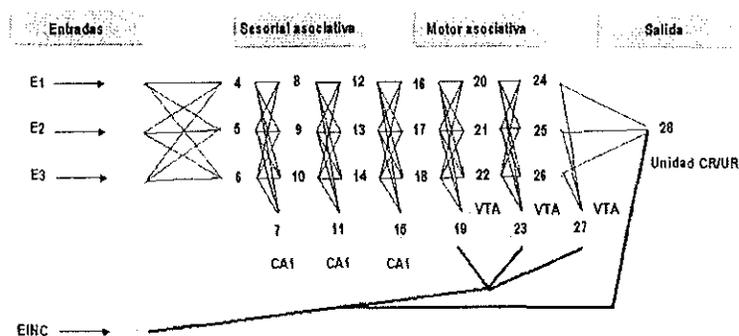


Figura 28. Arquitectura usada en las redes neurales de la condición experimental número 5 del experimento 1. Este grupo fue etiquetado con el nombre de conjunto de "Redes grandes para patrones pequeños".

Los parámetros empleados en todas las redes fueron los siguientes: las conexiones entre las unidades fueron inicializados con valores de 0.01, o también conocidos como valores "ingenuos". Conforme el entrenamiento se realizó, los pesos fueron modificados de acuerdo con la función de aprendizaje del modelo, excepto en los casos de las conexiones entre la unidad receptora de los estímulos incondicionados y las neuronas de tipo VTA, así como en la conexión entre esta unidad receptora y la unidad de

respuesta CR/UR; en ambos casos, los pesos de conexión son fijos y tienen un valor de 0.9. Los límites tanto de las activaciones, como de los pesos fueron, 0 y 1. En cada una de las capas sensoriales asociativas empleadas en la red, se empleó solamente una unidad del tipo ca1 y para cada capa motor asociativa usada se empleó solo una unidad del tipo VTA. Para la desviación estándar de la función logística, el valor predeterminado fue 0.1; el valor predeterminado para la sumación temporal fue 0.1, igual que el valor de decaimiento. Los valores predeterminados de las tasas de incremento y decremento fueron fijados en 0.5 y 0.1 respectivamente.

3.1.5.2. Aparatos

En todas las simulaciones se empleó una computadora personal portátil con procesador Pentium IV m, con velocidad de 3.5 gigahertz, 512 megabytes de memoria de acceso aleatorio, disco duro de 70 megabytes, unidad de lectura-escritura de discos compactos y video discos en formato DVD, sistema operativo Microsoft Windows Xp Professional Edition. Equipada con los siguientes paquetes: entorno de programación Microsoft Visual Studio, versión 6.0; procesador de palabras Microsoft Word Xp; hoja de cálculo Microsoft Excel Xp; utilería para dibujo de diagramas Microsoft Visio.NET; el graficador Sigma Plot versión 2001 y el simulador del modelo DBP, creado por el Dr. José E. Burgos Triano, "Selnet".

3.1.5.3. Procedimientos

Durante la primera fase de esta condición experimental se realizó una sesión de 300 ensayos, cada uno de ellos compuesto por 8 momentos temporales. En cada ensayo del experimento se tomó registro del nivel de activación de la unidad de salida (unidad número 28 en la Figura 28), en el momento temporal 7.

Los sujetos fueron entrenados con el conjunto de patrones de entrada que se presenta en la Tabla 9, cuya primer columna contiene el número de la entrada para las redes de este grupo y las siguientes

ocho columnas muestran los valores de activación correspondientes en cada uno de los momentos temporales del ensayo. La fila marcada como Entrada S* muestra el valor de la activación en la entrada correspondiente al estímulo incondicionado. En todos los casos, el valor .01 representa un estado mínimo de actividad y el valor 0.9 el estado máximo.

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.9	.9	.9	.9	.9	.9	.9	.9
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.01	.01	.01	.01	.01	.01	.01	.01
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.9

Tabla 9. Patrones de entrada en cada momento temporal presentados a los sujetos en la primera fase de la condición experimental número 5 del experimento 1.

Aunque el simulador empleado permite manipular el intervalo entre ensayos (IEEn), no se incluyó en las simulaciones realizadas, asumiendo que se estaba manejando el valor ideal para el tipo de procedimiento empleado.

Los sujetos fueron entrenados durante 300 ensayos, debido a que la experiencia reportada con el trabajo con el modelo, sugiere que 300 ensayos son suficientes para que se llegue a una activación estable en la neurona de salida (Donahoe, Burgos y Palmer, 1993; Donahoe, Palmer y Burgos, 1997)

Durante la segunda fase de esta condición experimental se sometió a los sujetos a una nueva sesión de entrenamiento ante los patrones que se muestran en la Tabla 10. Esta segunda sesión contempló también 300 ensayos

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.9	.9	.9	.9	.9	.9	.9	.9
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.9

Tabla 10. Patrones de entrada en cada momento temporal presentados en la segunda fase de la condición experimental número 5 del experimento 1.

La segunda fase de esta condición experimental se realizó sin reestablecer los pesos de conexión de las redes neurales empleadas, es decir, se realizó a partir de la configuración lograda en la primera fase.

3.1.6. CONDICIÓN EXPERIMENTAL 6

3.1.6.1. Sujetos

Se emplearon 20 Redes neurales completamente interconectadas, con: ocho entradas para estímulos condicionados y una entrada para estímulo incondicionado; una unidad de salida CR/UR; tres unidades del tipo Ca1 y tres unidades del tipo VTA; tres capas sensoriales asociativas con tres unidades cada una y tres capas motoras asociativas con tres unidades cada una (ver Figura 29).

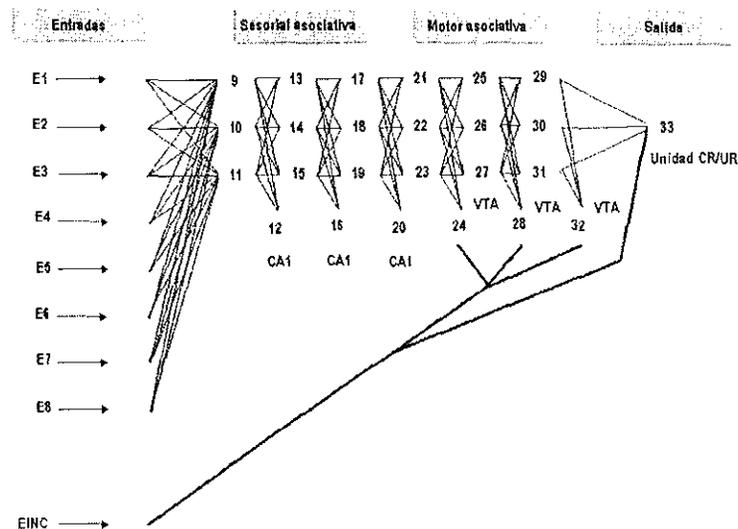


Figura 29. Arquitectura usada en las redes neurales de la condición experimental número 6 del experimento I. Este grupo fue etiquetado con el nombre de conjunto de "Redes grandes para patrones grandes".

Los parámetros empleados en todas las redes fueron los siguientes: las conexiones entre las unidades fueron inicializados con valores de 0.01, o también conocidos como valores "ingenuos". Conforme el entrenamiento se realizó, los pesos fueron modificados de acuerdo con la función de aprendizaje del

modelo, excepto en los casos de las conexiones entre la unidad receptora de los estímulos incondicionados y las neuronas de tipo VTA, así como en la conexión entre esta unidad receptora y la unidad de respuesta CR/UR; en ambos casos, los pesos de conexión son fijos y tienen un valor de 0.9. Los límites tanto de las activaciones, como de los pesos fueron, 0 y 1. En cada una de las capas sensoriales asociativas empleadas en la red, se empleó solamente una unidad del tipo ca1 y para cada capa motor asociativa usada se empleó solo una unidad del tipo VTA. Para la desviación estándar de la función logística, el valor predeterminado fue 0.1; el valor predeterminado para la sumación temporal fue 0.1, igual que el valor de decaimiento. Los valores predeterminados de las tasas de incremento y decremento fueron fijados en 0.5 y 0.1 respectivamente.

3.1.6.2. Aparatos

En todas las simulaciones se empleó una computadora personal portátil con procesador Pentium IV m, con velocidad de 3.5 gigahertz, 512 megabytes de memoria de acceso aleatorio, disco duro de 70 megabytes, unidad de lectura-escritura de discos compactos y video discos en formato DVD, sistema operativo Microsoft Windows Xp Professional Edition. Equipada con los siguientes paquetes: entorno de programación Microsoft Visual Studio, versión 6.0; procesador de palabras Microsoft Word Xp; hoja de cálculo Microsoft Excel Xp; utilería para dibujo de diagramas Microsoft Visio.NET; el graficador Sigma Plot versión 2001 y el simulador del modelo DBP, creado por el Dr. José E. Burgos Triano, "Selnet".

3.1.6.3. Procedimientos

Durante la primera fase de esta condición experimental se realizó una sesión de 300 ensayos, cada uno de ellos compuesto por 8 momentos temporales. En cada ensayo del experimento se tomó registro del nivel de activación de la unidad de salida (unidad número 33 en la Figura 29), en el momento temporal 7.

Los sujetos fueron entrenados con el conjunto de patrones de entrada que se presenta en la Tabla 11, cuya primer columna contiene el número de la entrada para las redes de este grupo y las siguientes ocho columnas muestran los valores de activación correspondientes en cada uno de los momentos temporales del ensayo. La fila marcada como Entrada S* muestra el valor de la activación en la entrada correspondiente al estímulo incondicionado. En todos los casos, el valor .01 representa un estado mínimo de actividad y el valor 0.9 el estado máximo.

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.9	.9	.9	.9	.9	.9	.9	.9
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 4	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 5	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 6	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 7	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 8	.01	.01	.01	.01	.01	.01	.01	.01
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.9

Tabla 11. Patrones de entrada en cada momento temporal presentados a los sujetos en la primera fase de la condición experimental número 6 del experimento 1.

Aunque el simulador empleado permite manipular el intervalo entre ensayos (IEEn), no se incluyó en las simulaciones realizadas, asumiendo que se estaba manejando el valor ideal para el tipo de procedimiento empleado.

Los sujetos fueron entrenados durante 300 ensayos, debido a que la experiencia reportada con el trabajo con el modelo, sugiere que 300 ensayos son suficientes para que se llegue a una activación estable en la neurona de salida (Donahoe, Burgos y Palmer, 1993; Donahoe, Palmer y Burgos, 1997)

Durante la segunda fase de esta condición experimental se sometió a los sujetos a una nueva sesión de entrenamiento ante los patrones que se muestran en la Tabla 12. Esta segunda sesión contempló también 300 ensayos.

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.9	.9	.9	.9	.9	.9	.9	.9
Entrada 4	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 5	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 6	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 7	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 8	.01	.01	.01	.01	.01	.01	.01	.01
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.9

Tabla 12. Patrones de entrada en cada momento temporal presentados en la segunda fase de la condición experimental número 6 del experimento 1.

La segunda fase de esta condición experimental se realizó sin reestablecer los pesos de conexión de las redes neurales empleadas.

3.2. EXPERIMENTO 2

En el segundo experimento se contempló una sola condición experimental en tres fases: una en la que se entrenó a cada uno de los sujetos para responder ante un estímulo determinado, otra en la que se trató de entrenar a los sujetos a que respondieran ante un estímulo distinto al de la primera fase y una tercera, en la que se entrenó a los sujetos nuevamente con el patrón empleado en la fase número 1. Para el inicio de la segunda y tercera fase, los sujetos no sufrieron modificación alguna en los pesos finales resultantes de la fase anterior inmediata.

3.2.1. CONDICIÓN EXPERIMENTAL 1

3.2.1.1. Sujetos

Se emplearon 20 Redes neurales completamente interconectadas, con: tres entradas para estímulos condicionados y una entrada para estímulo incondicionado; una unidad de salida CR/UR; una sola unidad del tipo Ca1 y una sola unidad del tipo VTA; solamente una capa sensorial asociativa con tres unidades y una capa motora asociativa con tres unidades (ver Figura 30).

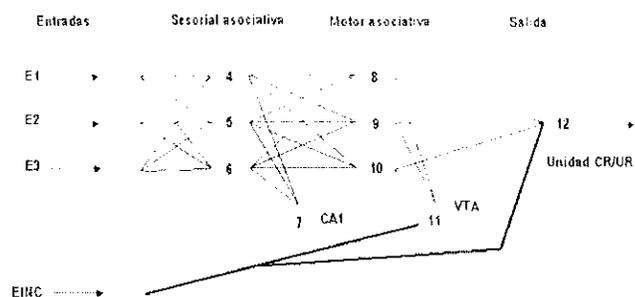


Figura 30. Arquitectura usada en las redes neurales de la condición experimental número 1 del experimento 2. Este grupo fue etiquetado con el nombre de conjunto de "Redes pequeñas para patrones pequeños".

Los parámetros empleados en todas las redes fueron los siguientes: las conexiones entre las unidades fueron inicializados con valores de 0.01, o también conocidos como valores "ingenuos". Conforme el entrenamiento se realizó, los pesos fueron modificados de acuerdo con la función de aprendizaje del modelo, excepto en los casos de las conexiones entre la unidad receptora de los estímulos incondicionados y las neuronas de tipo VTA, así como en la conexión entre esta unidad receptora y la unidad de respuesta CR/UR; en ambos casos, los pesos de conexión son fijos y tienen un valor de 0.9. Los límites tanto de las activaciones, como de los pesos fueron, 0 y 1. En cada una de las capas sensoriales asociativas empleadas en la red, se empleó solamente una unidad del tipo ca1 y para cada capa motor asociativa usada se empleó solo una unidad del tipo VTA. Para la desviación estándar de la función logística, el valor predeterminado fue 0.1; el valor predeterminado para la sumación temporal fue 0.1, igual que el valor de decaimiento. Los valores predeterminados de las tasas de incremento y decremento fueron fijados en 0.5 y 0.1 respectivamente.

3.2.1.2. Aparatos

En todas las simulaciones se empleó una computadora personal portátil con procesador Pentium IV m, con velocidad de 3.5 gigahertz, 512 megabytes de memoria de acceso aleatorio, disco duro de 70 megabytes, unidad de lectura-escritura de discos compactos y video discos en formato DVD, sistema

operativo Microsoft Windows Xp Professional Edition. Equipada con los siguientes paquetes: entorno de programación Microsoft Visual Studio, versión 6.0; procesador de palabras Microsoft Word Xp; hoja de cálculo Microsoft Excel Xp; utilería para dibujo de diagramas Microsoft Visio.NET; el graficador Sigma Plot versión 2001 y el simulador del modelo DBP, creado por el Dr. José E. Burgos Triano, "Selnet".

