

NÍCHOLAS BENDER HAYDU

**UMA ABORDAGEM BASEADA EM SELEÇÃO PELAS CONSEQÜÊNCIAS
PARA APRENDIZAGEM DE REDES NEURAIIS MULTI-CAMADAS
VOLTADAS À CONCEPÇÃO DE SISTEMAS
AUTÔNOMOS INTELIGENTES**

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Informática pelo Curso de Pós-Graduação em Informática, do Setor de Ciências Exatas da Universidade Federal do Paraná, em convênio com o Departamento de Informática da Universidade Estadual de Maringá.

Orientador: Prof. Dr. Maurício F. Figueiredo

CURITIBA

2003



Ministério da Educação
Universidade Federal do Paraná
Mestrado em Informática

PARECER

Nós, abaixo assinados, membros da Banca Examinadora da defesa de Dissertação de Mestrado em Informática, da aluno *Nicholas Bender Haydu*, avaliamos o trabalho intitulado, "*Uma Abordagem Baseada em Seleção pelas Conseqüências para Aprendizagem de Redes Neurais Multi-Camadas Voltadas à Concepção de Sistemas Autônomos Inteligentes*", cuja defesa foi realizada no dia 15 de agosto de 2003, às dez e trinta horas, no Auditório da Informática da Universidade Federal do Paraná. Após a avaliação, decidimos pela aprovação do candidato. (Convênio número 279-00/UFPR de Pós-Graduação entre a UFPR e a UEM - ref. UEM número 1331/2000-UEM).

Curitiba, 15 de agosto de 2003.

Prof. Dr. Maurício Fernandes Figueiredo
DIN/UEM (Orientador)

Prof. Dr. Fernando José Von Zuben
FEED/UNICAMP – Membro Externo



Prof.^a Dra. Aurora Trinidad Ramirez Pozo
DINF/UFPR

*Ao espírito científico e ao desejo do
homem por saber mais.*

AGRADECIMENTOS

Ao meu orientador, Prof. Dr. Maurício Fernandes Figueiredo, por ter me guiado ao longo de mais essa etapa com dedicação, conhecimento e, principalmente, imaginação.

À minha família pelo apoio, incentivo e conversas inteligentes. Em especial para o meu pai, com sua imaginação fértil, e para minha mãe, pela ajuda teórica, dicas e conselhos extremamente úteis.

Aos professores do programa de Pós-Graduação, em especial ao Prof. Dr. Luís Augusto Consularo, pela ajuda e também pela versatilidade em conversar sobre diversos assuntos interessantes.

Aos meus amigos e colegas pela amizade, companheirismo e distração nos momentos de lazer.

Ao resto do mundo por não interferir de forma decisiva.

LISTA DE FIGURAS

Figura 1 – Esquema de neurônios e suas ligações.....	13
Figura 2 – Esquema da sinapse.....	14
Figura 3 – Modelo básico de neurônio artificial.....	14
Figura 4 – Função sigmóide.....	15
Figura 5 – Arquitetura de rede neural de uma única camada.....	16
Figura 6 – Arquitetura de rede neural de quatro camadas.....	17
Figura 7 – Arquitetura de rede neural recorrente.....	17
Figura 8 – Estruturas neurais de seleção pelas conseqüências.....	34
Figura 9 – Exemplo de condicionamento operante na rede de seleção pelas conseqüências.....	36
Figura 10 – Contingência respondente e operante.....	37
Figura 11 – Arquitetura de rede neural para quatro camadas.....	40
Figura 12 – Conexões sinápticas típicas e respectivas distribuições.....	41
Figura 13 – Representação da rede.....	50
Figura 14 – Propagação da atividade neural por conexões excitatórias inter-camadas (com cone de alcance mínimo).....	53
Figura 15 – Propagação da atividade neural por conexões excitatórias inter-camadas (sem cone de alcance mínimo).....	53
Figura 16 – Atividade neural com conexões laterais inibitórias.....	54
Figura 17 – Atividade neural com conexões laterais inibitórias e excitação predominante.....	55
Figura 18 – Atividade inicial ainda sem grupos neurais definidos.....	57
Figura 19 – Atividade intermediária com surgimento dos grupos neurais.....	57
Figura 20 – Atividade final com grupos neurais formados.....	58
Figura 21 – Atividade neural do estímulo CS anterior ao treinamento respondente.....	61
Figura 22 – Atividade neural do estímulo US anterior ao treinamento respondente.....	61
Figura 23 – Resposta do estímulo CS após treinamento respondente.....	61
Figura 24 – Atividade neural anterior à extinção do comportamento respondente.....	63
Figura 25 – Atividade neural após extinção do comportamento respondente.....	63
Figura 26 – Atividade neural anterior ao treinamento operante.....	64
Figura 27 – Atividade neural após treinamento operante.....	64
Figura 28 – Atividade neural após treinamento operante para um estímulo não reforçado.....	65
Figura 29 – Atividade neural anterior ao treinamento operante (exemplo 2).....	66
Figura 30 – Atividade neural após treinamento operante (exemplo 2).....	66
Figura 31 – Atividade neural anterior à extinção do comportamento operante.....	67
Figura 32 – Atividade neural após extinção do comportamento operante.....	67
Figura 33 – Atividade neural anterior ao treinamento operante (exemplo de generalização).....	68
Figura 34 – Atividade neural após treinamento operante com estímulo original.....	68
Figura 35 – Atividade neural após treinamento operante com degradação do estímulo (exemplo 1).....	68
Figura 36 – Atividade neural após treinamento operante com degradação do estímulo (exemplo 2).....	69
Figura 37 – Atividade neural após treinamento operante com degradação do estímulo (exemplo 3).....	69
Figura 38 – Atividade neural anterior ao treinamento de controle da intensidade da resposta (estímulo 1).....	70

Figura 39 – Atividade neural anterior ao treinamento de controle da intensidade da resposta (estímulo 2).	70
Figura 40 – Atividade neural após treinamento de controle da intensidade da resposta (estímulo 1).	71
Figura 41 – Atividade neural após treinamento de controle da intensidade da resposta (estímulo 2).	71
Figura 42 – Atividade neural anterior ao treinamento de controle de múltiplas respostas (estímulo 1 – resposta 1).	72
Figura 43 – Atividade neural anterior ao treinamento de controle de múltiplas respostas (estímulo 2 – resposta 2).	72
Figura 44 – Atividade neural após treinamento de controle de múltiplas respostas (estímulo 1 – resposta 1).	73
Figura 45 – Atividade neural após treinamento de controle de múltiplas respostas (estímulo 2 – resposta 2).	73
Figura 46 – Atividade neural anterior ao treinamento de fusão sensorial (estímulo discriminativo composto).	75
Figura 47 – Atividade neural anterior ao treinamento de fusão sensorial (componente 1).	75
Figura 48 – Atividade neural anterior ao treinamento de fusão sensorial (componente 2).	75
Figura 49 – Atividade neural após treinamento de fusão sensorial (componente 1).	76
Figura 50 – Atividade neural após treinamento de fusão sensorial (componente 2).	76
Figura 51 – Atividade neural após treinamento de fusão sensorial (ocorrência conjunta dos estímulo 1 e 2).	76
Figura 52 – Atividade neural anterior ao treinamento operante em rede com camadas 20x20.	78
Figura 53 – Atividade neural após treinamento operante em rede com camadas 20x20.	78
Figura 54 – Gráfico de evolução de um treinamento operante típico.	79
Figura 56 – Tela do simulador para ajuste de parâmetros da rede neural.	92
Figura 57 – Tela do simulador para configurar o ambiente de treinamento.	93

SUMÁRIO

1 INTRODUÇÃO.....	1
1.1 Motivação.....	1
1.2 Objetivos e Contribuições.....	7
1.3 Organização do Trabalho	9
2 FUNDAMENTAÇÃO TEÓRICA.....	10
2.1 Redes Neurais Artificiais	10
2.1.1 Modelo Básico Biológico	12
2.1.2 Fundamentos	14
2.1.3 Modelos Clássicos de Redes Neurais Artificiais	18
2.2 Aprendizagem e Comportamento Inteligente	23
2.2.1 Reflexos e Reflexos Condicionados.....	23
2.2.2 Comportamento Operante.....	25
2.3 Revisão Bibliográfica.....	29
2.3.1 Autonomia.....	29
2.3.2 Aprendizagem por Reforço e Inteligência Artificial.....	30
2.3.3 Aprendizagem por Reforço e Redes Neurais Artificiais.....	31
3 MODELO DE REDE NEURAL ARTIFICIAL.....	39
3.1 Arquitetura.....	39
3.2 Raciocínio.....	42
3.3 Aprendizagem	45
4 RESULTADOS	47
4.1 Experimentos Básicos.....	51
4.1.1 Atividade Neural sem Conexões Laterais	52

4.1.2 Atividade Neural com Conexões Laterais	54
4.1.3 Formação Dinâmica de Grupos Neurais	56
4.1.4 Considerações Sobre a Evolução da Atividade Neural	58
4.2 Experimentos de Aprendizagem.....	59
4.2.1 Experimento 1 – Convergência da Aprendizagem em Treinamento Respondente.....	60
4.2.2 Experimento 2 – Extinção de Comportamento Respondente.....	62
4.2.3 Experimento 3 – Convergência da Aprendizagem em Treinamento Operante	63
4.2.4 Experimento 4 – Extinção de Comportamento Operante.....	66
4.2.5 Experimento 5 – Capacidade de Generalização.....	67
4.2.6 Experimento 6 – Controle de Intensidade da Resposta	69
4.2.7 Experimento 7 – Controle de Múltiplas Respostas.....	71
4.2.8 Experimento 8 – Fusão de Sensores	73
4.2.9 Experimento 9 –Treinamento Operante em uma Rede com Camadas de Dimensão 20x20.	77
4.2.10 Experimento 10 – Análise da Aprendizagem Operante.....	78
5 CONCLUSÕES.....	80
REFERÊNCIAS BIBLIOGRÁFICAS.....	86
ANEXO	89
Anexo I – Ferramenta de Simulação.....	90

HAYDU, Nicholas Bender. **Uma Abordagem Baseada em Seleção pelas Conseqüências para Aprendizagem de Redes Neurais Multi-Camadas Voltadas à Concepção de Sistemas Autônomos Inteligentes**. 2003. Dissertação (Mestrado em Informática) – Universidade Federal do Paraná, Curitiba.

RESUMO

Um modelo de rede neural artificial é proposto. A rede neural possui múltiplas camadas. Cada camada da rede neural é formada por uma grade quadrangular de neurônios (em um espaço toroidal). As conexões sinápticas de cada neurônio abrangem três tipos: excitatórias inter-camadas, laterais inibitórias intra-camada e laterais excitatórias intra-camada. A disposição espacial das conexões é do tipo Gaussiana e específica para cada tipo de sinapse. Cada neurônio estabelece um número restrito de conexões. O modelo de arquitetura contribui para eliminar restrições apresentadas por arquiteturas em que entradas e conexões são distribuídas a todos os neurônios de cada camada. O modelo do neurônio apresenta dinâmica interna, proporcionando uma memória da atividade recente e assumindo papel importante na aprendizagem. A aprendizagem é baseada na seleção pelas conseqüências, conforme princípios de aprendizagem por reforço. Em particular, a de aprendizagem por reforço utilizada é do tipo clássico. Os experimentos definidos para investigação e confirmação das capacidades da rede neural consideram um ambiente simulado, condizente com o modelo de Seleção pelas Conseqüências. Os resultados obtidos em simulações mostram que o modelo é capaz de reproduzir diversos fenômenos comportamentais, que são: aquisição de comportamento respondente, extinção de comportamento respondente, aquisição de comportamento operante, extinção de comportamento operante, capacidade de generalização de estímulos, habilidade no controle da intensidade das respostas, capacidade de controle de múltiplas respostas e fusão de sensores. Experimentos também ilustram o importante papel das conexões laterais inibitórias e das conexões laterais excitatórias na modelagem da formação de grupos neurais em nível operante. Entende-se que a capacidade de aprendizagem alcançada pela rede neural proposta torna-a viável para a concepção de sistemas autônomos inteligentes com potencialidades superiores àqueles divulgados na literatura especializada.

HAYDU, Nicholas Bender. **A Selection by Consequences Approach for Learning in Multi-Layer Neural Networks Directed to the Conception of Intelligent Autonomous Systems**. 2003. Dissertação (Mestrado em Informática) – Universidade Federal do Paraná, Curitiba.

ABSTRACT

A model of an artificial neural network is proposed. The neural network has multiple layers. Each network layer is formed by a quadrangular grid of neurons (on a toroidal space). The synaptic connections that every neuron has are defined between three types: inter-layer excitatory, lateral intra-layer inhibitory and lateral intra-layer excitatory. The spatial disposition of connections is of a Gaussian type and specific for each type of synapse. Each neuron has a limited number of connections. The model contributes to eliminate restrictions presented by other architectures in which connections are distributed to all of the neurons of each layer. The neuron model presents an internal dynamic, working as a memory of its recent activity and having important role in the learning process. The learning procedure is based on the selection by consequences according to reinforcement learning principles. Particularly the reinforcement learning approach used is of the classical type. The experiments defined for the investigation and confirmation of the capacities of the neural network consider a simulated environment that works according to the Selection by Consequences model. The simulation results show that the model is capable of reproducing several behavioral phenomena that are: acquisition of respondent behavior, extinction of respondent behavior, acquisition of operant behavior, extinction of operant behavior, stimulus generalization capacity, ability to control the response intensity, capacity to control multiple responses and sensor fusion. Besides that, the experiments also illustrate the important role of the lateral inhibitory and lateral excitatory connections for a correct shaping of operant level responses and neural groups. It is understood that the learning capacities that the proposed neural network exhibits make it viable for the conception of intelligent autonomous systems with potentialities superior to those already presented in the specialized literature.

1 INTRODUÇÃO

1.1 *Motivação*

Autonomia

Os progressos no campo da Inteligência Artificial vêm demarcando grandes ciclos, caso o ponto de vista adotado leve em conta as metas, expectativas e resultados alcançados.

Entende-se que um primeiro ciclo estaria associado aos esforços de pesquisa motivados pela possibilidade de reproduzir capacidades inteligentes humanas sofisticadas em máquinas computacionais. Após se constatar as imensas disparidades presentes entre as potencialidades das máquinas desenvolvidas e das anunciadas a partir das expectativas iniciais, este ciclo se extingue. Os resultados da pesquisa neste ciclo podem não ter sido satisfatórios, entretanto serviram para gerar uma noção mais realista das dificuldades gigantes que seriam enfrentadas para reproduzir um sistema artificial inteligente imaginado.

As frustrações segundo a perspectiva mais ficcionista não impediu que os progressos e resultados alcançados merecessem o reconhecimento por todo o esforço despendido. As potencialidades dos sistemas inteligentes propostos poderiam não satisfazer as metas iniciais, mas se tinham tornado suficientes para produzirem resultados relevantes em aplicações práticas. Esta nova leitura para o papel dos resultados alcançados pela pesquisa em Inteligência Artificial caracterizaria um segundo ciclo. A pesquisa no campo da Inteligência Artificial se consolidou e foi reconhecida como extremamente relevante para o progresso tecnológico, invadindo áreas que antes se mostravam avessas, e.g., controle automático. Neste ciclo os sistemas inteligentes passam a assumir uma importância crescente junto ao segmento tecnológico. O segundo ciclo desencadeou avanços teóricos notáveis e talvez seja aceitável admitir que, enquanto o primeiro ciclo esteve motivado por expectativas calcadas na ficção, o segundo foi fortemente guiado pelas demandas presentes no segmento tecnológico.

É possível que a pesquisa no campo da Inteligência Artificial esteja iniciando um terceiro ciclo. As demandas tecnológicas continuam exercendo

pressões constantes no que diz respeito às demandas cada vez mais sofisticadas e exigentes em desempenho. Entretanto também se nota uma forte preocupação dos pesquisadores em tornar os sistemas inteligentes artificiais mais semelhantes aos sistemas biológicos sofisticadamente inteligentes. Observa-se que, em vista das potencialidades demonstradas pelos sistemas biológicos e artificiais, as semelhanças que podem ser admitidas são pouco significativas. Desta forma, argumenta-se que a insatisfação presente seja motivo suficiente para incrementar pesquisas e reduzir as diferenças entre estes dois sistemas.

Assim, atualmente, de um lado observa-se a consolidação e origem de teorias ousadas, tais como: sistemas imunológicos artificiais (de Castro e Timmis, 2002), sistemas de inteligência coletiva (Bonabeau, Dorigo e Theraulaz, 1999), sistemas classificadores (Lanzi, 2000), e da evolução de teorias mais tradicionais relacionadas às redes neurais, sistemas nebulosos, computação evolutiva, aprendizagem artificial (Vapnik, 1999)(Sutton e Barto, 1998), sistemas simbólicos inteligentes bem representados pelos agentes inteligentes (Russel e Norvig, 1995).

De outro lado ressurgem as reflexões em termos das limitações dos atuais sistemas inteligentes. As críticas podem ser resumidas no problema do “a priori”, abordado por vários autores (Brooks, 1990)(Brooks, 1991)(Figueiredo, 1999) e em especial por Verschure (Verschure, 1993)(Verschure, 1996).

O problema do “a priori” diz respeito ao fato de um sistema simbólico requerer uma representação do mundo em termos de símbolos discretos para operar, que só pode ser fornecida pelos projetistas do sistema. Mesmo que o sistema possua a capacidade de aprender por meio da incorporação de novos símbolos e regras à sua base de conhecimentos, é necessário um conjunto básico inicial que permita a operação do sistema. Esse conjunto inicial é definido, portanto, “a priori” e requer dos projetistas a definição da essência do sistema. A partir deste problema principal Verschure reuniu quatro outros também de grande importância, que serão vistos resumidamente a seguir:

- 1. Problema da Falta de Fundamento Simbólico (Symbol Grounding Problem):** está relacionado à questão do valor verdade de uma proposição ser derivado de outras proposições, criando uma teia de relações que, no entanto, não possuem base com relação ao mundo real supostamente descrito pelo sistema;

- 2. Problema do Enquadramento (Frame Problem):** aborda a questão do tempo necessário para a atualização do modelo lógico que cresceria exponencialmente e logo tornaria impossível a atuação do sistema, ficando este confinado a rotinas de atualização;
- 3. Problemas da Visão de Referência (Frame-of-reference Problem):** trata do fator subjetivo quando da escolha de um modelo lógico para a representação de um domínio. Cada pessoa pode escolher um modelo lógico e uma forma diferente de relacionar os símbolos do sistema lógico com o problema que se deseja abordar. No entanto, nada garante que essa escolha seja a melhor;
- 4. Problema da Contextualização (Situatedness Problem):** sistemas inteligentes deveriam ser contextualizados, isto é, deveriam levar em consideração apenas fatores de importância imediata, caso contrário sofreriam do problema do enquadramento. Entretanto os sistemas existentes abordam o conhecimento de uma maneira global e que ainda depende de interpretação humana.

Uma interpretação simplificada (no entanto poderosa) do problema do "a priori" bem como dos problemas subseqüentes pode ser feita da seguinte forma: um sistema simbólico capaz de exibir inteligência necessita que em sua base existam símbolos representando verdades absolutas e atômicas sobre a inteligência, a partir dos quais pudessem ser deduzidas todas as implicações necessárias para um sistema inteligente. Como muito bem abordado por Penrose (1997), tais verdades absolutas provavelmente escondem-se em locais como o mundo subatômico entre outros. Assim sendo, extraí-las seria como extrair os átomos de conhecimentos necessários para deduzir todo o universo.

Tendo em vista as dificuldades existentes no contexto apresentado é possível justificar a crescente atenção voltada para o conceito de autonomia no contexto de sistemas artificiais inteligentes.

Autonomia e inteligência têm sido considerados conceitos fortemente relacionados (Steels, 1995). Propõe-se que a autonomia seja uma

capacidade que se manifeste segundo uma escala contínua, proporcionando, então, uma medida para o grau de inteligência de um sistema artificial (Figueiredo, 1999). Além disso, sistemas artificiais, nos quais a ausência de autonomia tenha sido constatada, tornam-se passíveis de críticas severas no que tange a sua natureza inteligente (Vershure, 1996).

Por sua vez, a autonomia é vinculada à capacidade de aprendizagem, particularmente quando a aquisição de conhecimento ocorre a partir da interação com o ambiente (Steels, 1995) (Figueiredo, 1997) (Figueiredo, 1999) (Russell&Norvig, 1995). Um sistema autônomo é capaz de aprender através da interação com o ambiente, recebendo estímulos e gerando comportamentos (respostas) que alteram este ambiente, aprimorando suas capacidades e adaptando-se às características do mesmo. Tanto a aprendizagem quanto os comportamentos e ou respostas geradas podem ser independentes de auxílios externos.

Assim, dentre as diversas possibilidades de aplicação de sistemas inteligentes, a presença de autonomia torna-se particularmente interessante naquelas em que o sistema atua em ambientes desconhecidos, e.g., navegação autônoma de robôs. Tais áreas de aplicação têm sido intensamente investigadas, não só devido ao claro interesse tecnológico, mas também por proporcionar um ambiente profícuo para a pesquisa dos sistemas autônomos inteligentes.

Redes Neurais

É entre os sistemas biológicos que exclusivamente se encontram os mais representativos sistemas autônomos inteligentes. Reconhece-se que o poder de autonomia de tais sistemas advém preponderantemente de seus sistemas nervosos. Por esta razão, alguns pesquisadores defendem que a melhor estratégia para a concepção de tais sistemas inteligentes é concentrar esforços na modelagem dos sistemas biológicos (Edelman, 1987).

Coerentemente com esta argumentação, as redes neurais artificiais são adotadas neste trabalho, dentre as diversas técnicas presentes na área de Inteligência Artificial, visto que são inspiradas em seus pares biológicos. Particularmente, o presente trabalho adota o princípio da máxima plausibilidade biológica para as propostas apresentadas, ou seja, as trajetórias definidas durante o

desenvolvimento do trabalho estão associadas às escolhas que favorecessem o modelo de RNA que melhor espelhasse o modelo biológico.

As redes neurais artificiais (RNAs) geram de forma qualitativa muitas das características de seus pares biológicos, além da capacidade de aprendizagem, essencial para a autonomia: tolerância a falhas, generalização, memória acessada por conteúdo e robustez ao ruído.

Embora os modelos de redes neurais existentes têm demonstrado tais características em variadas áreas de aplicação, muitas deficiências gerais podem ser apontadas, dentre elas: limitações de capacidade da memória associativa e recuperação de seu conteúdo, restrições quanto à convergência em aprendizagem supervisionada, limitações quanto à aprendizagem em ambientes desconhecidos. Esta última deficiência está intimamente relacionada ao foco de estudo do presente trabalho, pois reduzindo-a é possível alcançar potencialidades satisfatórias para a concepção de sistemas autônomos.

Aprendizagem

Sendo a aprendizagem algo fundamental para se alcançar a autonomia, é desejável que o modelo possua uma estratégia para realizar tal tarefa, que seja coerente com as idéias propostas.

Nesse contexto utiliza-se o modelo da Seleção pelas Conseqüências (Skinner, 1981) como base para um princípio de aprendizagem. A aprendizagem segundo essa visão ocorre a partir de variações no repertório de comportamentos dos organismos, as quais por sua vez são selecionados pelo ambiente com base nas conseqüências que produzem. No caso de produzir conseqüências favoráveis ao organismo, o comportamento é selecionado e passa a fazer parte de seu repertório. Caso produza conseqüências desfavoráveis a tendência é que o comportamento seja extinguido, deixando de existir. A variabilidade de comportamentos e a forma específica como estes são selecionados ocorre segundo três estratégias básicas, produzindo três tipos de comportamentos: comportamento reflexo, comportamento respondente e comportamento operante.

A primeira estratégia aborda os comportamentos do tipo reflexo que são filogeneticamente selecionados, ou seja, segundo sua capacidade ou não de permitir ao organismo transmitir seus genes com sucesso. É razoável assumir que a

variabilidade para este tipo de comportamento é introduzida por fatores genéticos como mutações e cruzamentos.

O segundo tipo de comportamento é conhecido pelo termo respondente e trata de associar estímulos quaisquer (S), presentes no ambiente, aos comportamentos do tipo reflexo. Para tanto, o processo de seleção depende da existência de uma correlação temporal entre estímulos S com os estímulos responsáveis por eliciar o comportamento reflexo. A seleção é mediada por um estímulo reforçador que é normalmente o próprio eliciador (estímulo que inicia) do reflexo. Note que comportamentos do tipo reflexo são selecionados segundo a história de vida do organismo e não mais ao longo de inúmeras gerações.

O último tipo de aprendizagem diz respeito ao assim denominado comportamento operante. Comportamentos operantes são selecionados segundo a consequência que suas respostas produzem no ambiente e não apenas de acordo com uma correlação temporal, como é o caso dos respondentes. No entanto, assim como o comportamento respondente, o comportamento operante também é selecionado através da apresentação, por parte do ambiente, de um estímulo reforçador. Assim a consequência que a emissão da resposta produz no ambiente é a apresentação do estímulo reforçador. O efeito do reforçador neste caso é selecionar as respostas emitidas pelo organismo em situações adequadas, tornando-as mais prováveis no futuro. Também no caso de comportamentos operante o repertório é desenvolvido ao longo da história de vida do organismo.

Devido ao fato da aprendizagem ser mediada por um estímulo reforçador, este princípio de aprendizagem é denominado de aprendizagem por reforço. Existem, no entanto, diversas teorias consideradas como modelos de aprendizagem por reforço. Para este trabalho, a não ser quando indicado o contrário, a interpretação deve ser de acordo com o modelo de Seleção pelas Conseqüências.

Diferentemente de outras estratégias, a estratégia de aprendizagem por reforço oferece mecanismos para a adaptação do sistema autônomo enquanto este interage com um ambiente desconhecido. Por esta razão torna-se muito útil para o treinamento de tais sistemas, em especial para aplicações associadas a controle autônomo. Assim, em todos os instantes em que o sistema se depara com eventos discrepantes (situações desconhecidas, situações de sucesso ou falha); o sistema recebe um sinal de avaliação (reforço) associado ao seu desempenho, sinal

esse no qual é baseado o processo de aprendizagem (Donahoe & Palmer, 1994) (Skinner, 1981).

1.2 Objetivos e Contribuições

Um sistema artificial capaz de operar de forma satisfatória em ambientes do cotidiano requer características comportamentais complexas que seriam comumente denominadas de inteligentes. As capacidades de tal sistema devem ser as mais diversas possíveis, por exemplo, a habilidade de navegar de forma autônoma por um ambiente em constante modificação, o uso de ferramentas em situações específicas, a coordenação e cooperação com outros sistemas artificiais e também com seres humanos. Com o objetivo de dotar sistemas artificiais com tais capacidades, são investigados os processos fundamentais que proporcionam o surgimento de tais comportamentos em sistemas biológicos, para então buscar reproduzi-los em um sistema artificial.

