



UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS MATEMÁTICAS E DE COMPUTAÇÃO
Departamento de Ciências de Computação

SCC-5809 - Capítulo 9

Sistemas Híbridos Simbólico-Conexionistas

João Luís Garcia Rosa¹

¹SCC-ICMC-USP - joaoluis@icmc.usp.br

2011

Sumário

- 1 Sistemas Híbridos
 - RNA Baseadas em Conhecimento
 - Extração de regras

- 2 PLN Conexionista
 - Representação
 - Solução connexionista
 - Estruturas

Sumário

- 1 **Sistemas Híbridos**
 - RNA Baseadas em Conhecimento
 - Extração de regras

- 2 PLN Conexionista
 - Representação
 - Solução conexionista
 - Estruturas

A Inteligência Artificial

- A Inteligência Artificial divide-se basicamente em dois paradigmas em princípio opostos:
 - simbólico, baseado na lógica, e
 - conexionista, baseado na propagação da atividade de processadores elementares.
- Vantagens da abordagem simbólica:
 - Facilidade de representação do conhecimento,
 - Poder expressivo das representações lógicas gerais,
 - Entendimento através da inferência lógica.
- Vantagens da abordagem conexionista (sub-simbólica):
 - aprendizado,
 - generalização,
 - tolerância a falhas.

Algumas considerações iniciais

- As RNA são uma arquitetura baseada em atributos e não têm o poder expressivo das representações lógicas gerais, já que parecem não ser adequadas para manipulação de símbolos de alto nível.
- Elas são bem adaptadas para entradas e saídas contínuas, ao contrário da maioria dos sistemas simbólicos de árvore de decisão.
- A classe de redes multicamadas, em geral, pode representar qualquer função de um conjunto de atributos desejada.
- Mas o projeto de uma boa topologia ainda é considerado uma arte.
- Uma vez que se consegue estabelecer uma arquitetura adequada para resolver determinado problema, o aprendizado da rede pode ser demorado (necessidade de muitos ciclos de ativação).

Algumas considerações iniciais

- Mas, em compensação, o reconhecimento é extremamente rápido (apenas um ciclo), muito mais rápido que qualquer sistema simbólico.
- Uma outra grande vantagem das RNAs é a sua capacidade de generalização.
- Esta característica está associada ao fato de a rede neural ser tolerante a falha, isto é, uma entrada imprecisa (incompleta) pode ativar a rede e fazer com que seja capaz de ativar parte de suas conexões e ainda assim, responder de forma apropriada.
- Isto se deve ao fato de que a RNA distribui a representação através dos pesos das conexões entre os seus elementos.
- Mas, a RNA tem uma desvantagem: normalmente por causa da falta de transparência, é muito difícil acompanhar o seu funcionamento.

Algumas considerações iniciais

- Mas as chamadas redes neurais baseadas em conhecimento (RNABC), que aproximam os até então opostos paradigmas da Inteligência Artificial, permitem a introdução e a extração de conhecimento simbólico em RNAs (abordagem híbrida).
- A extração de conhecimento simbólico a partir de RNAs treinadas permite a troca de informação entre representações de conhecimento conexionista e simbólico e tem sido de grande interesse para entender o que a RNA está realmente fazendo.
- O conhecimento simbólico pode ser inserido em RNAs e então refinado após o treinamento.
- Uma melhora significativa no tempo de aprendizado pode ser conseguida treinando RNAs com conhecimento inicial.

A abordagem híbrida

- Em resumo, a expressão “sistema híbrido” nesse contexto, é usada para designar sistemas que objetivam combinar todos os méritos da abordagem simbólica, redes neurais localizadas e processamento distribuído paralelo [6].
- É projetado para integrar mecanismos de baixo nível e computações de alto nível.
- A representação interna dos pesos de uma RNA é uma incógnita para os pesquisadores.
- Na RNABC, a rede não começa mais com pesos sinápticos aleatórios e sim com um conjunto de pesos que refletem regras de produção [5].

RNA baseadas em conhecimento

- Numa RNABC, dada uma teoria simbólica, cria-se a partir das regras dessa teoria, uma arquitetura de rede neural.
- Ou seja, de uma regra $A \rightarrow B$, cria-se uma conexão entre uma unidade que representa o conceito A e outra que representa o conceito B .
- Para uma regra $(A \wedge B) \rightarrow C$, tem-se duas unidades de entrada A e B , um neurônio na camada escondida faz o papel da conjunção e um neurônio na camada de saída representa o conceito C .
- Basta ligar essas unidades através de pesos de conexão grandes e tem-se uma rede que representa essa regra.
- A função da RNABC é revisar essa teoria.
- A teoria inicial (regras) está representada na rede.