3.2.1.3. Procedimientos

Durante la primera fase del experimento se realizó una sesión de 300 ensayos, cada uno de ellos compuesto por 8 momentos temporales. En cada ensayo del experimento se tomó registro del nivel de activación de la unidad de salida (unidad número 12 en la Figura 30), en el momento temporal 7.

Los sujetos fueron entrenados con el conjunto de patrones de entrada que se presenta en la Tabla 13, cuya primer columna contiene el número de la entrada para las redes de este grupo y las siguientes ocho columnas muestran los valores de activación correspondientes en cada uno de los momentos temporales del ensayo. La fila marcada como Entrada S* muestra el valor de la activación en la entrada correspondiente al estímulo incondicionado. En todos los casos, el valor .01 representa un estado mínimo de actividad y el valor 0.9 el estado máximo.

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.9	.9	.9	.9	.9	.9	.9	.9
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.01	.01	.01	.01	.01	.01	.01	.01
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.9

Tabla 13. Patrones de entrada en cada momento temporal presentados a los sujetos en la primera fase de la condición experimental número 1 del experimento 2.

Aunque el simulador empleado permite manipular el intervalo entre ensayos (IEEn), no se incluyó en las simulaciones realizadas, asumiendo que se estaba manejando el valor ideal para el tipo de procedimiento empleado.

Los sujetos fueron entrenados durante 300 ensayos, debido a que la experiencia reportada con el trabajo con el modelo, sugiere que 300 ensayos son suficientes para que se llegue a una activación estable en la neurona de salida (Donahoe, Burgos y Palmer, 1993; Donahoe, Palmer y Burgos, 1997)

Durante la segunda fase del experimento se sometió a los sujetos a una nueva sesión de entrenamiento ante los patrones que se muestran en la Tabla 14. Esta segunda sesión contempló también 300 ensayos

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.9	.9	.9	.9	.9	.9	.9	.9
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.9

Tabla 14. Patrones de entrada en cada momento temporal presentados en la segunda fase de la condición experimental número 1 del experimento 2.

La segunda fase del experimento se realizó sin reestablecer los pesos de conexión de las redes neurales empleadas. Es decir, la segunda fase de esta condición experimental se aplicó a partir de la configuración lograda en la primera fase.

En la tercera fase del experimento se sometió a los sujetos a una nueva sesión de entrenamiento ante los patrones que se muestran en la Tabla 15. Esta tercera sesión contempló también 300 ensayos

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.9	.9	.9	.9	.9	.9	.9	.9
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.01	.01	.01	.01	.01	.01	.01	.01
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.9

Tabla 15. Patrones de entrada en cada momento temporal presentados en la tercera fase de la condición experimental número 1 del experimento 2.

La tercera fase del experimento se realizó sin reestablecer los pesos de conexión de las redes neurales empleadas. Es decir, la segunda fase de esta condición experimental se aplicó a partir de la configuración lograda en la segunda fase.

3.3. EXPERIMENTO 3

En el tercer experimento se contempló una condición experimental de una sola fase, en la que se entrenó a cada uno de los sujetos para responder ante dos estímulos determinados presentados en la misma sesión, en forma aleatoria.

3.3.1. CONDICIÓN EXPERIMENTAL 1

3.3.1.1. Sujetos

Se emplearon 20 Redes neurales completamente interconectadas, con: tres entradas para estímulos condicionados y una entrada para estímulo incondicionado; una unidad de salida CR/UR; una sola unidad del tipo Ca1 y una sola unidad del tipo VTA; solamente una capa sensorial asociativa con tres unidades y una capa motora asociativa con tres unidades (ver Figura 31).

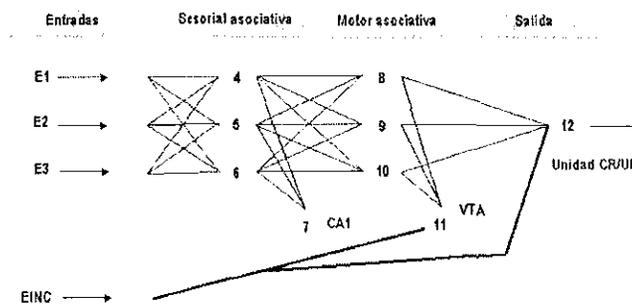


Figura 31. Arquitectura usada en las redes neurales de la condición experimental número 1 del experimento 3. Este grupo fue etiquetado con el nombre de conjunto de "Redes pequeñas para patrones pequeños".

Los parámetros empleados en todas las redes fueron los siguientes: las conexiones entre las unidades fueron inicializados con valores de 0.01, o también conocidos como valores "ingenuos". Conforme el entrenamiento se realizó, los pesos fueron modificados de acuerdo con la función de aprendizaje del modelo, excepto en los casos de las conexiones entre la unidad receptora de los estímulos incondicio-

dados y las neuronas de tipo VTA, así como en la conexión entre esta unidad receptora y la unidad de respuesta CR/UR; en ambos casos, los pesos de conexión son fijos y tienen un valor de 0.9. Los límites tanto de las activaciones, como de los pesos fueron, 0 y 1. En cada una de las capas sensoriales asociativas empleadas en la red, se empleó solamente una unidad del tipo ca1 y para cada capa motor asociativa usada se empleó solo una unidad del tipo VTA. Para la desviación estándar de la función logística, el valor predeterminado fue 0.1; el valor predeterminado para la sumación temporal fue 0.1, igual que el valor de decaimiento. Los valores predeterminados de las tasas de incremento y decremento fueron fijados en 0.5 y 0.1 respectivamente.

3.3.1.2. Aparatos

En todas las simulaciones se empleó una computadora personal portátil con procesador Pentium IV m, con velocidad de 3.5 gigahertz, 512 megabytes de memoria de acceso aleatorio, disco duro de 70 megabytes, unidad de lectura-escritura de discos compactos y video discos en formato DVD, sistema operativo Microsoft Windows Xp Professional Edition. Equipada con los siguientes paquetes: entorno de programación Microsoft Visual Studio, versión 6.0; procesador de palabras Microsoft Word Xp; hoja de cálculo Microsoft Excel Xp; utilería para dibujo de diagramas Microsoft Visio.NET; el graficador Sigma Plot versión 2001 y el simulador del modelo DBP, creado por el Dr. José E. Burgos Triano, "Selnet".

3.3.1.3. Procedimientos

Durante esta condición experimental se realizó una sesión de 300 ensayos, cada uno de ellos compuesto por 8 momentos temporales. En cada ensayo del experimento se tomó registro del nivel de activación de la unidad de salida (unidad número 12 en la Figura 31), en el momento temporal 7.

Los sujetos fueron entrenados con el conjunto de patrones de entrada que se presentan en las Tablas 16 y 17, que fueron presentados aleatoriamente en la sesión. En ambas tablas la primer columna con-

tiene el número de la entrada para las redes de este grupo y las siguientes ocho columnas muestran los valores de activación correspondientes en cada uno de los momentos temporales del ensayo. La fila marcada como Entrada S* muestra el valor de la activación en la entrada correspondiente al estímulo incondicionado. En todos los casos, el valor .01 representa un estado mínimo de actividad y el valor 0.9 el estado máximo.

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.9	.9	.9	.9	.9	.9	.9	.9
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.01	.01	.01	.01	.01	.01	.01	.01
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.9

Tabla 16. Primer patrón de entradas presentado a los sujetos en la condición experimental número 1 del experimento 3.

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.9	.9	.9	.9	.9	.9	.9	.9
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.9

Tabla 17. Segundo patrón de entradas presentado a los sujetos en la condición experimental número 1 del experimento 3.

Aunque el simulador empleado permite manipular el intervalo entre ensayos (IEEn), no se incluyó en las simulaciones realizadas, asumiendo que se estaba manejando el valor ideal para el tipo de procedimiento empleado.

Los sujetos fueron entrenados durante 300 ensayos, debido a que la experiencia reportada con el trabajo con el modelo, sugiere que 300 ensayos son suficientes para que se llegue a una activación estable en la neurona de salida (Donahoe, Burgos y Palmer, 1993; Donahoe, Palmer y Burgos, 1997)

3.4. EXPERIMENTO 4

En el cuarto experimento se contempló una sola condición experimental de una fase en la que se entrenó a cada uno de los sujetos para responder ante dos estímulos determinados presentados simultáneamente.

3.4.1. CONDICIÓN EXPERIMENTAL 1

3.4.1.1. Sujetos

Se emplearon 20 Redes neurales completamente interconectadas, con: tres entradas para estímulos condicionados y una entrada para estímulo incondicionado; una unidad de salida CR/UR; una sola unidad del tipo Ca1 y una sola unidad del tipo VTA; solamente una capa sensorial asociativa con tres unidades y una capa motora asociativa con tres unidades (ver Figura 32).

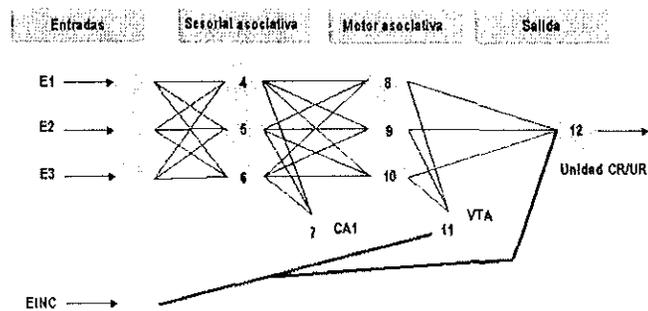


Figura 32. Arquitectura usada en las redes neurales de la condición experimental número 1 del experimento 3. Este grupo fue etiquetado con el nombre de conjunto de "Redes pequeñas para patrones pequeños".

Los parámetros empleados en todas las redes fueron los siguientes: las conexiones entre las unidades fueron inicializados con valores de 0.01, o también conocidos como valores "ingenuos". Conforme el entrenamiento se realizó, los pesos fueron modificados de acuerdo con la función de aprendizaje del modelo, excepto en los casos de las conexiones entre la unidad receptora de los estímulos incondicionados y las neuronas de tipo VTA, así como en la conexión entre esta unidad receptora y la unidad de respuesta CR/UR; en ambos casos, los pesos de conexión son fijos y tienen un valor de 0.9. Los límites tanto de las activaciones, como de los pesos fueron, 0 y 1. En cada una de las capas sensoriales asociativas empleadas en la red, se empleó solamente una unidad del tipo ca1 y para cada capa motor asociativa usada se empleó solo una unidad del tipo VTA. Para la desviación estándar de la función

logística, el valor predeterminado fue 0.1; el valor predeterminado para la sumación temporal fue 0.1, igual que el valor de decaimiento. Los valores predeterminados de las tasas de incremento y decremento fueron fijados en 0.5 y 0.1 respectivamente.

3.4.1.2. Aparatos

En todas las simulaciones se empleó una computadora personal portátil con procesador Pentium IV m, con velocidad de 3.5 gigahertz, 512 megabytes de memoria de acceso aleatorio, disco duro de 70 megabytes, unidad de lectura-escritura de discos compactos y video discos en formato DVD, sistema operativo Microsoft Windows Xp Professional Edition. Equipada con los siguientes paquetes: entorno de programación Microsoft Visual Studio, versión 6.0; procesador de palabras Microsoft Word Xp; hoja de cálculo Microsoft Excel Xp; utilería para dibujo de diagramas Microsoft Visio.NET; el graficador Sigma Plot versión 2001 y el simulador del modelo DBP, creado por el Dr. José E. Burgos Triano, "Selnet".

3.4.1.3. Procedimientos

Durante esta condición experimental se realizó una sesión de 300 ensayos, cada uno de ellos compuesto por 8 momentos temporales. En cada ensayo del experimento se tomó registro del nivel de activación de la unidad de salida (unidad número 12 en la Figura 32), en el momento temporal 7.

Los sujetos fueron entrenados con el conjunto de patrones de entrada que se presentan en las Tablas 18 y 19, que fueron presentados simultáneamente en la sesión. En ambas tablas la primer columna contiene el número de la entrada para las redes de este grupo y las siguientes ocho columnas muestran los valores de activación correspondientes en cada uno de los momentos temporales del ensayo. La fila marcada como Entrada S* muestra el valor de la activación en la entrada correspondiente al estímulo incondicionado. En todos los casos, el valor .01 representa un estado mínimo de actividad y el valor 0.9 el estado máximo.

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.9	.9	.9	.9	.9	.9	.9	.9
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.01	.01	.01	.01	.01	.01	.01	.01
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.9

Tabla 18. Primer patrón de entradas presentado a los sujetos en la condición experimental número 1 del experimento 3.

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.9	.9	.9	.9	.9	.9	.9	.9
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.9

Tabla 19. Segundo patrón de entradas presentado a los sujetos en la condición experimental número 1 del experimento 3.

Aunque el simulador empleado permite manipular el intervalo entre ensayos (IEEn), no se incluyó en las simulaciones realizadas, asumiendo que se estaba manejando el valor ideal para el tipo de procedimiento empleado.

Los sujetos fueron entrenados durante 300 ensayos, debido a que la experiencia reportada con el trabajo con el modelo, sugiere que 300 ensayos son suficientes para que se llegue a una activación estable en la neurona de salida (Donahoe, Burgos y Palmer, 1993; Donahoe, Palmer y Burgos, 1997)

3.5. EXPERIMENTO 5

En el quinto experimento se contempló una sola condición experimental con tres fases: una en la que se entrenó a cada uno de los sujetos para responder ante un estímulo determinado, una segunda fase en la que se sostuvieron presentaciones del estímulo sin refuerzo alguno; y una tercera fase en la que buscó entrenar al sujeto ante un nuevo estímulo.

Para el inicio de la segunda fase y la tercera de esta condición experimental, los sujetos no sufrieron modificación alguna en los pesos finales resultantes de la fase anterior inmediata.

3.5.1. CONDICIÓN EXPERIMENTAL 1

3.5.1.1. Sujetos

Se emplearon 20 Redes neurales completamente interconectadas, con: tres entradas para estímulos condicionados y una entrada para estímulo incondicionado; una unidad de salida CR/UR; una sola unidad del tipo Ca1 y una sola unidad del tipo VTA; solamente una capa sensorial asociativa con tres unidades y una capa motora asociativa con tres unidades (ver Figura 33).

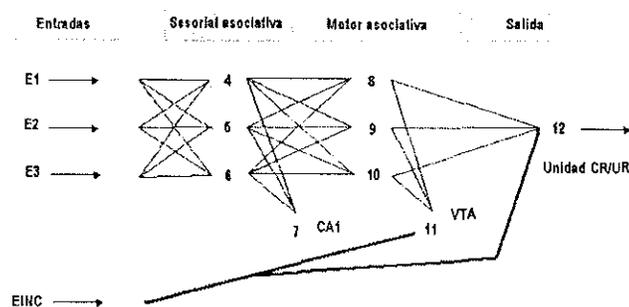


Figura 33. Arquitectura usada en las redes neurales de la condición experimental número 1 del experimento 1. Este grupo fue etiquetado con el nombre de conjunto de "Redes pequeñas para patrones pequeños".