O objetivo geral é desenvolver um sistema artificial inteligente capaz de adquirir conhecimento sobre o ambiente no qual este se encontra e utilizar tal conhecimento para agir no ambiente de forma autônoma, ou seja, um sistema autônomo inteligente. Entretanto tal meta deve ser alcançada seguindo-se alguns requisitos específicos, discutidos ao longo do texto (Capítulo 2), que são:

- Utilizar um sistema baseado em um modelo de redes neurais;
- A arquitetura neural deve buscar reproduzir tanto estruturalmente quanto funcionalmente sistemas biológicos nervosos;
- A aprendizagem do sistema deve seguir um modelo compatível com o modelo de Seleção pelas Conseqüências (Skinner, 1981);
- O sistema artificial deve ser capaz de reproduzir fenômenos comportamentais de que trata a teoria da Seleção pelas Conseqüências (Skinner, 1981).

O emprego de tais requisitos, como fica claro ao longo do trabalho, não se deve pelo fato de que se deseja provar a relevância ou não de cada um deles para o problema sugerido, mas sim por se acreditar que fornecem propostas concretas e importantes para a solução do mesmo.

Em termos práticos o trabalho propõe um modelo neural cuja arquitetura, operação e aprendizagem funcionam segundo os requisitos apresentados. A fundamentação do trabalho é feita com trabalhos específicos de cada área, bem como propostas que integram as mesmas, destaca-se neste ponto o trabalho de Donahoe, Burgos e Palmer (1993).

A principal contribuição do trabalho será mostrar que uma estratégia de aprendizagem por reforço pode ser empregada com sucesso no treinamento de redes multicamadas de diferentes dimensões. O modelo de aprendizagem por sua vez, atuando juntamente com a operação da rede é capaz de reproduzir, sinteticamente, fenômenos comportamentais observados em seres biológicos e que são normalmente tidos como indicadores de autonomia e inteligência. De forma secundária, mas não menos importante, o modelo apresenta princípios para a formação dinâmica de grupos neurais e a estabilização da atividade neural a partir de um estado inicial variante.

Especificamente, os resultados mostram que o modelo é capaz de reproduzir diversos fenômenos comportamentais particularmente interessantes para um sistema autônomo inteligente. São eles:

- Aquisição de comportamento respondente;
- Extinção de comportamento respondente;
- Aquisição de comportamento operante;
- Extinção de comportamento operante;
- Capacidade de generalização de estímulos;
- Habilidade de controle de intensidade da resposta emitida;
- Capacidade de controle de múltiplas respostas;
- Fusão de sensores.

Além dos fenômenos de aprendizagem, o modelo também apresenta resultados decorrentes da arquitetura neural empregada. Particularmente mostra-se a importância das conexões inibitórias para uma correta modelagem do comportamento de nível operante (Seção 2.2). Além disso, em conjunto com as conexões laterais excitatórias, as conexões inibitórias são responsáveis pela

formação dos grupos neurais. Os grupos neurais por sua vez formam a base para a análise da operação da rede.

Do ponto de vista das redes neurais artificiais o modelo contribui apresentando uma solução para o problema de atribuição de crédito que utiliza aprendizagem por reforço.

1.3 Organização do Trabalho

O Capítulo 2 trata de revisar os temas, conceitos e fundamentos que nos quais o trabalho é baseado. Primeiramente tratando de redes neurais artificiais (Seção 2.1), apresentando um breve histórico, discutindo os principais conceitos e modelos. Logo após discute-se o modelo de aprendizagem de seleção pelas conseqüências (Seção 2.2), como visto pela psicologia. Em seguida a Seção 2.3 traz uma revisão de conceitos como autonomia (Seção 2.3.1), aprendizagem por reforço segundo a visão atual e mais difundida da IA (Seção 2.3.2), aprendizagem por reforço e redes neurais (Seção 2.3.3) onde se busca um modelo que integre redes neurais e o modelo de Seleção pelas Conseqüências.

No Capítulo 3 descreve-se o modelo de rede neural proposto neste trabalho, apresentando sua arquitetura (Seção 3.1), os mecanismos de raciocínio (Seção 3.2) e o algoritmo de aprendizagem (Seção 3.3).

O Capítulo 4 apresenta os resultados obtidos em simulação do modelo proposto. Os resultados apresentados são de dois tipos: operacional (Seção 4.1) e de aprendizagem (Seção 4.2).

Ao final do trabalho são apresentadas as conclusões, seguida pelas referências e pelos anexos.

2 FUNDAMENTAÇÃO TEÓRICA

Devido às questões envolvidas, trabalhos com sistemas autônomos possuem a característica de serem multidisciplinares. Isso torna ainda mais importante uma revisão dos assuntos abordados.

Esta seção apresenta de forma resumida as principais teorias e conceitos necessários para um bom entendimento da proposta deste trabalho, entre elas, Redes Neurais Artificiais (Seção 2.1) e teoria da Seleção por Conseqüências (Seção 2.2). A Seção 2.3 aborda vários fundamentos e traz uma revisão bibliográfica de diversos trabalhos que investigam temas similares ao deste trabalho, fornecendo uma perspectiva atual da pesquisa na área.

2.1 Redes Neurais Artificiais

As redes neurais têm sido destaque em diversas pesquisas de sistemas autônomos. Autores como Brooks (1991), Verschure (1996) e Verschure (1993) destacam sua contribuição para o chamado problema do “a priori”. Tal questão, como destacado na introdução deste trabalho, aborda o fato de que é extremamente difícil definir a priori regras gerais capazes de descrever problemas como, por exemplo, o de navegação autônoma, nos quais o robô precisa deslocar-se de forma autônoma pelo ambiente, basicamente desviando de obstáculos e buscando alvos, além de outros objetivos mais complexos. Problemas como este requerem que o sistema adquira o conhecimento para a solução do problema diretamente através da interação com o ambiente, fazendo uso, portanto, de aprendizagem. É justamente na capacidade de aprendizagem que reside o poder das redes neurais.

Além da capacidade intrínseca de aprendizagem as redes neurais possuem outras características interessantes para a concepção de sistemas autônomos, entre as quais destacam-se:

- **Não-linearidade:** o modelo dos neurônios pode ser linear ou não-linear. Ser não-linear concede ao neurônio, e conseqüentemente à rede, a capacidade de processamento não-linear. Esta

capacidade é desejada sempre que o sistema deve adotar determinadas estratégias de interação com o ambiente de acordo com as variações dos estímulos (a estratégia para um estímulo mais intenso é diferente da estratégia para um menos intenso);

- **Processamento paralelo:** a atividade simultânea dos diversos elementos de uma rede neural faz com que essa seja capaz de processar grande quantidade de informação de forma veloz, algo indispensável para sistemas que devem operar em tempo real;
- **Memória acessada por conteúdo:** a forma particular com que uma rede neural armazena informação, e o próprio mecanismo de operação fazem com que esta seja um sistema de memória acessada por conteúdo. Assim a rede neural é capaz de recuperar um padrão previamente armazenado, mediante a apresentação de parte desse padrão;
- **Tolerância a falhas:** por se tratar de um sistema maciçamente paralelizado, cada elemento da rede é responsável apenas por uma pequena parte da computação realizada. Assim, a remoção ou falha de elementos isolados impacta de forma suave no desempenho da rede;
- **Generalização:** vários modelos de redes neurais têm por objetivo extrair características dos estímulos aos quais são apresentados. Devido ao seu modo de operação, mesmo quando um estímulo desconhecido for apresentado, a rede neural classifica segundo sua similaridade aos estímulos com os quais possui experiência. Essa é uma qualidade importante do comportamento dos seres vivos e reconhecidamente interpretada indispensável em sistemas inteligentes;
- **Robustez a ruídos:** uma rede neural possui grande robustez a ruídos contaminando os estímulos de interesse, conseguindo operar normalmente com desempenho satisfatório. Essa característica é muito útil, pois na prática, sensores inevitavelmente estão associados a ruído.

Este capítulo discorre de maneira breve a respeito dos principais tópicos relacionados às redes neurais, visando oferecer uma visão da área e alguns fundamentos que facilitem a compreensão dos capítulos seguintes. Textos mais aprofundados no tema podem ser encontrados em obras de referência como Haykin (2001) e Russel e Norvig (1995).

2.1.1 Modelo Básico Biológico

Uma rede neural é constituída por neurônios (Figura 1), as unidades básicas de computação, que interconectados formam a rede propriamente dita. Os neurônios operam recebendo os estímulos em suas sinapses (Figura 2), agregando tais estímulos no soma e quando o valor resultante da agregação ultrapassa um certo valor limiar, o neurônio emite (dispara) uma seqüência de pulsos elétricos, gerados no cone do axônio. O sinal elétrico emitido flui pelo axônio até atingir sua extremidade onde, através de trocas químicas nas sinapses, é transmitido aos próximos neurônios. Este esquema relativamente simples é repetido desde a entrada da rede até sua saída, tal qual acontece em algumas redes biológicas, por exemplo, em redes cujas entradas correspondem aos campos sensoriais (olhos, ouvidos, etc.) e aos órgãos motores (braços e músculos em geral) e/ou demais órgãos.

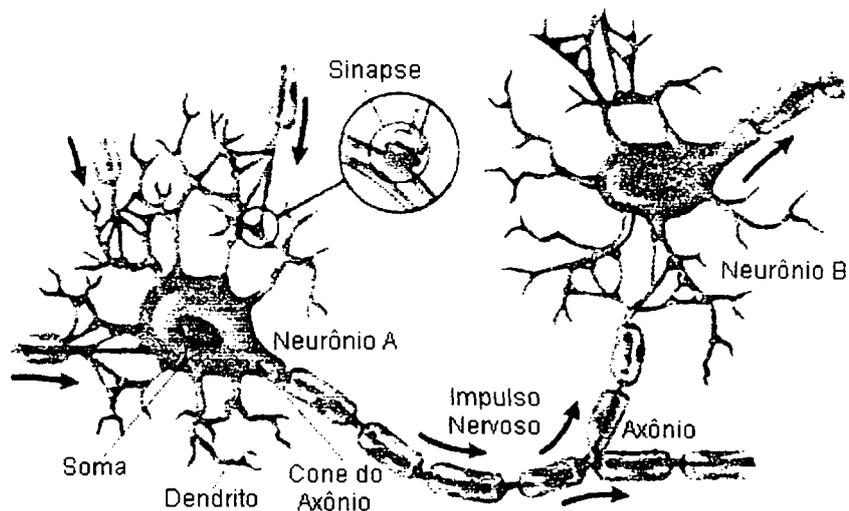


Figura 1 – Esquema de neurônios e suas ligações.¹

As sinapses (Figura 2) consistem de uma região de intercâmbio de informação, mais precisamente, são os locais onde os neurônios recebem e emitem neurotransmissores (substâncias químicas utilizadas para transmitir a informação) convertendo, portanto, um sinal elétrico (disparo do neurônio) em químico (neurotransmissores) e vice-versa. Através de processos complexos as sinapses podem alterar sua eficiência, tornando-se mais ou menos eficazes na propagação do estímulo. Dessa forma, cada sinapse pode alterar a intensidade do estímulo que recebe e passa para frente, intensificando-o ou suprimindo-o, sendo, portanto, a unidade básica de memória, onde a informação pode ser armazenada e manipulada. Esta característica das sinapses é tida como a principal responsável pela capacidade de aprendizagem das redes neurais.

Uma rede neural pode ser descrita como possuindo uma arquitetura, um mecanismo de processamento e sistema de aprendizado. A arquitetura de uma rede neural descreve como estão conectados os neurônios. A forma pela qual o sinal de entrada evolui (é processado) para um sinal de saída será tratada como o raciocínio². Já o processo responsável por modificar as sinapses é responsável pela aprendizagem. Operando de forma integrada e nas mais diversas configurações possíveis, estes elementos são capazes de processar e produzir resultados

¹ Adaptado de: <http://research.haifa.ac.il/~jmjaffe/mind/lecnotes/week7.html>. junho 2003

² A termo “raciocínio” segundo essa interpretação refere-se aos processos internos de uma rede neural, desencadeados por um estímulo, que evoluem até produzir uma resposta. De forma simplificado é o processamento da informação.

absolutamente não triviais. Os melhores exemplos são os próprios organismos biológicos.

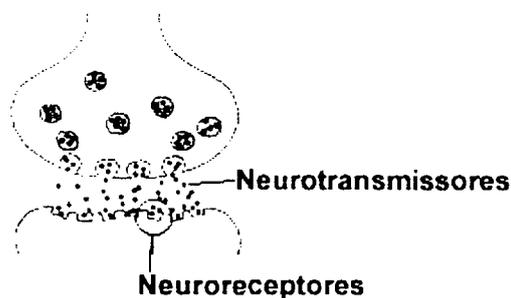


Figura 2 – Esquema da sinapse.¹

2.1.2 Fundamentos

Modelos Clássicos

O primeiro passo para o desenvolvimento da teoria de redes neurais artificiais surgiu em 1943 com o trabalho de McCulloch e Pitts que propuseram uma descrição matemática para um neurônio (Haykin, 2001). A Figura 3 apresenta um modelo atual de um neurônio artificial simples que embora mais recente retém as linhas gerais originais.

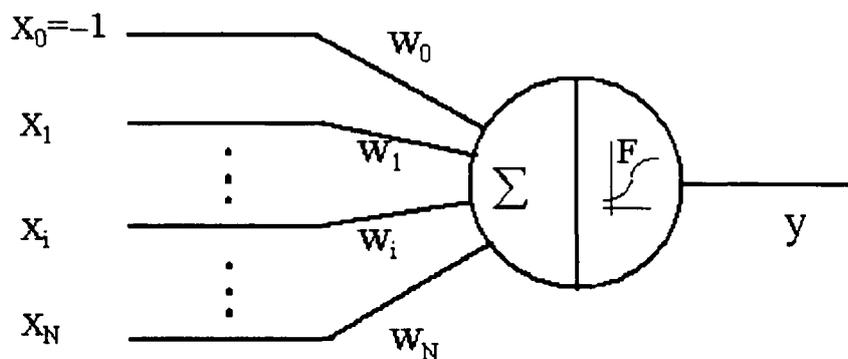


Figura 3 – Modelo básico de neurônio artificial.

¹ Adaptado de: <http://research.haifa.ac.il/~jmjaffe/mind/lecnotes/week7.html>, junho 2003.

$$y = F\left(\sum_{i=1}^N x_i w_i - w_0\right) \quad (1)$$

Neste modelo os sinais de entrada x_i ; $i = 1, \dots, N$; ao chegarem ao neurônio são ponderados pelos pesos sinápticos w_i ; $i = 1, \dots, N$; $x_i, w_i \in \mathfrak{R}$ (conjunto dos números reais). Estes são, então, enviados ao soma através dos dendritos onde são adicionados. Em seguida uma transformação não linear $F(\cdot)$ (função de ativação do neurônio) age sobre o sinal resultante da agregação gerando o sinal de saída y conforme a Equação 1. A função $F(\cdot)$ por sua vez é tipicamente qualquer função do tipo sigmóide, por exemplo (Figura 4):

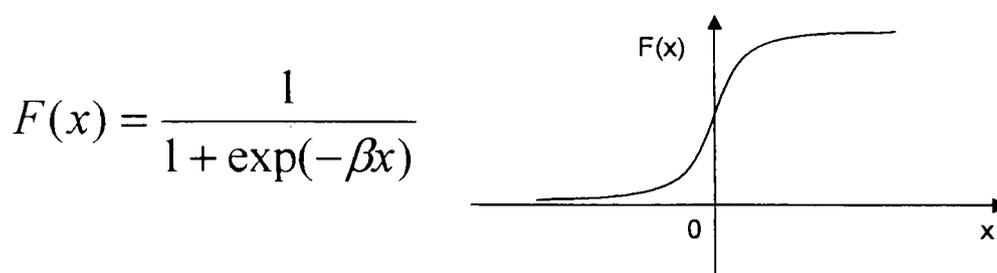


Figura 4 – Função sigmóide.

Outro importante avanço para as RNAs foi a descrição feita por Hebb em 1949 de um mecanismo fisiológico de modificação sináptica. Sua regra de aprendizagem prevê que a eficiência de uma sinapse deva ser aumentada quando há atividade em ambos os neurônios que a formam (neurônio pré e o pós-sináptico) e que tal atividade tenha se propagado através dessa sinapse (Haykin, 2001). Em sua forma mais simples, tal regra pode ser expressa matematicamente segundo a Equação 2.

$$\Delta w_{ji}(T) = \eta a_j(T) a_i(T) \quad (2)$$

onde: $\Delta w_{ji}(T)$ é o ajuste da sinapse no tempo T ; η é a taxa de aprendizagem; $a_j(T)$ é a atividade (sinal de saída) no neurônio pré-sináptico j no tempo T e $a_i(T)$ é a atividade do neurônio pós-sináptico i na iteração T .

Arquiteturas Básicas

A arquitetura de uma rede neural consiste de um conjunto de neurônios e suas conexões. Há várias combinações possíveis para estes elementos o que por sua vez dá origem a várias topologias possíveis. Entretanto, três configurações são básicas. A primeira delas é a rede alimentada adiante de uma única camada (Figura 5).

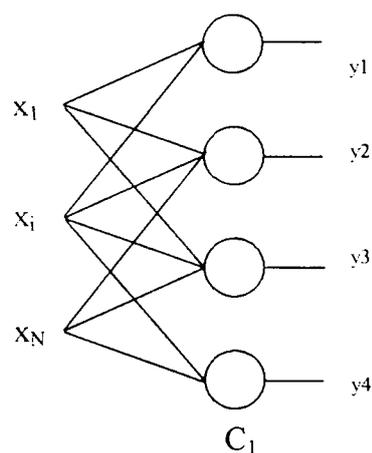


Figura 5 – Arquitetura de rede neural de uma única camada.

A segunda arquitetura, uma das mais difundidas, é a multicamadas, nas quais os neurônios são conectados por conjuntos de neurônios (camada). Neste tipo de topologia, as camadas podem ser numeradas naturalmente desde a entrada da rede (x_i) até sua saída (y), de tal forma que as conexões ocorrem somente entre camadas sucessivas. A Figura 6 ilustra um exemplo de uma rede neural com este tipo arquitetura, de fato, uma arquitetura de 4 camadas: C_1 , C_2 , C_3 e C_4 .

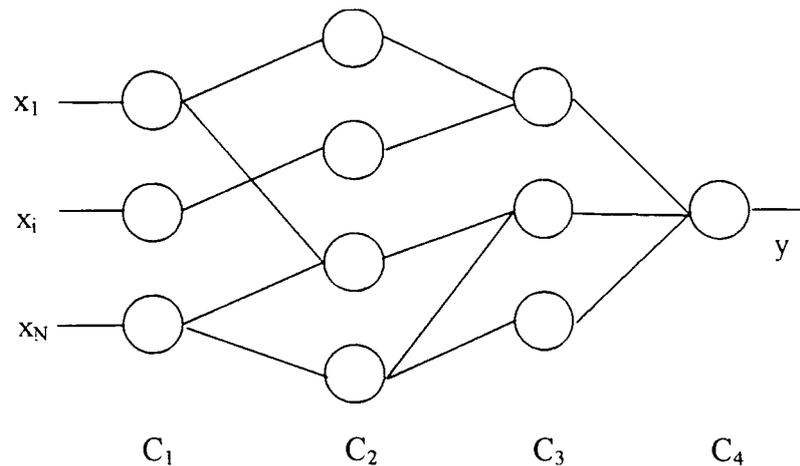


Figura 6 – Arquitetura de rede neural de quatro camadas.

Outro esquema de arquitetura bem conhecido é o recorrente. São consideradas recorrentes as redes que possuam conexões provenientes de um neurônio que possam, direta (conexão vermelha) ou indiretamente (conexão azul), influenciar a entrada do mesmo (Figura 7).

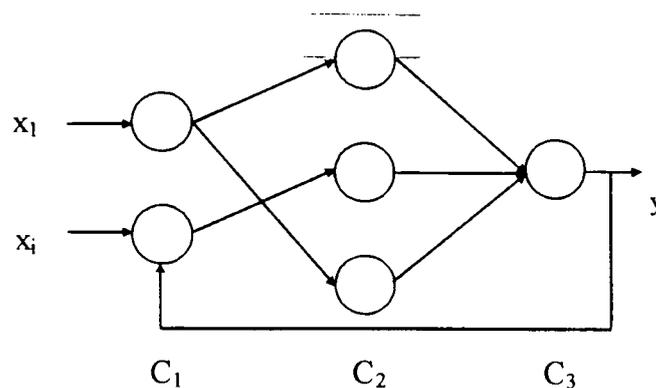


Figura 7 – Arquitetura de rede neural recorrente.

Em contraste com as redes de arquitetura em camadas, a saída da rede recorrente evolui ao longo do tempo indefinidamente (a menos que a rede esteja em um estado estável) mesmo com ausência de estímulos.

Estratégias de Aprendizagem

Para a concepção de redes neurais existem três principais estratégias de aprendizagem, a saber: não-supervisionada, supervisionada e por

reforço. Cada uma destas estratégias está associada a classes específicas de sinais, por meio das quais a aprendizagem se processa: (a) apenas entradas, (b) pares entrada-saída; e (c) entradas e sinais de avaliação (emitidos pelo ambiente), respectivamente.

A estratégia de aprendizagem não-supervisionada diz respeito somente a entradas, ou seja, o sistema aprendiz se auto-organiza de acordo com as características das entradas. Desde que aplicações de controle são associadas com um mapeamento entrada-saída, a estratégia de aprendizagem não-supervisionada não é adequada, a menos que juntamente com esta seja considerada outra estratégia.

A aprendizagem supervisionada depende do conhecimento do problema, pois é necessário um conjunto de pares de entrada-saída, representantes da solução desejada para o problema, para que a aprendizagem se processe. Esta estratégia pode ser adequada para diversas aplicações, mas não associadas a controle autônomo. Dificuldades surgem se o ambiente altera suas características, tornando inválidos os pares de entrada-saída usados durante o período de aprendizagem.

Na estratégia por reforço, entra em jogo um outro sinal (sinal de reforço) apresentado pelo ambiente. O ambiente apresenta o sinal somente após a rede ter apresentado sua saída. Este sinal corresponde a uma avaliação da rede, feita pelo ambiente, com respeito à resposta apresentada e é utilizado no processo de aprendizagem da rede. Na verdade o que ocorre é que as respostas (ações) executadas pela rede modificam o ambiente, uma vez modificado o ambiente pode conter estímulos reforçadores, que se tornam disponíveis para os elementos sensoriais da rede. Não existe, portanto, nenhum módulo de avaliação ou qualquer outro dispositivo similar que funciona como um avaliador. Vale lembrar que a aprendizagem por reforço está diretamente ligada à teoria de Seleção por Conseqüências (Donahoe e Palmer, 1994), tema da Seção 2.2.

2.1.3 Modelos Clássicos de Redes Neurais Artificiais

Uma introdução sobre redes neurais não estaria completa sem alguns exemplos de modelos de redes que aplicam os conceitos de arquitetura,

raciocínio e aprendizagem apresentados anteriormente. De fato, devido à sua ampla divulgação, tais exemplos são muitas vezes tidos como sinônimos dos próprios conceitos que aplicam.

Perceptron de Múltiplas Camadas

Exemplo típico de uma rede de várias camadas, o perceptron de múltiplas camadas (MLP, *multilayer perceptron*) tornou-se famoso principalmente devido ao algoritmo de aprendizagem por correção de erro utilizado em seu treinamento. Através de um processo conhecido por retropropagação de erro, o MLP tornou-se um dos primeiros modelos de RNA a possuir um método eficiente de treinamento para redes de múltiplas camadas. O algoritmo de retropropagação tornou-se famoso pela publicação de Rumelhart e McClelland (1985), *Parallel Distributed Processing*. O algoritmo também foi responsável por reacender o ânimo na pesquisa de RNA ao demonstrar que as redes de múltiplas camadas não eram tão limitadas como haviam sugerido Minsky e Papert (1969).

Utilizando uma estratégia supervisionada, o algoritmo de retropropagação consiste basicamente de dois passos:

- 1. Passo para frente:** um padrão de entrada é aplicado à camada sensorial da rede e a atividade da mesma se desenvolve (propaga) a partir destes, produzindo ao final um vetor de respostas na camada de saída;
- 2. Passo para trás:** a saída computada pela rede é comparada com a saída desejada (estratégia supervisionada) e um sinal de erro é produzido de acordo com a diferença entre os dois. O sinal de erro é então utilizado para ajustar os pesos sinápticos da rede em um processo de retropropagação (do final para o começo da rede) de forma a minimizar o erro, aproximando a saída computada da saída desejada. O processo de retropropagação envolve o cálculo da parcela de contribuição de cada neurônio para gerar a saída da rede, tornando possível o ajuste (através de derivadas parciais) de cada sinapse segundo sua contribuição específica.

O perceptron de múltiplas camadas utilizado em conjunto com o algoritmo de retropropagação é comumente empregado, entre outras, em tarefas de classificação de padrões e aproximação de funções. No entanto o fato de utilizar uma aprendizagem supervisionada consiste em um grave ponto negativo para sua aplicação em sistemas autônomos, tendo em vista que seu treinamento seria impossível pelo fato do ambiente de operação não ser totalmente conhecido.

Embora os algoritmos de correção de erro (incluindo o de retropropagação) aparentemente tenham sido originados de trabalhos matemáticos, sem qualquer inspiração biológica, existe forte evidência de que pelo menos uma área do cérebro opere de forma semelhante. O cerebelo está envolvido no controle de precisão de movimentos e é responsável, entre outros, por produzir movimentos suaves e bem coordenados. Evidências neurobiológicas (Rolls e Treves, 1998) (Rolls, 1999) demonstram que o cerebelo recebe realimentação neural dos sistemas motores que controla e que utiliza esse sinal (na prática um sinal de erro) para corrigir eventuais imprecisões.

Mapa de Kohonen

Quando se fala de sistemas auto-organizáveis (aprendizagem não-supervisionada) a referência principal é o modelo de Kohonen (Kohonen, 1982). Dentre suas aplicações, talvez aquela de maior destaque seja sua utilização como ferramenta para modelagem e compreensão de mapas corticais no cérebro. Estes resultados remetem aos objetivos iniciais que levaram ao desenvolvimento de toda a teoria de redes neurais e justificam grande parte do entusiasmo em relação ao modelo.

Tipicamente o Mapa de Kohonen consiste de uma grade uni ou bidimensional de elementos computacionais (neurônios) cujos pesos sinápticos estão, inicialmente, distribuídos de forma aleatória sobre o espaço de entrada. Durante a fase de treinamento, são selecionadas ao acaso elementos do espaço de entrada e estes são então apresentados à rede. A partir de então, para cada apresentação ocorrem três processos distintos:

1. **Competição:** os neurônios da grade competem entre si pelo direito de responder ao sinal de entrada apresentado, através do

cálculo de uma função discriminante. O neurônio cujo vetor de pesos for mais similar (ex: menor distância Euclidiana) ao vetor de entrada é o vencedor. Este processo é conhecido como o “vencedor-leva-tudo” ou *winner-take-all*;

2. **Cooperação:** o neurônio vencedor estabelece uma vizinhança topológica dentro da qual os neurônios possuem o direito de ajustar suas sinapses;
3. **Adaptação Sináptica:** os neurônios localizados dentro da vizinhança estabelecida no passo 2 têm suas sinapses ajustadas de modo a melhorar o valor de sua função discriminante. Na prática, isso resulta em aproximar o vetor de pesos do vetor de entrada. Tal ajuste é ponderado segundo a proximidade do neurônio em questão com o neurônio vencedor. Estes ajustes melhoram a eficácia da resposta dos neurônios ao estímulo.

