UMA REDE NEURAL SIMPLES COM ALGUMAS PROPRIEDADES SEMIÓTICAS INTERESSANTES

Angelo C. Loula, Hiata A.O. Coelho, Ricardo R. Gudwin

{angelocl, hiata, gudwin}@dca..fee.unicamp.br
DCA-FEEC-UNICAMP
Av. Albert Einstein 400
13.083-970 - Campinas - SP

Resumo: Nesse trabalho apresentamos uma rede neural simples com algumas propriedades semióticas interessantes. Resultados comprovam a capacidade da rede de aprender três tipos de signos: ícones, índices e símbolos, dependendo de um protocolo de ensino que apresenta à rede sequências de padrões com o propósito de ensiná-la diferentes tipos de signos. Esses resultados abrem perspectivas interessantes para a área de sistemas inteligentes.

Palavras Chave: rede neural, semiótica, ícone, índice, símbolo.

Abstract: In this work, we present a simple neural network with some interesting semiotic properties. Results demonstrate the network's capacity of learning three types of signs: icons, indexes and symbols, depending on a teaching protocol that presents the network to sequences of patterns with the purpose of teaching them different kinds of signs. These results open interesting perspectives to the field of intelligent systems.

Keywords: neural network, semiotics, icon, index, symbol.

1 INTRODUÇÃO

As aplicações tradicionais de redes neurais normalmente se restringem a problemas de classificação ou problemas em que a rede neural funciona como um interpolador universal de funções (Haykin 1998). Em problemas de classificação, deseja-se que a rede neural seja capaz de agrupar padrões de sinais de entrada em classes distintas, passíveis de discriminação. De maneira típica, as entradas destas redes são vetores de padrões e as saídas correspondem às classes discriminadas. Em problemas em que a rede funciona como um interpolador universal de funções, tais como em controladores, deseja-se que a rede possa correlacionar vetores de padrões de entrada com vetores de padrões de saída, de tal forma que a função mapeando entrada e saída seja aprendida pela rede neural, por meio de alguma regra de aprendizagem. Estes vetores de padrões de entrada e saída podem representar em princípio qualquer coisa, qualquer variável de interesse que o sistema representado pela rede neural se proponha a modelar. Neste trabalho, investigamos o uso de redes neurais para o processamento de signos. Mostraremos que, mesmo uma rede neural muito simples, pode apresentar um comportamento semiótico muito interessante. Entretanto, antes de começarmos a detalhar a estrutura da rede e suas propriedades, é necessário que façamos um pequeno arrazoado sobre o porquê de querermos processar signos, e porquê uma rede que apresente um comportamento semiótico seria de interesse, de um modo geral, na área de sistemas inteligentes. Assim, este trabalho está dividido da seguinte maneira: na sessão 2 abordamos brevemente o que é semiótica e qual a sua importância para a área de sistemas inteligentes. Na sessão 3 explicamos a implementação da rede neural com propriedades semióticas que desenvolvemos, e a sessão 4 mostra os resultados dos experimentos realizados. Ao final, temos as conclusões sobre os estudos realizados.

2 SISTEMAS INTELIGENTES E SEMIÓTICA

A área de estudos que conhecemos hoje sob o nome de "Sistemas Inteligentes" envolve um grande número de sub-áreas diferentes, que se desenvolveram também em períodos históricos diferentes. Temos por exemplo o estudo de técnicas de representação do conhecimento, tais como redes semânticas, ontologias, sistemas especialistas, sistemas de programação em lógica, resolvedores gerais de problemas, redes bayesianas, etc., que de uma certa maneira se agrupam dentro de uma área que convencionamos chamar de "inteligência artificial", por questões históricas (este termo foi cunhado em um workshop histórico no Dartmouth College, para abranger os estudos acima relacionados), e uma outra área, que convencionou-se chamar de "inteligência computacional", mais recente, que agrupa o estudo das redes neurais, dos sistemas e lógica fuzzy, da computação evolutiva, que envolve os algoritmos genéticos, os sistemas classificadores, a programação evolutiva. Outra área correlata que apresenta contribuições aos sistemas inteligentes é a área dos agentes inteligentes, as vezes também conhecida como a área dos sistemas multi-agentes. Percebe-se pois que a área de Sistemas Inteligentes, que agrega um tal leque de diferentes técnicas e metodologias, é uma área vasta e complexa, que tem diversas frontes de desenvolvimento. O que une entretanto todas estas sub-áreas diferentes, com técnicas diferentes e com objetivos diferentes, é a tentativa de simular computacionalmente atributos da mente, tais como a inteligência e a aprendizagem. O estudo da mente e de seus atributos não é privativo da área de sistemas inteligentes. Nas ciências biológicas, temos as neuro-ciências, a biologia evolutiva, os estudos com colônias de insetos, que já serviram de inspiração a um grande número de técnicas e metodologias em sistemas inteligentes. Nas ciências humanas, o estudo da mente é objeto de análise de diversas áreas de estudo, tais como as ciências cognitivas, a filosofia da mente, e a semiótica. Destas três, a menos conhecida dos pesquisadores envolvidos com sistemas inteligentes é a semiótica, ou o estudo dos signos e da significação, dos processos de representação e comunicação (Gudwin, 1996). Entretanto, como veremos a seguir, apesar desse aparente desconhecimento formal, a área de sistemas inteligentes sempre esteve influenciada, direta ou indiretamente, por modelos teóricos que emergiram de alguma forma da semiótica. O estudo tradicional da semiótica é direcionado para o processamento de signos em sistemas