RNA baseadas em conhecimento

- A rede começa a aprender.
- No final do aprendizado pode-se extrair de volta as regras da rede.
- A teoria representada pelas regras foi revisada pelo aprendizado.
- Fu [1, 2], Towell and Shavlik [10], Setiono and Liu [9] e outros apresentam sistemas neurais baseados em conhecimento.
- Numa RNABC, através do conhecimento inicial simbólico baseado em regras de produção, constróem-se conexões fortes entre unidades que representam esses conceitos.
- No restante da rede, atribuem-se pesos pequenos, mas não nulos, de tal forma que uma regra ainda inexistente possa eventualmente se estabelecer.

RNA baseadas em conhecimento

- Através do treinamento ao que a rede neural é exposta, vários padrões diferentes são apresentados à rede.
- Se um determinado padrão ocorre várias vezes, supondo que o treinamento é coerente com o que acontece no mundo, significa que deveria haver uma regra relacionando os conceitos envolvidos por esse padrão.
- Isto é, a teoria simbólica inicial deveria prever esse tipo de comportamento.
- Se a teoria realmente já contava com esse acontecimento, sua regra já foi implementada, e o treinamento se encarregará de fortalecer ainda mais os pesos sinápticos que relacionam esses conceitos.
- Caso contrário, é desejável que a teoria seja revista, ou seja, que essa “nova” regra seja adicionada ao conjunto de regras da teoria simbólica.
- Isso é conseguido através da extração de regras da RNA.

Sumário

- 1 **Sistemas Híbridos**
 - RNA Baseadas em Conhecimento
 - **Extração de regras**

- 2 **PLN Conexionista**
 - Representação
 - Solução conexionista
 - Estruturas

Procedimento de Fu [2]

- Segundo Fu [2], o procedimento completo é o seguinte: primeiro atribuem-se os pesos altos, relativos às regras da teoria inicial, à rede.
- Aos demais pesos são atribuídos valores pequenos.
- A seguir a rede passa pelo processo de treinamento, através do algoritmo *backpropagation* ou outro algoritmo conexionista qualquer.
- Depois, a rede sofre um processo de anulação de seus pesos pequenos, pois estima-se que pesos muito pequenos não vão contribuir para o conhecimento da rede.
- Com a anulação, a rede é simplificada.
- Uma outra simplificação ocorre quando a rede é “clusterizada,” ou seja, formam-se grupos de neurônios com vetores de pesos próximos entre si.

Procedimento de Fu [2]

- Depois, novamente a rede é submetida a um treinamento, sendo que no final desse processo, ocorre a extração das regras.
- Fu propõe um algoritmo de extração de regras chamado *KT*, que trabalha com subconjuntos de pesos das entradas de um determinado neurônio, que, se alcançado o limiar de ativação, forma uma regra com o neurônio alvo representando o conceito de saída.
- O grande problema do algoritmo *KT* é a grande quantidade de regras geradas.

Procedimento de Towell and Shavlik [10]

- Já Towell and Shavlik [10] propõem um algoritmo de extração de regras (*MofN*), um pouco diferente do *KT*, que reduz bastante o número de regras obtidas.
- Basicamente, o algoritmo *MofN* consiste em seis passos:
 - 1 *Agrupar*: criação de classes equivalentes, para agrupar as conexões da rede em grupos;
 - 2 *Tirar a média*: depois de agrupados, faz com que os pesos de todas as conexões dentro de um grupo tenham o valor médio de todas as conexões desse grupo;
 - 3 *Eliminar*: eliminação dos grupos com valores insignificantes, que não contribuem para o cálculo;
 - 4 *Otimizar*: com os grupos sem importância eliminados no passo 3, otimiza-se os limiares da unidade;
 - 5 *Extrair*: formam-se regras que expressam a rede, tal que uma regra é verdadeira se a soma dos seus antecedentes ponderados exceder o limiar;
 - 6 *Simplificar*: as regras são simplificadas quando possível para eliminar pesos e limiares.