Ao longo de várias iterações este processo irá “posicionar” os pesos sinápticos sobre o espaço de entradas, de forma que os respectivos neurônios correspondem a um domínio ou sub conjunto de característica particular do espaço de entrada (Kohonen, 1992). Operando dessa forma, o Mapa de Kohonen realiza uma redução de dimensionalidade no espaço de entrada, extraindo suas características principais e assemelhando-se tanto em forma como em função aos primeiros estágios da maquinaria cerebral, responsável pelo processamento sensorial.

Embora o Mapa de Kohonen ilumine o caminho para a construção de máquinas semelhantes ao cérebro e conseqüentemente ao desenvolvimento de sistemas autônomos, ele consiste de apenas uma parte do sistema necessário. A redução de dimensionalidade e extração de características proporcionada pelo modelo serve para tornar o problema da percepção algo tratável. No entanto, ainda faz-se necessário um mecanismo que associe as características identificadas com ações a serem tomadas.

Rede de Hopfield

A rede de Hopfield é um sistema dinâmico, ou seja, possui uma representação na forma de um sistema de equações diferenciais; sendo, portanto um sistema em que entradas provocam a evolução de estados internos e suas saídas ao longo do tempo. Algumas características deste sistema atraem a atenção dos pesquisadores, inclusive fazendo com que contribuíssem fortemente com a revitalização das redes neurais artificiais (Haykin, 2001) (Hopfield, 1982):

1. O sistema é estável no sentido de Liapunov (Haykin, 2001);
2. O sistema exhibe plausibilidade biológica, ou seja, o sistema de equações pode ser deduzido de modelos (simplificados) de neurônios e de redes neurais biológicas;
3. Existem procedimentos simples para configuração dos parâmetros do sistema, visando definir as características do seu espaço de fases. Estes procedimentos podem ser interpretados como o processo de aprendizagem.

A rede neural de Hopfield exhibe interessantes resultados em duas aplicações principais: na modelagem de uma memória associativa e na solução de problemas de otimização. Os instigantes resultados alcançados nestas aplicações podem ser compreendidos a partir de suas características.

Considere a aplicação de modelagem de uma memória associativa. O espaço de fase do sistema representado pela rede de Hopfield apresenta diversas bacias de atração. Supondo que cada padrão a ser registrado pela memória seja um ponto crítico estável no espaço de fase, então este seria alcançado sempre que o estado inicial estivesse na bacia associada ao padrão memorizado. Os procedimentos de configuração de parâmetros possuem o objetivo de fazer com que os padrões memorizados sejam exatamente os pontos críticos estáveis do sistema.

A estabilidade da rede de Hopfield é garantida por uma função de Liapunov associada. Problemas de otimização são solucionados pela rede caso sua função de Liapunov possa descrever a função a ser otimizada. Neste caso, iguala-se a função a ser otimizada à função de Liapunov associada à rede. A manipulação dessa igualdade resulta em um conjunto de pesos sinápticos para a rede neural. A

evolução da rede neural, configurada com tais pesos, leva o sistema a estabilizar-se em estados correspondentes à solução do problema de otimização.

2.2 Aprendizagem e Comportamento Inteligente

Conforme visto, a aprendizagem é de essencial importância para a autonomia. É, portanto, desejável que o modelo de aprendizagem empregado utilize uma estratégia que seja coerente com as idéias propostas.

De acordo com o modelo de Seleção pelas Conseqüências (Skinner, 1981) a aprendizagem ocorre a partir de variações no repertório de comportamentos dos organismos. Os comportamentos produzidos por tais variações desencadeiam modificações no ambiente, que atuam de forma a selecionar os mesmos. No caso de produzir conseqüências favoráveis ao organismo, o comportamento é retido e passa a fazer parte do repertório de ação. Caso produza conseqüências desfavoráveis, a tendência é que o comportamento seja extinguido.

Visando reunir conceitos fundamentais relacionados ao modelo de Seleção pela Conseqüências, o presente capítulo dedica-se ao estudo de alguns de seus aspectos segundo a visão da Análise do Comportamento, área de pesquisa da psicologia. Segundo o ponto de vista dessa área, aprendizagem e comportamento são totalmente dependentes um do outro (Millenson, 1967) e, portanto, torna-se possível investigar a aprendizagem estudando-se o comportamento.

2.2.1 Reflexos e Reflexos Condicionados

O tipo de comportamento mais simples exibido por um ser vivo é o reflexo. Este tipo de comportamento é definido filogeneticamente, ou seja, já nasce com o organismo e é resultado de uma história de seleção pelo ambiente ao longo de várias gerações (seleção natural). Um reflexo é uma contingência simples, envolvendo um estímulo específico e uma resposta específica. Diz-se que a presença do estímulo elicia a resposta, ou seja, desencadeia um processo cuja conseqüência é a emissão da resposta. Exemplos desse tipo de comportamento são: o reflexo patelar, o reflexo pupilar e o reflexo salivar.

Embora de extrema importância para os seres vivos em seus primeiros momentos de existência, os reflexos são insuficientes para formar um repertório comportamental adequado para toda a sua vida. Eles são frutos de uma história de seleção genética ao longo da evolução e, por este motivo, limitam-se a comportamentos essenciais relacionados à manutenção da espécie e sobrevivência dos indivíduos.

Para a formação de um repertório comportamental adequado à diversidade ambiental, os seres vivos exibem outros mecanismos de aprendizagem capazes de gerar comportamentos adicionais (além do comportamento reflexo), ou seja, capazes de construir um repertório mais amplo e que permita ao ser vivo uma melhor adaptação ao meio. Um destes mecanismos é chamado condicionamento respondente (ou Pavloviano, em homenagem ao pesquisador, Ivan Pavlov, que primeiro descreveu cientificamente tal fenômeno (Millenson, 1967)). Condicionamento neste contexto, e ao longo do restante do texto, se refere ao mecanismo responsável por formar o comportamento. A principal característica destes mecanismos é a capacidade de estabelecer gradualmente associações entre estímulos e respostas, sendo um deles parte componente de um comportamento reflexo. Tais associações ocorrem quando existe uma correlação temporal adequada entre a ocorrência dos estímulos. Após várias ocorrências de um estímulo neutro (que inicialmente não está associado a nenhuma resposta) seguido do estímulo que dispara o reflexo, o estímulo neutro torna-se capaz de produzir o comportamento reflexo (passando a ser chamado de estímulo condicionado). Desta forma, a resposta reflexa original, filogeneticamente especificada, passa a estar condicionada a um novo estímulo, sendo esta nova relação estímulo-resposta conhecida como reflexo condicionado (Millenson, 1967).

Um exemplo típico de um condicionamento respondente é a associação que se estabelece, por exemplo, entre a visão de um alimento e o reflexo salivar. O reflexo salivar é filogeneticamente relacionado à mucosa bucal, quando do contato desta com o alimento. No entanto, é comum a ocorrência deste reflexo mediante a simples visão do alimento, algo que é consequência de um comportamento respondente que se estabelece.

De forma semelhante aos reflexos, o condicionamento respondente é consequência de uma seleção. A diferença é que neste caso a seleção é realizada ao longo da vida do organismo e se desenvolve através de ocorrências repetidas e

correlacionadas (emparelhamento) de um estímulo neutro (que virá a se tornar o estímulo condicionado) e o estímulo eliciador (responsável pela ocorrência do comportamento reflexo incondicional).

O condicionamento respondente é responsável por fornecer aos organismos uma forma de adaptar seu repertório comportamental frente a um ambiente dinâmico, visivelmente dotando-os de capacidades autônomas. No entanto, um comportamento respondente ainda está limitado às mesmas respostas reflexas com as quais o indivíduo nasceu e, conseqüentemente, não produz novos comportamentos complexos.

2.2.2 Comportamento Operante

Uma forma de aprendizagem adicional e relevante para a adaptação dos organismos ao ambiente é o chamado condicionamento operante. No caso de comportamentos do tipo reflexo e reflexo condicionado, estímulos do ambiente eliciam respostas do organismo (respostas reflexas e respondentes, respectivamente) sendo que as conseqüências dessas respostas no ambiente não influem nas próprias respostas ou comportamentos correspondentes. Já o condicionamento operante se desenvolve por meio do efeito da ação do organismo sobre o ambiente. Quando as conseqüências de um comportamento são favoráveis ao desempenho, torna-se maior a probabilidade de que este comportamento ocorra novamente no futuro. O comportamento operante é formado pela seleção das ações do organismo para interagir com o ambiente. Essa seleção é feita através da apresentação de um reforçador gerado pelo ambiente. Assim, sempre que o organismo emitir a resposta desejada, o reforçador associado é apresentado. Esta relação causal tende a selecionar a resposta em questão. Este processo é denominado de fortalecimento operante (Skinner, 1981) (Millenson, 1976) e caracteriza-se pelo fato do comportamento específico desejado ser gradualmente selecionado, pela apresentação de um estímulo reforçador, em meio a todo o repertório de respostas que o organismo apresenta.

Um exemplo de fortalecimento operante ocorre quando um animal de estimação, ao realizar um truque específico, recebe de seu dono uma recompensa. Isso faz com que o animal tenda a repetir o truque no futuro.

Um aspecto que faz parte importante do mecanismo de condicionamento operante reside na capacidade de gerar uma gama extremamente grande de ações possíveis. Segundo essa teoria, os sistemas biológicos apresentam um nível operante caracterizado por um conjunto indiferenciado de respostas aleatórias emitidas pelo organismo, sem qualquer expectativa sobre suas conseqüências associadas (Millenson, 1967).

Exemplos de respostas em nível operante são: o conjunto de fonemas que recém nascidos são capazes de balbuciar; e os movimentos indiferenciados de braços e pernas. É a partir destas respostas simples, inicialmente aleatórias, que comportamentos complexos são selecionados.

O condicionamento operante não só possibilita o aumento na freqüência da resposta, mas também a sua redução. Quando um comportamento produz no ambiente conseqüências aversivas, a probabilidade de que ele ocorra novamente é reduzida (Skinner, 1981).

Assim, quando o mesmo animal de estimação entrar em casa e mastigar algum objeto de seu dono, este comportamento provavelmente produzirá uma reação do dono desagradável. Dessa maneira, a probabilidade de ocorrência do comportamento indesejado é reduzida, sendo esse procedimento denominado de punição. É importante ressaltar que o termo reforço é usado de forma abrangente para designar tanto o processo de fortalecimento quanto o de redução. Existe, no entanto, uma terminologia mais específica para identificar diferentes formas (procedimentos) de fortalecer e enfraquecer comportamentos. Assim, tem-se o reforço positivo e negativo e a punição positiva e negativa. A descrição destas classes de procedimentos está fora do escopo deste texto e pode ser encontrada em Baum (1999).

Uma outra forma de reduzir a probabilidade de ocorrência de respostas operantes é a extinção. A extinção ocorre quando um comportamento anteriormente reforçado passa a não produzir mais a conseqüência reforçadora (Skinner, 1981), (Millenson, 1967). Voltando ao exemplo do animal de estimação, considere que este tenha aprendido a executar um truque sempre quando do retorno de seu dono, sempre recebendo uma recompensa como reforço. No entanto, considere que as recompensas passam a serem negadas. Não existindo recompensas, gradualmente o truque deixa de ser reproduzido.

Embora didaticamente estudados em separado, os comportamentos reflexo, respondente e operante não estão sujeitos a esse tipo de separação. Assim sendo, estes comportamentos ocorrem paralelamente e seus mecanismos são interdependentes. Suponha, por exemplo, o comportamento de pressionar uma alavanca, reforçado com comida. Embora esse comportamento seja um operante, inevitavelmente algum estímulo (e.g. o som da alavanca) torna-se presente no momento que a cobaia obtiver o alimento, o que fornece as condições necessárias para o condicionamento do tipo respondente.

Os processos de condicionamento respondente e operante atuam em conjunto e em vários níveis, produzindo comportamentos complexos. Assim, os movimentos indiferenciados são gradualmente modificados pelas conseqüências, passando a se caracterizar como ações coordenadas que permitem apanhar objetos, caminhar etc. As vocalizações são diferenciadas de tal forma que balbucios acabam se tornando palavras e finalmente em um complexo repertório verbal. As palavras e ações em geral não ocorrem, no entanto, de forma aleatória. Elas tendem a ter uma relação precisa com eventos antecedentes. Estes eventos são chamados de estímulos discriminativos (S^D). Estímulos deste tipo permitem ao ser vivo distinguir a situação na qual a resposta adequada (R), produz o estímulo reforçador (S^+).

Um esquema geral de como funciona o paradigma operante é o seguinte:

$$S^D \rightarrow R \rightarrow S^+$$

Inicialmente os estímulos que têm função de reforço ou de punição são reduzidos em número e são conhecidos como reforçadores primários. Sabor doce, calor (conforto térmico), dor, sabor amargo são exemplos de reforçadores primários, isto é, têm esta propriedade filogeneticamente determinada. Outros eventos do ambiente podem adquirir as características de um reforçador, ou seja, a capacidade de fortalecer o comportamento por meio do emparelhamento com reforçadores primários. Um estímulo deste tipo é chamado de reforçador adquirido ou reforçador condicional. É exatamente neste ponto que os condicionamentos respondente e operante estão relacionados, já que o condicionamento respondente

faz justamente com que se estabeleçam relações entre estímulos inicialmente neutros e estímulos reforçadores primários. Talvez o melhor exemplo deste tipo de estímulo seja o dinheiro que é altamente reforçador para muitos indivíduos de diversas sociedades, mas insignificante para outros, demonstrando que sua capacidade reforçadora é adquirida e que isso depende da história de vida de cada indivíduo.

Reforçadores adquiridos estão envolvidos em um fenômeno conhecido como encadeamento de resposta, em que as conseqüências de um determinado comportamento exibido pelo organismo, são estímulos que indicam a ocasião para um outro comportamento. Voltando ao exemplo anterior, trabalha-se para se obter dinheiro apenas para que então seja possível, entre outras coisas, comprar alimento. Dessa forma, estabelece-se uma cadeia de estímulos e respostas que no caso do exemplo fornecido é: trabalho⇒dinheiro⇒compras⇒alimento. Atuando desse modo, o fenômeno de encadeamento é capaz de formar cadeias comportamentais extremamente complexas e conseqüentemente tem grande participação na construção de repertórios comportamentais complexos.

Uma conseqüência importante do condicionamento operante é o fato deste aumentar a estereotipia da resposta, ou seja, o padrão de respostas que compõe o comportamento tende a se tornar bem específico e peculiar. Considere que inicialmente o organismo emite um conjunto variado de respostas, ou seja, esteja em nível operante. À medida que esta ou aquela resposta em particular for reforçada, a probabilidade de sua ocorrência aumenta enquanto a probabilidade das demais respostas, comparativamente, diminui. Além disso, a resposta em questão passa a ocorrer de uma maneira particularmente específica. Por exemplo, considere uma cobaia que aprende a pressionar uma alavanca para obter alimento. Inicialmente ela pressiona a alavanca ao acaso e de diversas formas diferentes, isto é, com variações na topografia da resposta. À medida que o procedimento de reforço é realizado a cobaia passa a pressionar a alavanca com uma freqüência cada vez maior e restringindo a topografia da mesma (passando a executar a ação de forma bem específica, por exemplo, apenas com a pata esquerda).

2.3 Revisão Bibliográfica

2.3.1 Autonomia

Brevemente introduzido no primeiro capítulo, o conceito de autonomia é fundamental para este trabalho e define a essência do que se entende por um sistema artificial inteligente. Entretanto a interpretação adotada de tal conceito não é consenso entre pesquisadores de Inteligência Artificial, inclusive existindo controvérsias sobre seu entendimento, motivo pelo qual tal tema é aprofundado nesta seção.

Um dicionário, (Fernandes et al., 1991), traz o termo autônomo como sendo algo regido por leis próprias, um sistema independente. No caso específico de robôs autônomos, esse termo possui um significado preciso, como foi destacado por Russell e Norvig (1995) e Figueiredo (1999). Segundo os autores, um robô é autônomo quando possui a capacidade de alterar seu comportamento, baseando-se em suas próprias experiências e com o objetivo de aprimorar seu desempenho. Esta definição é perfeitamente aplicável a sistemas autônomos em geral; e é semelhante à abordagem de autores como Dorigo (1996), Krose (1995), bem como Steels (1995).

Autonomia, assim definida, introduz diversos pontos importantes a serem analisados. Uma das primeiras implicações é que para "alterar seu comportamento, baseando-se em suas próprias experiências" é necessário que o sistema possua capacidade de aprendizagem, ou seja, que ele seja capaz de alterar seu comportamento (autogoverno) de forma a maximizar seu desempenho.

De outra forma, também é comum na literatura a utilização do termo autônomo para referir-se a robôs que simplesmente não precisam de auxílio externo, sem que eles possuam necessariamente capacidades de aprendizagem. Robôs que não dependem de auxílio externo podem prescindir da capacidade de aprendizagem em casos nos quais o ambiente e demais fatores permitam. Entretanto, neste caso, suas ações seriam repetitivas e/ou totalmente programadas. Segundo o ponto de vista adotado, tais robôs são denominados de automáticos (Steels, 1995).

2.3.2 Aprendizagem por Reforço e Inteligência Artificial

As teorias sobre aprendizagem descritas são de grande interesse para pesquisadores de Inteligência Artificial, particularmente aquela sobre comportamento operante. Com o objetivo de fornecer um panorama geral da pesquisa de aprendizagem por reforço apresenta-se a seguir uma síntese da questão, bem como de duas técnicas bastante conhecidas da área: *Temporal Differences* e *Q-Learning*. Uma abordagem bem mais aprofundada pode ser encontrada no trabalho de Sutton e Barto (Sutton e Barto, 1988).

De forma geral, o modelo formal visto na IA sobre a aprendizagem por reforço consiste de:

1. Um número discreto de estados ambientais;
2. Um número discreto de ações possíveis;
3. Um conjunto de sinais de reforço (e.g. $\{0, 1\}$);
4. Uma política de ações.

Assim, o objetivo do sistema é encontrar uma política de ações que mapeie as entradas do sistema (estados ambientais) para as saídas (ações possíveis) de modo a maximizar os sinais de reforço obtidos.

O problema da aprendizagem também pode ser visto como um problema de previsão. Deste ponto de vista, o sistema artificial inteligente toma uma determinada ação com base em uma previsão das conseqüências (reforço) dessa ação. Para escolher a melhor ação é necessário algum tipo de avaliação das possibilidades. Uma das técnicas utilizada para realizar tal avaliação é chamada *Temporal Differences* (TD) (Sutton e Barto, 1988). Enquanto a maioria dos métodos avalia o desempenho do sistema com base na diferença do resultado previsto e daquele realmente obtido, o método TD atribui crédito às possibilidades a partir de uma seqüência de previsões sucessivas no tempo. A título de ilustração, tome-se o exemplo extraído de (Sutton e Barto, 1988). Imagine uma previsão do tempo feita durante toda a semana sobre a possibilidade de chover no sábado. Enquanto a maioria das técnicas compara a previsão feita em cada dia da semana com o resultado observado no sábado, usando um valor obtido dessa comparação para realizar a aprendizagem do sistema; o método TD utiliza uma comparação

incremental em que a previsão de cada dia influi na do dia seguinte. Assim, uma previsão de, por exemplo, 75% de probabilidade de chuva na terça, influi na probabilidade de que chova na quarta.

A técnica de TD fundamentalmente visa avaliar os custos de uma determinada política de ações (Ribeiro, 2002). No entanto, ela não aborda a questão da aprendizagem dessas políticas. Uma técnica que aborda tal aprendizagem, inclusive bastante utilizada no desenvolvimento de robôs autônomos, é a chamada Q-learning. Basicamente a aprendizagem consiste de um método iterativo que se baseia em avaliar a ação (Q) segundo seu valor ($Q(x,a)$) e que é definido como o desconto obtido no custo ao adotar-se a ação "a", quando no estado "x", seguindo uma política ideal de ação (Ribeiro, 2002).

Obviamente, esses assuntos são extensos e abordá-los mais a fundo foge ao escopo deste trabalho. O objetivo aqui é somente oferecer uma breve síntese sobre a visão da IA sobre comportamento operante e, com isso, denunciar as relações entre as teorias da Análise do Comportamento e a da Inteligência Artificial.

2.3.3 Aprendizagem por Reforço e Redes Neurais Artificiais

Em se tratando de aprendizagem por reforço aplicada a redes neurais, existem basicamente dois tipos de abordagens possíveis (Haykin, 2001):

1. **Clássica:** a aprendizagem acontece através de sinais de recompensas e punições, ou seja, segundo a teoria da Seleção por Conseqüências;
2. **Moderna:** trata de uma técnica matemática que visa prever e avaliar as possíveis ações para então escolher qual delas é a melhor a ser tomada, enfatizando, portanto, o planejamento.

A Seção anterior (2.3.2) fornece um panorama geral da teoria e dos trabalhos que utilizam técnicas que se enquadram em uma abordagem moderna. Tal revisão é interessante para construir um embasamento teórico que favorece a discussão e possibilita a comparação de resultados. No entanto, o presente trabalho

segue uma linha diferente e suas propostas enquadram-se em um panorama clássico.

Embora a teoria moderna, assim como a teoria clássica, tenha se originado de estudos em psicologia, em certa altura sua formulação desviou-se, deixando de lado sua inspiração, evoluindo em direção a solucionar problemas específicos da IA. Por outro lado, pesquisas experimentais no contexto da aprendizagem clássica demonstram que a aprendizagem por reforço ocorre em organismos para os quais não existe evidência alguma sugerindo o planejamento das ações futuras (Brooks, 1990) (Brooks, 1991). Dessa forma impossibilita-se explicar os comportamentos complexos exibidos por tais organismos com base na teoria moderna. Pesquisas como essas levam a crer que a teoria original possui potencial para produzir resultados concretos. Além disso, sua utilização favorece todo o processo de aprendizagem à medida em que torna mais intuitivo o processo de interação do sistema com o ambiente.

Apresenta-se a seguir um modelo de rede neural que utiliza uma abordagem clássica para sua aprendizagem por reforço, servindo de base para o desenvolvimento das propostas do Capítulo 3.

Fundamentos Neurobiológicos

O modelo proposto neste trabalho baseia-se em conceitos característicos derivados da neurobiologia. Devido à natureza pouco comum desse tema, pelo menos no âmbito da Inteligência Artificial, primeiramente estes conceitos serão introduzidos para na sequência ser apresentado o modelo.

Sabe-se que o cérebro dos animais, incluindo o humano, possui regiões nas quais a atividade elétrica é desencadeada quando estímulos específicos são apresentados. Sabe-se também que tais estímulos são reforçadores primários e que o circuito neural envolvido é capaz de operar logo após o nascimento sendo, portanto, filogeneticamente determinado. Uma dessas regiões, como destacam autores como Donahoe e Rolls, é a *Ventral Tegmental Area* (Área Ventral Tegmentar ou *VTA*) (Figura 8) (Donahoe, Burgos e Palmer, 1993) (Rolls, 1999).

A partir da *VTA*, projetam-se, de forma distribuída, várias ramificações em direção ao córtex motor e lóbulos frontais, que são do tipo dopaminérgico, isto é, provocam a liberação de dopamina quando estimuladas. A

dopamina é um neuromodulador capaz de fortalecer as sinapses entre neurônios pré e pós-sinápticos, imediatamente após a ativação da sinapse em questão. Estímulos reforçadores primários possuem capacidade filogeneticamente pré-determinada de iniciar atividade na *VTA*, assim a *VTA* é capaz de mediar relações entre ações geradas no córtex motor.

Além dos caminhos neurais filogeneticamente definidos, capazes de iniciar atividade na região *VTA* a partir de reforçadores primários, alguns dos neurônios do córtex frontal associativo também são capazes de estimular a *VTA* pois possuem axônios que se projetam, através do "*medial forebrain bundle*" (*MFB*) (Figura 8), até a própria *VTA*. Tal estrutura possibilita o desenvolvimento de reforçadores adquiridos, tendo em vista que estímulos discriminativos, cuja informação viaja até os lóbulos frontais através dos córtices sensoriais, podem adquirir, por meio de realimentação pelo *MFB*, a capacidade de estimular o sistema de fortalecimento dopaminérgico. Concluindo, a Figura 8 ilustra, portanto, um mecanismo neural capaz de ajustar conexões sinápticas com base em reforçadores primários, cujos estímulos viajam por caminhos neurais filogeneticamente selecionados, e reforçadores adquiridos, cujos estímulos viajam através da região conhecida como *MFB*.

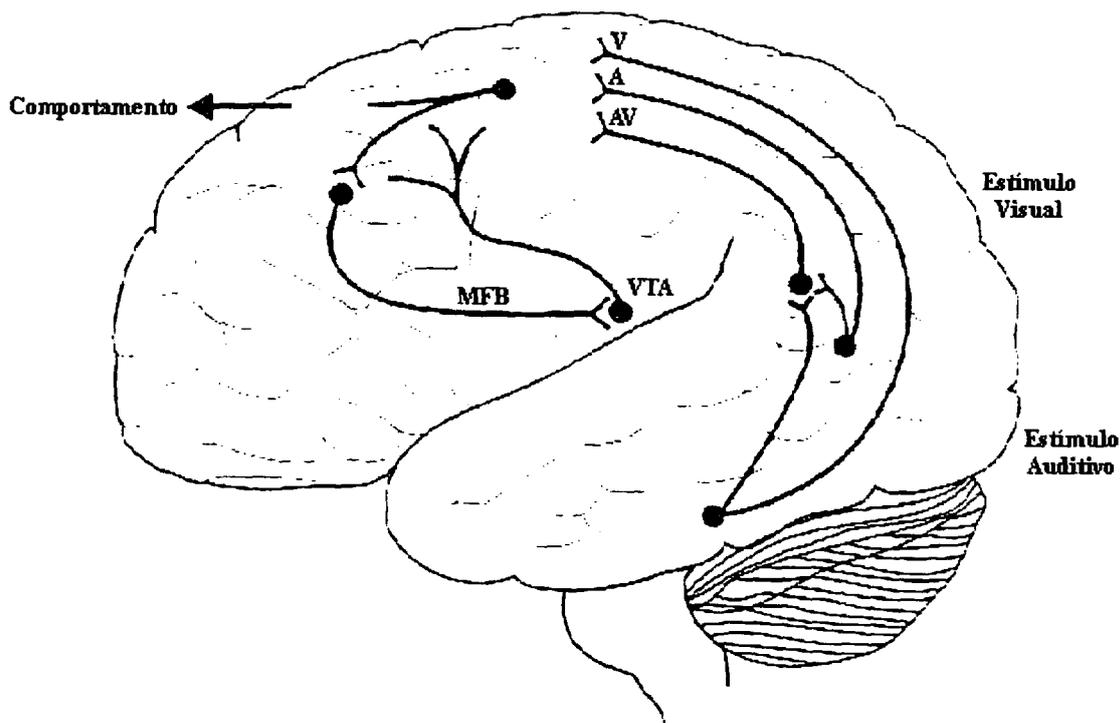


Figura 8 – Estruturas neurais de seleção pelas conseqüências¹.