naturais, tais como seres humanos e animais, ou fenômenos físicos e químicos. Nestes, entende-se o conceito de signo como algo que é capaz de representar alguma outra coisa. Assim, o conceito de signo está diretamente envolvido com o conceito de significado, sendo que diversos modelos para estes conceitos foram propostos na semiótica. Como bem ilustra Nöth em (Nöth, 1995), a semiótica é uma ciência muito antiga, remontando aos gregos clássicos. Na semiótica moderna, temos duas grandes correntes, a semiótica estruturalista e a semiótica Peirceana, que apresentam diferentes modelos para o signo e para o processo de significação. Uma grande parte das metodologias de sistemas inteligentes (principalmente aquelas relacionadas às sub-áreas que convencionamos chamar de "inteligência artificial"), utilizam de maneira implícita o modelo de signo estruturalista (também conhecido como modelo diádico, ou semiótica diádica, por modelar o signo por meio de uma relação diádica). Este modelo diádico do signo, privilegia o símbolo como o modelo de signo por excelência. Para muitos pesquisadores de sistemas inteligentes, o único tipo reconhecido de signo é o símbolo, como bem fica evidente em trabalhos como o de Newell (Newell 1980). Entretanto, essa visão de signo diádica exclui os chamados signos naturais, tais como os ícones e os índices (Nöth, 1990). Na prática, isso leva a um problema que tem se mostrado insolúvel na área de sistemas inteligentes, que é o problema da falta de fundamento simbólico (ou *symbol grounding problem*) (Harnad, 1990). Felizmente, a corrente Peircena da semiótica produziu um modelo mais elaborado para o conceito de signo, que é chamado de modelo triádico, por utilizar uma relação triádica para modelar o signo. A consideração de um modelo triádico para o signo resolve um grande número de problemas com os quais a área de sistemas inteligentes vem se degladiando já há décadas. Dentre outros, a possibilidade de trabalharmos com outros tipos de signos, tais como os ícones e índices, e a possibilidade de resolvermos, talvez de maneira histórica, o problema da falta de fundamento simbólico. Assim, consideramos que o estudo da semiótica, e em especial a semiótica triádica Peirceana, poderá nos legar, pesquisadores de sistemas inteligentes, uma nova direção a seguir, uma nova concepção do fenômeno da inteligência, que só o futuro poderá revelar. Essa é portanto, a razão de nosso interesse em estudarmos semiótica, dentro do contexto dos sistemas inteligentes.

Até recentemente, o estudo da semiótica era restrito à análise semiótica, ou seja, o uso da teoria semiótica para analisar o processo sígnico em diferentes atividades nos seres vivos. À medida que pesquisadores da área de sistemas inteligentes passaram a considerar modelos formais de signos em seus desenvolvimentos, abriu-se a oportunidade de experimentar o caminho oposto, ou seja, a síntese ou a re-criação do processo de semiose em sistemas computacionais.

Este foi o início da Semiótica Computacional, uma área de pesquisa interdisciplinar dedicada à criação de sistemas inteligentes baseados na teoria semiótica. Assim, a síntese semiótica é a tentativa de emular a interpretação de signos (ciclo semiótico ou semiose) em um computador digital. Entre outras coisas, isto é feito com o objetivo de construir sistemas inteligentes autônomos capazes de desenvolver comportamento inteligente, que inclui percepção, criação de modelo do mundo, julgamento de valores e geração de comportamento. Há uma declaração implícita aqui de que a maior parte do comportamento inteligente é devido ao processo semiótico, de modo que um sistema inteligente deve ser comparável com um sistema semiótico (Gudwin 2002). Uma das vertentes mais recentes no estudo dos sistemas inteligentes, identifica a capacidade de processar signos como a fonte da inteligência, ou seja, um sistema é inteligente porque processa signos, e sua inteligência dependerá da quantidade e dos tipos de signos que está apto a processar.