Abordagem Simbólica

- Baseada em regras:
 - Implicação da lógica:

$$(A \wedge B) \rightarrow C$$

- Cláusula do Prolog:

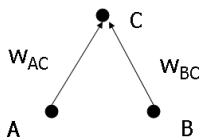
$$C : -A, B.$$

- Maiores críticas aos sistemas conexionistas:
 - Falta de transparência (funciona mas não se sabe como),
 - Demora no treinamento.
- **Solução:** Redes Neurais Baseadas em Conhecimento (abordagem híbrida simbólico-conexionista).

Abordagem Híbrida

- Vantagens da abordagem híbrida:
 - Extração do conhecimento simbólico a partir da rede neural permite a troca de informação entre representações de conhecimento conexionista e simbólico,
 - O conhecimento simbólico pode ser inserido na rede neural e depois refinado,
 - Uma melhora significativa no tempo de aprendizado.
- Abordagem Híbrida:
 - 1 'Conhecimento' inicial simbólico \rightarrow Pesos de conexão,
 - 2 Aprendizado,
 - 3 Pesos de conexão \rightarrow 'Conhecimento' final simbólico.

Abordagem Híbrida



$$((W_{AC} * A) + (W_{BC} * B)) \rightarrow C$$

- As redes neurais podem aprender associações entrada-saída.
- As redes neurais são capazes de recuperar uma saída baseado em itens de entrada incompletos.
- As redes neurais mostram degradação suave.
- Na abordagem híbrida, a IA simbólica e a IA conexionista se encontram.

Sumário

- 1 Sistemas Híbridos
 - RNA Baseadas em Conhecimento
 - Extração de regras
- 2 PLN Conexionista
 - Representação
 - Solução conexionista
 - Estruturas

Representação em RNA: Um Exemplo de PLN conexionista

- “Uma boa solução depende de uma boa representação,”
- Representação utilizada nos sistemas CPPro [8] e HTRP [7]: Microcaracterísticas Semânticas, de McClelland e Kawamoto [4].

Papéis de caso temáticos

- Um aspecto da compreensão da sentença: a atribuição dos constituintes de uma sentença aos papéis de caso temáticos corretos.
- Várias sentenças com o verbo *quebrar*:
 1. (O menino)_{SN} quebrou (a vidraça)_{SN}.
 2. (A pedra)_{SN} quebrou (a vidraça)_{SN}.
 3. (A vidraça)_{SN} quebrou.
 4. (O menino)_{SN} quebrou a vidraça (com (a pedra)_{SN})_{SP}.
 5. (O menino)_{SN} quebrou (a vidraça)_{SN} (com (a cortina)_{SN})_{SP}.
- O primeiro *SN* pode ser:
 - o AGENTE (sentenças 1, 4 e 5),
 - o INSTRUMENTO (sentença 2),
 - o PACIENTE (sentença 3).
- O *SN* do *SP* pode ser:
 - o INSTRUMENTO (sentença 4),
 - modificador do segundo *SN* (sentença 5).

Papéis de caso temáticos

- Ambigüidade da atribuição de papéis de *com-SN*:
 6. O menino comeu o macarrão com molho.
 7. O menino comeu o macarrão com garfo.
- Restrições da ordem da palavra
 8. O vaso quebrou a vidraça.
 9. A vidraça quebrou o vaso.
 10. O lápis chutou a vaca.
- Restrições na ordem das palavras (como em 10) são muito fortes em português, mas não é universal.
- Restrições semânticas: é possível (no italiano) atribui o papel de AGENTE a *a vaca* em 10.

Contexto

- Contexto global em que a sentença é apresentada:
11. O menino viu a menina com os binóculos.
- 1º. Contexto: Um menino olha pela janela tentando descobrir quanto ele consegue ver com vários instrumentos ópticos,
- 2º. Contexto: Duas meninas tentam identificar alguns pássaros quando um menino se aproxima. Uma menina tem um par de binóculos e a outra não.

Um Exemplo de PLN conexionista

- Enquanto o fato de que a ordem da palavra e as restrições semânticas influenciam a atribuição de caso tem sido reconhecido, existem alguns poucos modelos que vão além e propõem um mecanismo para explicar a causa desses efeitos.
- Entretanto, existem alguns pesquisadores em PLN que têm tentado encontrar formas de trazer as considerações semânticas para o processamento sintático de uma forma ou outra.