É interessante notar que grande parte da teoria da Seleção pelas Conseqüências, desenvolvida a partir de observações comportamentais feitas em animais, possui uma correspondência com estudos da neurobiologia. Tudo indica que tais correspondências tendem a aumentar conforme novas descobertas forem sendo feitas em cada área.

Modelo de Redes Neurais Artificiais

O modelo proposto por Donahoe, Burgos & Palmer (1993) busca espelhar os mecanismos neurais que acabam de ser apresentados (no tópico anterior), utilizando-se técnicas de RNAs. O objetivo é reproduzir fenômenos comportamentais de que trata a teoria de Seleção por Conseqüências, fornecendo ao mesmo tempo um modelo neural e uma explicação biológica correspondente.

Neste modelo, a questão da aprendizagem é tratada sob o ponto de vista do princípio unificado do reforço, ou seja, considerando que os condicionamentos respondente e operante são tratados em conjunto sob as mesmas

¹ Adaptado de: Donahoe e Palmer (1994).

regras. Entretanto, as conseqüências de ambos os tipos de aprendizagem permanecem distintas. Tal como enfatizado anteriormente, essa divisão entre condicionamento respondente e operante é mais uma questão didática do que prática e a abordagem do modelo ajuda a consolidar tal proposta.

Uma simulação do sistema é ilustrada na Figura 9. O exemplo refere-se ao mecanismo de aprendizagem baseado no condicionamento operante, tendo em vista que a apresentação do estímulo US (*unconditioned stimulus* ou reforçador primário) é dependente da emissão da resposta R. Considere que o ambiente não tenha provocado qualquer reforço sobre o indivíduo. Desta forma, as respostas emitidas pelo mesmo ocorrem de forma equiprovável, em conseqüência do estado de nível operante em que se encontra. Considere que a resposta R tenha sido emitida a partir do estímulo S1. A resposta R altera o ambiente que, por sua vez, gera um estímulo US. O estímulo US dispara a resposta UR (*unconditioned response* ou resposta reflexa). A resposta UR por sua vez está ligada ao sistema difusor de dopamina e, portanto, provoca um reforço de todas as conexões sinápticas recentemente utilizadas. Assim, ao longo de várias apresentações de S1, emissões de R e apresentações de US, as conexões favoráveis à emissão de R e UR, a partir do estímulo S1, são gradualmente estabelecidas. Quando as ligações estiverem fortemente estabelecidas, o estímulo S1 passa, seguramente, a evocar a resposta R e também a resposta UR, passando a ser um estímulo discriminativo.

A resposta UR em particular, quando emitida a partir da estimulação de S1 e não do estímulo US é chamada de CR (*conditioned response* ou resposta condicionada) e é conseqüência do condicionamento respondente. Lembre-se de que este condicionamento é capaz de estabelecer relações entre diversos estímulos e respostas reflexas. Além disso, a saída CR também representa um reforçador adquirido, pois torna possível a ativação do mecanismo neural de reforço, por meio de um estímulo qualquer que não o US.

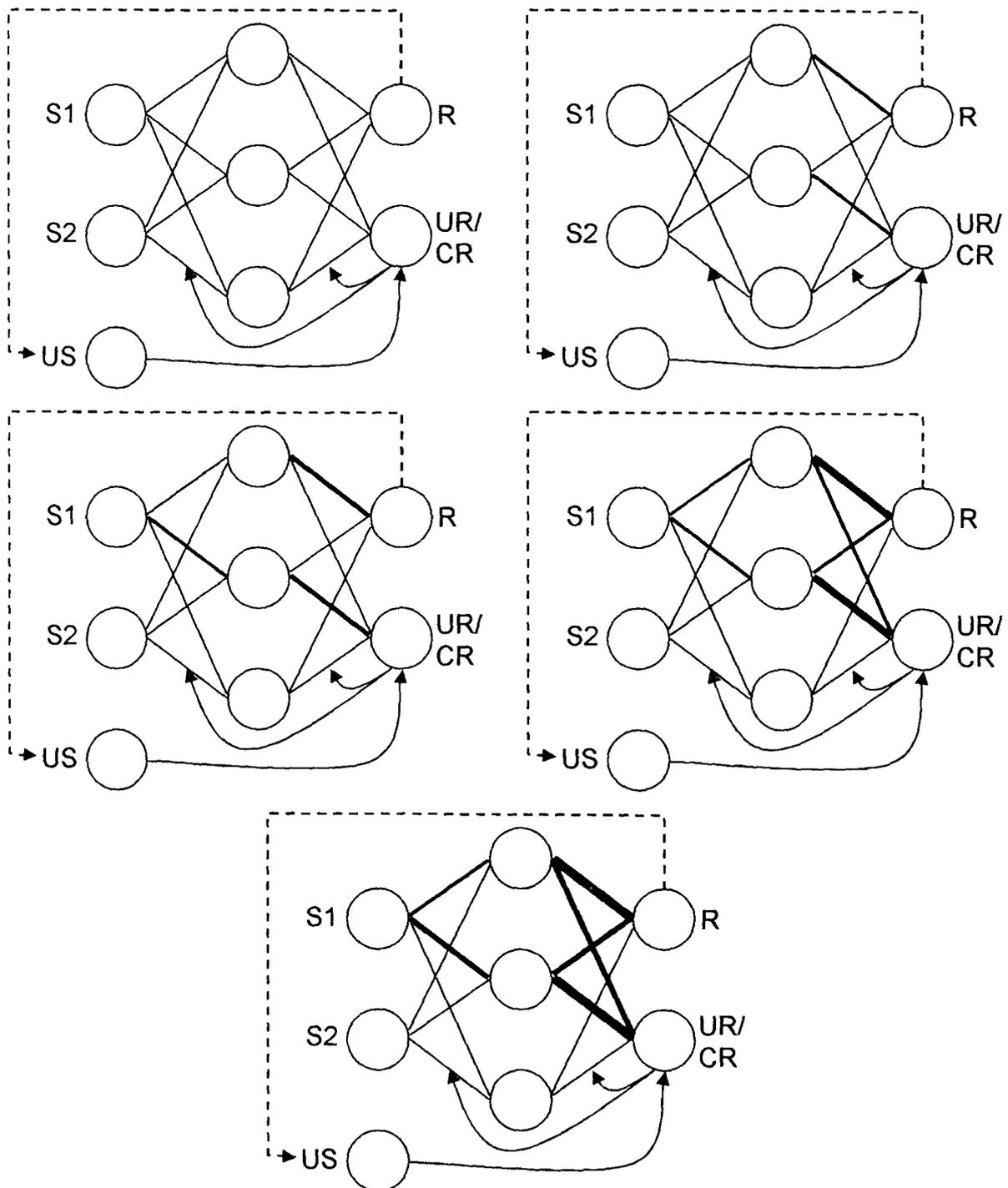


Figura 9 – Exemplo de condicionamento operante na rede de seleção pelas consequências¹.

Embora a Figura 9 ilustre o mecanismo de aprendizagem associado ao condicionamento operante, o modelo também ilustra o mecanismo de aprendizagem associado ao condicionamento respondente (fortalecimento das ligações com UR/CR). De fato, ambos ocorrem juntos. A Figura 10 mostra que a

¹ Fonte: Donahoe e Palmer (1994).

diferença entre os dois tipos de condicionamento reside apenas na dependência da ocorrência do estímulo US (linha tracejada). O desenvolvimento da rede (fortalecimento das conexões) acontece de forma similar em ambos os casos. Diferenças ocorrem apenas no caso de um condicionamento puramente respondente, pois nenhuma conexão com a saída R é estabelecida, uma vez que sua importância para a ocorrência do estímulo US é nula, ao passo que um condicionamento operante sempre envolve um condicionamento respondente associado.

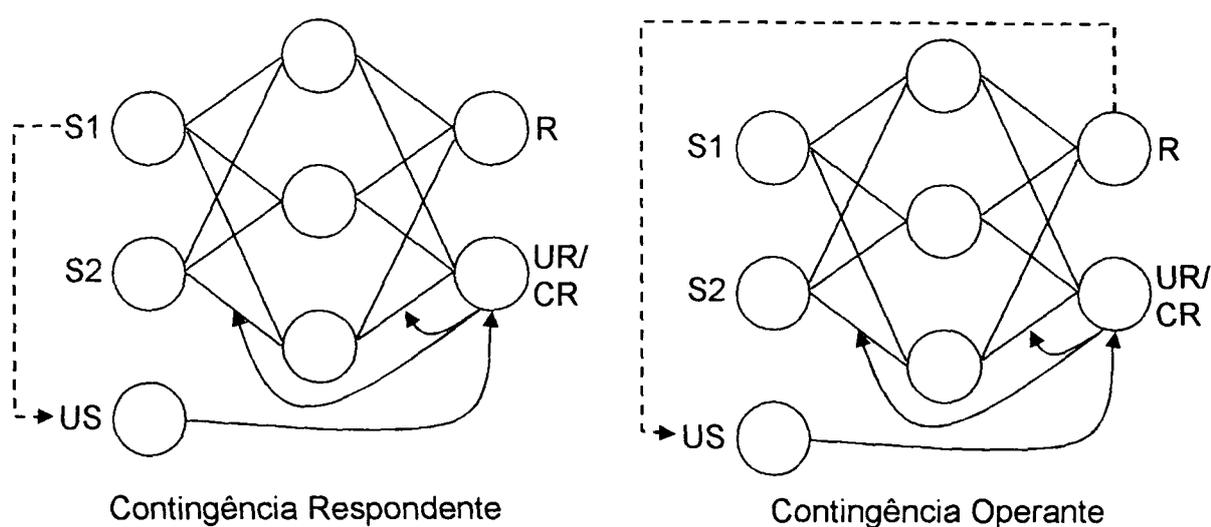


Figura 10 – Contingência respondente e operante¹.

Note-se que a liberação de dopamina depende de um reforço do ambiente em relação à resposta emitida pelo organismo, sendo que tal reforço não é imediato. Assim, utiliza-se um traço de atividade que decai com o tempo e que permite ao sistema registrar quais sinapses foram utilizadas na emissão do comportamento e que, conseqüentemente, devem ser fortalecidas quando o reforço for aplicado, podendo ser modelado por meio de uma memória.

Os resultados apresentados por Donahoe, Burgos e Palmer (1993) em simulações foram bastante promissores ao reproduzirem diversos princípios comportamentais, entre eles:

1. Aquisição de comportamento respondente;

¹ Fonte: Donahoe e Palmer (1994).

2. Extinção de comportamento respondente;
3. Aquisição de comportamento operante;
4. Extinção de comportamento operante;
5. Reaquisição de comportamento operante;

Embora apresentem resultados interessantes, as simulações utilizaram estímulos e respostas bem simples, sendo que apenas um ou dois neurônios ativos eram utilizados como estímulo e apenas um neurônio de saída era considerado como resposta. As próprias dimensões da rede também são bem reduzidas, pouco diferente, em número de neurônios, daquelas exibidas nas Figura 9 e Figura 10. Além disso, apesar de descrever em teoria o funcionamento de unidades inibitórias, nas simulações tais unidades não foram empregadas, o que obviamente deixa muitas questões em aberto com relação à sua importância e possíveis contribuições.

Assim embora demonstre o funcionamento de princípios interessantes, a capacidade reduzida do modelo em termos do número de contingências que pode armazenar, bem como possíveis contribuições de fatores como inibição lateral, representam fatores limitantes para a implementação de sistemas autônomos mais complexos. Para o desenvolvimento de tais sistemas, os princípios que o modelo introduz devem primeiramente ser verificados em uma arquitetura mais complexa e com maior capacidade de armazenamento.

3 MODELO DE REDE NEURAL ARTIFICIAL

Neste capítulo, apresenta-se a rede neural cujo modelo é a proposta principal desse trabalho. O modelo é baseado no já discutido trabalho de Donahoe, Burgos e Palmer (1993) e Donahoe e Palmer (1994) cuja teoria foi apresentada no Capítulo 2 (Seção 2.2.3). O capítulo divide-se em três partes, apresentando arquitetura, raciocínio e aprendizagem.

3.1 Arquitetura

A rede neural consiste de M camadas. Existe uma camada sensorial ou camada de entrada (a camada inferior). Os neurônios sensores desta camada podem pertencer a diferentes classes, definidas segundo a natureza do estímulo que detectam. Há uma camada de saída que define alterações no estado do ambiente. As demais camadas são internas. As camadas consistem de neurônios dispostos em uma estrutura retangular $N \times N$ (Figura 11). Existe ainda uma estrutura adjunta, formada por um conjunto de neurônios sensores denominados US. Os neurônios US são conectados diretamente a neurônios da camada de saída, denominados respondentes (na Figura 11 os neurônios US são representados por um único elemento). Os neurônios respondentes são aqueles que codificam a resposta reflexa. Tal resposta, segundo a teoria do condicionamento, quando emitida devido ao estímulo US é definida como UR (*unconditioned response*). Caso ela seja emitida devido às ligações sinápticas que se formaram desde a camada de entrada até os neurônios respondentes, passando pelo interior da rede, é então chamada de CR (*conditioned response*).

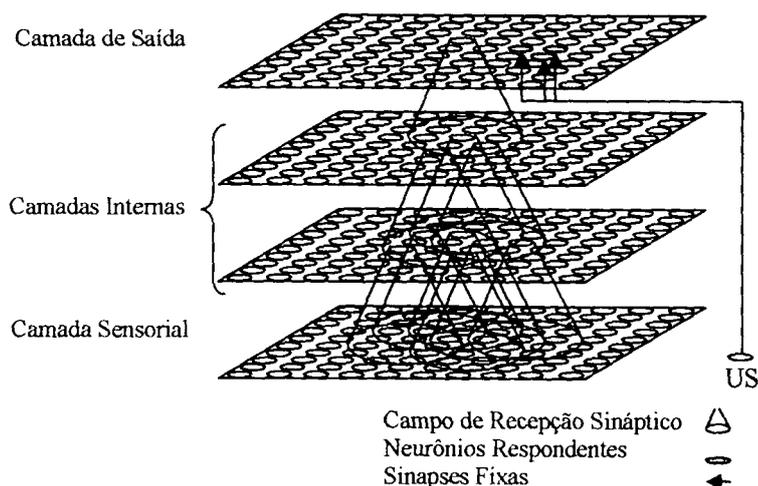


Figura 11 – Arquitetura de rede neural para quatro camadas.

Cada neurônio estabelece conexões sinápticas segundo três classes distintas, a saber: excitatórias inter-camadas, inibitórias e excitatórias intra-camadas. Conforme as denominações já deixam claro, as classes definem sinapses de características específicas. Para um dado neurônio (não pertencente a camada de entrada) as sinapses excitatórias inter-camadas ocorrem com neurônios pré-sinápticos da camada anterior, situados em uma vizinhança cujo centro corresponde ao neurônio de posição relativa idêntica ao neurônio pós-sináptico. Assim, as conexões excitatórias inter-camadas, a partir de neurônios pré-sinápticos para um único neurônio pós-sináptico, definem uma figura espacial em forma de cone, tal como ilustra a Figura 11. A distribuição espacial de neurônios pré-sinápticos que estabelecem tais conexões segue uma distribuição de probabilidades Gaussiana (Figura 12). Por meio das sinapses excitatórias inter-camadas, os sinais fluem de camada em camada. As sinapses inibitórias estão presentes em conexões estabelecidas entre neurônios da mesma camada de forma que a distância ao neurônio pós-sináptico é maior que uma distância mínima (vizinhança distante). Sua função primordial é regular a atividade dos neurônios, impedindo que ocorra uma explosão de atividade na rede, tendo também papel crucial na formação de grupos neurais. As sinapses excitatórias intra-camadas existem em conexões entre neurônios da mesma camada de forma que a distância ao neurônio pós-sináptico é menor que uma distância máxima (vizinhança próxima). Tais sinapses, em conjunto com as sinapses inibitórias, têm como função estabilizar a atividade da rede. Assim

como as sinapses excitatórias inter-camadas, as conexões excitatórias intra-camadas são escolhidas aleatoriamente segundo uma distribuição Gaussiana na região de vizinhança. A distribuição das conexões inibitórias se dá segundo uma diferença de distribuições Gaussianas (Figura 12). Uma restrição impede que os neurônios possuam conexões com eles mesmos (auto-sinapse), além de conexões repetidas (duas ou mais) com outro neurônio qualquer. De uma forma geral, a organização das conexões sinápticas favorece a formação de grupos neurais (Edelman, 1987).

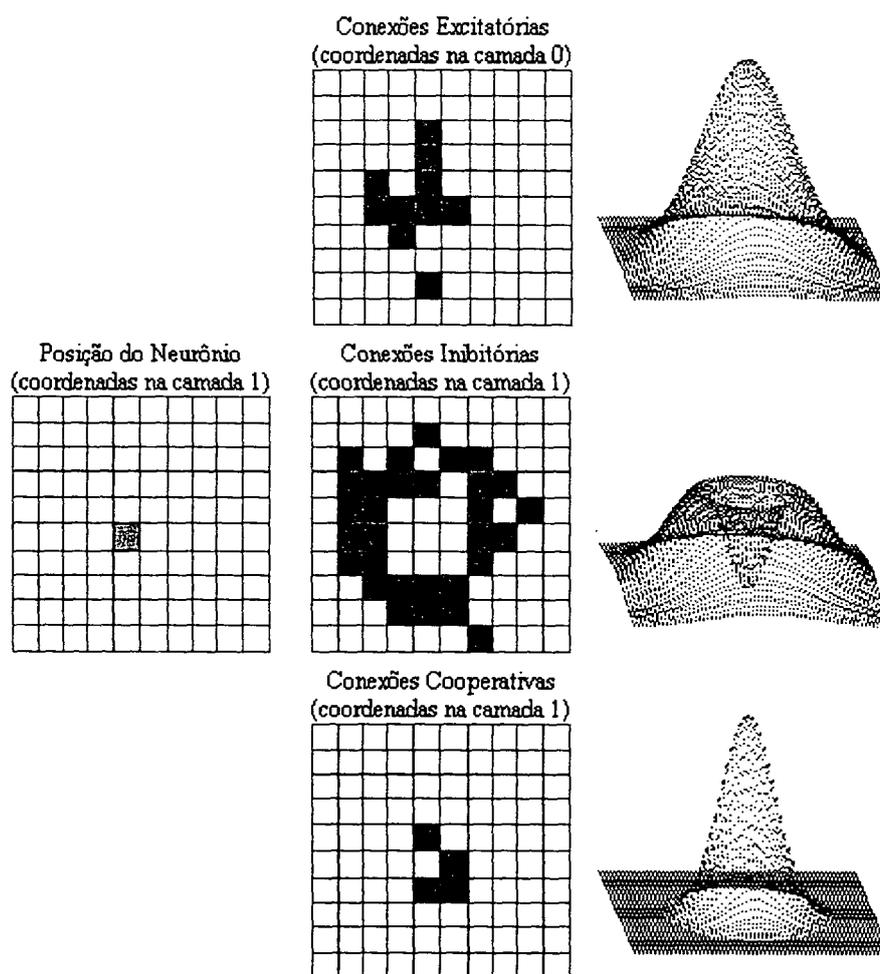


Figura 12 – Conexões sinápticas típicas e respectivas distribuições.

Na Figura 12, a área vermelha indica a posição do neurônio dentro da camada (camada-1\linha-6\coluna-5, neste caso). Os pontos pretos indicam as coordenadas dos neurônios em cada camada com os quais o “neurônio vermelho” faz conexões, ou seja, são os neurônios pré-sinápticos. As conexões excitatórias

inter-camada (quantidade = 10) são feitas com neurônios da camada imediatamente anterior (camada 0). As demais conexões, inibitórias (quantidade = 30) e excitatórias intra-camada (quantidade = 4), são estabelecidas com neurônios da mesma camada (camada 1). Note a inexistência de sinapses repetidas bem como de auto-sinapses.

Outra característica da arquitetura (não ilustrada na Figura 11), a configuração toroidal das camadas, visa minimizar problemas de borda. Dessa forma, para efeitos de estabelecer conexões, os neurônios à extrema direita de cada camada são vizinhos daquelas à extrema esquerda da mesma camada, assim como os da parte superior o são para os da parte inferior (e vice-versa).

3.2 Raciocínio

Os neurônios possuem natureza dependente da camada a que pertencem. Para a camada de entrada, a atividade do neurônio sensor j , $a_j(T)$, na iteração T , é definida por $a_j(T) = e_j(T)$, em que $e_j(T)$ é o componente j do estímulo $e(T)$.

Observe que o estímulo $e(T)$ consiste de $N \times N$ componentes, existindo uma correspondência biunívoca com os neurônios da camada sensorial. O índice que especifica o neurônio é único, apesar da estrutura da camada ser retangular. No entanto, a notação favorece a simplicidade e não deve prejudicar o entendimento do texto.

Para os sensores presentes na estrutura adjunta, a atividade do neurônio sensor j , $a_j(T)$, na iteração T , é definida por $a_j(T) = US_j(T)$, em que $US_j(T)$ é o componente j do estímulo $US(T)$. Cada componente $US_j(T)$ corresponde a um sinal de avaliação (reforço) distinto, emitido pelo ambiente.

Para as camadas internas, seguindo a proposta de Donahoe, Burgos e Palmer (1993), a atividade do neurônio j , $a_j(T)$, é dada por:

$$a_j(T) = \begin{cases} \left[S(\text{exc}_j(T)) + \tau_j \cdot S(\text{exc}_j(T-1)) \cdot (1 - S(\text{exc}_j(T))) \right] - S(\text{inb}_j(T)), \\ \quad \text{se } (\text{inb}_j(T) < \text{exc}_j(T)) \text{ e } (\text{exc}_j(T) \geq \theta_j(T)); \\ \\ (a_j(T) - \kappa_j \cdot a_j(T-1)) - S(\text{inb}_j(T)), \\ \quad \text{se } \text{inb}_j(T) < \text{exc}_j(T) < \theta_j(T); \\ \\ 0, \text{ se } \text{inb}_j(T) \geq \text{exc}_j(T); \end{cases} \quad (1)$$

com,

$$\text{inb}_j(T) = \sum_i \alpha_i^I(T) w_{ji}(T); \text{ e}$$

$$\text{exc}_j(T) = \sum_i \alpha_i^{EA}(T) w_{ji}(T) + \sum_i \alpha_i^{EE}(T) w_{ji}(T);$$

em que: $\alpha_i^I(T)$ é a atividade do neurônio pré-sináptico i que estabelece conexão inibitória com o neurônio pós-sináptico j ; $\alpha_i^{EE}(T)$ ($\alpha_i^{EA}(T)$), de forma análoga, é a atividade do neurônio pré-sináptico i que estabelece conexão excitatória intercamada (intra-camada) com o neurônio pós-sináptico j ; $w_{ji}(T)$ é o peso sináptico correspondente à conexão entre os neurônios pré-sináptico i e o pós-sináptico j ; θ_j (0.0, 0.5) é um limiar de ativação sorteado para cada neurônio j a cada iteração T ; κ_j é a taxa de decaimento de atividade do neurônio j (traço de atividade); τ_j (0,1) é uma constante; e $S(\cdot)$ é a função sigmóide como a da Figura 4 (parâmetros na Tabela 3).

Para a camada de saída, a atividade do neurônio j , que seja respondente, é dada por:

$$a_j(T) = \begin{cases} a_j(T), & \text{se } US_j(T) = 0; \\ US_j(T), & \text{se } US_j(T) \neq 0. \end{cases} \quad (2)$$

em que: $US_j(T)$ é o componente j do estímulo $US(T)$. Caso o neurônio não seja respondente, sua atividade $a_j(T)$ segue a mesma definição apresentada para os neurônios de camadas internas (Equação 1).

Um detalhe importante da Equação 1 diz respeito ao termo θ_j . Seu objetivo é simular uma certa imprevisibilidade ou variância no nível de sensibilidade do neurônio. Em um neurônio natural, essa sensibilidade representa, entre outros, o tempo necessário para a membrana plasmática se estabilizar após um disparo. Assim, a cada disparo, o neurônio está em um estado um pouco diferente e um mesmo estímulo pode provocar respostas ligeiramente distintas de uma iteração para outra, dependendo do valor do limiar θ_j naquela iteração. Matematicamente, o limiar θ_j é uma variável aleatória com distribuição Gaussiana (média = 0.2, variância = 0.2) assumindo valores no intervalo [0.0 , 0.5]. O objetivo principal desse esquema é produzir variabilidade de comportamentos, criando na rede um padrão de disparos variante que busca refletir o que seria um comportamento de nível operante. Tal comportamento é vital para o funcionamento da rede. De fato, é possível estabelecer um paralelo deste com problemas de busca que são parte do desafio na operação dos sistemas inteligentes (Haykin, 2001).

O processo de atualização das atividades dos neurônios segue uma seqüência distinta. O valor de atividade $a(T)$ de cada neurônio da rede neural é definido de acordo com o procedimento de atualização da atividade, descrito em seguida na forma de uma seqüência de passos:

1. A camada sensorial, inclusive considerando a estrutura adjunta, é ativada de acordo com os estímulos apresentados $e(T)$ e $US(T)$;
2. Caso o componente $US_j(T)$ seja diferente de zero, o neurônio sensor j da estrutura adjunta e os neurônios respondentes a ele conectados são ativados, com o mesmo nível de intensidade do componente $US_j(T)$;
3. Escolhe-se um neurônio qualquer das camadas internas ou de saída (que ainda não tenha sido escolhido e não seja respondente) e determina-se sua atividade;

4. Repete-se o passo 3 até que todos os neurônios tenham sido atualizados.

Observe que o procedimento de atualização da atividade proporciona ao modelo da rede neural uma dinâmica paralela. Além disso, ele busca evitar qualquer favorecimento a um neurônio ou grupo de neurônios, algo que poderia ocorrer caso tal grupo fosse sempre o primeiro a ser atualizado e, conseqüentemente, fosse o primeiro a estabelecer seu campo de influência através das conexões inibitórias e excitatórias intra-camada.

3.3 Aprendizagem

O mecanismo de aprendizagem consiste de uma regra de Hebb adaptada. Assim sendo, as sinapses que sofrem modificação são apenas aquelas em que ambos os neurônios (pré e pós-sináptico) estão ativos. Além disso, a equação (3) ainda incorpora o reforço (US) já que se trata de um modelo de aprendizagem por reforço (Donahoe, Burgos e Palmer, 1993).

$$\Delta w_{ji}(T) = \begin{cases} \alpha_j a_j(T) d(T) p_i(T) r_j(T) & \text{se } d(T) > 0 \\ -\beta w_{ji}(T-1) a_i(T) a_j(T) & \text{se } d(T) \leq 0 \end{cases} \quad (3)$$

em que: α é a taxa de aquisição; β é a taxa de extinção;

$d(T)$ é uma função do estímulo US (ao longo do trabalho $d(T) = US(T)$ a menos que especificado o contrário, diferentemente da proposto inicial de Donahoe, Burgos e Palmer (1993) onde $d(T)$ é obtido segundo uma outra função distinta);

$r_j(T) = 1 - \sum_i^n w_{ji}(T)$ é o termo de competição entre as sinapses;

$p_i(T) = \frac{\alpha_i(T) w_{ji}(T-1)}{N}$ é o termo de contribuição de cada sinapse (sinapses que

contribuem mais para a atividade recebem mais peso) onde $N = exc_j(T)$ ou

$N = inj_j(T)$ conforme o tipo da sinapse.