2.1 Modelos de Signos

Vamos considerar inicialmente o modelo de signo diádico da semiótica estruturalista, conforme a figura 1 a seguir:

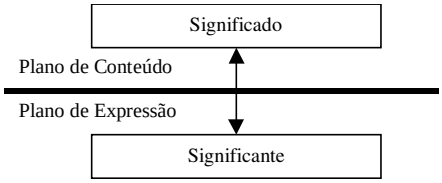


Figura 1 – Modelo de Signo Estruturalista (diádico)

Neste modelo, existem dois planos conectados bi-univocamente, os chamados planos de expressão e de conteúdo. No plano de expressão, encontramos unidades discrimináveis que chamamos de "significantes". No plano de conteúdo, encontramos unidadades discrimináveis que chamamos de "significados". Podemos relacionar as unidades do plano de expressão a unidades do plano de conteúdo, gerando pares cartesianos do tipo (significante, significado). A cada par (significante, significado), chamamos então de um signo. Um signo é portanto, neste modelo, constituído de uma relação diádica que une um significante a um significado. Este modelo tem origem nos trabalhos de Saussure (Nöth, 1990), tendo sido elaborado posteriormente por Hjelmslev e outros. Em sua determinação original, tanto significante quanto significado constituiam-se de unidades mentais. Assim, um exemplo de significante seria a palavra "carro", dentro de um plano de expressão das ondas sonoras, devidamente codificado em termos mentais, e seu significado seria a idéia que temos de um carro, devidamente também codificada em termos mentais. Variações desse modelo associam o plano de expressão ao plano mental, e o plano de conteúdo ao mundo exterior. Dessa maneira, a cada significante no plano mental, haveria um correlato natural no mundo exterior (plano de conteúdo). Entretanto, essa "correlação natural" é totalmente arbitrária. Essa é a visão dos chamados cognitivistas nas ciências cognitivas, que dizem que se em uma memória de computador um certo conjunto de símbolos representam a proposição "Sócrates é Mortal", a ligação destes símbolos com o Sócrates histórico (indivíduo) seria "automática". Essa conexão totalmente arbitrária entre significante e significado é aparentemente a fragilidade deste modelo, e a origem do problema da falta de fundamento simbólico na inteligência artificial. Neste modelo, também, não é possível modelarmos os signos naturais, os ícones e os índices, mas tão somente os símbolos.

Em contraposição ao modelo diádico de signo, Peirce desenvolveu um modelo mais elaborado, triádico (Nöth 1990), que faz uma divisão do significado em duas partes, uma delas ligada a algum elemento do mundo real, chamado de **objeto** do signo, e outra ligada ao efeito do signo na mente de um potencial intérprete, chamada de **interpretante** do signo. Um signo, segundo Peirce, é aquilo que, sob certo aspecto ou modo, representa algo para alguém. Esse signo criará na mente desse alguém um segundo signo equivalente a si mesmo, isto é, um signo mais desenvolvido, que é chamado de interpretante. Tanto o signo como seu interpretante referem-se da mesma forma a um objeto (Nöth 1990). Assim, um processo de semiose (ou um processo de significação), é uma relação

triádica que une um signo, um objeto e um interpretante, conforme ilustrado na figura 2 a seguir:

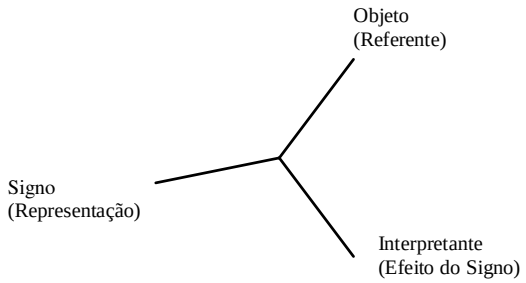


Figura 2: Modelo de Signo Peirceano (triádico)

Neste modelo, o objeto, por meio de sua relação com o signo, confere e ele um poder de representá-lo. Esse poder se consolida na geração do interpretante, na mente de um potencial intérprete (embora a figura do intérprete, segundo os Peirceanos, seja dispensável). Assim, não existe uma ligação "automática" entre significante e significado, como no modelo diádico, mas um processo, em que um signo só se torna signo, quando possui a capacidade de gerar um interpretante, e só terá essa capacidade, em virtude da relação que possui com seu objeto. Isso nos dá margem para a apreciação dos signos ditos "naturais", ou seja, os ícones e os índices, que não eram passíveis de modelagem, no modelo diádico.

O modelo Peirceano de signo é elaborado e cheio de detalhes, que omitimos neste artigo por fugirem a seus propósitos. Vamos nos ater à descrição da relação entre o signo e seu objeto, que é o que dá o caráter de interpretabilidade do signo, e que usaremos em nossa rede neural semiótica, a seguir. Assim, considerando-se a relação de um signo com seu objeto, os signos podem ser divididos em ícones, índices e símbolos, conforme a natureza dessa relação.