Preferência lexical

- Uma abordagem recente depende do léxico para influenciar o processamento sintático e a construção de representações funcionais básicas, que consideram casos como:
 12. A mulher (queria (o vestido do armário)_{SN})_{SV}.
 13. A mulher (colocou (o vestido)_{SN} no armário)_{SV}.
- A leitura preferida para a primeira dessas sentenças (12) tinha *do armário* como um modificador de *o vestido*, enquanto que a leitura preferida para a segunda (13) tinha *no armário* como um argumento de local de *colocou*.
- Para explicar essa diferença na atribuição de caso, foi proposto dois princípios:
 - preferência lexical,
 - argumentos finais.

Preferência lexical

- Basicamente, a **preferência lexical** estabelece uma estrutura de argumento esperada (por ex. Sujeito-Verbo-Objeto Direto no caso de *querer*; Sujeito-Verbo-Objeto Direto-Objeto Indireto no caso de *colocar*) consultando uma lista ordenada de possíveis estruturas de argumentos associadas com cada verbo.
- Se um constituinte que poderia preencher uma posição na estrutura de argumento esperada é encontrado, o mesmo é tratado como um argumento do verbo.
- Portanto, se um constituinte que aparece para satisfazer as condições do **argumento final** da estrutura de argumento esperada é encontrado, sua colocação na sentença é atrasada para permitir a incorporação dos constituintes subseqüentes.

Preferência lexical

- Portanto, com *querer*, o *SN o vestido* é um candidato para argumento final e não é colocado diretamente como um constituinte do Sintagma Verbal (*SV*); antes, uma estrutura *SN* superordenada contendo *o vestido do armário* é finalmente colocada no *SV*.
- Com *colocar*, entretanto, *o vestido* não poderia ser o argumento final, e portanto, ele é colocado diretamente ao *SV*.
- *No armário* está então disponível para a colocação como argumento final do *SV*.

Um Exemplo de PLN conexionista

- De qualquer forma, está claro que um mecanismo é necessário no qual todos os constituintes de uma sentença possam trabalhar simultaneamente para influenciar a atribuição de casos aos constituintes.
- Esse modelo toma como entrada uma análise parcial superficial e gera a partir disto uma representação de nível de caso.

Sumário

- 1 Sistemas Híbridos
 - RNA Baseadas em Conhecimento
 - Extração de regras

- 2 **PLN Conexionista**
 - Representação
 - **Solução connexionista**
 - Estruturas

Um Exemplo de PLN connexionista

- Meta principal:
 - prover um mecanismo que considere o papel da ordem da palavra e restrições semânticas na atribuição de papel (de caso):
 - que seja capaz de **aprender** a fazer isto baseado na experiência com sentenças e suas representações de caso.
 - que seja capaz de **generalizar** o que aprendeu para novas sentenças construídas a partir de novas combinações de palavras.

Um Exemplo de PLN connexionista

- Outras metas:
 - que seja capaz de selecionar contextualmente leituras apropriadas de palavras ambíguas,
 - que selecione o verbo (*verb frame*) apropriado baseado no padrão de argumentos e em suas características semânticas,
 - que preencha os argumentos ausentes em sentenças incompletas com valores plausíveis,
 - que seja capaz de generalizar seu conhecimento de atribuição de papel correto a sentenças que contenham uma palavra nunca vista antes, dado apenas uma especificação de algumas das propriedades semânticas da palavra.

Arquitetura do Modelo

- O modelo consiste de dois conjuntos de unidades:
 - um para representar a estrutura de superfície (ES) da sentença,
 - um para representar a estrutura de caso (EC) da sentença.
- O modelo aprende através de apresentações de pares corretos ES/EC.
- Sentenças:
 - um verbo e de um a três *SNs*.
 - Há sempre um *SN* sujeito e opcionalmente, um *SN* objeto.
 - Pode haver também um *com-SN*, isto é, um *SN* num *SP* começando com *com*.
- Formato de Entrada das Sentenças:
 - representação canônica da sentença,
 - forma produzida por um *parser* de superfície e um léxico simples.

Microcaracterísticas Semânticas

- No formato de entrada canônico, as palavras são representadas como listas de microcaracterísticas semânticas.
- Para verbos e substantivos, os traços (características) são agrupados em várias dimensões.
- Cada dimensão consiste de um conjunto de valores mutuamente exclusivos.
- Em geral, cada palavra é representada por um vetor no qual apenas um único valor em cada dimensão está ligado (representado por “1”).
- Valores que estão desligados são representados por pontos (“.”).