Dessa forma, caso o reforço seja maior que zero, a sinapse é fortalecida (processo de aquisição de comportamentos) e caso o reforço seja menor ou igual a zero a sinapse é enfraquecida (processo de extinção de comportamentos). Note que a Equação (3) é válida para ambos os tipos de sinapse (excitatória e inibitória).

O método de atualização dos pesos sinápticos é semelhante em abordagem e objetivo àquele utilizado para a atualização da atividade e segue os seguintes passos:

Para todo instante T :

1. É sorteado um neurônio equiprovavelmente e sem reposição;
2. Sorteia-se equiprovavelmente uma sinapse (sem reposição);
3. Atualiza-se o peso sináptico correspondente, aplicando a Equação (3) e utilizando as entradas no instante T ;
4. Repete-se o passo 2 e 3 para todas as sinapses;
5. Repetem-se os passos 1, 2, 3 e 4 para todos os neurônios.

4 RESULTADOS

Este capítulo apresenta os resultados obtidos com o modelo proposto no Capítulo 3 em diversos experimentos de simulação. Os experimentos visam demonstrar as características de funcionamento bem como as capacidades de aprendizagem do modelo. Diversos experimentos têm fundamento na teoria da seleção pelas conseqüências e objetivam avaliar o modelo segundo os fenômenos comportamentais que a teoria descreve.

Para efeitos da interpretação dos resultados de experimentos apresentados na forma de figuras faz-se necessária uma observação. A atividade da rede é um fator dinâmico dependente de diversas regras e extremamente variante no tempo, não apenas quando essa se encontra em nível operante (exibindo respostas aleatórias), mas também depois de realizada a aprendizagem (onde normalmente a atividade estabiliza-se ao redor de um certo padrão, mas continua variante). Devido a essa natureza, cada figura representa uma amostra do padrão de atividade dos neurônios em cada camada em um determinado instante T e não um comportamento fixo.

Os resultados apresentados são divididos em duas partes, na Seção 4.1 os exemplos visam mostrar os mecanismos e princípios de funcionamento da rede, fornecendo um melhor entendimento de como determinadas características e parâmetros influenciam seu funcionamento. A Seção 4.2 traz experimentos cujo objetivo é demonstrar as capacidades de aprendizagem da rede bem como a gama de comportamentos que ela é capaz de adquirir.

Parâmetros das Simulações

Ao longo de todos os experimentos realizados nas próximas seções, procurou-se manter o mesmo conjunto de parâmetros para o modelo. A não ser quando especificado diferentemente, tal conjunto segue a descrição fornecida nas Tabelas 1, 2 e 3. Obteve-se os parâmetros apresentados através de observações preliminares de funcionamento do modelo. Muitos dos parâmetros visam fornecer condições mínimas para o funcionamento de características esperadas do modelo.

Por exemplo, a dimensão das camadas e a distribuição sináptica visam fornecer condições para o surgimento de grupos neurais.

Tabela 1 – Propriedades da Rede

Número de camadas	4
Número de neurônios por camada	100 (grade de 10x10)

Tabela 2 – Propriedades das Conexões Sinápticas

	Excitatórias	Inibitórias	Exc. Intra-Camada
Nº de sinapses por neurônio	10	30	4
Raio de distribuição¹	[0.0 , 2.5]	[1.5 , 3.0]	[0.0 , 1.5]
Valor inicial²	[0.0095 , 0.0105]	[0.095 , 0.105]	[0.095 , 0.105]

Faz-se aqui necessária uma observação com relação as conexões do tipo inibitórias e excitatórias intra-camadas, a respeito do ajuste sináptico das mesmas. Por motivos discutidos nas conclusões, o ajuste desses tipos de sinapses é desativado durante as simulações apresentadas neste capítulo, ou seja, estas sinapses não aprendem. No entanto, sua função permanece e sua importância também, sendo que apenas sua influência deve ser considerada como constante.

Tabela 3 – Parâmetros das Equações

Parâmetros de Atividade Neural	Parâmetros de Aprendizagem
$\theta_j = [0.0 , 0.5]$	$\alpha = 0.5$
$\kappa_j = 0.8$	$\beta = 0.1$
$\tau_j = 0.1$	$d(T) = US(T)$
Inclinação da Sigmóide = 0.1	
Limiar da Sigmóide = 0.3	

¹ Medido em posições na grade a partir do neurônio (Ex: raio de intervalo [1.5 , 3.0] significa que a sinapse pode ser sorteada e efetuada com qualquer neurônio que esteja no mínimo a 1.5 posições de distância na grade e no máximo a 3.0). O sorteio é feito segundo uma distribuição Gaussiana com média = 0.0 e variância = valor médio do intervalo definido para o tipo da sinapse.

² Variável aleatória com distribuição uniforme no intervalo especificado.

Avaliando Respostas Desejadas

O procedimento básico de operação da rede neural envolve a apresentação de um estímulo, a computação por parte da rede de uma resposta e, caso a resposta seja desejada, o retorno apresentado pelo ambiente na forma de um reforçador.

A resposta desejada é definida antes do início da simulação e envolve a especificação, para cada estímulo definido para o ambiente, de quais neurônios da camada de saída são considerados resposta (neurônios resposta) bem como a faixa de atividade considerada válida. Além disso, é definida a porcentagem de participação necessária para que a resposta seja considerada válida e também o valor do US associado à resposta.

A avaliação é feita segundo o grau de participação dos neurônios resposta na atividade geral da camada de saída. Para tanto, após cada iteração do procedimento de atualização da atividade (ver Seção 3.2), é feito um somatório da atividade de todos os neurônios da camada de saída, bem como um somatório apenas daqueles neurônios especificados como resposta válida e que apresentaram atividade dentro da faixa definida. A relação entre esses dois somatórios, isto é, a porcentagem de participação dos neurônios resposta na atividade geral da última camada é comparada com o valor definido para a resposta. Valores maiores ou iguais são considerados respostas válidas (e são seguidos pela apresentação do US associado). Do contrário, a resposta é considerada inválida.

Interpretando as Ilustrações

Nos experimentos que se seguem, são utilizadas figuras para ilustrar as observações realizadas e estas seguem algumas definições. As figuras representam tanto a arquitetura como a atividade neural e de acordo com o padrão

dos experimentos, ilustram quatro (4) camadas, cada camada consistindo de uma grade retangular de 10x10 (dez por dez) neurônios (Figura 13).

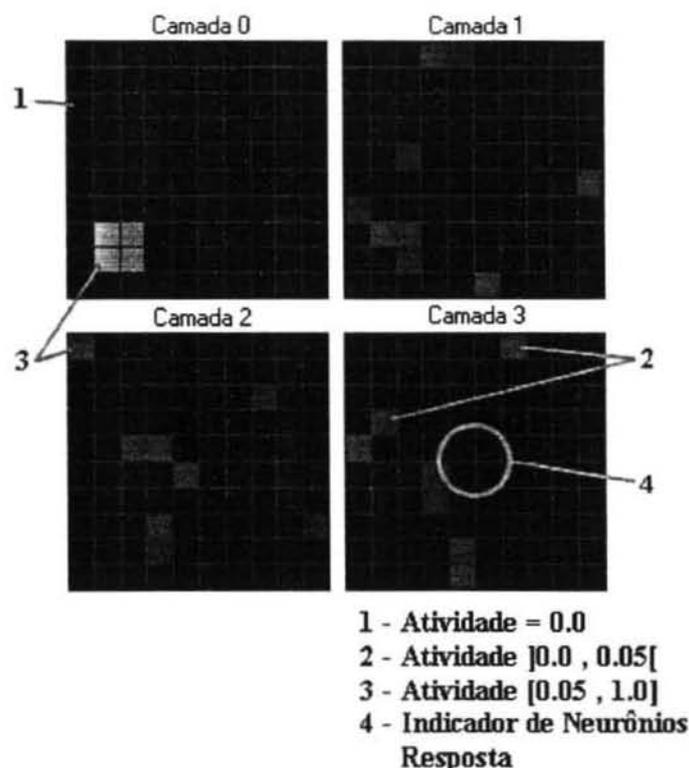


Figura 13 – Representação da rede.

A camada de entrada é referida como “Camada 0”, a “Camada 1” e “Camada 2” são intermediárias e a última camada ou camada de saída é a “Camada 3”. Cada neurônio é representado por um quadrado de borda clara. A atividade do neurônio é indicada pela cor de preenchimento do quadrado. Preto para atividade nula (igual a zero), marrom para indicar uma atividade muito reduzida (entre zero e 0.05) e uma escala de cor azulada para atividades acima de 0.05 até 1.0. Neste último caso, quanto maior a atividade mais intensa a cor. Nos experimentos em que existe treinamento efetivo da rede, existe ainda um círculo vermelho na última camada que destaca os neurônios resposta.

Definindo o Treinamento

A aprendizagem em muitas redes neurais acontece em períodos bem definidos denominados treinamento. O mecanismo de aprendizagem por

reforço segundo uma perspectiva clássica (fundamentada na Análise do Comportamento) permanece indefinidamente habilitado (não existindo distinção entre uma fase de aprendizagem e operação), portanto não podendo ser adotado o conceito de treinamento com o mesmo significado.

Nos experimentos deste capítulo, a interpretação do conceito de treinamento deve, portanto, ser feita segundo o seguinte significado:

1. O início do treinamento é considerado a partir do momento em que a rede é inserida no ambiente;
2. Considera-se que o treinamento chegou ao fim quando, a partir do estímulo CR (contingência respondente) ou S (contingência operante), a atividade neural convergir de forma a representar a resposta desejada (adquirindo comportamentos) ou atingir o estado de nível operante (extinguindo comportamentos).

Um estado estável que represente uma convergência da atividade é identificado quando não se observa alterações significativas no padrão de atividades dos neurônios.

4.1 Experimentos Básicos

Os diversos elementos componentes do modelo proposto, entre eles o padrão de conexões sinápticas característico, interagem de forma particularmente interessante e pouco comum durante o funcionamento da rede neural. As consequências de tais interações são processos indispensáveis para uma correta modelagem das características comportamentais, as quais deseja-se que o modelo seja capaz de reproduzir.

Uma análise independente de tais processos emergentes é feita na seqüência. O foco do estudo não é o treinamento, capacidade de aprendizagem ou o comportamento que a rede é capaz de desenvolver. O objetivo é somente demonstrar características de funcionamento do modelo proposto. A partir de tais exemplos fica claro o papel que cada tipo de sinapse, juntamente com suas configurações específicas, desempenha dentro da rede.

4.1.1 Atividade Neural sem Conexões Laterais

O fato de possuírem alcance limitado bem como serem limitadas em número faz com que as sinapses excitatórias inter-camadas tornem a propagação da atividade neural bem característica.

A partir de um estímulo em uma determinada região da camada de entrada da rede (Camada 0 da Figura 14), a atividade neural propaga-se de camada em camada através das conexões excitatórias inter-camadas para um número limitado de neurônios das camadas posteriores (Camada 1, Camada 2, etc.). A distância que a atividade neural percorre dentro da rede, isto é, a difusão que ela sofre ao longo das camadas (por exemplo, comparando-se a Camada 0 com a Camada 1) é no máximo o alcance permitido pelas conexões sinápticas (aproximadamente ilustrado pelo círculo tracejado na Camada 1) (um parâmetro especificado quando da inicialização da rede).

Para que um estímulo consiga, eventualmente, produzir atividade em qualquer neurônio da camada de saída, é necessário que o alcance mínimo das conexões, acumulado ao longo das camadas, seja grande o suficiente para englobar toda a extensão da última camada (Figura 14). Para este experimento são utilizadas apenas conexões excitatórias inter-camadas, os demais tipos de conexões não são empregados. A atividade neural não converge, pois não se trata de um experimento de aprendizagem, conseqüentemente a atividade neural é oscilante.

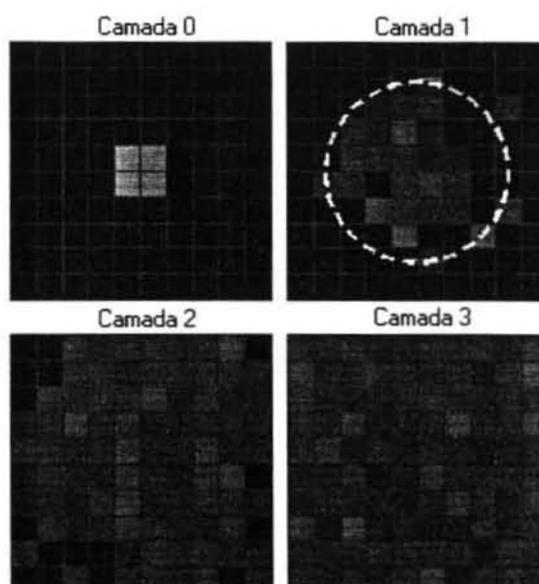


Figura 14 – Propagação da atividade neural por conexões excitatórias inter-camadas (com cone de alcance mínimo).

Caso o alcance mínimo das conexões não seja respeitado não é possível garantir que exista, com certeza, um caminho neural capaz de estabelecer uma ligação entre qualquer estímulo e qualquer resposta possível (Figura 15). Note, na Figura 15, que os neurônios na periferia da Camada 3, são incapazes de serem estímulos pelo estímulo apresentado na Camada 0.

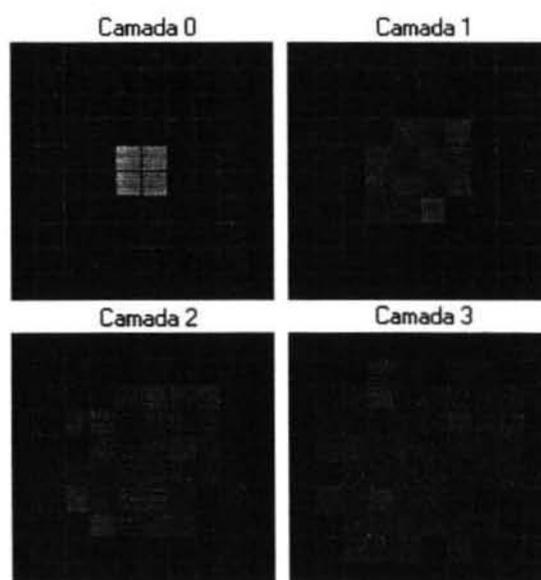


Figura 15 – Propagação da atividade neural por conexões excitatórias inter-camadas (sem cone de alcance mínimo)

4.1.2 Atividade Neural com Conexões Laterais

Na Seção 4.1.1, fica claro como a atividade neural propaga-se pelas várias camadas da rede. No entanto, o resultado final é pouco útil tendo em vista que a resposta produzida abrange toda a extensão da camada de saída. Não existe diferenciação nítida entre possíveis respostas, apenas uma explosão de atividade.

Para obter padrões distintos de atividade que possam ser considerados respostas específicas, válidas ou não, é necessário, de alguma forma, limitar a atividade da rede. É justamente esse o papel das conexões laterais inibitórias.

O exemplo da Figura 16 trata exatamente da mesma rede apresentada no exemplo da seção 4.1.1 apenas com as conexões inibitórias habilitadas e operando. Neste experimento, não foram utilizadas as conexões excitatórias intra-camada. Nota-se uma redução drástica na atividade geral da rede bem como uma maior especificidade, com melhor definição de padrões locais. Os padrões 1 e 2 da Figura 16 ilustram duas respostas emitidas em instantes de tempo arbitrários. Tais respostas não representam um estado estável tendo em vista que não é realizada a aprendizagem, ou seja, a rede não converge.

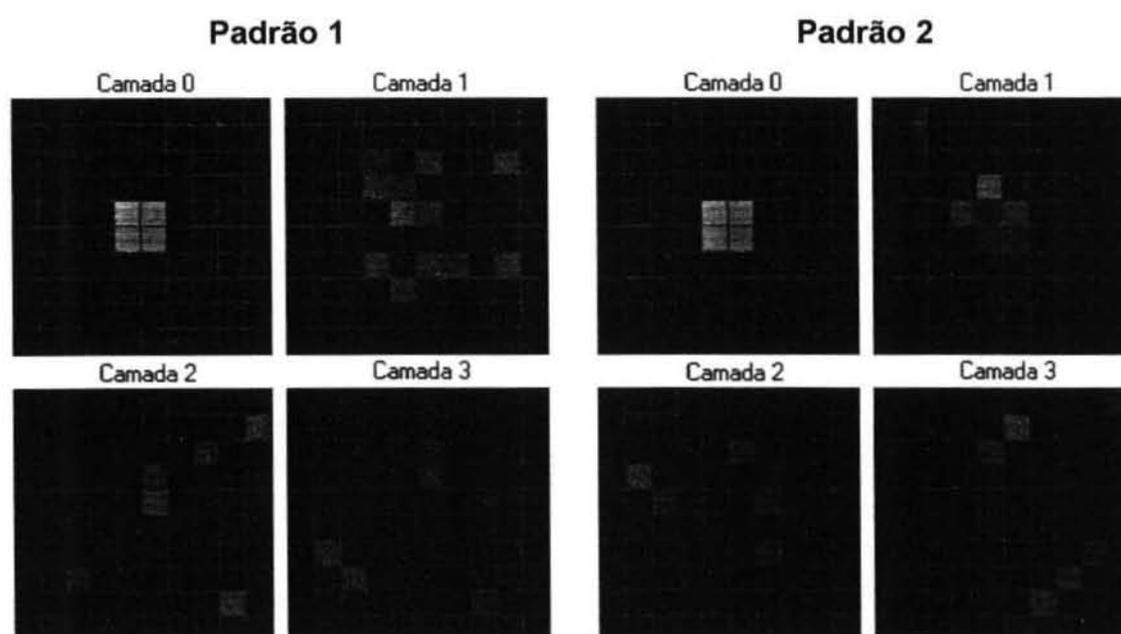


Figura 16 – Atividade neural com conexões laterais inibitórias.

A forma como as conexões inibitórias contribuem para a operação da rede não é tão simples como pode parecer em princípio. Uma das contribuições do uso de conexões inibitórias é um aumento no contraste da atividade neural. Observe que no experimento anterior (Figura 14) as atividades dos neurônios assumiam valores numa ampla faixa, enquanto neste experimento (Figura 16) poucos neurônios atingem o limiar necessário para exibir atividade. Outro fator de grande importância devido às conexões inibitórias é o comportamento aleatório das respostas específicas (nível operante). Com tais conexões existe um certo equilíbrio entre excitação e inibição ao longo de toda a rede e assim surgem padrões específicos de disparo erráticos.

Considere o exemplo ilustrado na Figura 17 em que os pesos sinápticos são configurados para $[0.295, 0.305]$ (maiores que nos demais experimentos). Definindo maiores valores para os pesos a excitação torna-se predominante sobre a inibição. Observe que a excitação mais intensa, causa desequilíbrio entre excitação e inibição, a atividade assume valores altos e a resposta deixa de ser específica. Respostas não específicas deixam de ser interessantes, pois impossibilitam a definição de padrões de resposta específicos desejados.

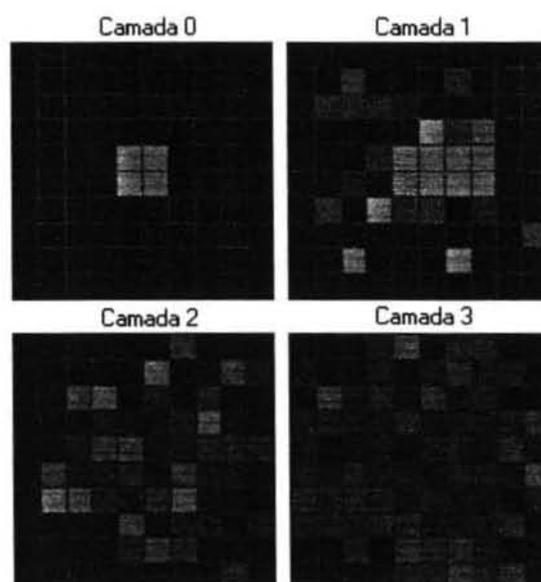


Figura 17 – Atividade neural com conexões laterais inibitórias e excitação predominante.

É interessante notar que, sob a influência do algoritmo de aprendizagem, uma rede com tal atividade rapidamente evoluiria para o equilíbrio excitação/inibição. Conseqüentemente a rede retornaria ao estado de nível operante (Figura 16). Isso acontece uma vez que é pequena a probabilidade de que respostas pouco específicas, como as ilustradas na Figura 17, sejam de fato respostas desejadas, implicando na ausência do sinal de reforço. Conseqüentemente, ocorre o enfraquecimento das conexões inter-camadas, o que tende a equilibrar o nível de excitação com o de inibição e produzir o comportamento de nível operante. No entanto, isso pode não ocorrer, bastando para isso que a atividade ilustrada na Figura 17 seja considerada uma resposta válida de algum tipo.

Os experimentos representados nas Figura 14, Figura 16 e Figura 17 oferecem indícios de que as sinapses inibitórias possuem um papel essencial para que a rede estabeleça um estado de nível operante em que respostas desejadas são geradas.

4.1.3 Formação Dinâmica de Grupos Neurais

A distribuição espacial das conexões sinápticas favorece a existência de grupos neurais, isto é, de grupos localizados de neurônios que respondem a um mesmo estímulo ou conjunto de estímulos. No entanto, apenas a distribuição das sinapses não é suficiente para definir grupos dessa natureza. Para tanto considere a discussão em seguida.

Um grupo neural é formado por neurônios centrais, localizados em seu centro, e por neurônios periféricos, posicionados nos limites espaciais do grupo. Neurônios centrais possuem um maior número de sinapses que participam ativamente da operação do grupo, uma vez que ao seu redor existem mais neurônios que também participam do grupo, ao passo que neurônios periféricos possuem algumas sinapses não utilizadas, estabelecidas com neurônios próximos a ele e que, no entanto, não participam do grupo. As conexões sinápticas por si só permitem que cada neurônio seja o centro de seu grupo (Figura 12).

A formação dos grupos neurais é uma característica emergente, dependente não apenas da distribuição, mas também de adaptação sináptica. Ao longo do processo de aprendizagem, os neurônios que inicialmente disparam com

maior intensidade tendem a se tornar neurônios centrais, promovendo a definição dos grupos. A configuração evolui gradualmente à medida que a atividade nas camadas vai se estabelecendo (Figura 18 e Figura 19) e os grupos se definindo em consequência do treinamento (Figura 20).

O tamanho dos grupos depende da distribuição espacial das sinapses. Um padrão de conexão com poucas sinapses excitatórias intra-camadas e com estas estabelecidas com neurônios próximos produz grupos pequenos. As conexões inibitórias intra-camadas também participam na definição do tamanho do grupos. Um raio de alcance pequeno para as conexões inibitórias pode significar que neurônios que possivelmente estabeleceriam um grupo, por estarem próximos, podem na verdade estar inibindo uns aos outros, conseqüentemente diminuindo o número de neurônios que efetivamente cooperam entre si. O alcance das conexões inibitórias também define quantos grupos serão permitidos em cada camada. Um alcance moderado permite que dois ou mais grupos isolados surjam (Camada 2 da Figura 20).

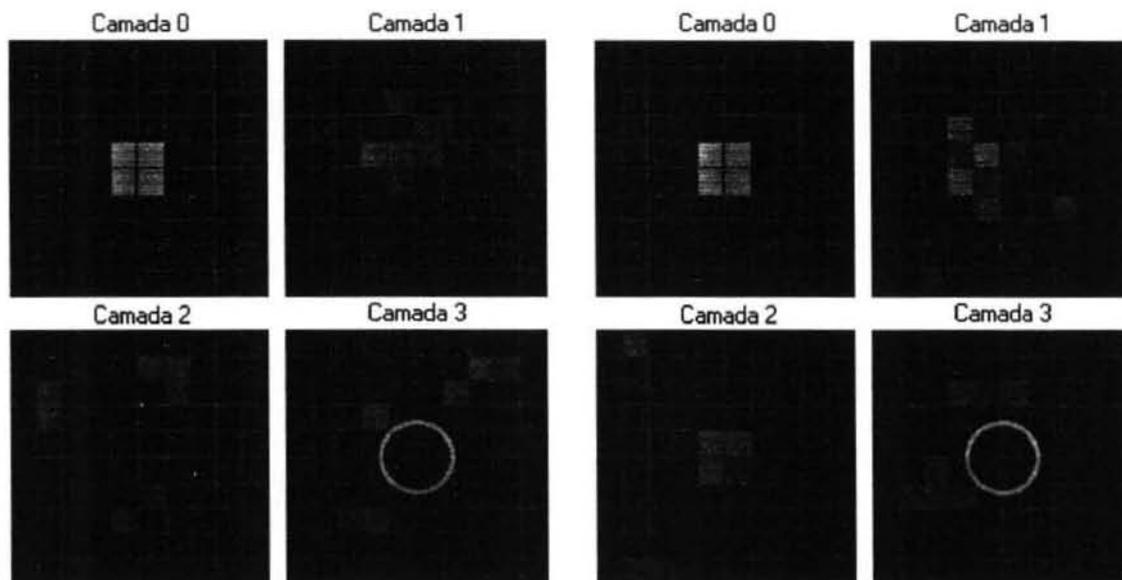


Figura 18 – Atividade inicial ainda sem grupos neurais definidos.

Figura 19 – Atividade intermediária com surgimento dos grupos neurais.

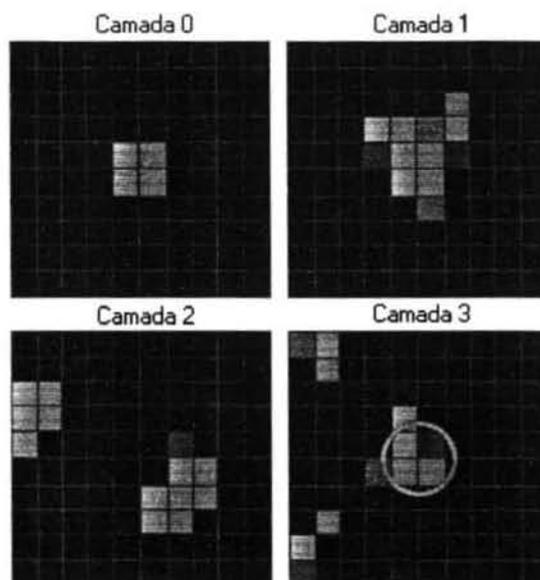


Figura 20 – Atividade final com grupos neurais formados.

Note uma diferença clara entre uma atividade com predominância geral das conexões excitatórias inter-camadas, no caso da Figura 17, e um fortalecimento específico das mesmas através de treinamento (Figura 18, Figura 19 e Figura 20). Fica claro, portanto, que durante o treinamento as conexões sinápticas são fortalecidas segundo um esquema particular, evoluindo uma rede com padrões distintos daqueles que são obtidos simplesmente aumentando-se os pesos sinápticos.

Outra consequência do fato dos grupos neurais desenvolverem-se de forma dinâmica é que, após o treinamento, o grupo que define a resposta não necessariamente abrange todos os neurônios definidos como parte da mesma (indicados nas figuras pelo círculo vermelho). Como o grupo é dinâmico ele pode incorporar mais ou menos neurônios do que os especificados.