Os ícones são signos que têm em sua estrutura uma relação de similaridade com seu objeto designado. Ou seja, guardam em si próprios, as mesmas qualidades (ou um sub-conjunto delas) que os próprios objetos possuem, e exatamente por isso podem ser utilizados para representar seu objeto. Os ícones podem ser divididos em três categorias, as imagens, os diagramas e as metáforas. As imagens guardam em si próprias as mesmas qualidades de seus objetos. Isso ocorre, por exemplo, quando utilizamos uma fotografia para representar algo que foi fotografado. Assim, observando a fotografia de uma casa, podemos saber de que casa se trata. Esse significado é dividido em duas partes. De um lado temos a casa real, existente no mundo, e de outro a idéia que temos dessa mesma casa, que é acionada pela presença do signo. Os diagramas, ao contrário das imagens, não possuem diretamente as mesmas qualidades de seus objetos, mas apresentam relações entre suas partes, que correspondem a relações reais existentes em seu objeto. Assim, as relações existentes entre as partes de um diagrama são equivalentes às relações existentes em partes do objeto no mundo real. Esse constitui um outro tipo de ícone. As metáforas, ainda, estão ligadas a seus objetos por meio de abstrações que podemos fazer tanto no signo como no objeto. Assim, apesar de o signo não apresentar em si próprio as qualidades do objeto, quando fizermos uma abstração do signo e uma abstração do objeto, veremos que as qualidades de ambas as abstrações são equivalentes. Trata-se portanto, de um tipo mais sofisticado ainda de ícone. De um modo geral, dizemos que os ícones independem de seus objetos (ou seja, não dependem da presença simultânea destes para sua interpretação), pois carregam em si as qualidades do objeto, e dessa forma, permitem sua interpretação como signos deste. Veremos, ao contrário, que os índices e os símbolos dependem de seu objeto e de seu interpretante, respectivamente, para serem interpretáveis (Nöth 1990).

O símbolo, do mesmo modo que o ícone, carrega em si mesmo um significado absoluto, mas ao contrário deste, o significado não está vinculado a uma relação de similaridade com o objeto, mas sim a uma convenção externa, ou lei, que une o signo ao objeto. A interpretação de um símbolo decorre em duas fases. Em uma primeira fase, sobrevem a interpretação de um ícone, correspondendo à própria identificação do signo (como algo conhecido, em si). Em seguida, ocorre a interpretação em si, quando é invocada a ligação entre este signo preliminar e seu objeto designado final. Vemos que um importante elemento na interpretação de um símbolo é a convenção, que une o signo ao objeto. Essa convenção ocorre como uma definição do intérprete por si só (uma decisão pessoal, tal como ... a partir de agora vou chamar isso de xyzt ...), ou a partir de um acordo com outros intérpretes, com os quais o intérprete deseja se comunicar. A celebração desse acordo envolve um protocolo sofisticado, ainda pouco conhecido em toda sua complexidade, envolvendo ícones e índices.

O terceiro tipo de signo, o índice, é o mais complexo dos três, e o mais difícil de ser modelado. O índice, ao contrário do ícone e do símbolo, não é considerado um signo caracterizante, pois não carrega em si, o objeto que denota. Uma boa definição para um índice seria de uma "chave" para um procedimento, que após ser executado, definirá seu objeto designado. Esta definição é um tanto quanto geral, mas é necessária, de modo a incorporar todos os signos que podem ser índices. Um exemplo bem simples de se entender é o exemplo dos índices referenciais. Neste caso, um índice referencia um interpretante anterior do contexto, tomando seu significado. Este tipo de índice, ocorre usualmente na linguagem humana. Por exemplo, considere o final de uma frase, engatada com o início de uma outra, como em "... a casa. Esta, ...". Imagine agora que um sistema inteligente esteja processando este texto. O último signo da primeira frase é a palavra casa. Este signo é um símbolo para o conceito de "casa". Em seguida, aparece o signo "Esta". Este signo, em si, não carrega nenhum significado. No caso o interpretante do signo "Esta", será uma cópia do interpretante do signo "casa" (Gudwin & Gomide, 1996).

3 REDE NEURAL SEMIÓTICA

Em um trabalho anterior, foi proposta por Gudwin e Gomide (1996), uma rede neural com algumas propriedades que modelavam esse processamento semiótico, permitindo o processamento de ícones e símbolos. Esta proposta foi ampliada neste trabalho, visando entre outras coisas permitir o aprendizado de índices.

3.1 Interpretando Ícones e Símbolos

O módulo neural simplificado proposto trabalha com protótipos (signos internos) contra os quais as entradas (signos externos) são confrontadas para determinar o interpretante. Além disso, os signos internos podem possuir ligações entre eles para representar vínculos simbólicos entre signos.

Esta representação permite a interpretação de ícones e símbolos, caracterizados, respectivamente, por uma relação de semelhança e por uma relação arbitrária, conforme explicado na seção 2.