Los parámetros empleados en todas las redes fueron los siguientes: las conexiones entre las unidades fueron inicializados con valores de 0.01, o también conocidos como valores "ingenuos". Conforme el entrenamiento se realizó, los pesos fueron modificados de acuerdo con la función de aprendizaje del modelo, excepto en los casos de las conexiones entre la unidad receptora de los estímulos incondicionados y las neuronas de tipo VTA, así como en la conexión entre esta unidad receptora y la unidad de respuesta CR/UR; en ambos casos, los pesos de conexión son fijos y tienen un valor de 0.9. Los límites tanto de las activaciones, como de los pesos fueron, 0 y 1. En cada una de las capas sensoriales asociativas empleadas en la red, se empleó solamente una unidad del tipo ca1 y para cada capa motor asociativa usada se empleó solo una unidad del tipo VTA. Para la desviación estándar de la función logística, el valor predeterminado fue 0.1; el valor predeterminado para la sumación temporal fue 0.1,

igual que el valor de decaimiento. Los valores predeterminados de las tasas de incremento y decremento fueron fijados en 0.5 y 0.1 respectivamente.

3.5.1.2. Aparatos

En todas las simulaciones se empleó una computadora personal portátil con procesador Pentium IV m, con velocidad de 3.5 gigahertz, 512 megabytes de memoria de acceso aleatorio, disco duro de 70 megabytes, unidad de lectura-escritura de discos compactos y video discos en formato DVD, sistema operativo Microsoft Windows Xp Professional Edition. Equipada con los siguientes paquetes: entorno de programación Microsoft Visual Studio, versión 6.0; procesador de palabras Microsoft Word Xp; hoja de cálculo Microsoft Excel Xp; utilería para dibujo de diagramas Microsoft Visio.NET; el graficador Sigma Plot versión 2001 y el simulador del modelo DBP, creado por el Dr. José E. Burgos Triano, "Selnet".

3.5.1.3. Procedimientos

Durante la primera fase de esta condición experimental se realizó una sesión de 300 ensayos, cada uno de ellos compuesto por 8 momentos temporales. En cada ensayo del experimento se tomó registro del nivel de activación de la unidad de salida (unidad número 12 en la Figura 33), en el momento temporal 7.

Los sujetos fueron entrenados con el conjunto de patrones de entrada que se presenta en la Tabla 20, cuya primer columna contiene el número de la entrada para las redes de este grupo y las siguientes ocho columnas muestran los valores de activación correspondientes en cada uno de los momentos temporales del ensayo. La fila marcada como Entrada S* muestra el valor de la activación en la entrada correspondiente al estímulo incondicionado. En todos los casos, el valor .01 representa un estado mínimo de actividad y el valor 0.9 el estado máximo.

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.9	.9	.9	.9	.9	.9	.9	.9
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.01	.01	.01	.01	.01	.01	.01	.01
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.9

Tabla 20. Patrones de entrada en cada momento temporal presentados a los sujetos en la primera fase de la condición experimental número 1 del experimento 5.

Aunque el simulador empleado permite manipular el intervalo entre ensayos (IEEn), no se incluyó en las simulaciones realizadas, asumiendo que se estaba manejando el valor ideal para el tipo de procedimiento empleado.

Los sujetos fueron entrenados durante 300 ensayos, debido a que la experiencia reportada con el trabajo con el modelo, sugiere que 300 ensayos son suficientes para que se llegue a una activación estable en la neurona de salida (Donahoe, Burgos y Palmer, 1993; Donahoe, Palmer y Burgos, 1997)

Durante la segunda fase de esta condición experimental se sometió a los sujetos a una nueva sesión de entrenamiento ante los patrones que se muestran en la Tabla 21. Esta segunda sesión contempló también 300 ensayos.

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.9	.9	.9	.9	.9	.9	.9	.9
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.01	.01	.01	.01	.01	.01	.01	.01
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.01

Tabla 21. Patrones de entrada en cada momento temporal presentados en la segunda fase de la condición experimental número 1 del experimento 5.

La segunda fase de esta condición experimental se realizó sin reestablecer los pesos de conexión de las redes neurales empleadas. Es decir, la segunda fase de esta condición experimental se aplicó a partir de la configuración lograda en la primera fase.

Durante la tercera fase de esta condición experimental se sometió a los sujetos a una nueva sesión de entrenamiento ante los patrones que se muestran en la Tabla 22. Esta sesión contempló también 300 ensayos.

Numero de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.9	.9	.9	.9	.9	.9	.9	.9
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.01

Tabla 22. Patrones de entrada en cada momento temporal presentados en la tercera fase de la condición experimental número 1 del experimento 5.

La tercera fase de esta condición experimental se realizó sin reestablecer los pesos de conexión de las redes neurales empleadas.

3.6. EXPERIMENTO 6

En el sexto experimento se contempló una condición experimental con cuatro fases: una en la que se entrenó a cada uno de los sujetos para responder ante un estímulo determinado, una segunda fase en la que se sostuvieron presentaciones del estímulo sin refuerzo alguno; una tercera fase en la que buscó entrenar al sujeto ante un nuevo estímulo, y una cuarta fase en la que se buscó entrenar nuevamente a la red para responder ante el estímulo empleado en la primera fase de esta misma condición experimental. Para el inicio de la segunda, tercera y cuarta de esta condición experimental, los sujetos no sufrieron modificación alguna en los pesos finales resultantes de la fase anterior inmediata.

3.6.1. CONDICIÓN EXPERIMENTAL 1

3.6.1.1. Sujetos

Se emplearon 20 Redes neurales completamente interconectadas, con: tres entradas para estímulos condicionados y una entrada para estímulo incondicionado; una unidad de salida CR/UR; una sola unidad del tipo Ca1 y una sola unidad del tipo VTA; solamente una capa sensorial asociativa con tres unidades y una capa motora asociativa con tres unidades (ver Figura 34).

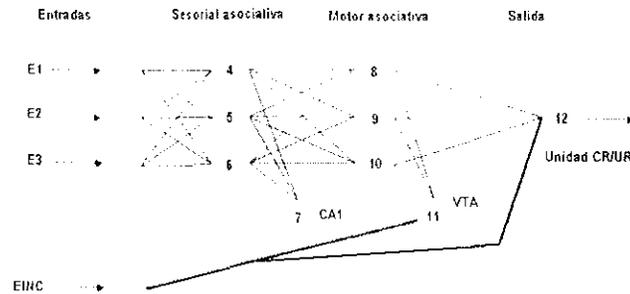


Figura 34. Arquitectura usada en las redes neurales de la condición experimental número 1 del experimento 1. Este grupo fue etiquetado con el nombre de conjunto de "Redes pequeñas para patrones pequeños".

Los parámetros empleados en todas las redes fueron los siguientes: las conexiones entre las unidades fueron inicializados con valores de 0.01, o también conocidos como valores "ingenuos". Conforme el entrenamiento se realizó, los pesos fueron modificados de acuerdo con la función de aprendizaje del modelo, excepto en los casos de las conexiones entre la unidad receptora de los estímulos incondicionados y las neuronas de tipo VTA, así como en la conexión entre esta unidad receptora y la unidad de respuesta CR/UR; en ambos casos, los pesos de conexión son fijos y tienen un valor de 0.9. Los límites tanto de las activaciones, como de los pesos fueron, 0 y 1. En cada una de las capas sensoriales asociativas empleadas en la red, se empleó solamente una unidad del tipo ca1 y para cada capa motor asociativa usada se empleó solo una unidad del tipo VTA. Para la desviación estándar de la función logística, el valor predeterminado fue 0.1; el valor predeterminado para la sumación temporal fue 0.1, igual que el valor de decaimiento. Los valores predeterminados de las tasas de incremento y decremento fueron fijados en 0.5 y 0.1 respectivamente.

3.6.1.2. Aparatos

En todas las simulaciones se empleó una computadora personal portátil con procesador Pentium IV m, con velocidad de 3.5 gigahertz, 512 megabytes de memoria de acceso aleatorio, disco duro de 70 megabytes, unidad de lectura-escritura de discos compactos y video discos en formato DVD, sistema

operativo Microsoft Windows Xp Professional Edition. Equipada con los siguientes paquetes: entorno de programación Microsoft Visual Studio, versión 6.0; procesador de palabras Microsoft Word Xp; hoja de cálculo Microsoft Excel Xp; utilería para dibujo de diagramas Microsoft Visio.NET; el graficador Sigma Plot versión 2001 y el simulador del modelo DBP, creado por el Dr. José E. Burgos Triano, "Selnet".

3.6.1.3. Procedimientos

Durante la primera fase de esta condición experimental del experimento 5 se realizó una sesión de 300 ensayos, cada uno de ellos compuesto por 8 momentos temporales. En cada ensayo del experimento se tomó registro del nivel de activación de la unidad de salida (unidad número 12 en la Figura 34), en el momento temporal 7.

Los sujetos fueron entrenados con el conjunto de patrones de entrada que se presenta en la Tabla 23, cuya primer columna contiene el número de la entrada para las redes de este grupo y las siguientes ocho columnas muestran los valores de activación correspondientes en cada uno de los momentos temporales del ensayo. La fila marcada como Entrada S* muestra el valor de la activación en la entrada correspondiente al estímulo incondicionado. En todos los casos, el valor .01 representa un estado mínimo de actividad y el valor 0.9 el estado máximo.

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.9	.9	.9	.9	.9	.9	.9	.9
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.01	.01	.01	.01	.01	.01	.01	.01
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.9

Tabla 23. Patrones de entrada en cada momento temporal presentados a los sujetos en la primera fase de la condición experimental número 1 del experimento 6.

Aunque el simulador empleado permite manipular el intervalo entre ensayos (IEEn), no se incluyó en las simulaciones realizadas, asumiendo que se estaba manejando el valor ideal para el tipo de procedimiento empleado.

Los sujetos fueron entrenados durante 300 ensayos, debido a que la experiencia reportada con el trabajo con el modelo, sugiere que 300 ensayos son suficientes para que se llegue a una activación estable en la neurona de salida (Donahoe, Burgos y Palmer, 1993; Donahoe, Palmer y Burgos, 1997)

Durante la segunda fase de esta condición experimental se sometió a los sujetos a una nueva sesión de entrenamiento ante los patrones que se muestran en la Tabla 24. Esta segunda sesión contempló también 300 ensayos.

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.9	.9	.9	.9	.9	.9	.9	.9
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.01	.01	.01	.01	.01	.01	.01	.01
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.01

Tabla 24. Patrones de entrada en cada momento temporal presentados en la segunda fase de la condición experimental número 1 del experimento 6.

La segunda fase de esta condición experimental se realizó sin reestablecer los pesos de conexión de las redes neurales empleadas.

Durante la tercera fase de esta condición experimental se sometió a los sujetos a una nueva sesión de entrenamiento ante los patrones que se muestran en la Tabla 25. Esta sesión contempló también 300 ensayos.

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.9	.9	.9	.9	.9	.9	.9	.9
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.01

Tabla 25. Patrones de entrada en cada momento temporal presentados en la tercera fase de la condición experimental número 1 del experimento 6.

La tercera fase de esta condición experimental se realizó sin reestablecer los pesos de conexión de las redes neurales empleadas.

Durante la cuarta fase de esta condición experimental se sometió a los sujetos a una nueva sesión de entrenamiento ante los patrones que se muestran en la Tabla 26. Esta sesión contempló también 300 ensayos.

Número de entrada	Momento 1	Momento 2	Momento 3	Momento 4	Momento 5	Momento 6	Momento 7	Momento 8
Entrada 1	.9	.9	.9	.9	.9	.9	.9	.9
Entrada 2	.01	.01	.01	.01	.01	.01	.01	.01
Entrada 3	.01	.01	.01	.01	.01	.01	.01	.01
Entrada S*	.01	.01	.01	.01	.01	.01	.01	.01

Tabla 26. Patrones de entrada en cada momento temporal presentados en la cuarta fase de la condición experimental número 1 del experimento 6.

La cuarta fase de esta condición experimental se realizó sin reestablecer los pesos de conexión de las redes neurales empleadas.

4. RESULTADOS

4.1. EXPERIMENTO 1

4.1.1. CONDICIÓN EXPERIMENTAL 1

En esta condición experimental, las redes pequeñas procesando patrones pequeños de entrada, la neurona CR/UR comienza a tener un nivel de activación superior a 0.7 cerca del ensayo 110. Una vez que el nivel de activación alcanzó un nivel superior a 0.8, se mantuvo relativamente estable (ver la Figura 35).

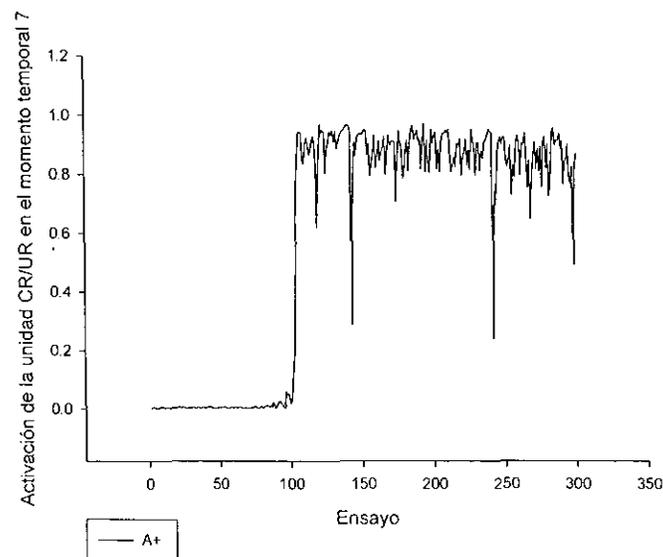


Figura 35. Nivel de activación de la unidad CR/UR en la primera fase de la condición experimental 1 del experimento 1.

Sin embargo, en la segunda fase de la condición experimental 1, la neurona CR/UR no alcanzó valores de activación superiores al 0.008 (ver la Figura 36).