4.1.4 Considerações Sobre a Evolução da Atividade Neural

A rede inicia sua operação exibindo padrões de disparos equiprováveis, pois inicialmente seu comportamento deve simular o nível operante. À medida em que a rede recebe sinais de reforço como retorno a respostas desejadas, emitidas na presença de estímulos específicos (discriminativos), suas

conexões sinápticas são fortalecidas gradualmente e, como consequência, o padrão de disparos aleatórios vai se tornando restrito.

Este panorama torna-se ainda mais interessante devido à restrição de alcance das sinapses, a inibição lateral e ao padrão específico de progressão da atividade produzida por tais fatores. Propagando-se a partir da camada de entrada, a atividade neural pode percorrer no máximo a distância que o raio das conexões permite. Deste modo, a distância total que a atividade desencadeada por um estímulo pode atravessar dentro da rede, desde a camada de entrada até a camada de saída, é no máximo a soma do raio das conexões sinápticas em cada camada, levando-se em conta que pequenas variações são permitidas dado o fator aleatório ao se estabelecerem as conexões.

Durante a aprendizagem, as sinapses são mais intensamente reforçadas nas camadas inferiores da rede (pois estão mais próximas aos estímulos que possuem nível de atividade expressivo e também por possuírem atividade mais freqüente). Quando a atividade torna-se estável na primeira camada, na forma de grupos neurais, automaticamente ela passa a definir uma região mais restrita (que é menor que o cone total de propagação possível da atividade) da rede a partir da qual a atividade poderá propagar-se, uma vez que o cone de dispersão da atividade (criado pelo padrão de conexões) possui agora uma camada a menos para se expandir. Conseqüentemente, o espaço de busca torna-se mais limitado, pois a cada camada ele limita-se a um número de neurônios menor que o total de neurônios possíveis. Tal fato pode ser comparado ao fenômeno da especificação da resposta, visto na Seção 2.2.2.

4.2 Experimentos de Aprendizagem

Nesta seção o objetivo dos experimentos realizados passa ser a demonstração das capacidades de aprendizagem do modelo. Os dois tipos de aprendizagem envolvem contingências do tipo respondente e operante, com destaque para este último. Lembrando que os princípios mostrados na Seção 4.1 permanecem presentes e atuando.

Apesar de ter sido ressaltado durante a discussão da teoria que durante um treinamento operante inevitavelmente também ocorre condicionamento

do tipo respondente, nos experimentos apresentados a seguir preferiu-se separar os dois em benefício da compreensão. Isso, no entanto, não invalida o que foi discutido até o momento. A separação dos dois tipos de aprendizagem foi realizada simplesmente não especificando neurônios respondentes durante os treinamentos de contingências operantes, impossibilitando assim a associação do tipo clássica entre o estímulo US e tais neurônios.

4.2.1 Experimento 1 – Convergência da Aprendizagem em Treinamento Respondente

O primeiro experimento visa demonstrar as capacidades da rede neural em associar um estímulo qualquer (estímulo condicionado - CS) a um estímulo incondicionado US (reforçador). Considera-se que existe uma correlação temporal do primeiro com o segundo. Após o treinamento, o estímulo condicionado, que inicialmente não produz qualquer resposta particularmente interessante, passa a eliciar a mesma resposta ou uma resposta muito próxima àquela produzida inicialmente apenas pelo estímulo reforçador.

Para este experimento são definidos o estímulo incondicionado US (Camada 0 da Figura 21), os neurônios respondentes (círculo vermelho na Camada 3 da Figura 21), a intensidade do US (igual a 1.0 ou intensidade total) e a correlação temporal entre US e CS (o US é apresentado 3 iterações após o CS).

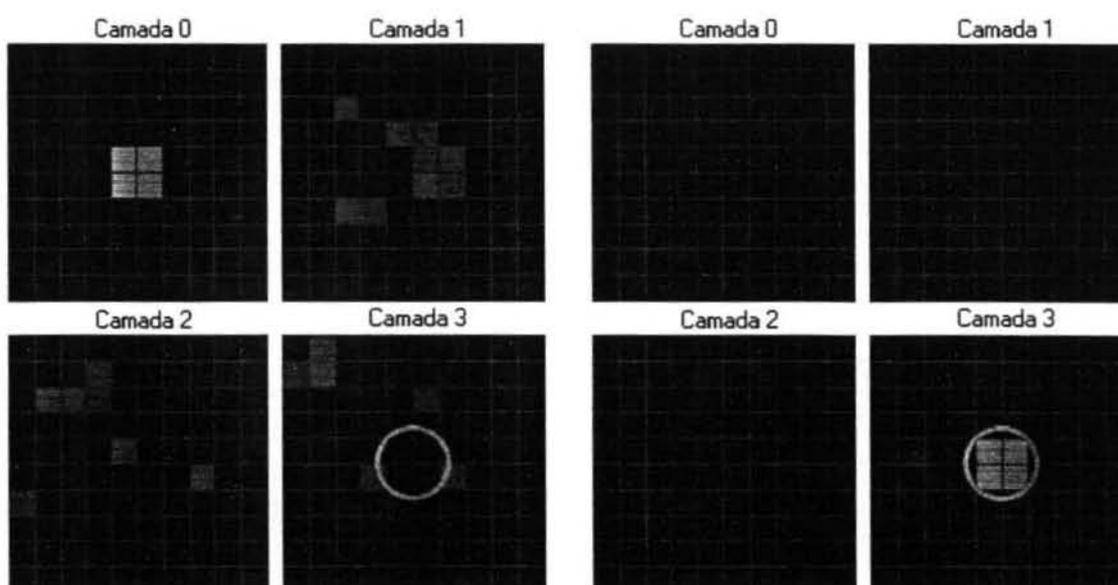


Figura 21 – Atividade neural do estímulo CS anterior ao treinamento respondente.

Figura 22 – Atividade neural do estímulo US anterior ao treinamento respondente.

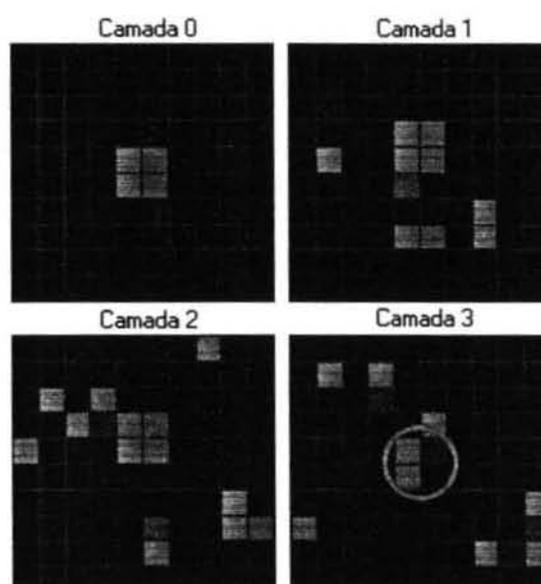


Figura 23 – Resposta do estímulo CS após treinamento respondente.

Inicialmente, como era de se esperar, o estímulo CS não produz uma resposta específica na rede (Figura 21). O estímulo US, como também é de se esperar, é capaz de produzir atividade nos neurônios respondentes mesmo antes do treinamento (Figura 22). Note que a presença do estímulo US afeta apenas os neurônios respondentes (ver Seção 3.1).

Após o treinamento, o estímulo CS adquire a capacidade de produzir na rede uma resposta semelhante àquela produzida pelo estímulo US (resposta condicionada), mesmo, agora, sem a presença deste último (Figura 23). Portanto, o treinamento respondente foi bem sucedido.

Na Figura 23, além de atividade neural na região esperada (região dos neurônios respondentes), a rede também exibe atividade em outros neurônios. Isso se deve a vários fatores, entre eles o padrão de conexões sinápticas. O interessante é que a própria teoria do condicionamento respondente, assim como visto, prevê como consequência do mesmo o aumento na frequência de qualquer comportamento que esteja ocorrendo nas proximidades temporais da ocorrência do estímulo US. Sendo assim, a existência de atividade neural fora da região definida como resposta indica a existência de uma correlação entre os neurônios, provavelmente devido a um padrão de conexões característico, o que é perfeitamente condizendo com a teoria.

4.2.2 Experimento 2 – Extinção de Comportamento Respondente

Para um sistema artificial ser capaz de comportar-se de forma autônoma ele precisa ser capaz de se adaptar a um ambiente mutável. Caso uma contingência ambiental mude de característica e deixe de estar associada a um estímulo reforçador, consequentemente o sistema artificial deve deixar de emitir a resposta que aprendeu a associar até então. Adaptar-se nesse sentido não limita-se a adquirir comportamentos, mas também a desfazer-se de comportamentos.

Para exemplificar esse processo, utiliza-se como ponto de partida o Experimento 1, onde foi estabelecida uma contingência clássica entre um CS e um US (Figura 23 e Figura 24 que ilustram o mesmo caso). A partir dessa configuração, uma vez removido o estímulo US (eliminando a relação US – CS), a rede neural cessa de emitir a resposta condicionada à presença do CS (Figura 25), extinguindo a resposta anteriormente adquirida.

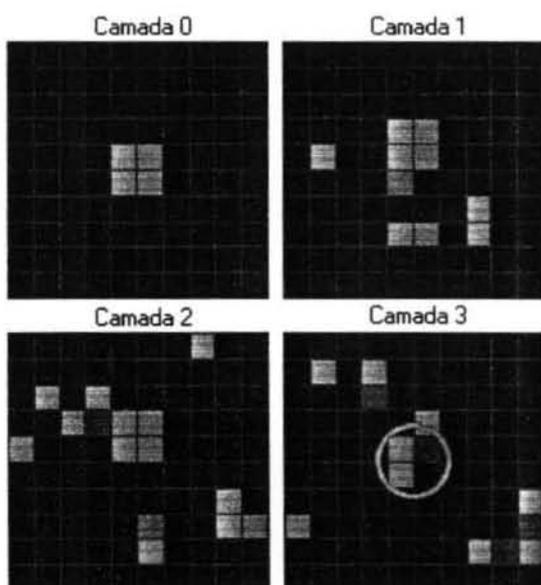


Figura 24 – Atividade neural anterior à extinção do comportamento respondente.

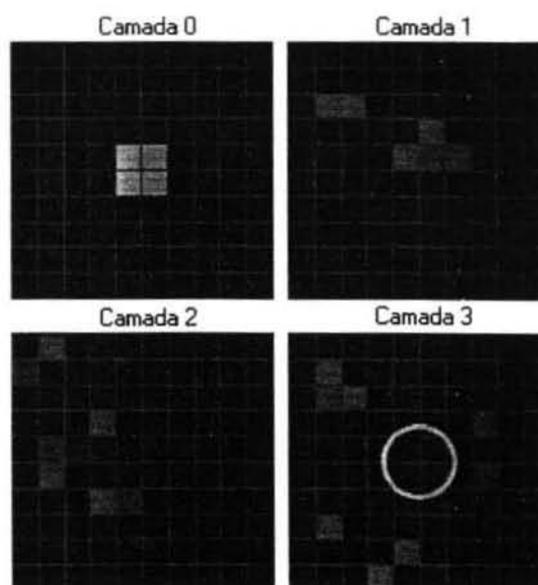


Figura 25 – Atividade neural após extinção do comportamento respondente.

4.2.3 Experimento 3 – Convergência da Aprendizagem em Treinamento Operante

De acordo com os conceitos deste trabalho, um dos maiores desafios para um sistema autônomo consiste em investigar o ambiente, explorando-o, e aprender a partir dessa exploração quais as contingências presentes no ambiente. Ou seja, o sistema deve aprender quais respostas são reforçadas pelo ambiente e também deve aprender a emitir tais respostas nas situações em que elas são esperadas.

O objetivo desse experimento é demonstrar a capacidade do modelo proposto em aprender uma contingência operante simples. A simulação envolve a presença de um estímulo (estímulo discriminativo), uma resposta desejada e um estímulo reforçador (US), sendo apresentado apenas quando, na presença do estímulo discriminativo, a rede emite a resposta desejada.

Inicialmente, a resposta da rede para o estímulo discriminativo compreende o padrão de disparos de nível operante (Figura 26). Durante o treinamento a rede recebeu o sinal de reforço (US) sempre quando exibia a resposta correta (círculo vermelho). Desta forma, a atividade neural evoluiu para a configuração exibida pela Figura 27, adquirindo a capacidade de responder

corretamente (apresentando a resposta desejada) sempre que o estímulo discriminativo é apresentado.

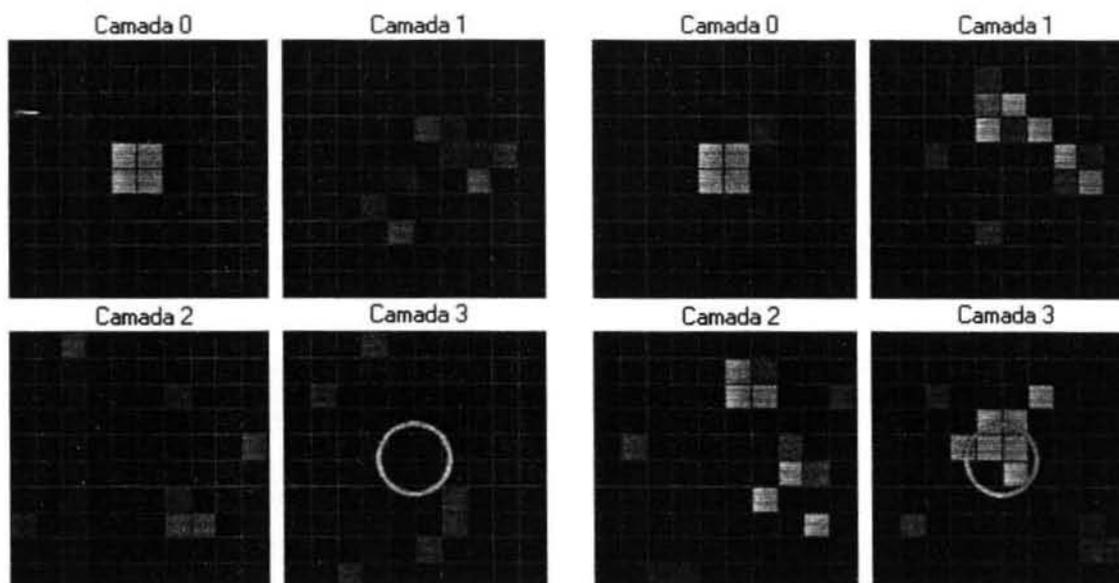


Figura 26 – Atividade neural anterior ao treinamento operante.

Figura 27 – Atividade neural após treinamento operante.

É interessante notar que, para um outro estímulo qualquer diferente do estímulo discriminativo (estímulo que não possui uma relação operante associada) a atividade da rede permanece em nível operante. Para ilustrar essa característica, durante o condicionamento exibido nas Figura 26 e Figura 27, um estímulo não discriminativo é apresentado em alternância com o estímulo discriminativo. Após o treinamento, o estímulo não discriminativo não produz uma resposta específica, permanecendo a rede em nível operante (Figura 28). Isto ocorre pelo fato de não existir uma resposta desejada associada ao mesmo, ou seja, o ambiente não possui uma contingência relacionada a este estímulo em particular (tal estímulo não está associado a um estímulo reforçador US).

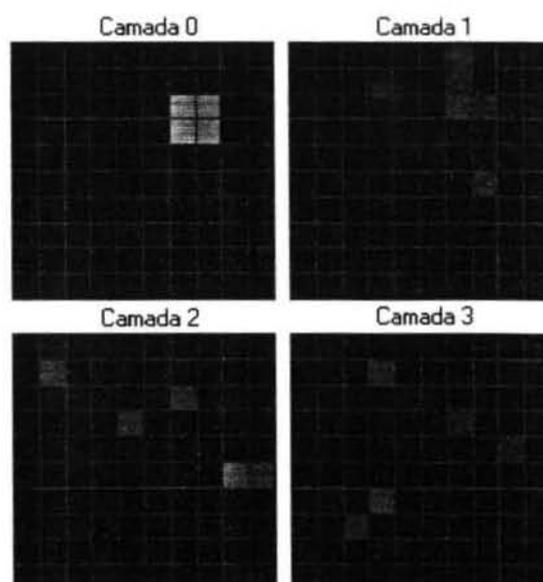


Figura 28 – Atividade neural após treinamento operante para um estímulo não reforçado.

A Figura 29 e a Figura 30 apresentam uma configuração alternativa para este experimento. No entanto, visam demonstrar o mesmo princípio. Neste caso, a resposta desejada é definida em neurônios que estão localizados em uma posição relativa diferente daquela assumida pelos neurônios que são estimulados. Dessa forma, a atividade neural dentro da rede é obrigada a propagar-se em um padrão diagonal, considerando-se as várias camadas. O objetivo aqui é mostrar que o treinamento é capaz de estabelecer conexões e de fato consolidar um padrão de disparos entre qualquer região da primeira e última camadas.

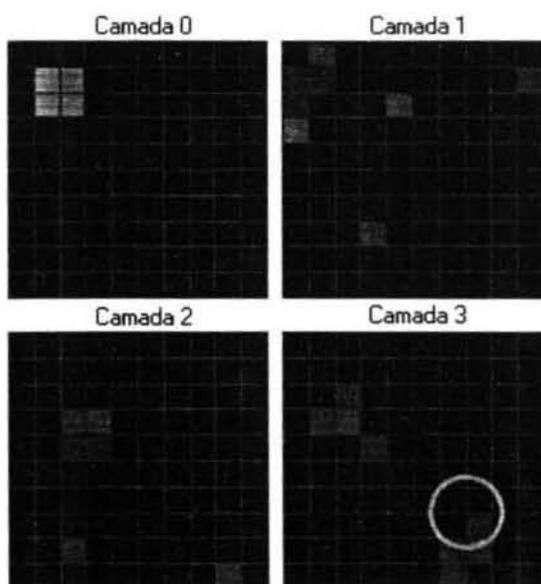


Figura 29 – Atividade neural anterior ao treinamento operante (exemplo 2).

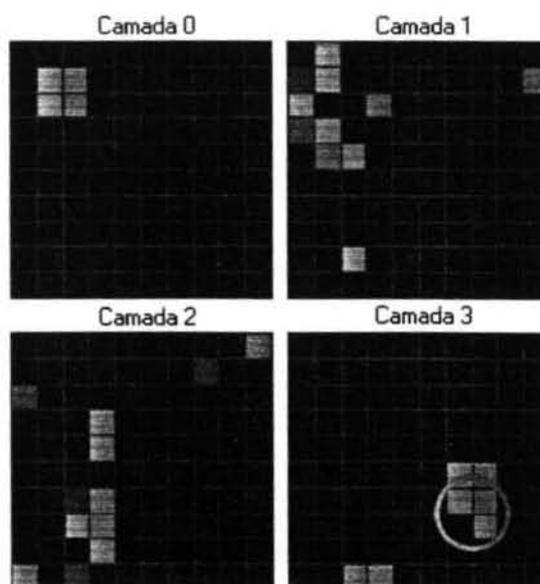


Figura 30 – Atividade neural após treinamento operante (exemplo 2).

Essa característica é interessante não apenas do ponto de vista comportamental, mas também do ponto de vista arquitetural. Pensando em termos de arquitetura, existe um favorecimento por parte das conexões sinápticas iniciais (devido ao padrão Gaussiano de distribuição) de que a atividade neural propague-se ascendentemente em linha reta, camada após camada. No entanto, o mecanismo de aprendizagem e a própria forma de operar da rede são capazes de superar essa tendência inicial, fortalecendo as conexões de forma adequada.

4.2.4 Experimento 4 – Extinção de Comportamento Operante.

Assim como é importante esquecer contingências clássicas não mais existentes (Experimento 2), também é fundamental para um sistema autônomo extinguir contingências operantes previamente estabelecidas que não são mais válidas.

Utilizando como ponto de partida o segundo treinamento realizado no experimento anterior (Figura 31), é feita uma pequena modificação na definição do ambiente, removendo-se a apresentação do US quando a resposta desejada é emitida na presença do estímulo discriminativo. Após algumas poucas iterações, ao longo das quais as conexões sinápticas são gradualmente enfraquecidas devido à ausência do reforço, a resposta da rede diminui e retorna ao padrão de nível

operante (Figura 32). Dessa forma a contingência previamente estabelecida é eliminada.

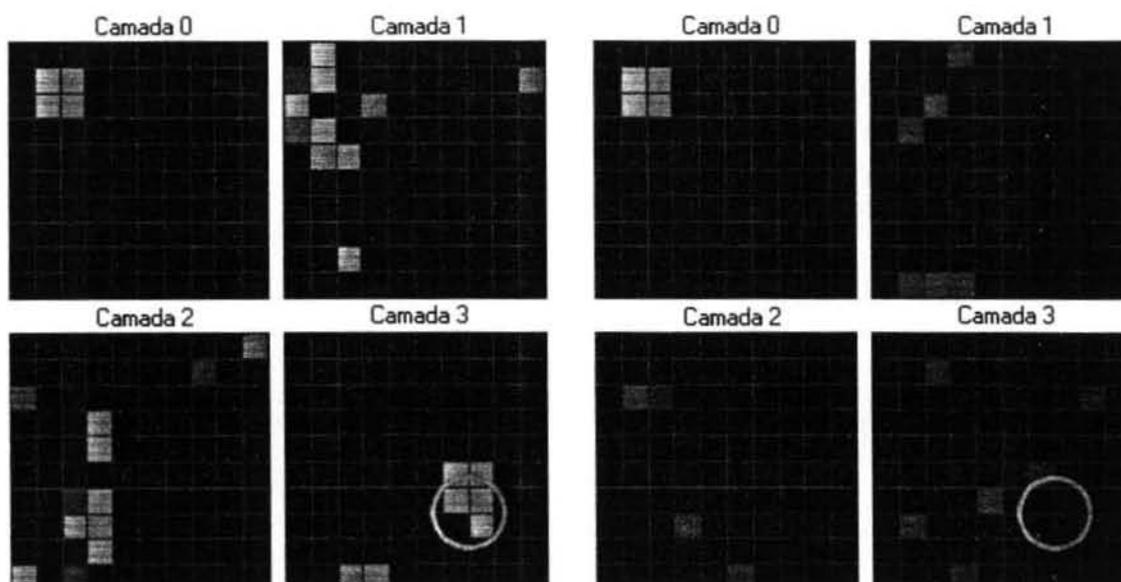


Figura 31 – Atividade neural anterior à extinção do comportamento operante.

Figura 32 – Atividade neural após extinção do comportamento operante.

4.2.5 Experimento 5 – Capacidade de Generalização

Uma das contribuições do modelo proposto para o desenvolvimento de sistemas autônomos inteligentes, que é herdada do modelo de redes neurais, é a capacidade de generalização. Através desta capacidade, uma rede neural pode emitir uma resposta particular para um estímulo, mesmo que não tenha feito parte do processo de aprendizagem, categorizando-o segundo sua semelhança com um estímulo que de fato tenha participado do treinamento.

Este experimento demonstra um treinamento operante simples. A Figura 33 apresenta a resposta inicial, apresentada pela rede antes do treinamento. Após o treinamento, o estímulo discriminativo (Camada 0 da Figura 34) adquire controle sobre a resposta desejada (Camada 3 da Figura 34). Até esse ponto o experimento não é diferente dos apresentados anteriormente. A

Figura 35, a Figura 36 e a Figura 37 ilustram a resposta da rede para estímulos semelhantes ao estímulo de treinamento que, no entanto, não fizeram parte do mesmo. Note que a resposta da rede permanece praticamente inalterada para o estímulo que é diferente apenas na atividade de um neurônio (

Figura 35). À medida que o estímulo apresentado torna-se cada vez mais discrepante em relação ao estímulo de treinamento (Figura 36 e Figura 37), a

atividade torna-se menos similar. No entanto, mesmo para um estímulo consideravelmente degradado (Figura 37), a resposta da rede ainda assemelha-se bastante à resposta original obtida com o treinamento.

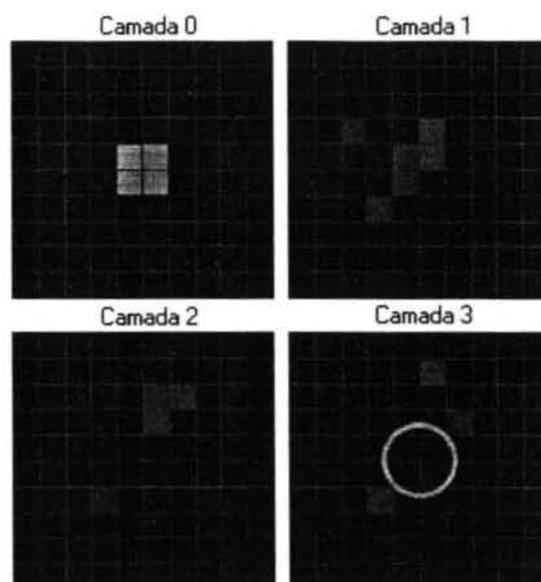


Figura 33 – Atividade neural anterior ao treinamento operante (exemplo de generalização).

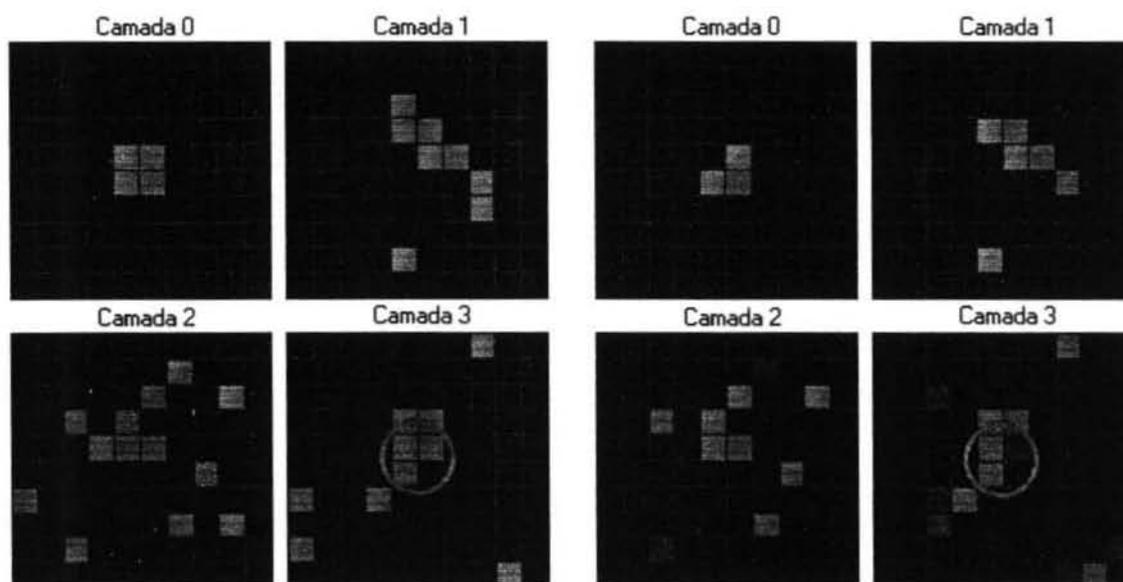


Figura 34 – Atividade neural após treinamento operante com estímulo original.