Quando um signo externo é apresentado, o módulo compara-o com todos os signos internos e determina qual o mais similar. O signo interno com maior similaridade, isto é, a menor

distância em relação ao signo externo determinaria o interpretante. Desta forma o sistema seria capaz de reconhecer ícones, uma vez que leva em consideração aspectos qualitativos para gerar o interpretante (Figura 3). Isso corresponderia, por exemplo, a apresentar a figura de um círculo para o sistema, que seria interpretado como 'bola' por serem similares.

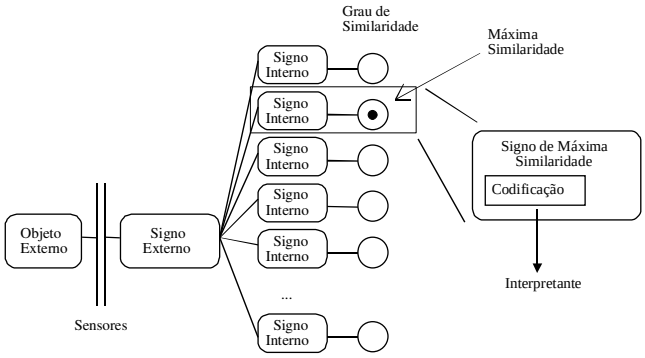


Figura 3: Reconhecimento de Ícones

Já a interpretação de um símbolo não leva em conta relações de similaridade, mas sim uma relação convencional ou arbitrária. O processo iniciaria com uma interpretação icônica do signo, que possuiria uma ligação com outro signo que seria o interpretante desejado (Figura 4). Como exemplo, se a palavra "bola" fosse apresentado como signo externo, este signo seria interpretado inicialmente como o signo interno do protótipo da palavra "bola", e em seguida a imagem do objeto bola seria recuperada.

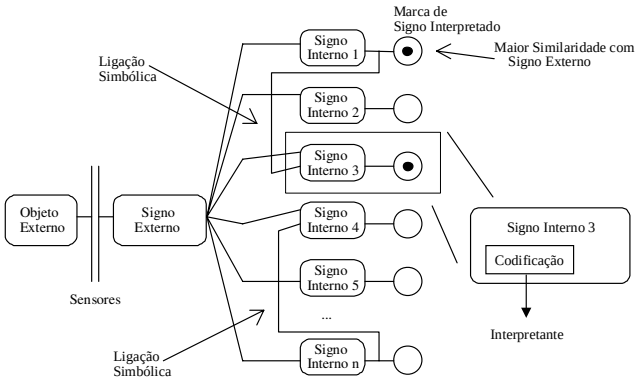


Figura 4: Reconhecimento de Símbolos

3.2 Descrição da Rede Neural

A rede neural proposta inicialmente em (Gudwin & Gomide, 1996) possuia algumas falhas e limitações que foram corrigidas e aperfeiçoadas neste trabalho. Esta nova rede implementa o reconhecimento icônico, as ligações simbólicas e também as relações indiciais, uma nova funcionalidade. A rede será capaz também de aprender ícones e índices, que podem transformar-se em símbolos.

Os signos externos são caracterizados como um vetor de dados **X** que o caracteriza, segundo os vários sensores mediante uma fusão sensorial. Os protótipos (signos internos) são codificados pelos pesos de entrada **V** e as relações simbólicas pelos pesos de recorrência **W**. As saídas dos neurônios **Z** definem o grau de ativação de cada um. (Figura 5)

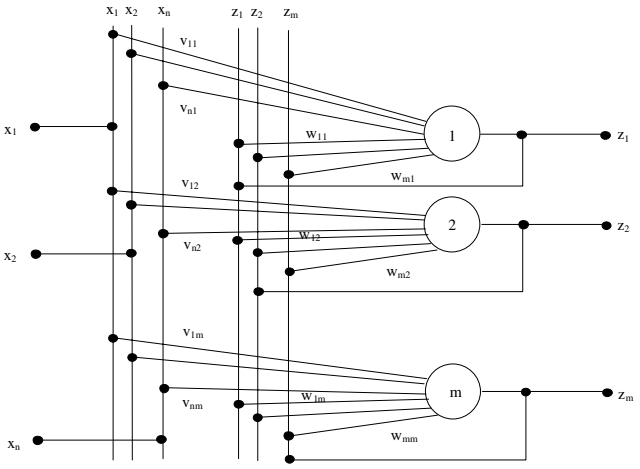


Figura 5: Rede Neural Semiótica

3.2.1 Operação

A ativação do neurônio *i* na *k*-ésima iteração é calculada para duas situações: nenhum neurônio se identifica o suficiente com o signo externo (ativação de cada neurônio é menor que um limiar de dissimilaridade); ou algum neurônio se identifica o suficiente com signo externo (ativação do neurônio mais ativo é maior que o limiar de dissimilaridade).