Microcaracterísticas Semânticas

- As dimensões e os valores em cada dimensão foram escolhidos para capturar dimensões importantes de variação semântica nos significados das palavras com implicação na atribuição de papel de caso.
- Dimensões e valores de características dos **substantivos**:

HUMANO	humano/não humano
SOFTNESS	<i>soft/hard</i>
GÊNERO	masculino/feminino/neutro
VOLUME	pequeno/médio/grande
FORMA	compacto/1-D/2-D/3-D
PONTA	pontudo/arredondado
FRAGILIDADE	frágil/inquebrável
TIPO DE	alimento/brinquedo/ferramenta/utensílio/
OBJETO	mobília/animado/natural inanimado

Microcaracterísticas Semânticas

- Dimensões e valores de características dos **verbos**:

ATOR	sim/não
CAUSA	sim/sem causa/ sem mudança
TOQUE	agente/instrumento/ambos/nenhum/Agente é Paciente
NATUREZA DA MUDANÇA	pedaços/lascas/químico/nenhum/não usado
MOVIMENTO DO AGENTE	transformação/ participação/ nenhum/ não se aplica
MOVIMENTO DO PACIENTE	transformação/ participação/ nenhum/ não se aplica
INTENSIDADE	baixa/alta

Microcaracterísticas Semânticas: Exemplo

- Microcaracterísticas do substantivo *bola*:
 - não humano; *soft*; neutro; pequeno; compacto; arredondado; inquebrável; brinquedo.
- Microcaracterísticas do verbo *quebrar*:
 - ator; causa; toque instrumento; troca em pedaços; movimento agente participação; nenhum movimento paciente; intensidade alta.

Ambigüidade

- Um das metas para o modelo é mostrar como ele pode seleccionar o significado apropriado no contexto para uma palavra ambígua.
- Para palavras ambíguas (*cravo*, por exemplo) o padrão de entrada é a média dos padrões de características de cada uma das duas leituras da palavra.
- Isto significa que nos casos onde as duas concordam com o valor de uma dimensão de entrada particular, essa dimensão tem o valor acordado na representação de entrada.
- Nos casos onde os dois discordam, a característica tem o valor de 0.5 (representado por “?”) na representação de entrada.

Sumário

- 1 Sistemas Híbridos
 - RNA Baseadas em Conhecimento
 - Extração de regras
- 2 PLN Conexionista
 - Representação
 - Solução connexionista
 - Estruturas

Unidades de Estrutura de Superfície - ES

- Um dos objetivos das simulações é ver se o modelo pode corretamente preencher esses valores não especificados, efetivamente recuperando os valores perdidos do contexto no processo da atribuição da palavra ao caso apropriado.
- A representação do nível de Estrutura de Superfície de uma sentença de entrada não é o conjunto de vetores de características constituintes, mas sim o padrão de ativação que esses vetores produzem sobre as unidades que correspondem a pares de características.
- Essas unidades são chamadas unidades de estrutura de superfície (ES).
- Cada unidade ES representa a conjunção de duas microcaracterísticas do preenchedor de um caso de superfície particular.

Unidades de Estrutura de Superfície - ES

- Como há quatro casos de estruturas de superfície, existem quatro conjuntos de unidades ES.
- Dentro de cada conjunto existe uma unidade que representa a conjunção de todo valor de microcaracterística em cada dimensão com todo valor de microcaracterística em qualquer outra dimensão.
- Considera-se vários esquemas para a ativação das unidades de ES.
- Um esquema possível seria usar uma regra de ativação determinística, tal que uma unidade ES particular seria ligada apenas se ambas as características que a unidade representa estivessem dentro do vetor de características.

Unidades de Estrutura de Superfície - ES

- Esse uso das unidades ES permitiria que o modelo aprendesse a responder de uma forma mais apurada a determinadas conjunções de microcaracterísticas.
- Entretanto, deseja-se ver o quão bem o modelo poderia funcionar usando uma representação de entrada com ruídos.
- Contudo sabe-se que a generalização é facilitada quando as unidades que unificam apenas parcialmente com a entrada têm alguma chance de serem ativadas.
- No presente caso, considera-se isto importante para ser capaz de generalizar para palavras com significados similares.
- Portanto, as unidades ES são tratadas como unidades binárias estocásticas.