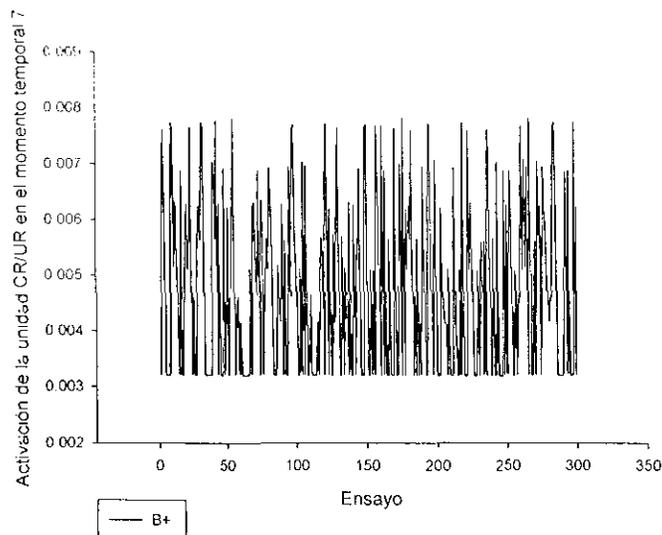


Figura 36. Nivel de activación de la unidad CR/UR en la segunda fase de la condición experimental 1 del experimento 1.

En la Figura 37 se muestra el valor promedio de la activación de la unidad de salida de la red. Como puede observarse, en la fase de entrenamiento ante el estímulo A+, el valor promedio durante los primeros 100 ensayos es muy cercano al cero, mientras que durante los siguientes 200 ensayos es cercano al 0.9. Para la segunda fase, es decir, durante el entrenamiento ante el estímulo B+, a lo largo de los 300 ensayos el valor decae a valores cercanos a cero.

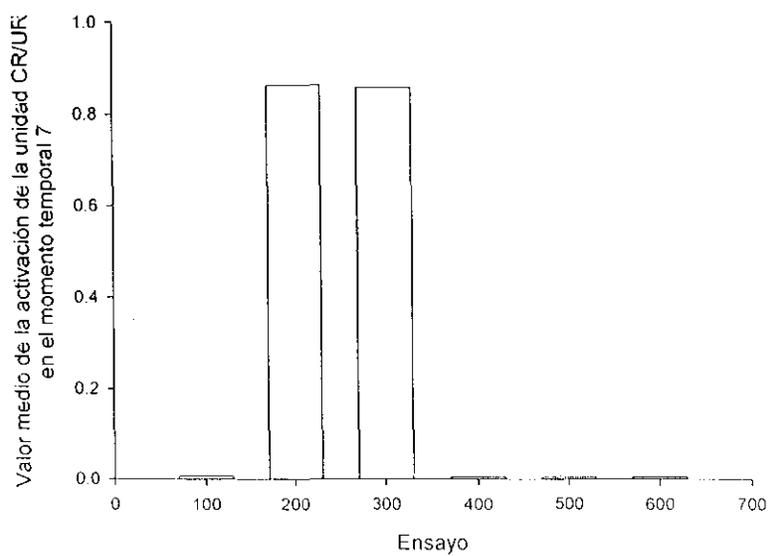


Figura 37. Nivel promedio de activación de la unidad CR/UR durante los 600 ensayos que implicó la condición experimental 1 del experimento 1.

4.1.2. CONDICIÓN EXPERIMENTAL 2

En esta condición experimental, las redes pequeñas procesando patrones grandes de entrada, la neurona CR/UR comienza a tener un nivel de activación superior a 0.7 cerca del ensayo 120. Una vez que el nivel de activación alcanzó un nivel superior a 0.8, se mantuvo relativamente estable (ver la Figura 38).

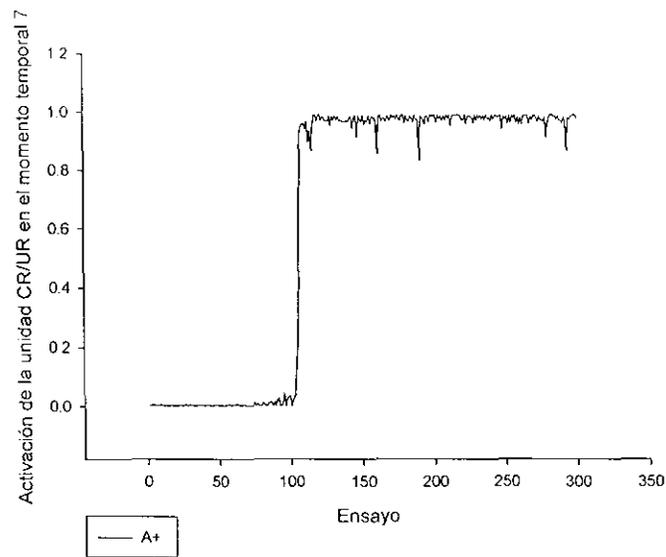


Figura 38. Nivel de activación de la unidad CR/UR en la primera fase de la condición experimental 2 del experimento 1.

Sin embargo, en la segunda fase de la condición experimental 2, la neurona CR/UR no alcanzó valores de activación superiores al 0.0083 (ver la Figura 39).

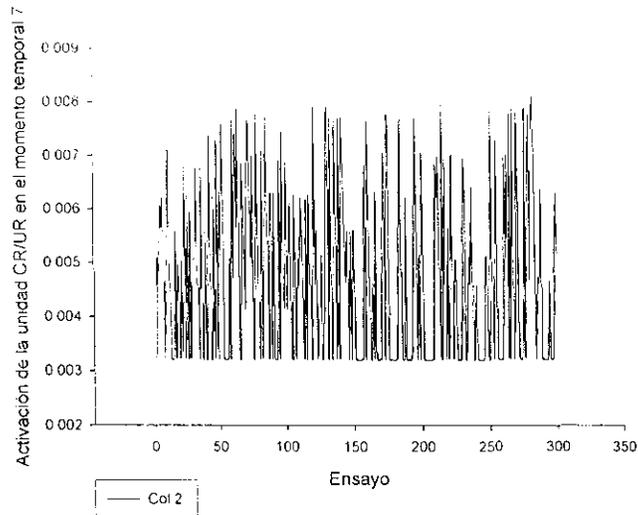


Figura 39. Nivel de activación de la unidad CR/UR en la segunda fase de la condición experimental 2 del experimento 1.

En la Figura 40 se muestra el valor promedio de la activación de la unidad de salida de la red. Como puede observarse, en la fase de entrenamiento ante el estímulo A+, el valor promedio durante los primeros 100 ensayos es muy cercano al cero, mientras que durante los siguientes 100 ensayos es cercano al 0.95 y en los siguientes 100 es muy cercano a 1.0. Para la segunda fase, es decir, durante el entrenamiento ante el estímulo B+, a lo largo de los 300 ensayos el valor decae a valores cercanos a cero.

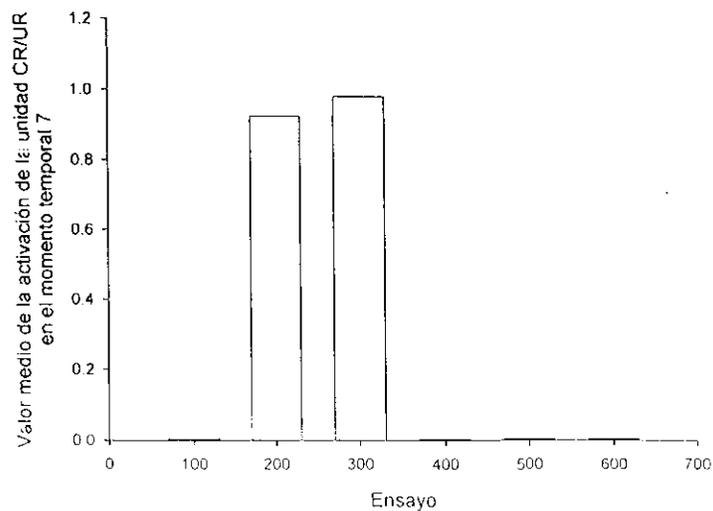


Figura 40. Nivel promedio de activación de la unidad CR/UR durante los 600 ensayos que implicó la condición experimental 2 del experimento 1.

4.1.3. CONDICIÓN EXPERIMENTAL 3

En esta condición experimental, las redes medianas procesando patrones pequeños de entrada, la neurona CR/UR comienza a tener un nivel de activación superior a 0.95 cerca del ensayo 97. Una vez que el nivel de activación alcanzó un nivel superior a 0.96, se mantuvo relativamente estable (ver la Figura 41).

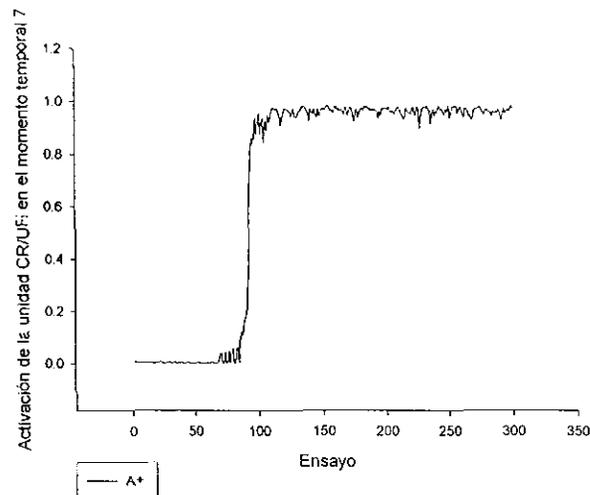


Figura 41. Nivel de activación de la unidad CR/UR en la primera fase de la condición experimental 3 del experimento 1.

Sin embargo, en la segunda fase de la condición experimental 3, la neurona CR/UR no alcanzó valores de activación superiores al 0.0077 (ver la Figura 42).

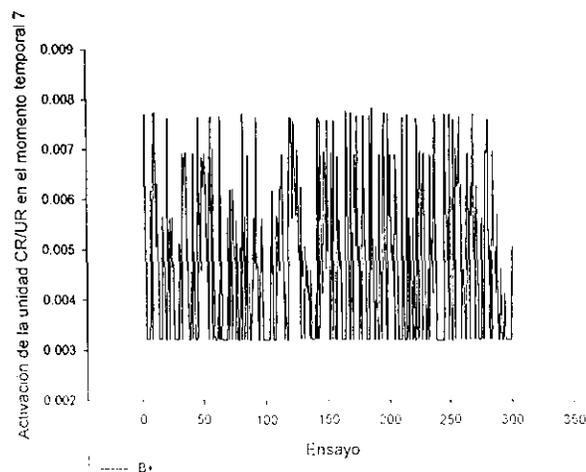


Figura 42. Nivel de activación de la unidad CR/UR en la segunda fase de la condición experimental 3 del experimento 1.

En la Figura 43 se muestra el valor promedio de la activación de la unidad de salida de la red. Como puede observarse, en la fase de entrenamiento ante el estímulo A+, el valor promedio durante los primeros 100 ensayos es muy cercano al 0.1, mientras que durante los siguientes 200 ensayos es cercano al 0.97. Para la segunda fase, es decir, durante el entrenamiento ante el estímulo B+, a lo largo de los 300 ensayos el valor decae a valores cercanos a cero.

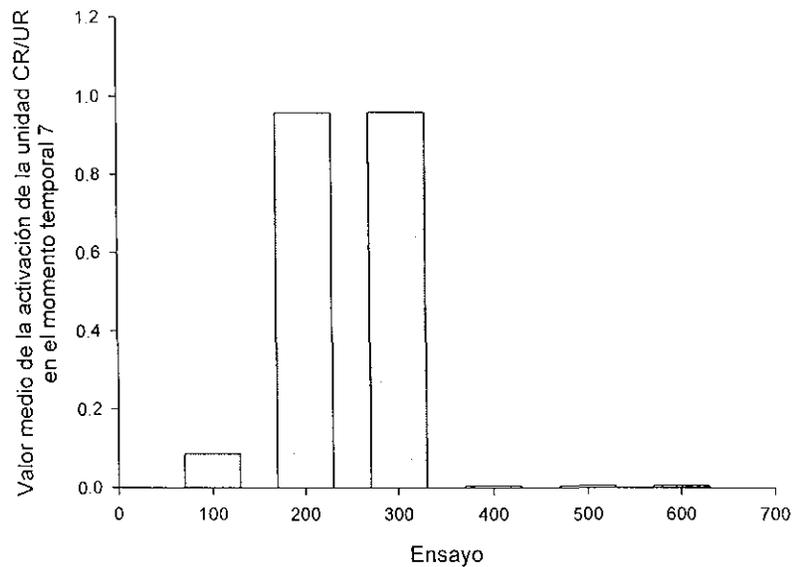


Figura 43. Nivel promedio de activación de la unidad CR/UR durante los 600 ensayos que implicó la condición experimental 1 del experimento 1.

4.1.4. CONDICIÓN EXPERIMENTAL 4

En esta condición experimental, las redes medianas procesando patrones pequeños de entrada, la neurona CR/UR comienza a tener un nivel de activación superior a 0.97 cerca del ensayo 93. Una vez que el nivel de activación alcanzó un nivel superior a 0.96, se mantuvo relativamente estable (ver la Figura 44).

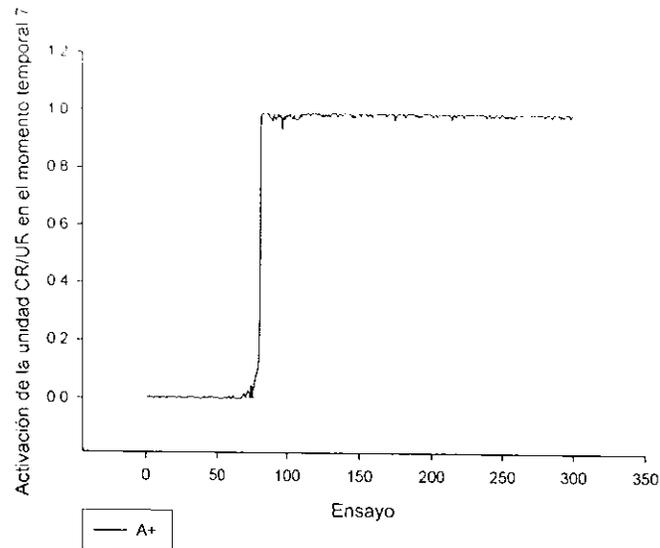


Figura 44. Nivel de activación de la unidad CR/UR en la primera fase de la condición experimental 4 del experimento 1.

Sin embargo, en la segunda fase de la condición experimental 4, la neurona CR/UR no alcanzó valores de activación superiores al 0.0078 (ver la Figura 45).