Figura 35 – Atividade neural após treinamento operante com degradação do estímulo (exemplo 1).

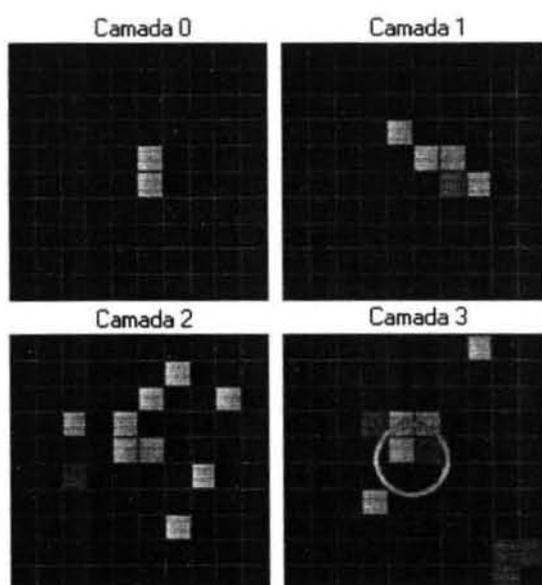


Figura 36 – Atividade neural após treinamento operante com degradação do estímulo (exemplo 2).

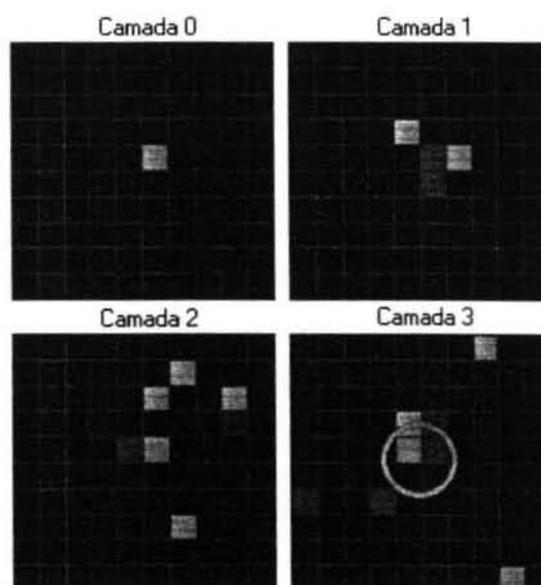


Figura 37 – Atividade neural após treinamento operante com degradação do estímulo (exemplo 3).

Sem dúvida, a capacidade de generalização é fundamental para o desenvolvimento de sistemas autônomos inteligentes e o fato do modelo proposto incorporar tal qualidade acrescenta valor ao mesmo, principalmente diante das expectativas de sua utilização.

4.2.6 Experimento 6 – Controle de Intensidade da Resposta

Certas situações exigem que um sistema autônomo seja capaz, além de emitir uma resposta correta em uma determinada situação, emitir essa mesma resposta com uma certa intensidade particular para a circunstância atual. Considere um exemplo em que um robô autônomo cujo repertório comportamental inclui o controle de sua velocidade de deslocamento. Em certas situações ele pode desenvolver velocidade máxima, por exemplo, quando se desloca em linha reta, acelerando até o limite de sua capacidade. Em outras ocasiões pode ser necessária uma velocidade diferente, por exemplo, em uma curva onde, a fim de evitar derrapagens, a velocidade desenvolvida deva ser menor.

O experimento a seguir foi montado tendo-se em mente situações como a descrita acima e visa avaliar a capacidade do modelo em associar uma mesma resposta a dois estímulos diferentes, sendo que, a resposta envolve os

mesmos neurônios, porém, disparando com intensidades diferentes para os diferentes estímulos. Para tanto o estímulo número 1 (Camada 0 da Figura 38) foi associado à resposta em questão (Camada 3 da Figura 38) sem nenhuma restrição, assim como o treinamento operante do Experimento 3. O estímulo número 2 (Camada 0 da Figura 39) por sua vez está associado à mesma resposta. No entanto, neste caso, o estímulo reforçador só é apresentado caso a atividade nos neurônios resposta não ultrapassem a marca de 0.5.

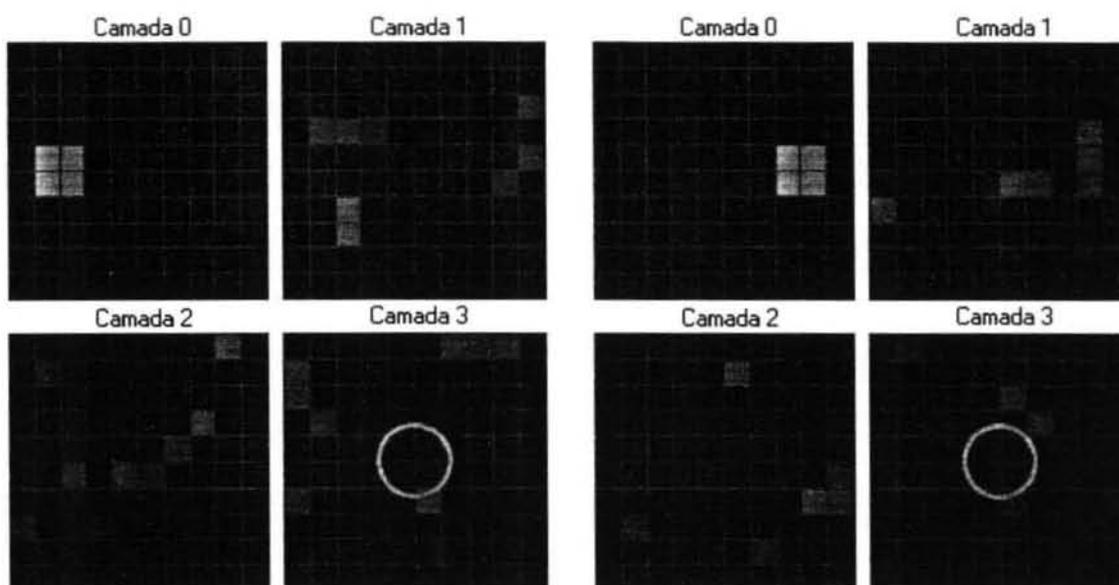


Figura 38 – Atividade neural anterior ao treinamento de controle da intensidade da resposta (estímulo 1).

Figura 39 – Atividade neural anterior ao treinamento de controle da intensidade da resposta (estímulo 2).

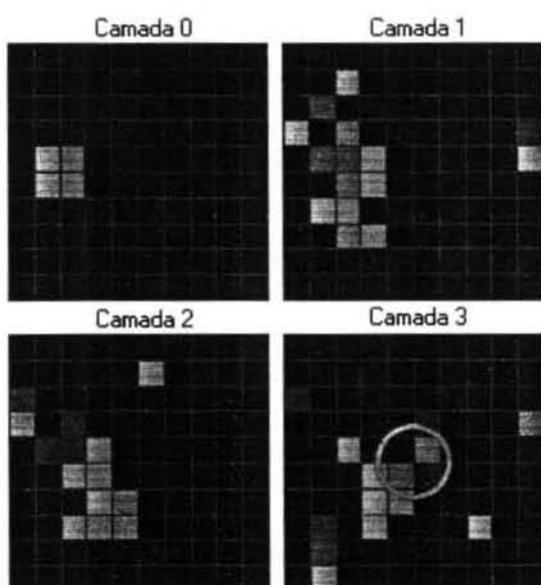


Figura 40 – Atividade neural após treinamento de controle da intensidade da resposta (estímulo 1).

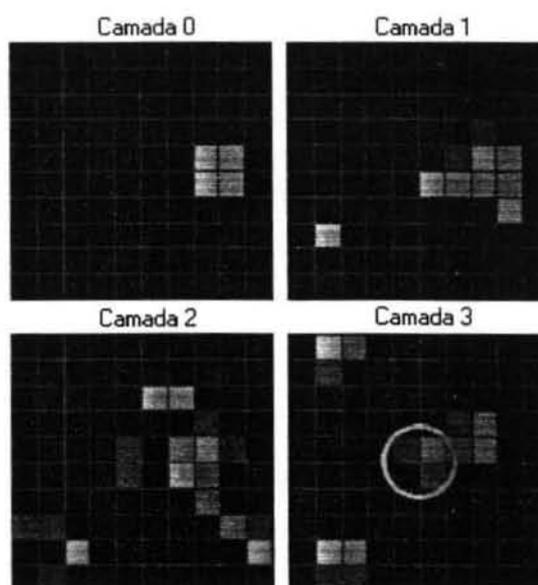


Figura 41 – Atividade neural após treinamento de controle da intensidade da resposta (estímulo 2).

Como consequência do treinamento a resposta neural que o estímulo número 1 produz é bem mais intensa (Figura 40) que a provocada pelo estímulo número 2 (Figura 41) (cerca de 50%), levando-se em conta a atividade do grupo que define a resposta para cada um dos estímulos. Na prática, o que ocorre para o segundo estímulo é que a atividade, ao longo do treinamento, gradualmente eleva-se. À medida que a atividade neural passa do limite de 0.5 ela começa a sofrer extinção, pois não recebe US. Com a extinção, a atividade volta a cair abaixo do limite de 0.5, o que propicia um novo fortalecimento e nova elevação do nível de atividade. Esse ciclo repete-se pelo tempo que durar a simulação.

Treinada dessa maneira, a rede neural pode aprender a distinguir situações para as quais a resposta desejada deve ser emitida com intensidade particular.

4.2.7 Experimento 7 – Controle de Múltiplas Respostas

Na maioria dos casos, situações distintas requerem do sistema autônomo inteligente a emissão de comportamentos distintos. Portanto, o modelo deve permitir a associação de estímulos a respostas distintas, sem que haja interferência de uma relação na outra.

O treinamento, para este experimento, consistiu da apresentação aleatória de dois estímulos distintos, sendo que cada um deles está associado a uma resposta diferente. O sinal de reforço é apresentado sempre que a resposta correta para o estímulo corrente for apresentada. Assim, quando estimulada pelo estímulo 1 (Camada 0 da Figura 42), a rede recebe o reforço apenas se emitir a resposta 1 (Camada 3 da Figura 42). Da mesma forma, quando o estímulo corrente é o estímulo 2 (Camada 0 da Figura 43), o reforço se torna disponível apenas mediante emissão da resposta 2 (Camada 3 da Figura 43).

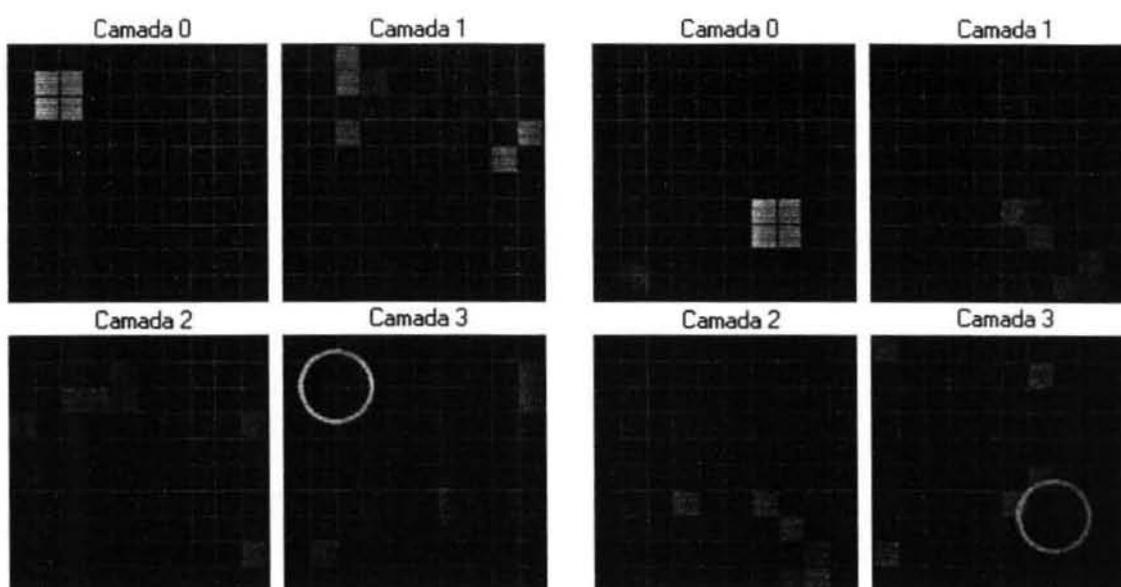


Figura 42 – Atividade neural anterior ao treinamento de controle de múltiplas respostas (estímulo 1 – resposta 1).

Figura 43 – Atividade neural anterior ao treinamento de controle de múltiplas respostas (estímulo 2 – resposta 2).

Após o treinamento (Figura 44 e Figura 45) a rede é capaz de emitir as respostas corretas, para ambos os estímulos, sem que haja interferência (por exemplo, sem que o estímulo 1 produza a resposta 2).

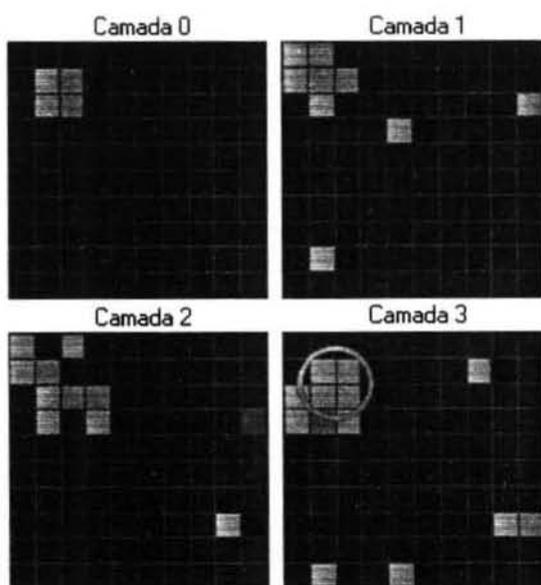


Figura 44 – Atividade neural após treinamento de controle de múltiplas respostas (estímulo 1 – resposta 1).

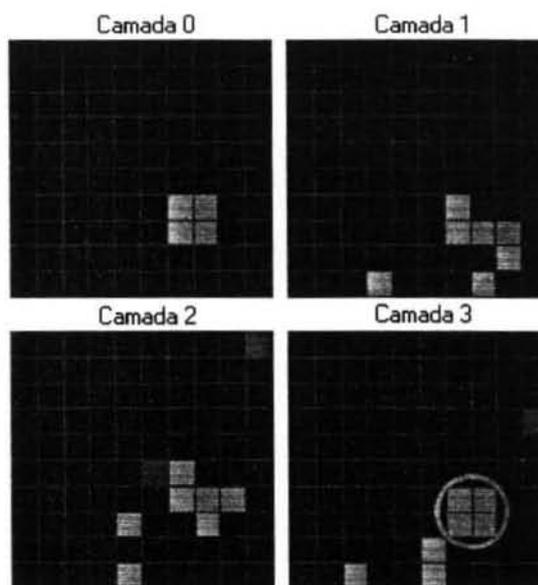


Figura 45 – Atividade neural após treinamento de controle de múltiplas respostas (estímulo 2 – resposta 2).

Este experimento aborda uma questão interessante quanto à capacidade de memória do modelo. Entretanto, o objetivo aqui não é analisar quantitativamente a capacidade da rede neural em armazenar diversas relações distintas, o objetivo é somente demonstrar que o modelo possui essa funcionalidade. Em investigações futuras será analisada a capacidade de memória do modelo.

4.2.8 Experimento 8 – Fusão de Sensores

Um sistema autônomo inteligente deve aprender a emitir comportamentos relevantes (respostas) em situações relevantes (estímulos discriminativos). Para tanto, deve ser capaz de identificar tais situações, discriminando-as das demais. Essa capacidade não é trivial, como pode parecer em princípio. Estímulos discriminativos podem apresentar-se de forma bastante complexa, envolvendo a combinação de diferentes estímulos em sua configuração. Um sistema autônomo inteligente deve ser capaz de distinguir um estímulo discriminativo complexo de situações envolvendo apenas alguns de seus componentes.

Em seres biológicos, situações complexas normalmente envolvem a associação de estímulos provenientes de diferentes campos sensoriais. Assim, um

organismo biológico deve ser capaz de distinguir uma situação reforçada, envolvendo, por exemplo, um estímulo visual e um auditivo, das situações em que estes estímulos aparecem isoladamente. Por normalmente envolver diferentes tipos de informação sensorial, esse problema é usualmente conhecido por fusão de sensores.

O experimento apresentado tem por objetivo avaliar as capacidades da rede neural proposta em associar estímulos complexos a uma resposta desejada, sem que, no entanto, tal associação também ocorra com cada parte independente do estímulo. Embora não seja feita qualquer especificação quanto à natureza dos estímulos ou sua origem, se estes são ou não provenientes de campos sensoriais diferentes, esse fator não limita os resultados obtidos uma vez que a definição do modelo e de sua camada sensorial não especifica qualquer parâmetro quanto à natureza da informação que recebe e processa. Assim sendo a informação sensorial que estimula a camada de entrada pode ser de qualquer tipo, inclusive proveniente de campos sensoriais diferentes.

O estímulo discriminativo a ser associado (Camada 0 da Figura 46) à resposta desejada (Camada 3 da Figura 46) envolve dois grupos bem distintos de neurônios estimulados. O treinamento consiste em apresentar o estímulo discriminativo para a rede e, sempre que a resposta emitida for a desejada, apresentar o reforço. No entanto, juntamente com o estímulo discriminativo, são apresentados seus componentes individuais (Camada 0 da Figura 47 e Camada 0 da Figura 48), de forma que o padrão de estimulação apresentado à rede a cada iteração é sorteado dentre estes três possíveis casos (estímulo discriminativo, componente 1 e componente 2). Os estímulos componentes 1 e 2, no entanto, não são estímulos discriminativos e não possuem um reforçador associado ou resposta desejada.

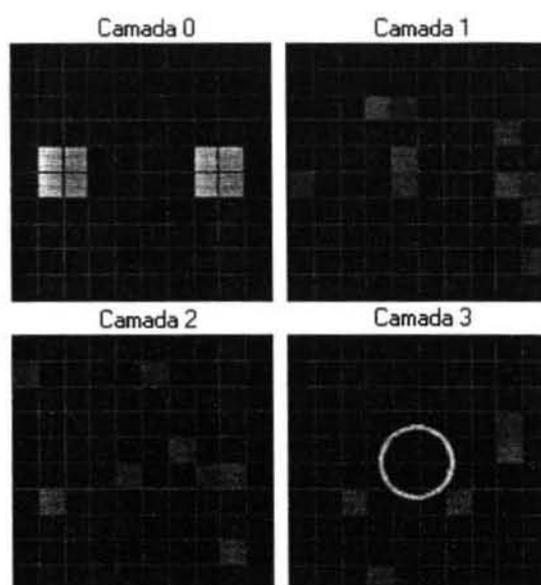


Figura 46 – Atividade neural anterior ao treinamento de fusão sensorial (estímulo discriminativo composto).

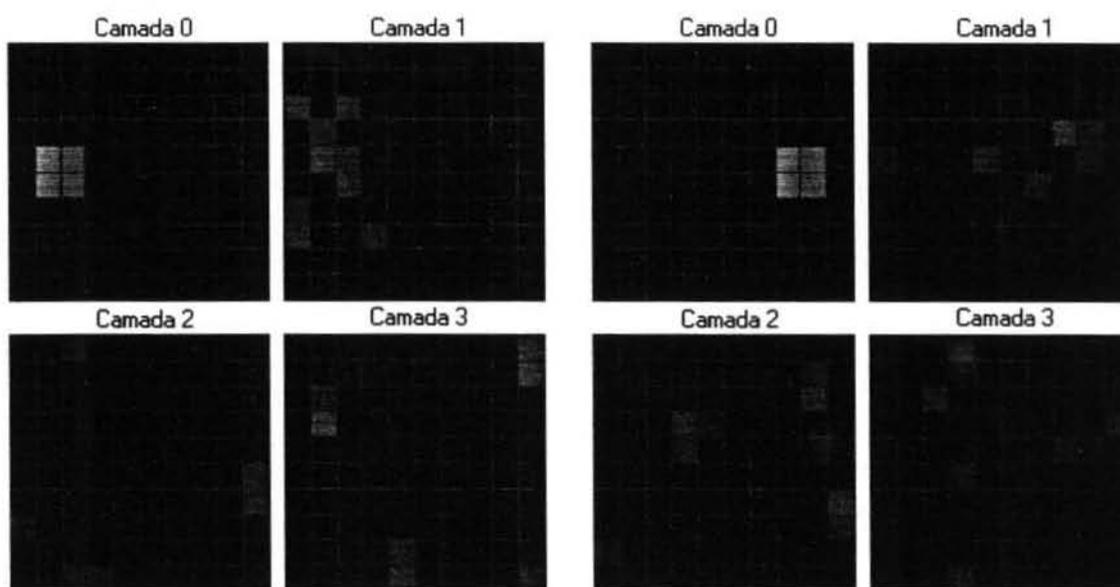


Figura 47 – Atividade neural anterior ao treinamento de fusão sensorial (componente 1).

Figura 48 – Atividade neural anterior ao treinamento de fusão sensorial (componente 2).

Após o treinamento, o estímulo discriminativo adquiriu a capacidade de produzir a resposta desejada (Figura 51), sem que seus componentes individuais também fossem associados (Figura 49 e Figura 50). Portanto, a rede é capaz de operar em uma situação envolvendo um problema de fusão sensorial.

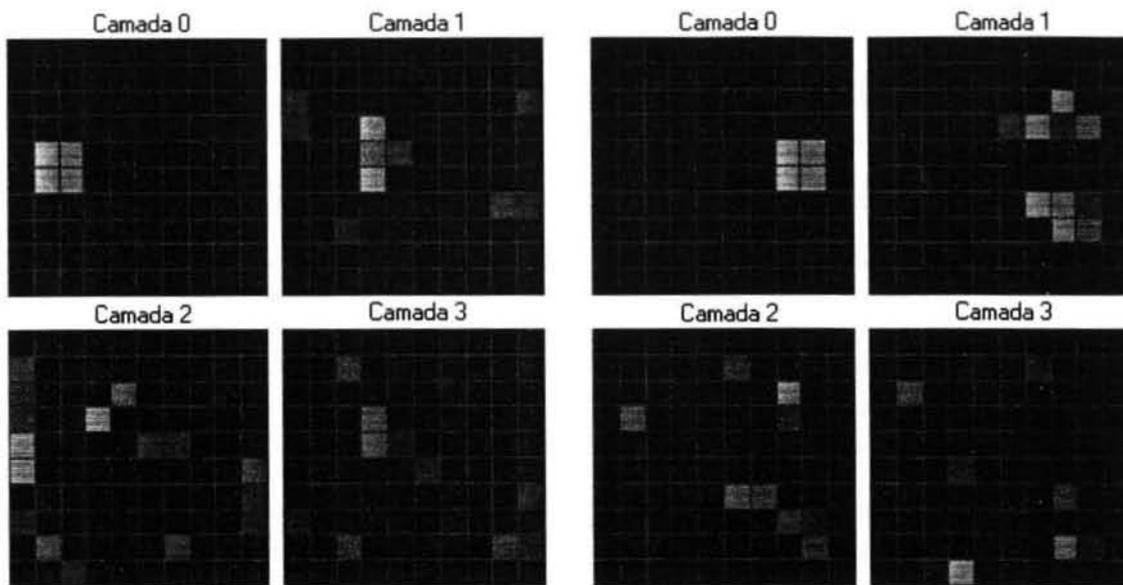


Figura 49 – Atividade neural após treinamento de fusão sensorial (componente 1).

Figura 50 – Atividade neural após treinamento de fusão sensorial (componente 2).

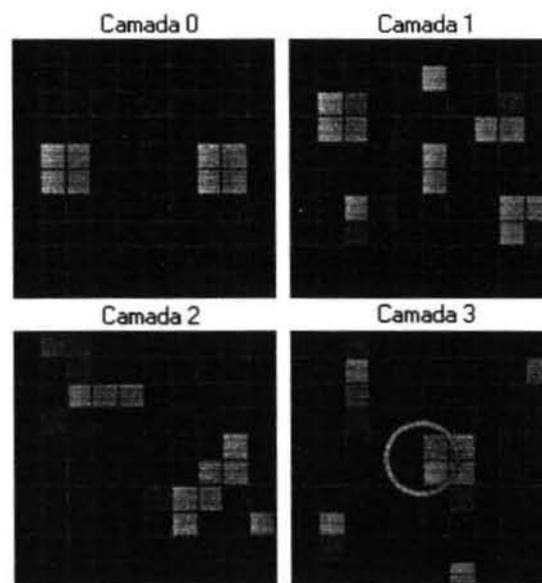


Figura 51 – Atividade neural após treinamento de fusão sensorial (ocorrência conjunta dos estímulo 1 e 2).

A importância dos estímulos componentes participarem do treinamento é fundamental. O fato de serem apresentados, sem que jamais sejam seguidos pela apresentação de um reforçador, faz com que o algoritmo de aprendizagem trabalhe em modo de extinção para os mesmos. Dessa forma, as sinapses utilizadas durante a apresentação individual dos estímulos componentes são enfraquecidas, restando apenas as sinapses utilizadas quando os dois

componentes aparecem juntos, ou seja, quando o estímulo é o discriminativo. Gradualmente esse processo tende a selecionar apenas os caminhos neurais ativos quando o estímulo discriminativo está presente, e não aqueles ativados pelos seus componentes.

4.2.9 Experimento 9 –Treinamento Operante em uma Rede com Camadas de Dimensão 20x20.

Nos experimentos apresentados até o momento, procurou-se manter a mesma arquitetura. No entanto, como foi destacado durante a apresentação do modelo, os parâmetros que definem a rede neural são suficientemente flexíveis e permitem diversas configurações possíveis. O atual experimento ilustra essa característica. Trata-se de um treinamento operante simples, como o do Experimento 3, apenas as dimensões das camadas foram alteradas, passando de uma grade de 10x10 neurônios para uma de 20x20 neurônios (Figura 52 e Figura 53). Todos os demais parâmetros permanecem iguais aos dos experimentos anteriores.

Para facilitar a visualização, não é ilustrada a separação em grade dos neurônios, uma vez que a observação das dimensões reduzidas dos neurônios seria prejudicada pela mesma.