Unidades de Estrutura de Superfície - ES

- Cada unidade ES recebe entrada excitatória de cada uma das duas características que ela representa e a polarização e a variância das unidades é colocada de tal forma que quando ambas as características das unidades ES estão ativas, a unidade “liga” com probabilidade 0.85; e quando nenhuma está ativa, ela “liga” com probabilidade 0.15.
- Esses casos são representados na tabela por “1” e “.”, respectivamente.
- As unidades que recebem apenas uma entrada excitatória “liga” com probabilidade .5; essas unidades são representadas por “?”.

Unidades de Estrutura de Superfície - ES

- Enquanto o modelo funciona bem com essa simulação, presume-se que simulações que usem um léxico maior requeiram maior diferenciação de algumas representações de substantivos e verbos.
- Para trabalhar com tais casos, acredita-se que seria necessário permitir o ajuste das conexões de entrada às unidades ES através do algoritmo *backpropagation* de tal forma que uma diferenciação maior possa ser obtida quando necessário.

Unidades de Estrutura de Caso - EC

- A representação de caso tem uma forma levemente diferente da representação de estrutura de sentença.
- Para entender essa representação, é útil voltar a um ponto de vista mais abstrato e considerar mais geralmente como se deve representar uma descrição estrutural numa representação distribuída.
- Em geral uma descrição estrutural pode ser representada por um conjunto de triplas da forma $(A R B)$ onde A e B correspondem aos nós na descrição estrutural e R representa a relação entre os nós.
- Por exemplo, uma hierarquia de inclusão de classes pode ser representada por triplas da forma $(X \text{ É-UM } Y)$, onde X e Y são nomes de categorias.

Unidades de Estrutura de Caso - EC

- Qualquer outra descrição estrutural, seja uma estrutura de constituinte sintático, uma estrutura de constituinte semântico ou qualquer outra coisa, pode ser representada dessa forma.
- Especificamente, a atribuição de caso dos constituintes da sentença *O garoto quebrou a vidraça com o martelo* pode ser representada como:
 - *Quebrou* AGENTE *Garoto*,
 - *Quebrou* PACIENTE *Vidraça*,
 - *Quebrou* INSTRUMENTO *Martelo*.
- A estrutura constituinte de uma sentença tal como *O garoto comeu o macarrão com molho* poderia ser representada por:
 - *Comeu* AGENTE *Garoto*,
 - *Comeu* PACIENTE *Macarrão*,
 - *Macarrão* Modificador *Molho*.

Unidades de Estrutura de Caso - EC

- Numa representação local, poder-se-ia representar cada uma dessas triplas por uma unidade única.
- Cada unidade dessas representaria então a conjunção de uma determinada cabeça ou lado esquerdo de uma tripla, uma determinada relação, e uma determinada cauda ou lado direito.
- Nessa abordagem mais distribuída seriam alocados grupos de unidades para representar cada uma das possíveis relações (ou casos), que seriam o AGENTE, PACIENTE, INSTRUMENTO e Modificador, e teriam unidades dentro de cada grupo representando conjunções de microcaracterísticas do primeiro e terceiro argumentos (a cabeça e a cauda) da tripla.

Unidades de Estrutura de Caso - EC

- Portanto, a tripla é representada não por uma única unidade ativa, mas por um padrão de ativação sobre um conjunto de unidades.
- Nessa implementação, existe um grupo de unidades para cada uma das quatro relações permitidas na estrutura de caso.
- Os grupos do AGENTE, PACIENTE, INSTRUMENTO e Modificador são colocados da esquerda para a direita.
- Dentro de cada grupo, as unidades individuais representam conjunções de uma microcaracterística da cabeça de cada relação com uma microcaracterística da cauda de cada relação.

Unidades de Estrutura de Caso - EC

- Então, por exemplo no slide 12, *Quebrou-AGENTE-Garoto* é representado por um padrão de ativação.
- A unidade na i -ésima linha e j -ésima coluna representa a conjunção da característica i do verbo com característica j do substantivo.
- Então, todas as unidades com a mesma característica de verbo estão alinhadas na mesma linha, enquanto que todas as unidades com a mesma característica de substantivo estão alinhadas na mesma coluna.
- Para o grupo do Modificador, a unidade na i -ésima linha e j -ésima coluna representaria a conjunção da característica i do *SN* modificado e a característica j do *SN* modificador.