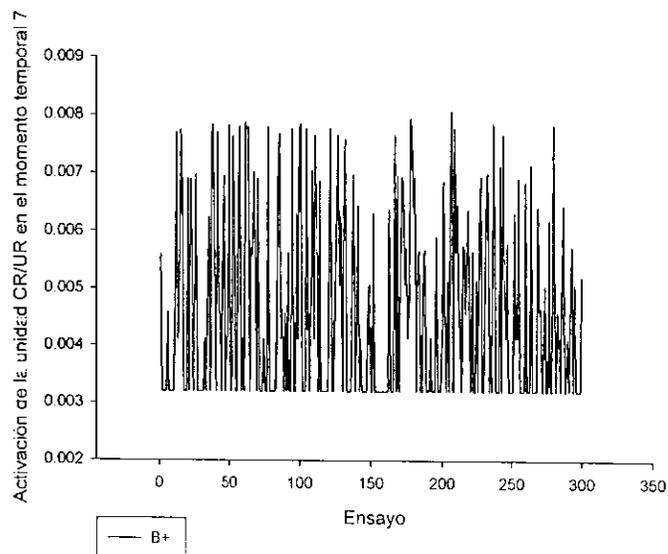


Figura 45. Nivel de activación de la unidad CR/UR en la segunda fase de la condición experimental 4 del experimento 1.

En la Figura 46 se muestra el valor promedio de la activación de la unidad de salida de la red. Como puede observarse, en la fase de entrenamiento ante el estímulo A+, el valor promedio durante los primeros 100 ensayos es muy cercano al 0.2, mientras que durante los siguientes 200 ensayos es cerca-

no al 1.0. Para la segunda fase, es decir, durante el entrenamiento ante el estímulo B+, a lo largo de los 300 ensayos el valor decae a valores cercanos a cero.

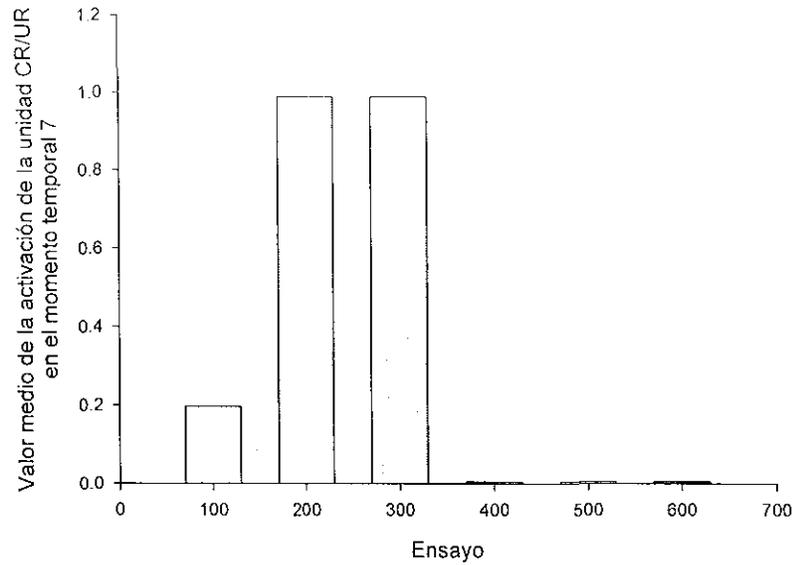


Figura 46. Nivel promedio de activación de la unidad CR/UR durante los 600 ensayos que implicó la condición experimental 1 del experimento 1.

4.1.5. CONDICIÓN EXPERIMENTAL 5

En esta condición experimental, las redes medianas procesando patrones pequeños de entrada, la neurona CR/UR comienza a tener un nivel de activación superior a 0.97 cerca del ensayo 90. Una vez que el nivel de activación alcanzó un nivel superior a 0.96, se mantuvo relativamente estable (ver la Figura 47).

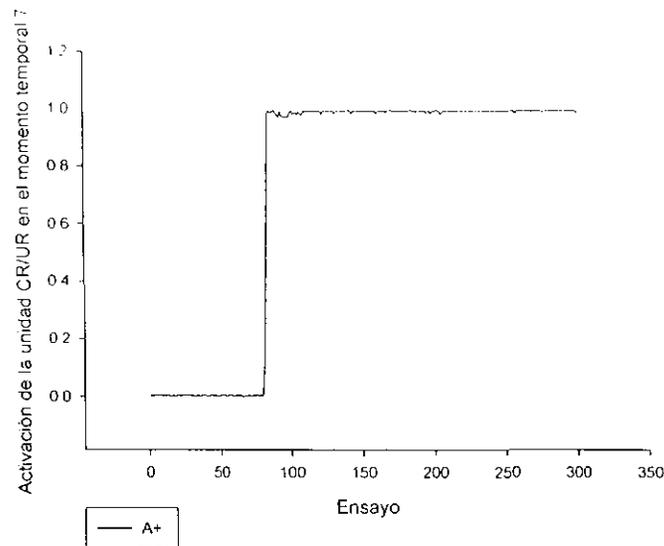


Figura 47. Nivel de activación de la unidad CR/UR en la primera fase de la condición experimental 5 del experimento 1.

Sin embargo, en la segunda fase de la condición experimental 5, la neurona CR/UR no alcanzó valores de activación superiores al 0.0078 (ver la Figura 48).

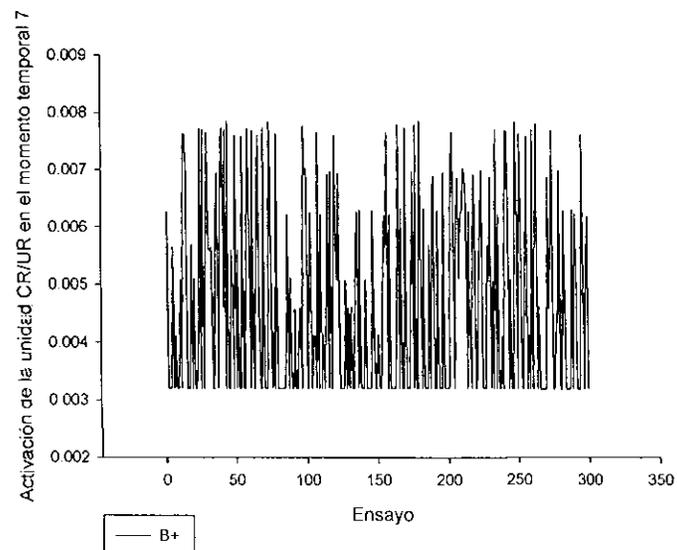


Figura 48. Nivel de activación de la unidad CR/UR en la segunda fase de la condición experimental 5 del experimento 1.

En la Figura 49 se muestra el valor promedio de la activación de la unidad de salida de la red. Como puede observarse, en la fase de entrenamiento ante el estímulo A+, el valor promedio durante los primeros 100 ensayos es muy cercano al 0.2, mientras que durante los siguientes 200 ensayos es cerca-

no al 1.0. Para la segunda fase, es decir, durante el entrenamiento ante el estímulo B+, a lo largo de los 300 ensayos el valor decae a valores cercanos a cero.

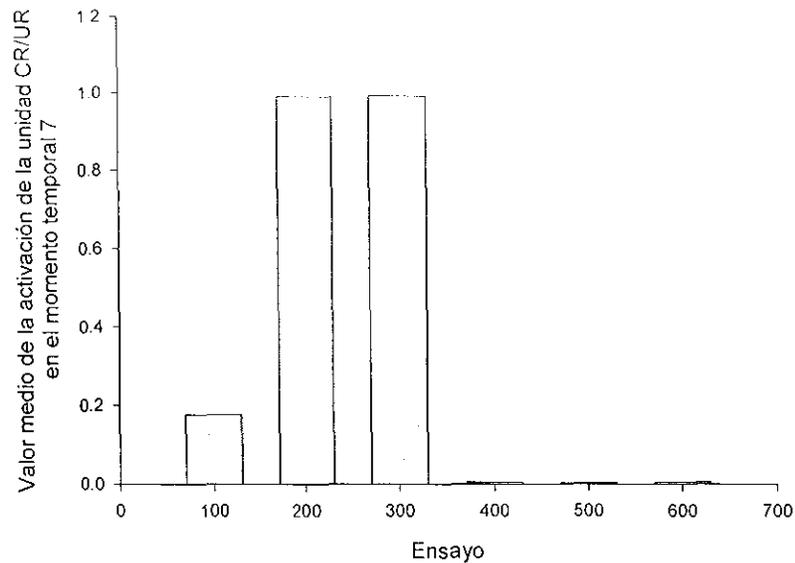


Figura 49. Nivel promedio de activación de la unidad CR/UR durante los 600 ensayos que implicó la condición experimental 1 del experimento 1.

4.1.6. CONDICIÓN EXPERIMENTAL 6

En esta condición experimental, las redes medianas procesando patrones pequeños de entrada, la neurona CR/UR comienza a tener un nivel de activación superior a 0.97 cerca del ensayo 88. Una vez que el nivel de activación alcanzó un nivel superior a 0.98, se mantuvo relativamente estable (ver la Figura 50).

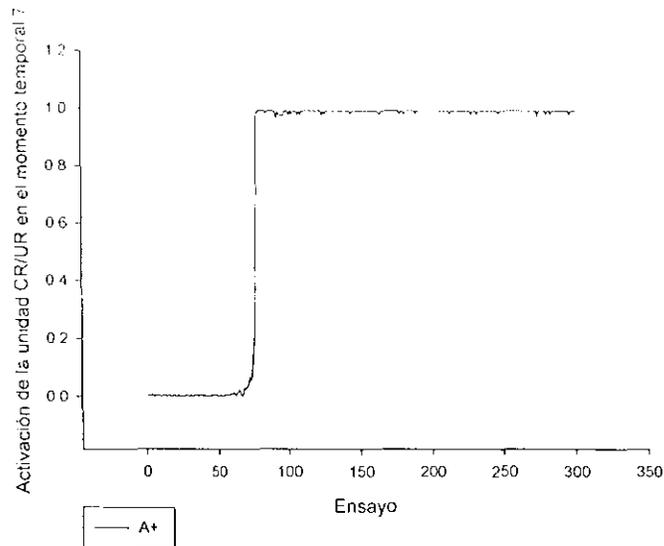


Figura 50. Nivel de activación de la unidad CR/UR en la primera fase de la condición experimental 6 del experimento 1.

Sin embargo, en la segunda fase de la condición experimental 6, la neurona CR/UR no alcanzó valores de activación superiores al 0.0078 (ver la Figura 51).

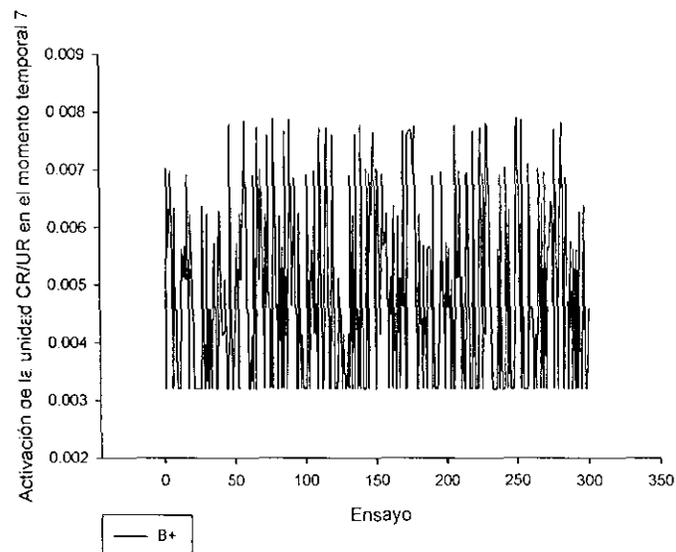


Figura 51. Nivel de activación de la unidad CR/UR en la segunda fase de la condición experimental 1 del experimento 1.

En la Figura 52 se muestra el valor promedio de la activación de la unidad de salida de la red. Como puede observarse, en la fase de entrenamiento ante el estímulo A+, el valor promedio durante los primeros 100 ensayos es muy cercano al 0.24, mientras que durante los siguientes 200 ensayos es cer-

cano al 0.99. Para la segunda fase, es decir, durante el entrenamiento ante el estímulo B+, a lo largo de los 300 ensayos el valor decae a valores cercanos a cero.

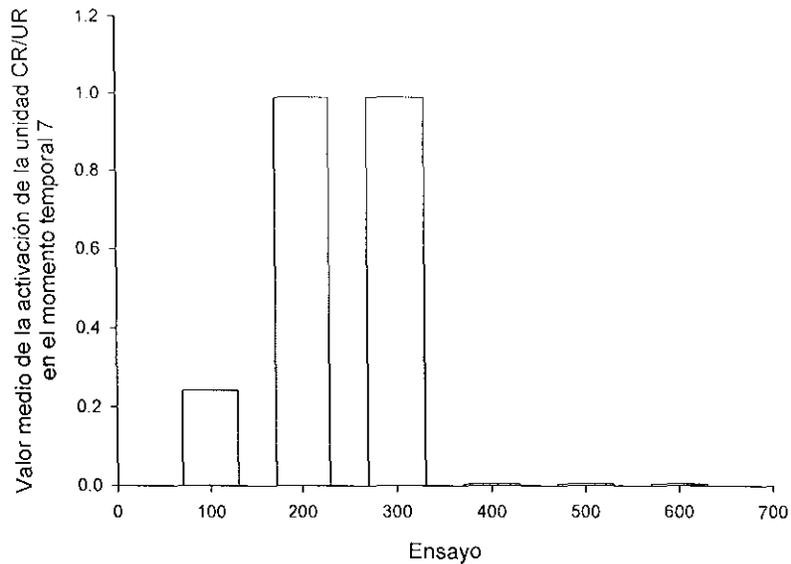


Figura 52. Nivel promedio de activación de la unidad CR/UR durante los 600 ensayos que implicó la condición experimental 1 del experimento 1.

4.2. EXPERIMENTO 2

4.2.1. CONDICIÓN EXPERIMENTAL 1

En esta condición experimental, las redes pequeñas procesando patrones pequeños de entrada, la neurona CR/UR comienza a tener un nivel de activación superior a 0.7 cerca del ensayo 105. Muy similar a lo obtenido en la primera fase de la condición experimental 1 del experimento 1. Una vez que el nivel de activación alcanzó un nivel superior a 0.8, se mantuvo relativamente estable (ver la Figura 53).