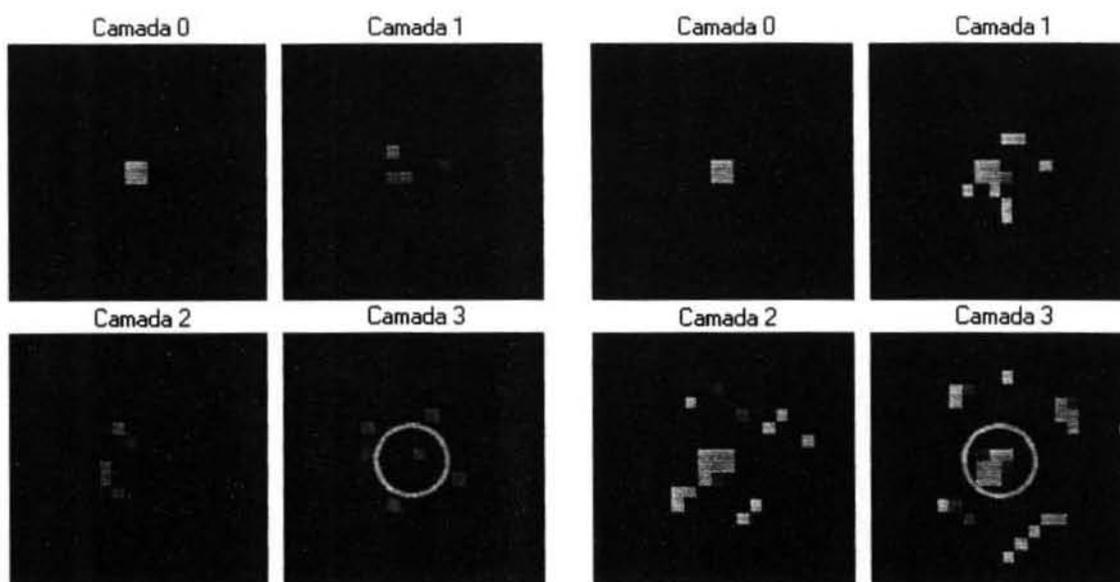


Figura 52 – Atividade neural anterior ao treinamento operante em rede com camadas 20x20.

Figura 53 – Atividade neural após treinamento operante em rede com camadas 20x20.

Especificou-se com neurônios resposta um conjunto de 16 neurônios no centro do círculo vermelho (Camada 3 da Figura 52 e Figura 53). O treinamento ocorre normalmente, produzindo ao final um grupo neural localizado na região da resposta desejada. Entretanto, nota-se ao redor do grupo que define a resposta a formação de alguns grupos neurais extras. Isso se deve ao fato de que, apesar das dimensões das camadas terem sido ampliadas, o mesmo não aconteceu para o número de sinapses e principalmente para o raio de distribuição das mesmas. Conseqüentemente, o tamanho dos grupos neurais e a região de influência de cada um não foi expandida na mesma proporção das camadas. Permitiu-se dessa maneira a formação de um maior número de grupos neurais, uma vez que o grupo principal não possui influência sobre toda a rede. Também é devido ao padrão de conexões o fato do grupo que define a resposta não ter se expandido para compreender todos os 16 neurônios resposta.

4.2.10 Experimento 10 – Análise da Aprendizagem Operante.

A fim de proporcionar uma análise mais formal da aprendizagem, a Figura 54 apresenta um gráfico da evolução de um treinamento operante típico (evolvendo um estímulo e uma resposta apenas).

O gráfico representa a atividade conjunta dos quatro neurônios resposta especificados. Assim, por exemplo, uma atividade de 100% indica que todos os quatro neurônios estão disparando e com intensidade máxima.

Existem três regiões distintas destacadas no gráfico. A primeira é anterior ao início da aprendizagem e visa mostrar o comportamento típico da rede em nível operante (região esquerda do gráfico). A segunda região (parte central do gráfico – Período de Aquisição) começa no marco zero e descreve o período onde o estímulo reforçador é apresentado em consequência da emissão da resposta correta. Nesta fase, nota-se um rápido aumento na atividade dos neurônios resposta (em torno das primeiras 14 iterações), o que mostra que a resposta correta está sendo associada ao estímulo correspondente. Também é interessante notar que a atividade sempre demonstra algum nível de variação. Próximo da iteração nº 64 as variações produziram uma melhora na resposta (provavelmente ativando um dos neurônios resposta que ainda não estava respondendo). Tal alteração, que significa uma melhora, foi rapidamente assimilada pela rede. A última região (parte direita do gráfico – Período de Extinção) mostra o comportamento da rede após a remoção do estímulo reforçador (US). Assim, uma vez que a recompensa pela emissão da resposta correta deixa de ser apresentada (iteração nº 174), a rede tende a extinguir a associação previamente estabelecida, retornando a um padrão de atividade semelhante ao de nível operante, o que ocorreu por volta da iteração nº 214.

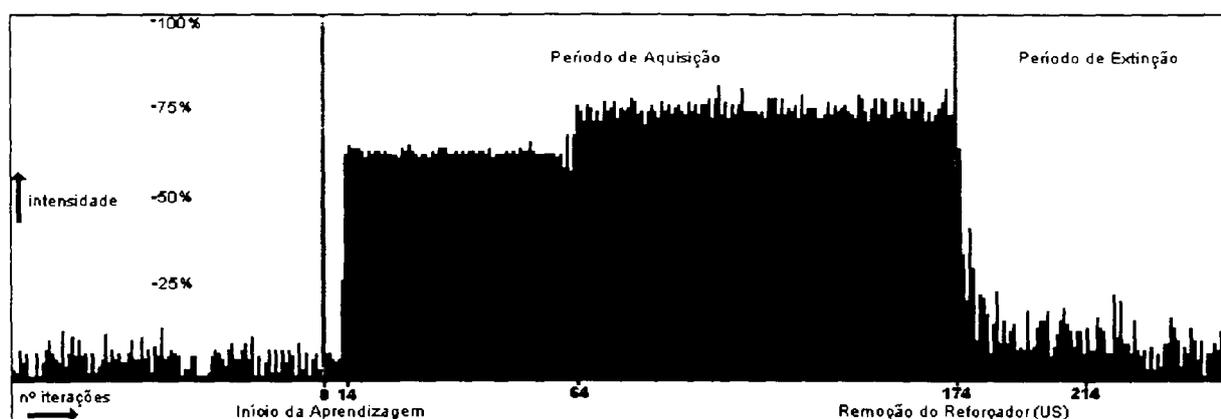


Figura 54 – Gráfico de evolução de um treinamento operante típico.

5 CONCLUSÕES

Sistemas autônomos inteligentes consistem em um avanço tecnológico que promete revolucionar a sociedade. Seu desenvolvimento é, no entanto, uma tarefa extremamente complexa. De fato, ainda é impossível criar sistemas artificiais capazes de operar de forma satisfatória em ambientes do cotidiano. Acredita-se que dotar um sistema artificial com características autônomas seja um passo chave no processo de aperfeiçoamento dos sistemas atuais.

Ser autônomo no sentido de auto-suficiente e capaz de operar em ambientes diversificados requer do sistema constante aprimoramento e atualização. Torna-se, portanto, indispensável a presença de mecanismos de aprendizagem que possibilitem ao sistema adquirir conhecimento de forma a aprimorar seu desempenho. A capacidade de aprendizagem é tida assim como fator decisivo para que um sistema seja verdadeiramente autônomo.

Considera-se que os únicos sistemas plenamente autônomos são os biológicos. Assim sendo, consistem em uma fonte segura de informação e inspiração. Reproduzir as características de tais sistemas em modelos artificiais análogos é uma linha de pesquisa interessante e promissora. Define-se assim o fator plausibilidade biológica, que motiva grande parte das idéias do trabalho.

Este trabalho propõe um modelo de redes neurais de múltiplas camadas cujas características são inspiradas em modelos biológicos. A escolha de um modelo de múltiplas camadas deve-se principalmente a três fatores:

1. Para que uma rede de múltiplas camadas tradicional, como um perceptron, possua a característica de aproximador universal de funções, é necessária a presença da camada intermediária;
2. Certos problemas, como o de fusão sensorial, parecem requerer a presença de múltiplas camadas;

3. O fator da plausibilidade biológica conduz a uma configuração de múltiplas camadas, uma vez que sistemas orgânicos assim o são.

Além desses fatores, observou-se durante os experimentos que uma configuração em múltiplas camadas permite que a rede evolua gradualmente para uma resposta específica.

Ainda com relação à arquitetura, o modelo emprega um modelo de conexões sinápticas que favorece a formação de grupos neurais (Edelman, 1987). Para tanto, são utilizados três tipos de conexões sinápticas: excitatórias intercamadas, laterais inibitórias intra-camada e laterais excitatória intra-camada.

Utiliza-se como modelo de aprendizagem uma abordagem da teoria da Seleção pelas Conseqüências, sendo, portanto, um modelo de aprendizagem por reforço. A utilização de um modelo de aprendizagem por reforço é importante uma vez que permite ao sistema adquirir comportamentos através da interação com o ambiente, sem a necessidade de uma supervisão. Além disso, o fato particular do modelo de aprendizagem ser baseado na teoria da Seleção pelas Conseqüências promove características interessantes e motivadoras, entre elas:

1. O modelo é perfeitamente condizente com o fator da plausibilidade biológica;
2. O princípio de aprendizagem é por reforço;
3. Torna-se possível avaliar as capacidades de aprendizagem do modelo segundo os fenômenos comportamentais descritos na teoria da Seleção pelas Conseqüências;
4. A interação com o ambiente de um sistema com tais características acontece de forma natural, visto que a teoria descreve e incorpora fatores do ambiente cotidiano. Por exemplo, considere o exemplo de um estímulo reforçado. Este é algo concreto existente no ambiente, não sendo necessária a criação de mecanismos adicionais que promovam sua integração com o sistema artificial. Basta existir um sensor adequado que capture o estímulo.

Para o uso conjunto das teorias de Redes Neurais Artificiais e da Seleção pelas Conseqüências, é necessário integrar as duas de alguma forma. Para tanto, se recorre à neurobiologia, em especial ao trabalho de Donahoe, Burgos e Palmer (1993) e Donahoe e Palmer (1994).

Resultados e Contribuições

Os resultados obtidos em simulação demonstram os mecanismos de funcionamento da rede, incluindo a forma de propagação da atividade neural ao longo das camadas, a importância das conexões laterais inibitórias e a formação dinâmica de grupos neurais.

Fica claro que a existência de conexões laterais inibitórias é fundamental para a correta modelagem de respostas em nível operante. Estas por sua vez formam a base do princípio de aprendizagem empregado.

Além de modelarem as respostas em nível operante, as conexões laterais inibitórias, juntamente com as conexões excitatórias intra-camadas, são responsáveis pela formação dinâmica de grupos neurais que ocorre durante a aprendizagem. São os grupos neurais (Edelman, 1987) que codificam as respostas emitidas pela rede, sendo que seu modo de funcionamento, por assim dizer, dita as regras de operação da rede.

Os resultados obtidos também incluem experimentos de aprendizagem, onde foi possível reproduzir diversos fenômenos comportamentais de que trata o modelo de Seleção pelas Conseqüências. Tais experimentos incluem a aquisição de comportamento respondente, extinção de comportamento respondente, aquisição de comportamento operante, extinção de comportamento operante, capacidade de generalização da rede, habilidade de controle de intensidade da resposta, capacidade de controle de múltiplas respostas e capacidade de fusão sensorial. Por fim foi apresentada uma análise gráfica da evolução de um treinamento operante típico.

Os experimentos de aquisição e extinção de comportamentos respondentes e operantes mostram as capacidades de convergência do modelo, reproduzindo tais fenômenos comportamentais.

O experimento de generalização demonstra que o modelo exibe de forma satisfatória a capacidade das redes neurais de generalizar estímulos,

classificando estímulos desconhecidos segundo sua similaridade com estímulos conhecidos.

No experimento de controle de intensidade de resposta, a rede demonstra a capacidade de controlar o nível de intensidade de emissão da resposta. Como destacado, tal habilidade pode ser interessante para aplicações como o controle de velocidade em robôs.

A fim de demonstrar que a rede é capaz de associar estímulos distintos a respostas também distintas, realizou-se um experimento de controle de múltiplas respostas, onde o objetivo é o treinamento simultâneo de duas relações distintas (estímulo 1 – resposta 1 e estímulo 2 – resposta 2). Assim como esperado, a rede foi capaz de aprender as devidas relações sem que uma associação interferisse na outra.

Um outro fenômeno interessante reproduzido em experimento é a fusão sensorial. Segundo a perspectiva desse problema, a rede é capaz de associar diferentes estímulos e é capaz de distinguir situações nas quais os estímulos aparecem individualmente das situações em que eles aparecem em conjunto, emitindo a resposta adequada para cada situação.

As contribuições do trabalho não são limitadas aos fenômenos reproduzidos em experimentos.

Existem poucos modelos de redes multi-camadas com aprendizagem por reforço que se prestam a problemas de controle. Menos ainda quando a aprendizagem por reforço é do tipo clássico.

O modelo apresenta uma solução para o problema da atribuição de crédito utilizando aprendizagem por reforço. Esse problema trata a dificuldade que existe, em uma rede de múltiplas camadas, em atribuir responsabilidade a neurônios de camadas intermediárias. Pode-se argumentar que a atribuição de crédito consiste no problema fundamental que uma rede de múltiplas camadas deve solucionar, sendo imprescindível para a validade do modelo. Uma solução que utiliza aprendizagem por reforço é ainda mais interessante em vista de ser pouco comum.

Trabalhos Futuros

A natureza multidisciplinar do trabalho abre um leque bastante amplo de possibilidades de trabalhos futuros. Praticamente todos os aspectos do trabalho podem ser aprimorados.

Além de alguns aprimoramentos menores, o próximo passo lógico é implementar os mecanismos que possibilitam o desenvolvimento de reforçadores adquiridos, com implicações cruciais para as capacidades comportamentais do sistema. Tais mecanismos já foram descritos por Donahoe e Palmer (1994), no entanto, não foram implementados de forma satisfatória.

Um ponto importante é solucionar uma dificuldade quanto à aprendizagem das conexões laterais, tanto inibitórias quanto excitatórias. Durante os experimentos realizados não foi permitido a estes tipos de conexões alterar o valor do peso sináptico associado. Decidiu-se assim, pois problemas de aprendizagem que impedem a correta atualização desses tipos de sinapses estavam atrasando o projeto demasiadamente.

Uma importante contribuição futura é possível através de um desenvolvimento formal de teorias que descrevam o processo de convergência da rede.

Uma análise interessante pode ser feita quanto à capacidade de memória do modelo. Respondendo perguntas tais como: quantas relações distintas estímulo-resposta à rede é capaz de aprender ao mesmo tempo?

Outra etapa prevista é o desenvolvimento de uma simulação de robôs e ambientes, utilizando o modelo proposto como controlador do robô. Espera-se que em uma simulação desse tipo o modelo seja capaz de solucionar questões como navegação autônoma, comportamento coletivo e até mesmo o uso de ferramentas. O projeto do simulador está bastante avançado, sendo que um protótipo chegou a ser implementado.

Com relação à modelagem neural, existem vários pontos para o trabalho evoluir. Por exemplo, o modelo dos neurônios utilizado é simplificado, assim como visto na maioria de trabalho com redes neurais artificiais. É possível, no entanto, realizar uma simulação muito mais precisa utilizando-se um modelo que represente, por exemplo, canais independentes de íons, fluxo de moléculas, etc. Este tipo de modelo é mais utilizado por neurocientistas para estudos na área.

Nessa mesma linha, existem trabalhos que modelam redes auto-organizadas com comportamento oscilatório, algo que provavelmente possui implicações para uma representação de comportamentos de nível operante.

REFERÊNCIAS BIBLIOGRÁFICAS

Brooks, R., A., **Elephants Don't Play Chess**, *Robotics and Autonomous Systems*, vol. 6, pp. 3-15, 1990.

Baum, M. W., **Compreender o Behaviorismo: Ciência, Comportamento e Cultura**, Editora Artes Médicas Sul, Porto Alegre, 1999.

Brooks, R., A., **Intelligence Without Reason**, Proceedings of International Joint Conference on Artificial Intelligence, pp. 569-595, Sydney, Australia, agosto, 1991

de Castro, L. e Timmis, J.; **Artificial Immune Systems: A New Computational Intelligence Approach**, Springer Verlag, 2002.

Donahoe, J., W., Burgos J. E., Palmer D. C., **A Selectionist Approach to Reinforcement**, *Journal of the Experimental Analysis of Behavior*, nº 60, pp. 17-40, 1993.

Donahoe, J., W. e Palmer, D., C., **Learning and Complex Behavior**, Allyn and Bacon, U.S.A, 1994.

Dorigo, M., **Introduction to the Special Issue on Learning Autonomous Robots**, *IEE Transactions on Systems, Man and Cybernetics-PartB: Cybernetics*, vol. 26, nº 3, 1996.

Edelman, G., **Neural Darwinism: The Theory of Neuronal Group Selection**, Basic Books, EUA, 1987.

Bonabeau, E., Dorigo, M., e Theraulaz, G., **From Natural to Artificial Swarm Intelligence**, Oxford University Press, 1999.

Fernandes, F., Luft, C.P., Guimarães, F.M., **Dicionário Brasileiro Globo**, 18.

ed., São Paulo, Globo, 1991.

Figueiredo, M., F., **Redes Neurais Nebulosas Aplicadas em Problemas de Modelagem e Controle Autônomo**, dissertação de doutorado, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, 1997.

Figueiredo, M., F., **Navegação Autônoma de Robôs**, Livro da VII ERI (Escola de Informática da SBC – Regional Sul), pp.74-106, maio, 1999.

Haykin, S., **Redes Neurais – Princípios e Prática**, Bookman, Porto Alegre, 2001.

Hopfield, J.J., **Neural Networks and Physical Systems with Emergent Collective Computational Abilities**, Proceedings of the National Academy of Sciences, USA, vol.79, pp. 2554-2558, 1982.

Kohonen, T., **Self-Organized Formation of Topologically Correct Feature Maps**, Biological Cybernetics, vol.43, pp.59-69, 1982.

Krose, B., J., A., **Learning From Delayed Rewards**, Robotics and Autonomous Systems, nº 15, pp. 233-235, 1995.

Lanzi, P., L., **Learning Classifier Systems From Foundations to Applications**, Springer-Verlag, 2000.

Millenson, J., R., **Princípios da Análise do Comportamento**, Coordenada, Brasil, 1967.

Minsky, M.L., S.A. Papert, **Perceptrons**, Cambridge, MA: MIT Press, 1969.

Penrose, R., **A Nova Mente do Rei**, Editora Campus, Rio de Janeiro, 1997.

Ribeiro, C., **Reinforcement Learning Agents**, Artificial Intelligence Review,

nº 17, pp. 223-250, 2002.

Rolls, E. T. e Treves A., **Neural Networks and Brain Function**, Oxford University Press, Oxford, Inglaterra, 1998.

Rolls, E., T., **The Brain and Emotion**, Oxford University Press”, Oxford, Inglaterra, 1999.

Rumelhart, D.E., and J.L. McClelland, **Parallel Distributed Processing**, MIT Press, 1986.

Russell, S., J. e Norvig, P., **Artificial Intelligence – A modern Approach**, Prentice Hall, 1995.

Skinner, B. F., **Ciência e Comportamento Humano**, 5.ed., Editora Martins Fontes, São Paulo, 1981.

Steels, L., **When Are Robots Intelligent Autonomous Agents?**, Journal of Robotics and Autonomous Systems, nº 15, pp. 3-9, 1995.

Sutton, R. e Barto, G., **Reinforcement Learning: an Introduction**, MIT Press, 1998.

Vapnik, V.; **The Nature of Statistical Learning Theory**, Springer, New York, 1999.

Verschure, P.,F.,M.,J., **Formal Minds and Biological Brains**, IEEE Expert, pp. 66-75, outubro, 1993.

Verschure, P.,F.,M.,J., **Minds, brains, and robots: Explorations in distributed adaptive control**, Proceedings of the Second Brazilian-International Conference on Cognitive Science, Universidade Estadual do Norte Fluminense, Campos, Brazil, 1996.

ANEXO

Anexo I – Ferramenta de Simulação

A fim de testar as capacidades do modelo foi desenvolvido um simulador para a rede neural sendo que a linguagem de programação utilizada é C++.

O simulador consiste basicamente de três (3) telas. A tela principal (Figura 55) agrupa os principais controles necessários durante a simulação (localizados na parte central da tela). Entre eles os que permitem habilitar e desabilitar a aprendizagem, o uso de estímulos definidos na configuração do ambiente ou estímulos especificados pelo usuário (que os especifica clicando na figura da Camada 0 o que permite “ligar” e “desligar” neurônios específicos), definir quais neurônios da última camada serão respondentes, habilitar o uso independente dos três tipos de conexões sinápticas existentes, iniciar e parar a simulação, executá-la passo a passo e por fim reconectar a rede (sorteando novamente as conexões).

Além dos controles da simulação a tela principal também fornece uma interface de visualização da atividade nas várias camadas da rede (lado direito da tela), do padrão de sinapses de cada neurônios (cuja visualização é feita clicando-se em um neurônio qualquer da rede) inclusive permitindo visualizar o peso de cada sinapse (posição inferior esquerda da tela), e também um histórico da atividade recente de alguns neurônios específicos (entre eles o US) bem como de sinais que influenciam a rede, entre eles o sinal de reforço $d(T) = US$ (posição esquerda superior da tela).

Em outra tela (Figura 56) é possível editar os principais parâmetros da rede neural para cada camada da mesma, entre eles: o número de conexões sinápticas de cada tipo, seu raio de conexões, os parâmetros das equações de aprendizagem e raciocínio e etc. Entretanto existem alguns dos parâmetros descritos no modelo que não podem ser definidos em tempo de execução, entre eles, o intervalo inicial para os valores dos pesos sinápticos.

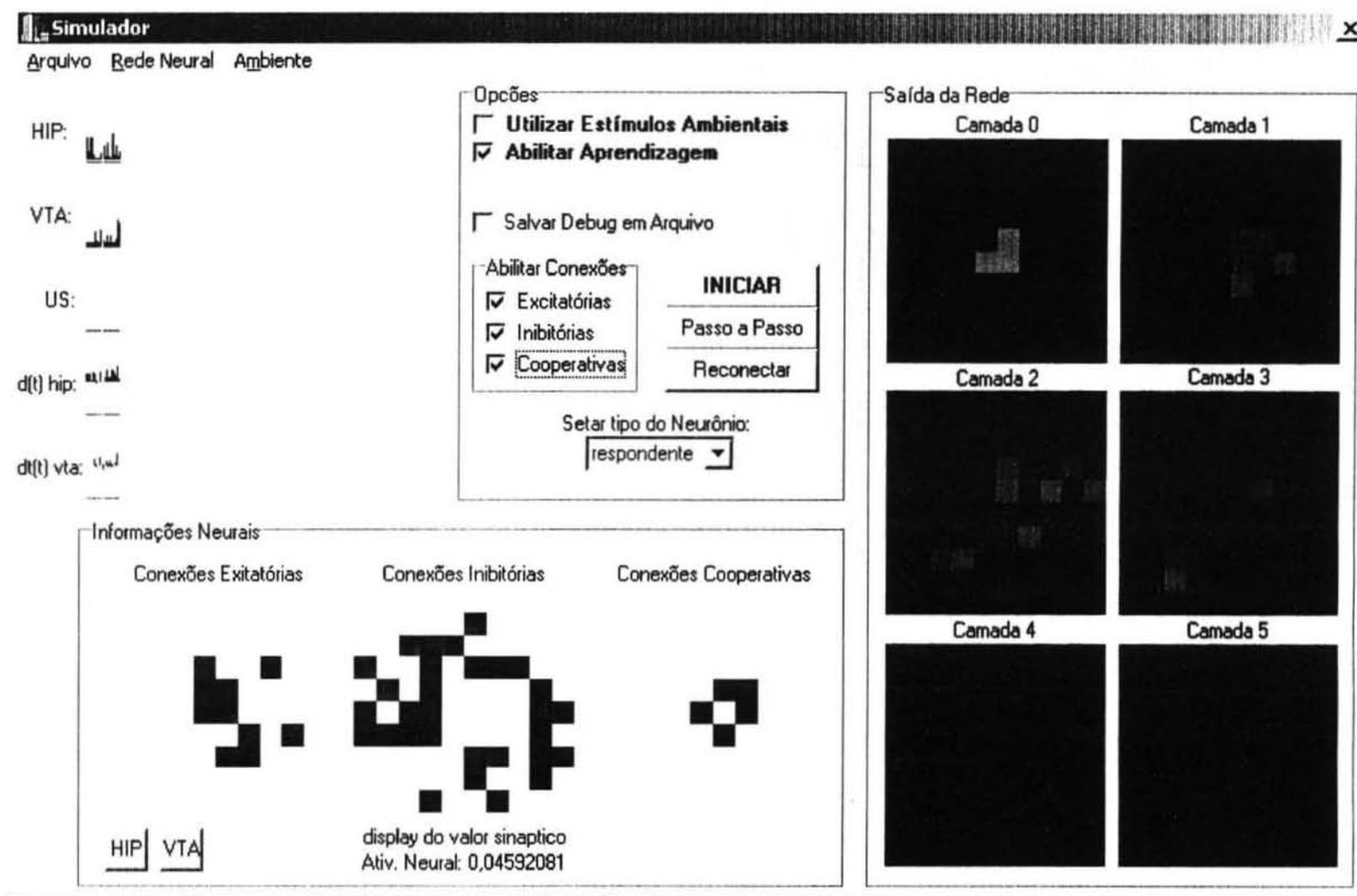


Figura 55 – Tela principal do simulador.

Parâmetros da Rede						OK
	Camada 1	Camada 2	Camada 3	Camada 4	Camada 5	Cancela
Taxa de Decaimento:	0,80000	0,80000	0,80000	0,80000	0,80000	
Limiar Sigmóide:	1,920924	0,30000	0,30000	0,30000	0,30000	
Inclinação Sigmóide:	0,10000	0,10000	0,10000	0,10000	0,10000	
Taxa Aquisicao E.:	0,5	0,5	0,5	0,5	0,5	
Taxa Extinção E.:	0,10000	0,10000	0,10000	0,10000	0,10000	
Taxa Aquisicao I.:	0,5	0,5	0,5	0,5	0,5	
Taxa Extinção I.:	0,10000	0,10000	0,10000	0,10000	0,10000	
Raio Conexão Sináptica E.:	2,5	2,5	2,5	2,5	2,5	
Raio Conexão Sináptica I.:	4	4	4	4	4	
Raio Conexão Sináptica C.:	1,5	1,5	1,5	1,5	1,5	
Numero Sinapses E.:	10	10	10	10	10	
Numero Sinapses I.:	30	30	30	30	30	
Numero Sinapses C.:	5	5	5	5	5	

Figura 56 – Tela do simulador para ajuste de parâmetros da rede neural.

A terceira tela do simulador é utilizada para definir as configurações do ambiente. A principal função realizada aqui é especificar o conjunto de estímulos que serão apresentados para a rede durante o treinamento bem como o conjunto de respostas desejadas que o ambiente deve monitorar. Além disso, pode-se especificar os estímulos reforçadores associados a cada estímulo ou respostas, o tempo de duração de cada estímulo, quanto tempo de atraso deve-se esperar para apresentar o US após a rede ter emitido uma resposta desejada e qual a intensidade da estimulação de cada neurônio.

Configurações do Ambiente

Ambiente:

Abrir | Salvar

Número de Estímulos: 3

Número de Respostas: 1

Tamanho das Trials: 10

Estímulos

Exibindo: Estímulos Respostas

Navegar

Índice: 0

- | +

Estímulo: 0

	0	1	2	3	4	5	6	7	8	9
0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0
4	0	1	1	0	0	0	0	1	1	0
5	0	1	1	0	0	0	0	1	1	0
6	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0

Associar Resposta

Resposta Associada: 0

Propriedades do Estímulo:

Tempo de Duração: 0

Associar US

Valor US: 0

Duração US: 0

Início US: 0

Figura 57 – Tela do simulador para configurar o ambiente de treinamento.