No primeiro caso para cada neurônio *i*,

$$a_i(k) = f\left(\sum_{k=0}^{n-1}|v_k(k) - x_k(k)| + \sum_{k=0}^{m-1}\left(w_{ki}(k).z_k(k)\right)\right) \quad (1)$$

E no segundo caso, para o neurônio com maior ativação,

$$a_{ativo}(k) = f\left(\sum_{k=0}^{n-1}|v_k(k) - x_k(k)| + \sum_{k=0}^{m-1}\left(w_{kativo}(k).z_k(k)\right)\right) \quad (2)$$

e para os demais neurônios,

$$a_{i\neq ativo}(k) = f\left(\sum_{k=0}^{m-1}\left(w_{ki}(k).z_k(k)\right)\right) \quad (3)$$

A ativação dos neurônios possui duas partes: uma icônica, caracterizada por uma medida de distância entre os pesos do neurônios **V** e a entrada **X**; e uma simbólica que permite que um neurônio seja ativado simbolicamente caso a saída de um outro neurônio Z_k esteja alta e a ligação através de W_{ki} seja alta. A divisão da ativação dos neurônios em duas partes permite aos neurônios competir pela similaridade com o signo externo, e se existir um bom candidato somente ele é ativado iconicamente.

A função de ativação f é decrescente, limitada à ativação dos neurônios entre 0 e 1, com $f(\theta)\leq 0.5$ sendo θ o limiar de dissimilaridade. Quanto mais próximo de 1 mais ativo estará o neurônio. Esta função é implementada por um sigmóide invertiva dada por

$$f_s(x) = \frac{e^{-\alpha.(x-\theta)}}{1+e^{-\alpha.(x-\theta)}} \quad (4)$$

onde θ define o ponto de transição da sigmóide entre 1 e 0, e α define a inclinação da curva sigmoidal.

Mas as saídas dos neurônios não são dadas simplesmente pela ativação, mas sim obedecem a seguinte relação:

$$z_i(k+1) = \begin{cases} a_i(k), \ se \ \ a_i(k) > z_i(k) \\ \gamma.z_i(k) + (1-\gamma).a_i(k), \ \ caso \ \ contrário \end{cases} \quad (5)$$

ou seja, a saída dos neurônios sobe instantaneamente para o valor da ativação se esta for maior que a saída atual ou desce lentamente, caso contrário, segundo um parâmetro γ, chamado de fator de decaimento da saída, com valor entre 0 e 1.

3.2.2 Aprendizado

O aprendizado da rede se dá pelo ajuste dos pesos icônicos **V** e dos pesos simbólicos **W**. O aprendizado é não-supervisionado e auto-organizável.

Inicialmente a rede é inicializada com os pesos **W** nulos e com os pesos **V** com valores aleatórios próximos de zero. Os pesos **V** não são inicializados com valores aleatórios entre 0 e 1, pois isto caracterizaria que os protótipos não estariam vazios e sim com "sujeira".

Quando um signo externo é apresentado a rede, o neurônio mais ativo terá seus pesos ajustados na direção do vetor de entrada, conforme

$$v_i(k+1) = (\beta - 1)v_i(k) + \beta x_i(k) \quad (6)$$

onde β é o fator de aprendizado icônico, com valor entre 0 e 1. Este fator é inicialmente alto β(0), favorecendo a um aprendizado rápido, mas converge ao longo das iterações para um fator baixo β_{min}, definindo uma maior inércia ao aprendizado. A variação do valor de β é dada por

$$\begin{cases} \beta(k) = \beta(0) - \beta(0).{}^{A(k)}\!/_{A_{max}}, se\ \ \beta(k) > \beta_{min} \\ \beta(k) = \beta_{min}, caso\ \ contrário \end{cases} \quad (7)$$

onde A(k) indica o número de ajustes feitos nos pesos icônicos do neurônio, e A_{max} é o número máximo de interações quando β alcançaria o valor nulo.

Já o aprendizado simbólico ocorre quando dois neurônios estão com suas ativações altas, acima de um limiar de esquecimento simbólico φ. A cada iteração, somente o neurônio mais ativo têm seus pesos simbólicos ajustados, e o reforço só acontece entre este e o segundo neurônio mais ativo:

$$w_{ij}(k+1) = \begin{cases} \rho.w_{ij}(k) - (1-\rho)n, \\ \quad se\ \ (\min(z_i(k), z_j(k)) > \varphi\ e\ \ (i = ativo_2)) \\ \quad\quad ou\ \ \ w_{ij}(k) \le -\lambda.n \\ \rho.w_{ij}(k), caso\ \ \ contrário \end{cases} ; j = ativo_1 \quad (8)$$

onde λ define o limiar simbólico, ρ é o fator de aprendizado simbólico, e ativo1 e ativo2 são os neurônios mais ativo e segundo mais ativo, respectivamente. O limiar simbólico define o nível após o qual o peso simbólico entre dois neurônios não decresce mais.