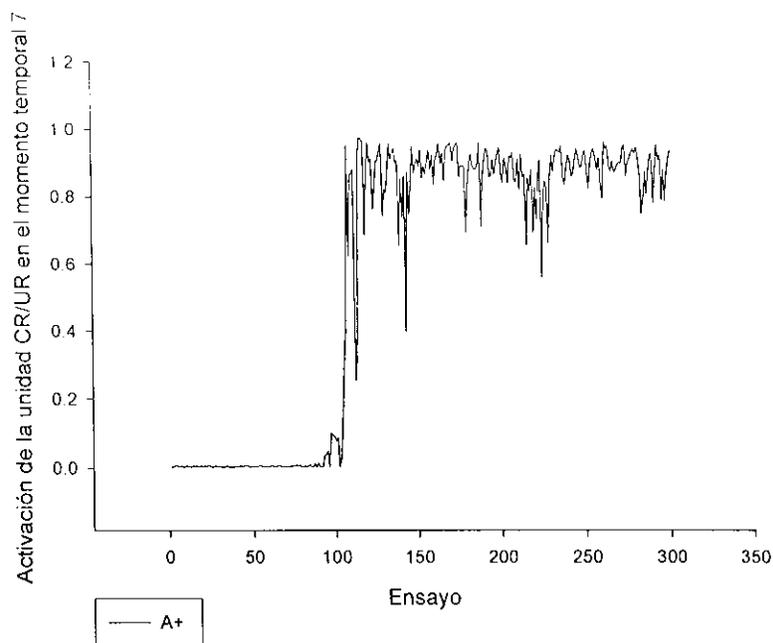


Figura 53. Nivel de activación de la unidad CR/UR en la primera fase de la condición experimental 1 del experimento 2.

Similar a la fase 2 de la condición experimental 1 del experimento 1, en esta segunda fase de la condición experimental 6, la neurona CR/UR no alcanzó valores de activación superiores al 0.0078 (ver la Figura 54).

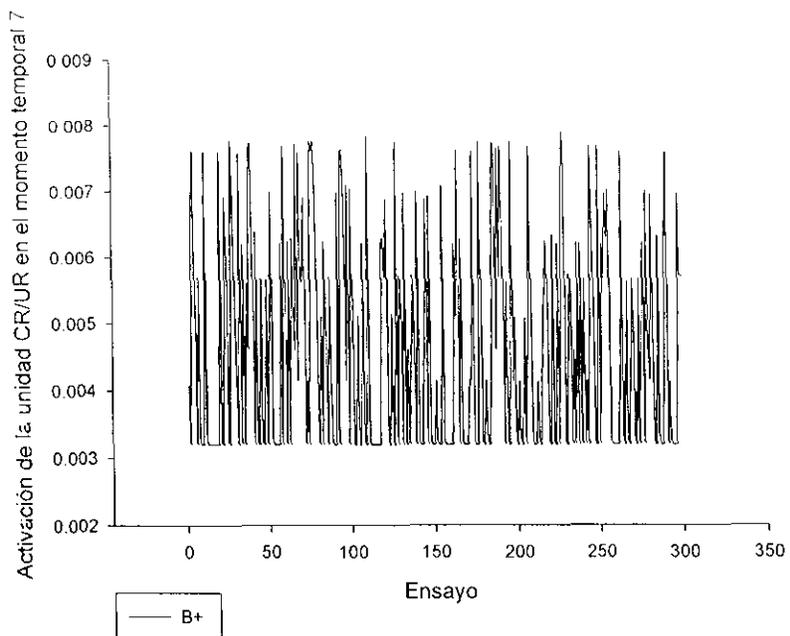


Figura 54. Nivel de activación de la unidad CR/UR en la segunda fase de la condición experimental 2 del experimento 2.

En la tercera fase de la condición experimental 1 del experimento 2, los niveles de activación de la neurona CR/UR regresaron a valores muy cercanos al 0.97 y se mantuvieron relativamente estables durante los 300 ensayos de esta fase (ver Figura 55).

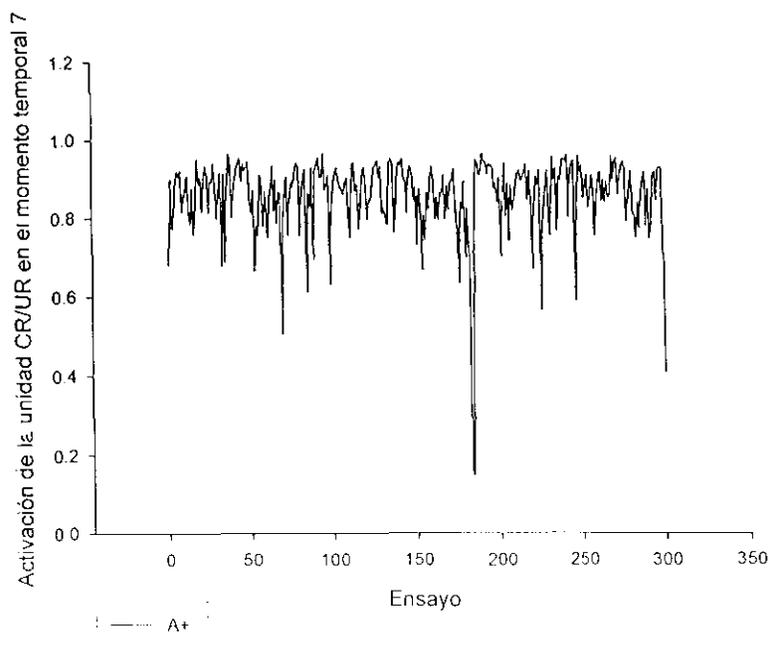


Figura 55. Nivel de activación de la unidad CR/UR en la segunda fase de la condición experimental 3 del experimento 2.

En la Figura 56 se muestra el valor promedio de la activación de la unidad de salida de la red. Como puede observarse, en la fase de entrenamiento ante el estímulo A+, el valor promedio durante los primeros 100 ensayos es muy cercano al cero, mientras que durante los siguientes 100 ensayos el valor se incrementa a .820; en los siguientes 100 ensayos el valor se incrementa aún más para llegar a 0.86. Para la segunda fase, es decir, durante el entrenamiento ante el estímulo B+, a lo largo de los 300 ensayos el valor decae a valores cercanos a cero. En la tercera fase, los valores regresan a niveles cercanos al 0.83 durante los últimos 300 ensayos.

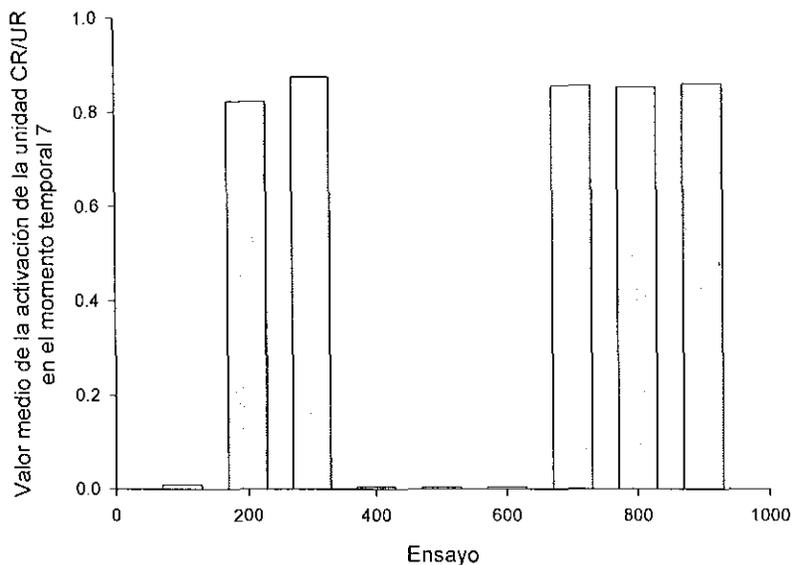


Figura 56. Nivel promedio de activación de la unidad CR/UR durante los 900 ensayos que implicó la condición experimental 1 del experimento 2.

4.3. EXPERIMENTO 3

4.3.1. CONDICIÓN EXPERIMENTAL 1

En esta condición experimental, las redes pequeñas procesando patrones pequeños de entrada, la neurona CR/UR comienza a tener un nivel de activación superior a 0.7 cerca del ensayo 128. Sin embargo, los niveles de activación alto se presentan ante solamente uno de los dos estímulos. En el caso

de la Figura 57 se trata del estímulo B, mientras que ante el estímulo A la red sigue manteniendo niveles bajos de activación. La figura 57 muestra solo un ejemplo del tipo de comportamiento que presenta la red. Parece no existir una tendencia hacia niveles de activación ante alguno de los dos estímulos específicos. Del total de sujetos empleados en esta condición experimental, el 51% presentó niveles altos de activación en la neurona CR/UR ante el estímulo A, mientras que el 49% de ellos presentó niveles altos en la misma unidad ante el estímulo B.

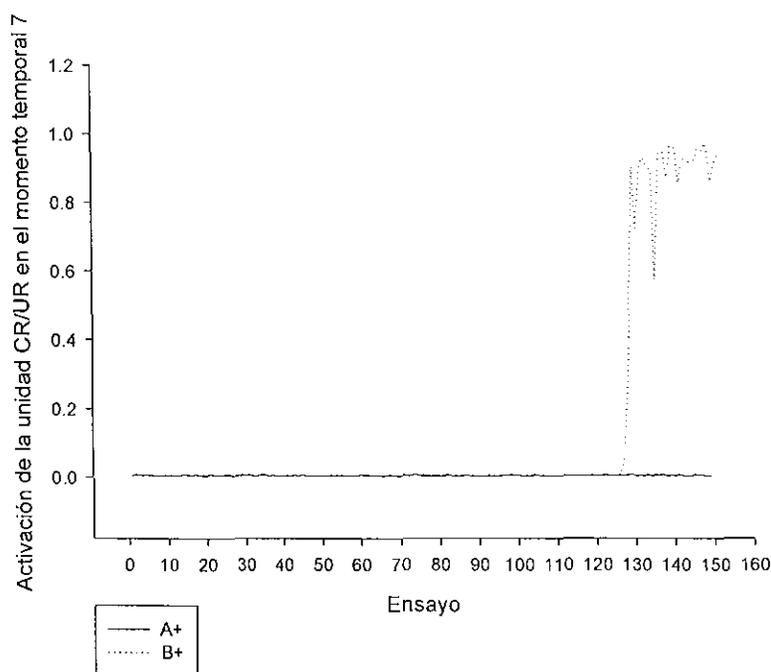


Figura 57. Ejemplo del Nivel de activación obtenido en la unidad CR/UR durante la condición experimental 1 del experimento 3.

4.4. EXPERIMENTO 4

4.4.1. CONDICIÓN EXPERIMENTAL 1

Los resultados obtenidos en esta condición experimental son similares a los obtenidos en la primera fase de la condición experimental 1 del experimento 1. En la Figura 58 se presentan los niveles de activación logrados por la unidad CR/UR.

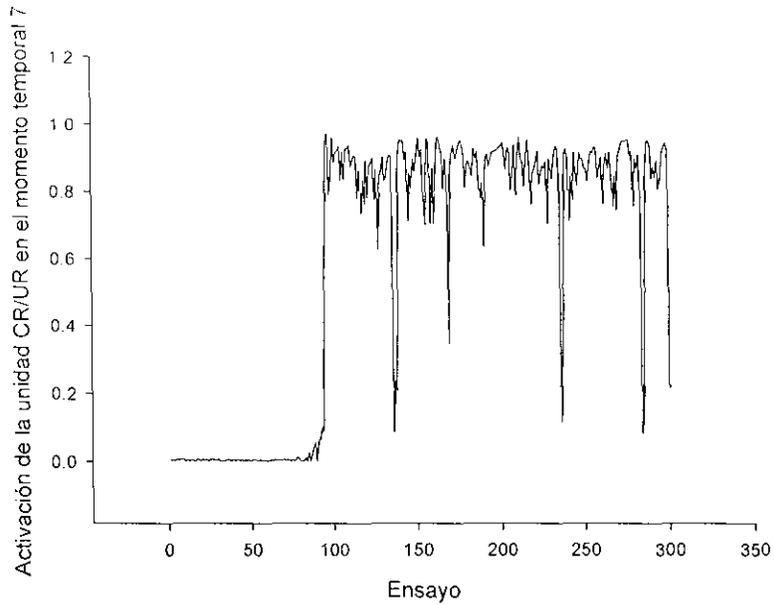


Figura 58. Nivel de activación obtenido en la unidad CR/UR durante la condición experimental 1 del experimento 4.

4.5. EXPERIMENTO 5

4.5.1. CONDICIÓN EXPERIMENTAL 1

Los resultados obtenidos en esta condición experimental son similares a los obtenidos en la primera fase de la condición experimental 1 del experimento 1. En la Figura 59 se presentan los niveles de activación logrados por la unidad CR/UR.

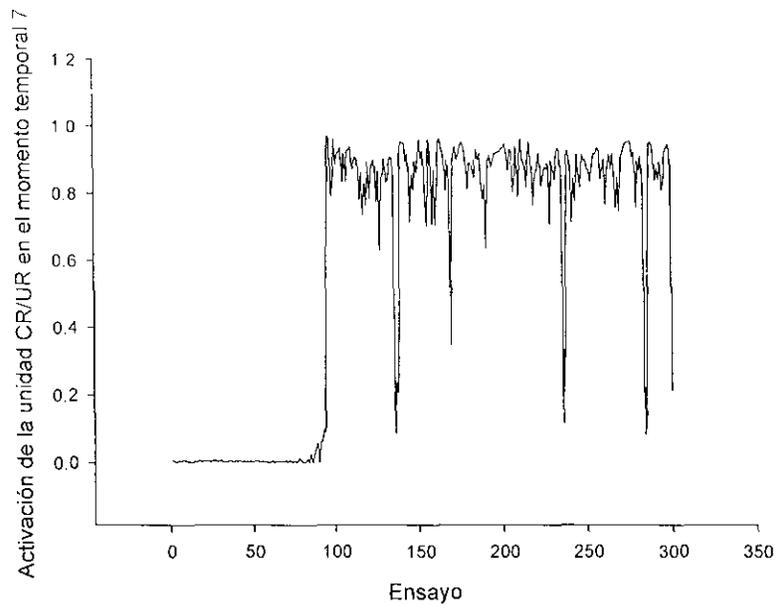


Figura 59. Nivel de activación obtenido en la unidad CR/UR durante la primera fase de la condición experimental I del experimento 5.

En la Figura 60 se muestran los niveles de activación logrados durante la fase de presentación del estímulo A no reforzado. El decremento en los niveles de actividad de la unidad CR/UR se da en los primeros 8 ensayos, para después mantenerse estable en niveles muy cercanos al cero.

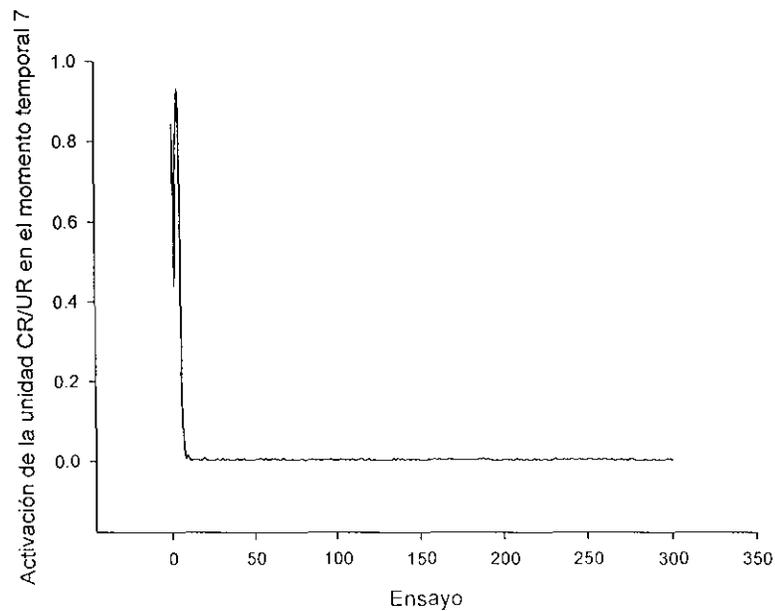


Figura 60. Nivel de activación obtenido en la unidad CR/UR durante la segunda fase de la condición experimental I del experimento 5.

Durante la fase del experimento que implicó el entrenamiento con el estímulo B, los niveles de activación de la unidad CR/UR se mantuvieron estables en valores muy cercanos al cero, como se puede observar en la Figura 61.

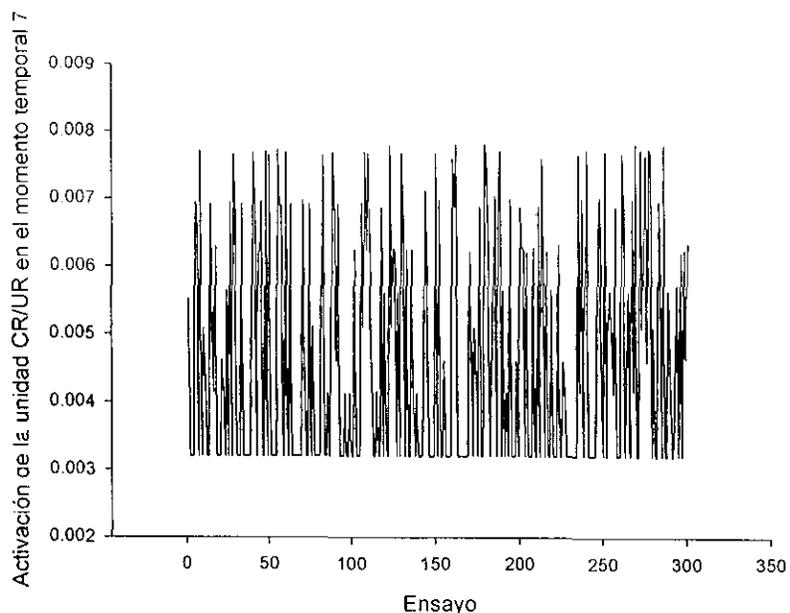


Figura 61. Nivel de activación obtenido en la unidad CR/UR durante la tercera fase de la condición experimental 1 del experimento 5.

4.6. EXPERIMENTO 6

4.6.1. CONDICIÓN EXPERIMENTAL 1

Los resultados obtenidos en esta condición experimental son similares a los obtenidos en la primera fase de la condición experimental 1 del experimento 1. En la Figura 62 se presentan los niveles de activación logrados por la unidad CR/UR. El incremento en el nivel de activación de la unidad se da a partir del ensayo 106.

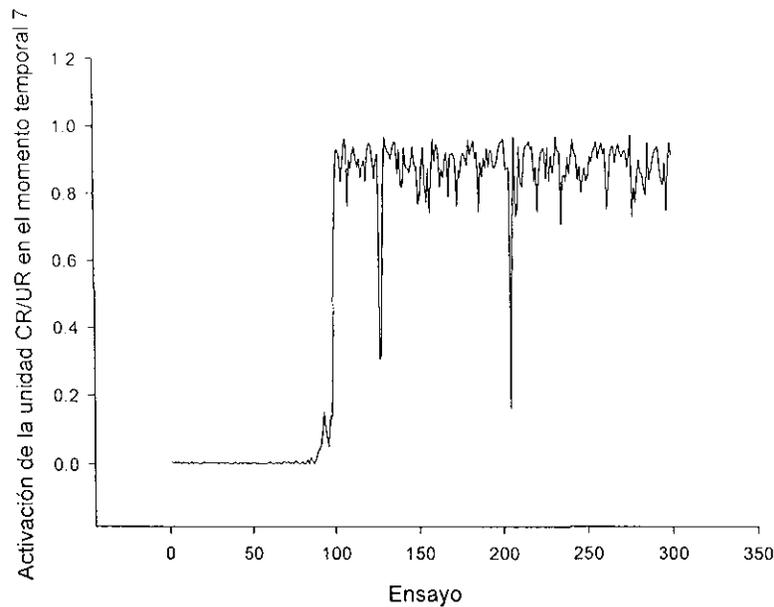


Figura 62. Nivel de activación obtenido en la unidad CR/UR durante la primera fase de la condición experimental 1 del experimento 6.

En la Figura 63 se muestran los niveles de activación logrados durante la fase de presentación del estímulo A no reforzado. El decremento en los niveles de actividad de la unidad CR/UR se da en los primeros 12 ensayos, para después mantenerse estable en niveles muy cercanos al cero.

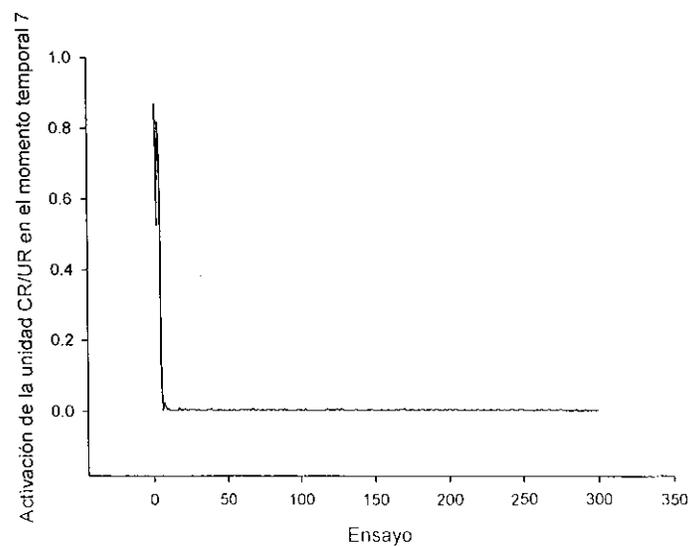


Figura 63. Nivel de activación obtenido en la unidad CR/UR durante la segunda fase de la condición experimental 1 del experimento 6.

Durante la fase del experimento que implicó el entrenamiento con el estímulo B, los niveles de activación de la unidad CR/UR se mantuvieron estables en valores muy cercanos al cero, como puede observarse en la Figura 64.

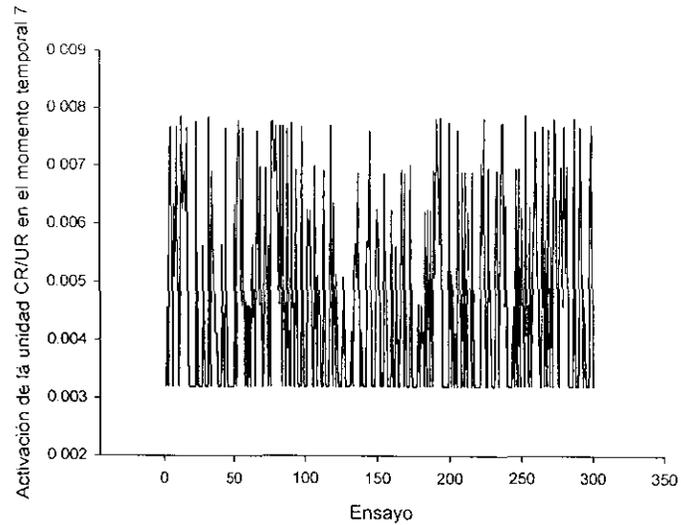


Figura 64. Nivel de activación obtenido en la unidad CR/UR durante la tercera fase de la condición experimental 1 del experimento 6.

En la Figura 65 se puede observar un incremento en los niveles de activación de la unidad CR/UR muy similares a los obtenidos en los últimos 100 ensayos de la fase 1 de este mismo experimento (ver Figura 62) los valores se sostuvieron entorno a un valor de 0.95.

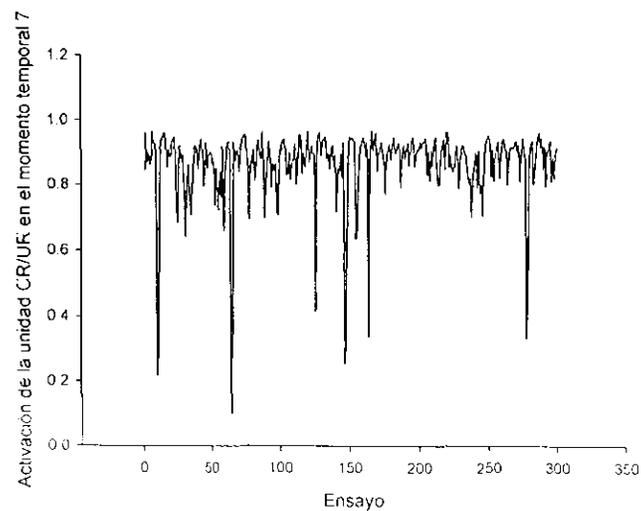


Figura 65. Nivel de activación obtenido en la unidad CR/UR durante la cuarta fase de la condición experimental 1 del experimento 6.

5. DISCUSIÓN

Los resultados obtenidos han mostrado que la implementación computacional del modelo DBP para que una red que muestra niveles altos de activación en la unidad CR/UR ante un estímulo determinado, llegue a mantener niveles altos de activación en esta misma unidad ante un patrón de estímulos de entrada distinto.

No parece haber evidencia de que el tamaño de la red, entendido como el número de capas ocultas tanto de unidades sensoriales asociativas, como de motoras asociativas; sea una variable que inter venga en esta imposibilidad para aprender un estímulo nuevo. Tampoco parece haber evidencia de que esta imposibilidad para aprender estímulos nuevos dependa del tamaño de los patrones de entrada que se presenten en la situación de entrenamiento.

En el experimento número 6 se sometieron las redes al olvido de un primer estímulo aprendido y al igual que en el experimento 5 se corroboró que aún después de que los niveles de activación de la unidad CR/UR bajaron a valores muy cercanos al cero; la red era incapaz de aprender a responder ante un nuevo estímulo. En el experimento 6 se llega más allá al comprobar que con estimulación, la red es capaz de regresar a los niveles originales de activación en la unidad CR/UR ante el mismo estímulo. En todos los casos de re-entrenamiento ante el patrón de estímulos original, la recuperación de los niveles de activación se dio dentro de los 20 primeros ensayos, tiempo que es relativamente corto, considerando el tiempo que lleva la adquisición en redes que no han sido expuestas previamente al estímulo, como las de los experimentos 1 al 3.

El experimento 4 demostró que tampoco es posible para las redes de este tipo el aprendizaje de dos estímulos presentados alternadamente o aleatoriamente en la misma sesión. La red termina aprendiendo a responder ante uno u otro estímulo, pero no ante ambos. En este tipo de casos, el comportamiento de la red ante el estímulo aprendido es muy similar al de la adquisición de un patrón único, solo que sufre una especie de retardo leve (no más de 30 ensayos) probablemente debido a la presentación del patrón alterno. El comportamiento de la red ante el patrón alterno es de mínima activación de la neurona CR/UR similar a la baja actividad presentada cuando se trata de entrenar a la red ante un nuevo estímulo una vez que se ha aprendido a responder ante otro previo.

Este problema adquiere especial importancia si se considera que gran parte de los fenómenos conductuales observados en organismos vivos implican la intervención de dos o más estímulos condicionados. Por ello, el sorteo de este obstáculo se presenta como el primer paso posible hacia la simulación de los mismos con el DBP. Es importante resaltar que, dados los alcances del modelo en la simulación de gran cantidad de fenómenos conductuales, la tarea se vuelve aún más compleja al tener que conservar los logros alcanzados hasta ahora al momento de avanzar en la resolución del problema. Ello implicaría buscar la solución del problema tratando de conservar hasta el máximo posible el propio modelo.

Con base en el trabajo realizado, se vislumbra que el primer paso hacia la solución de este problema implicaría la manipulación de las siguientes variables: a) arquitectura, que aunque en este trabajo se ha manipulado inicialmente, concibiendo redes con diferentes tamaños, queda aún por explorar el desempeño que tendría este tipo de redes en situaciones de interconectividad parcial o formas diferentes de conexión entre unidades de las diferentes capas; b) el manejo de conexiones inhibitorias en alguna o todas las capas de la red; c) los parámetros internos de la red; y finalmente, las reglas de aprendizaje y de activación que el modelo establece.

La apuesta inicial hacia la arquitectura de las redes, parte de lo siguiente: Si se considera que una red neural S1, bajo el Modelo DBP es capaz de manejar un estímulo condicionado CS1 y otra red neural S2, bajo el Modelo DBP, es capaz de manejar un estímulo condicionado CS2; y que en ambos casos las redes, una vez que han sido entrenadas para responder a los estímulos correspondientes, han logrado un nivel de configuración que les impide el aprendizaje de un nuevo patrón; entonces, habría que buscar una combinación tal en la que las dos redes neurales S1 y S2 fuesen consideradas como parte de un mismo sistema, con lo que darían a este sistema la capacidad de manejar dos estímulos condicionados CS1 y CS2 al mismo tiempo. En la Figura 66 se presenta el esquema mediante el cual se buscaría la nueva red fusionada a partir de dos redes neurales entrenadas previamente.

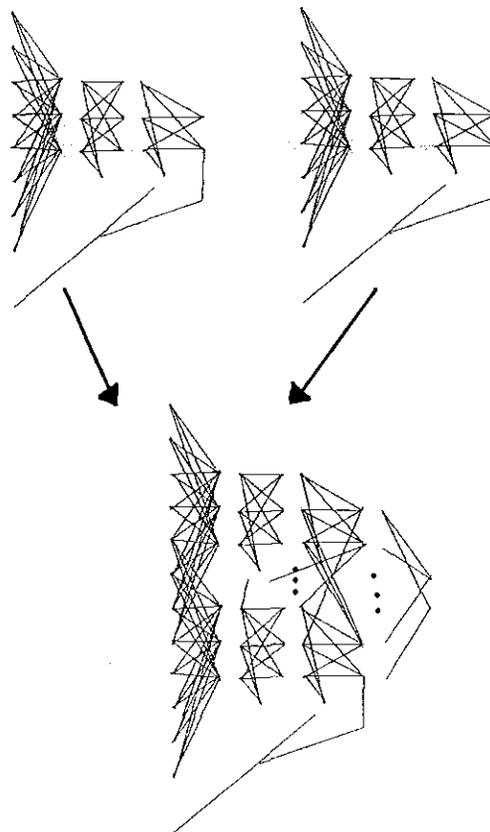


Figura 66. Búsqueda de una Arquitectura adecuada para que una Red DBP pueda manejar 2 CS's.