O aprendizado simbólico depende fortemente da existência de um decaimento lento na saída dos neurônios, pois isto transforma uma relação temporal (um signo apresentado após o outro) em uma relação numérica (saídas altas de dois neurônios). Para um símbolo ser formado, então, os signos que se relacionam dessa forma são apresentados em seqüência, e a rede aprenderá gradativamente esta relação temporal que existe entre eles, até o ponto onde uma convenção é estabelecida.

3.3 Aspectos Teóricos

A rede neural semiótica proposta foi inspirada inicialmente em princípios da semiótica (Gudwin & Gomide, 1996) procurando modelar, na forma de uma rede conexionista, processamentos sígnicos de ícones e símbolos. Neste trabalho, esta proposta foi aprimorada e estendida para um modelo mais amplo, contemplendo índices. Mas é também possível identificar além de propriedades semióticas, aspectos teóricos provenientes de sistemas auto-organizáveis e das redes neurais naturais.

O processamento icônico dos neurônios segue o princípio semiótico de iconicidade, segundo o qual um signo é reconhecido como um ícone quando este é classificado em virtude de seus atributos (Deacon 1997). Isso ocorre quando um signo externo é classificado por sua similaridade com um protótipo (no caso o neurônio), indicado pela ativação do neurônio. Já o processamento simbólico baseia-se nos relacionamentos referenciais do tipo indicial e simbólico entre signos. O aprendizado referencial entre signos permite que após um determinado limiar um índice se transforme em um símbolo. Até atingir este limiar, a relação indicial pode ser desfeita, sem caracterizar uma convenção (Deacon 1997).

Temos aqui, portanto, o protocolo que funciona como convenção para a formação de símbolos. Quando dois padrões diferentes são apresentados em sequência, a rede tenta fazer uma associação entre eles. Nesse caso, o primeiro padrão funciona como um índice do segundo. A medida que essa sequência é repetida, o peso w conectando os dois padrões é incrementado. Entretanto, por se tratar de uma conexão indicial, ela pode ser desfeita. Na medida em que a sequência deixa de se repetir, o peso w passa a decair, e o padrão inicial deixa de ser signo do seguinte. Entretanto, caso a repetição na sequência dos dois padrões faça com que w ultrapasse o limiar λ, o valor de w não mais decai, criando uma ligação simbólica definitiva entre os dois padrões. Assim, caso apresentemos uma repetição entre padrões de maneira sucessiva, um certo número de vezes (que dependerá do limiar λ), a rede "compreende" que queremos que o primeiro padrão seja o símbolo do segundo, aprendendo essa correlação.

4 EXPERIMENTO E RESULTADOS

O experimento realizado utilizou cinco ícones que deveriam ser aprendidos pela rede, assim como relações simbólicas e indiciais entre eles conforme a Figura 6. Estes signos são imagens preto e branco, 12 por 12 pixels.

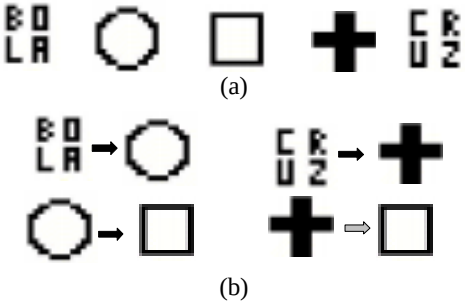
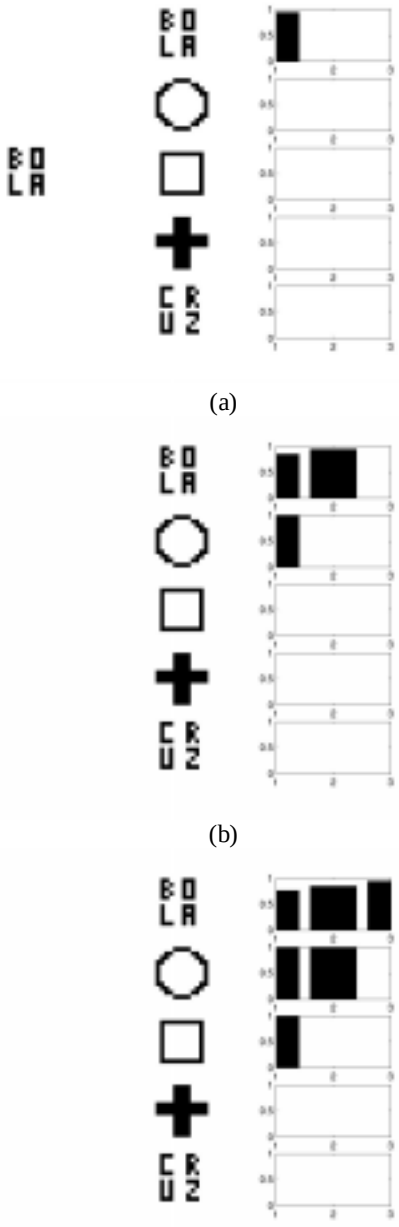


Figura 6: Signos e relações a serem aprendidas pela rede: (a) Ícones e (b) Relações simbólicas e indiciais. As setas pretas indicam ligações simbólicas e a cinza uma ligação indicial.