Si embargo, sería altamente criticable considerar como un mismo sistema a dos redes neurales completamente separadas, por ello, habría que probar con arquitecturas que busquen una posible solución al sistema dentro de este rango de arquitecturas posibles, que van desde los dos sistemas separados, que si pueden manejar a CS1 y CS2 simultáneamente y el sistema completamente integrado, que está impedido para manejarlos. El trabajo, manipulando esta variable, sería entonces, la búsqueda sistemática de una arquitectura que permita el manejo de dos estímulos, dentro de este rango de posibilidades.

El primer paso en este sentido, sería demostrar que existe una arquitectura tal, con una configuración de pesos determinada, que dadas las funciones de activación que el Modelo DBP, permitiría a la red responder ante dos estímulos completamente distintos. Aún a pesar de que esta arquitectura ideal haya sido encontrada con métodos manuales y no conformada a partir de la regla de aprendizaje del modelo.

Esta búsqueda manual de la arquitectura ideal se podría realizar con base en el análisis de los pesos finales de la red entrenada, examinando su distribución total. Este análisis permitiría establecer cuáles son los "pesos significativos" en cada sujeto para responder ante el patrón de entrada determinado. Entendiendo por pesos significativos a aquellos valores altos de conexión, resultantes del proceso de entrenamiento. En este momento, cabría esperar que la configuración final de pesos permita establecer por lo menos dos categorizaciones de pesos llegando a establecer diferencias entre los significativos y los no significativos para la identificación del patrón entrenado.

Si esta distinción no fuera posible tras un análisis rápido de la configuración de pesos, se podría recurrir a clasificadores automáticos de vectores, de manera que posibilitara una categorización de los mismos. Una vez establecidos los pesos significativos y los no significativos, se podrían conformar

nuevos sujetos integrados por las conexiones entre los nodos cuyos pesos de conexión en la fase anterior han resultado significativos; con lo que se posibilitaría la transformación de una red neural como se presenta en la Figura 67 y su integración en una nueva red neural compuesta como la que se presenta en la Figura 68.

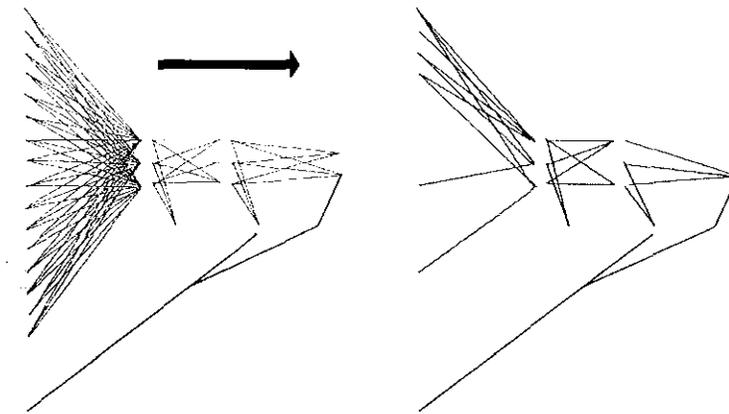


Figura 67. Transformación de una arquitectura de red, con base en la reconfiguración a partir de la selección de conexiones significativas.

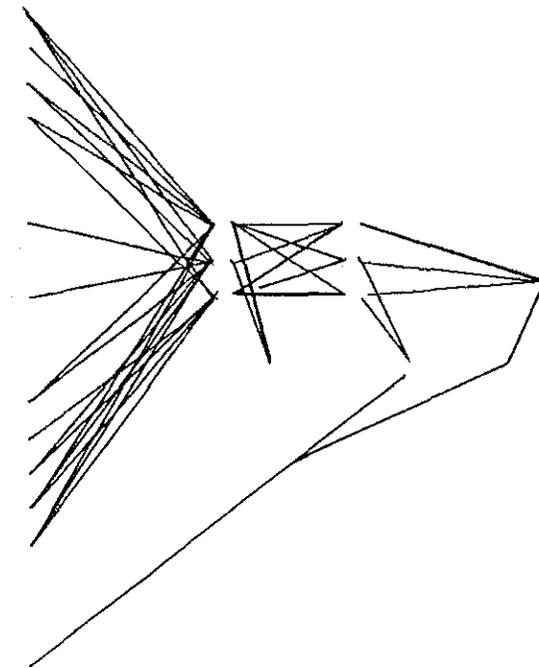


Figura 68. Nuevo sujeto conformado a partir de las dos redes con pesos no significativos podados

El siguiente paso sería la demostración de que la función de activación puede conservar las propiedades individuales de las redes originales en el interior de esta nueva red. Una vez demostrado que la función de activación puede manejar dos estímulos condicionados al mismo tiempo, entonces, se hace necesario el establecimiento de una estrategia que lleve al ajuste de pesos ante esta nueva configuración de interconexiones. Por supuesto, que el primer paso sería probar el algoritmo original del Modelo a fin de conocer, si es posible lograr el funcionamiento del Modelo ante esta nueva configuración. Dado que la experiencia con el modelo indica que no hay problema con redes parcialmente conectadas, se vislumbran pocos problemas llegado este momento.

Sin embargo, es necesario considerar que hasta este momento, todo son suposiciones, que aunque se hacen con base en algo de experiencia de trabajo con el Modelo, es necesario que sean corroboradas. Finalmente, este trabajo apunta la necesidad de explorar estas vías de solución en aras de poder resolver este problema.

En este trabajo se ha explorado un poco acerca de la posible intervención de algunas de las variables implicadas en el problema; o al menos de las que se creía intuitivamente que estarían implicadas en el problema. Finalmente se puede establecer con base en los resultados obtenidos, que la intervención de estas variables en el problema son nulas, y que efectivamente, existe una imposibilidad para que las redes neurales diseñadas actualmente bajo las especificaciones del Modelo DBP puedan aprender a responder ante un nuevo estímulo, una vez que han sido entrenadas para responder a uno diferente.

REFERENCIAS

- Aizawa K, *History of connectionism*, documento en internet, dirección: <http://artsci.wustl.edu/~philos/MindDict/connectionismhistory.html>. Revisión: 6 de Julio del 2002.
- Alkon, D L., T P. Vogl y D Tam, Memory Function in Neural and Artificial Networks, *Neural Network Models of Conditioning and Action*, Nueva York, Lawrence Erlbaum Associates Publishers, 1991, pp. 1-11.
- Baxter, D. A., D. V. Buonomano, J. L. Raymond, D. G. Cook, F. M. Kuenzi, T. J. Carew y J.H. Byrne, Empirically derived adaptative elements and networks simulate associative learning, *Neural Network Models of Conditioning and Action*, Nueva York, Lawrence Erlbaum Associates Publishers, 1991, pp. 13-52.
- Burgos, J E. Superstition in Artificial Neural Networks, *Revista Mexicana de Análisis de la Conducta*, Volumen 26, Número 2, 2000, pp. 159-190.
- Burgos, J. E., Evolving artificial neural networks in pavlovian environments, *Neural Networks Models of Cognition*, J. W. Donahoe y Packard Dorsel (Editores), 1997, Elsevier Science B.V., pp. 58-79.
- Cheng-Lian y Lin Chin-Teng, Reinforcement Learning for an ART-Based Fuzzy Adaptative Learning Control Network, *IEEE Transactions on Neural Networks*, Vol 7, No 3, Mayo de 1996, pp. 709- 731.
- Churchland, P. *Matter and Consciousness*, Revised Edition, The MIT-press, 1993.
- Cohen, Paul R. *Empirical Methods for Artificial Intelligence*, Cambridge Mass, The MIT Press, 1995.
- Donahoe, J W, J E. Burgos y D C. Palmer, A Selectionist approach to reinforcement, *Journal of The Experimental Analysis of Behavior*, Vol 60, Num 1, Julio de 1993, pp. 17- 40.
- Freeman J.A. y D.M. Skapura, *Neural Networks, Algorithms, Applications and Programming Techinques*, Addison-Wesley, 1991.
- Gallant, Stephen I. *Neural Network Learning and Expert Systems*, Cambridge, The MIT Press, 1994.
- Gluck, Mark A y Catherine E. Myers, A Neural Network Approach to Adaptative Similarity and Stimulus Representations in Cortico-hippocampal function *Neural Networks Models of Cognition*, edición de J. W. Donahoe y V. Packard Dorsel, 1997, Elsevier Science, B.V. pp. 220-243.
- Haykin, S. *Neural networks. A Comprehensive foundation*, Nueva York, McMillan College Publishing Company, 1994.
- Hebb, D. O. *The Organization of Behavior*, Nueva York, Wiley. Introducción, Capítulo III y Capítulo IV, reimpresos *Neurocomputing*, J. Anderson y E. Rosenfeld (editores), The MIT-Press.
- Hecht-Nielsen, R, *Neurocomputing*, Reading, Addison-Wesley Publishing Company, 1989.
- Hilera, J. R. y V J. Martinez, *Redes Neurales Artificiales, Fundamentos, modelos y aplicaciones*, México, Editorial Alfa Omega, 2000.
- Hopfield, J.J. Neural Networks and Physical systems with emergent collective computational habiliteís, *Proceedings of The National Academy of Sciences*, No. 79, 1982, págs. 2554- 2558.
- Humphrey, W. S. *Managing the software process*, Addison-wesley publishing company, 1990.
- James, W, *Principles of Psychology*, Henry Holt, New York, 1890
- Kaski, S. *Data exploration using self-organizing maps*, Número 82 de la serie Acta Polytechnica Scandinavica, Mathematics, Computing and Management in Engineering Series, (English version) Finnish Academy of Technology, Espoo, 1997
- Kehoe, E. James, A layered model of associative learning: learning to learn and configuration, *Psychological Review*, 1988, Vol 95, No. 4, págs 411-433.
- Kohonen, Teuvo, Self Organized formation of topologically correct feature maps, *Biological Cybernetics*, No. 43, 1982, págs. 59-69.

- Lpimann, R. P. An introduction to computing with Neural Nets, *IEEE ASSP Magazine*, IEEE, Abril de 1987, págs 4 -22.
- Maravall Gómez_Allende, D, *Reconocimiento de formas y visión artificial*, Wilmington, Delaware, Addison-Wesley Iberoamericana/RAMA, 1994.
- McClelland, J. L., D. E. Rummelhart y G. Hinton, Representaciones distribuidas, *Filosofía de la Inteligencia Artificial*, compilación de Margaret Boden, México, Fondo de Cultura Económica, 1994.
- McCulloch, Warren S. Y Walter Pitts, A logical Calculus of the ideas immanent in nervous activity, *Bulletin of Mathematical Biophysics*, Numero 5, 1943 págs 115-133 reimpresso en *Neurocomputing*, J. Anderson y E. Rosenfeld (editores), The MIT-Press.
- Mensik, Ger-Jan y Jeroen G.W.Raaijmakers, A model for interference and forgetting *Psychological Review*, 1988, Vol 95, No. 4, págs 434-455.
- Minsky, M y S. Papert, *Perceptrons*, 1969, reimpresso *Neurocomputing*, J. Anderson y E. Rosenfeld (editores), The MIT-Press.
- Minsky, M, Logical vs. Analogical or Symbolic vs. Connectionist or Neat vs. Scruffy, *Artificial Intelligence at MIT. Expanding Frontiers*, Patrick H. Winston (editor), Vol 1, Massachusetts, MIT Press, 1990. Reimpresso en *AI Magazine*, 1991
- Moore, J W. Implementing connectionist Algorithms for Classical Conditioning in the brain, *Neural Network Models of Conditioning and Action*, Nueva York, Lawrence Erlbaum Associates Publishers, 1991, pp. 181-199.
- Olmsted P, *Brief History of Neural Networks*, McMillan, Nueva York, 1999
- Pagels, H. R. *Los sueños de la razón. El ordenador y los nuevos horizontes de las ciencias de la complejidad*, México, CO-NACyT/GEDISA, 1991.
- Parker, D.B. *Learning Logic* Technical Report 47, Center for Computational Research in Economics and Management Science, MIT, Cambridge, 1985, reimpresso *Neurocomputing*, J. Anderson y E. Rosenfeld (editores), The MIT-Press.
- Pedrycz, W y J Waletzky, Neural Network front Ends in Unsupervised Learning *IEEE Transactions on Neural Networks*, Vol 8, No 2, 1997, pp. 390-402.
- Ponce, A y J Burgos, *Simulación de la Generalización Pavloviana con redes neurales*, ponencia presentada en la 25 Conferencia Anual de la SQAB, Tornado Canadá, mayo 24 y 25 del 2002.
- Rashevsky, N, *Mathematical Biophysics*, University of Chicago Press, Chicago, IL, 1938
- Rosenblatt, F. The perceptron: a probabilistic model for information storage and organization in the brain, *Psychological Review* No. 65, 1958, págs 386-408, reimpresso *Neurocomputing*, J. Anderson y E. Rosenfeld (editores), The MIT-Press.
- Rummelhart D. y J. McClelland (editores) *Paralell distributed processing: Explorations in the microstructure of cognition*, Vol 1 Foundations, The MIT-Press, 1986.
- Schmajuck, N. A. *Animal Learning and Cognition. A Neural Network approach*, Cambridge University Press, 1997.
- Skapura, D. M. *Building Neural Networks*, Nueva York, Nueva York, ACM Press, 1995.
- Thompson, R. F., N. H. Donegan, G. A. Clark, D. G. Lavond, J. S. Lincoln, J. Madden IV, L. Mamounas, M D. Mauk y D A. McCormick, Neural substrates of discrete, defensive conditioned reflexes, conditioned fear states and their interactions in the rabbit, *Classical Conditioning*, 3rd Edition, Isidore Gormezano, William F. Prokasy y Richard F. Thompson (editores), Hillsdale Nueva Jersey, Lawrence Erlbaum Associates, publishers, 1987, pp. 371-399
- Turban, E, *Expert Systems and Applied Artificial Intelligence*, Nueva York, McMillan Publishing Company, 1992.
- Wasserman, E. A., S. R. Franklin y E. Hearst, Pavlovian appetitive contingencies and approach vs. Withdrawal to conditioned stimuli in pigeons, *Journal of Comparative and Psysiological Psychology*, Num 86, 1974, 616-627.
- Werbos, P, *Beyond Regression: New Tools for Prediction and Analysis in Behavioral Sciences*, Cambridge, Ma. 1974, citado por Hecht-Nielsen, *Neurocomputing*, Reading, Addison-Wesley Publishing Company, 1989