Após o treinamento da rede, dividido em duas etapas, aprendizado dos ícones e aprendizado das ligações simbólicas, foram feitos testes para avaliar o aprendizado da rede. A rede neural semiótica foi capaz de reconhecer os ícones, ativar as ligações simbólicas e também as indiciais, que desaparecem posteriormente. (Figura 7)

Na figura 7 estão representadas a entrada da rede à esquerda, a representação da codificação icônica dos neurônios ao centro, e

a saída dos neurônios nas últimas três iterações a direita (1 é a mais recente em *t*, 2 é a iteração anterior em *t-1* e 3 é a iteração em *t-2*).



(a)

(b)

(c)

Figura 7: (a) Ativação do neurônio 1 pela apresentação do signo "bola" em *t*. (b) Ativação do neurônio 2, em momento posterior *(t+1)*, devido ao vínculo simbólico. (c) Ativação do neurônio 3, no terceiro momento *(t+2)*, devido a vínculo simbólico em cascata.

5 CONCLUSÕES

A rede neural semiótica inicialmente apresentada em (Gudwin & Gomide, 1996), e aperfeiçoada neste trabalho possui algumas características interessantes, como:

- Aprendizado icônico com refinamento por repetição;
- Aprendizado de vínculos simbólicos por repetição de co-ocorrências;
- Aprendizado indicial com destruição posterior;
- Preservação de signos reconhecidos por certo tempo com decaimento;

Redes neurais com comportamento icônico são bem comuns na área de redes neurais. Praticamente todas as redes que atuam como classificadores, de algum modo estão implementando um comportamento icônico. A grande novidade da rede aqui proposta, é exatamente ir além desse comportamento icônico, introduzindo algumas funcionalidades indiciais e outras simbólicas.

Apesar do experimento aqui relatado ser muito simples, cremos que pudemos demonstrar algumas propriedades interessantes, do ponto de vista semiótico:

- como podemos utilizar um padrão como índice de outro, em virtude da repetição de uma sequência, e de sua desconsideração, caso a sequência não se repita mais
- como um índice desse tipo possa se transformar em um símbolo, caso a repetição ultrapasse um limiar simbólico, e a partir daí a conexão entre ambos os padrões se torna perene.

Apesar da simplicidade, criamos um procedimento efetivo para a determinação de símbolos, fornecendo um *ground* para sua utilização. Não temos pretensões maiores neste estudo, do que demonstrar um novo caminho que poderia ser aberto, conectando todo o arsenal de metodologias para processamento simbólico fornecidos pela inteligência artificial. Utilizando outros tipos de signos além do símbolo, teríamos ampliado o leque de opções para a representação do conhecimento. Fazendo o *ground* desses símbolos em termos de índices e ícones, utilizando redes neurais como meio, abrimos caminhos para desenvolvimentos ulteriores na área de sistemas inteligentes. Sistemas híbridos simbólico-conexionistas, com inspiração na semiótica, poderiam ser desenvolvidos, trazendo certamente grandes contribuições à área de sistemas inteligentes em geral, e à automação inteligente, em particular.

REFERÊNCIA BIBLIOGRÁFICA

Deacon, T. W. (1997). *The Symbolic Species - The Co-evolution of Language and the Brain.* W.W. Norton & Company, Inc.

Gudwin, R.R. (1996). *Contribuições ao Estudo Matemático de Sistemas Inteligentes* – tese de Doutorado, DCA-FEEC-UNICAMP.

Gudwin, R & F. Gomide (1996). *Sistemas Inteligentes Semióticos segundo a Semiótica Behaviorista de Charles Morris*. Relatório Técnico, DCA-FEEC-UNICAMP

Gudwin, R. R. (2002). *Semiotic Synthesis and Semionic Networks,* S.E.E.D. Journal (Semiotics, Evolution, Energy, and Development), Volume 2, No. 2, August 2002, p. 55-83.

Harnad, S. (1990) The Symbol Grounding Problem, Physica D, 42:335-346.

Haykin, S.S., et al (1998). *Neural Networks: A Comprehensive Foundation*, Prentice Hall.

Newell, A. (1980). Physical symbol systems. Cognitive Science, 4:135--183.

Nöth, W. (1990) *Handbook of Semiotics*. Indiana University Press.

Nöth, W. (1995) – *Panorama da Semiótica – De Platão a Peirce* – Editora Annablume.