Microcaracterísticas Semânticas

- As letras indicando as especificações de dimensão das unidades são colocadas ao longo das bordas lateral e superior.
- Exemplo de uma unidade de ES para o verbo *quebrar* e EC para *Quebrou-AGENTE-Garoto* é mostrada abaixo.
- A matriz do verbo é diferente das matrizes dos três *SNs* pois características diferentes são juntadas nas representações de verbo e de *SN*.

Microcaracterísticas Semânticas

1.1..1..1....1...1.1	
rrccttttnnnnaapppii	hhssggvvvfffpddtttt
r	r 1.1.1..1..1.1.1.1.
r	r
c 1?	c 1.1.1..1..1.1.1.1.
c ?.	c
t ?.?.	t
t 1?1?	t 1.1.1..1..1.1.1.1.
t ?.?.	t
t ?.?.	t
n 1?1??1??	n 1.1.1..1..1.1.1.1.
n ?.?....?.	n
n ?.?....?.	n
n ?.?....?.	n
a ?.?....?.?.	a
a 1?1??1??1???	a 1.1.1..1..1.1.1.1.
a ?.?....?.?.	a
p ?.?....?.?....?.	p
p ?.?....?.?....?.	p
p 1?1??1??1??1???	p 1.1.1..1..1.1.1.1.
i ?.?....?.?....?....?	i
i 1?1??1??1??1??1???	i 1.1.1..1..1.1.1.1.

Experimentos de Simulação

- Geradores para sentenças usadas no treinamento e testes:
 - o humano quebrou o objeto-frágil.
- Categorias de substantivos para preencher a tabela do gerador:
 - humano: homem, mulher, menino, menina.
 - objeto-frágil: prato, vidraça, vaso.
- Resultado: Número médio de microcaracterísticas incorretas produzidas como uma função da quantidade de experiência de aprendizado.

Resultados Básicos

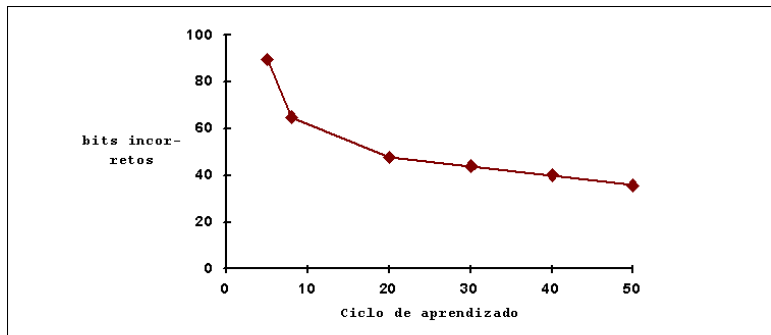


Figura 4. Experiência obtida durante o aprendizado. (McClelland e Kawamoto, 1986)

Métodos para Formar Representações Distribuídas

- Uma abordagem popular para formar representações distribuídas é a codificação de características (traços) semânticos usadas por McClelland e Kawamoto [4].
- Esse tipo de representação é significativo por si só.
- É possível extrair informação apenas examinando a representação, sem necessidade de treinar uma rede para interpreta-la.
- Vários sistemas diferentes podem processar as mesmas representações e se comunicar usando-as.

Métodos para Formar Representações Distribuídas

- Por outro lado, tais padrões devem ser pré-codificados e mantidos fixos.
- A performance não pode ser otimizada adaptando as representações às tarefas e dados reais.
- Porque todos os conceitos são classificados ao longo das mesmas dimensões, o número de dimensões se torna muito grande, e muitas delas são irrelevantes ao conceito particular (em Inglês, o gênero de substantivos).

Bibliografia I

- [1] L. M. Fu
Knowledge Base Refinement by Backpropagation.
Data and Knowledge Engineering 7, 35-46, 1991.
- [2] L. M. Fu
Knowledge-Based Connectionism for Revising Domain Theories.
IEEE Trans. on Systems, Man, and Cybernetics, Vol. 23, No.1, 173–182, 1993.
- [3] S. Haykin
Neural networks - a comprehensive foundation.
2nd. edition. Prentice Hall, 1999.

Bibliografia II

- [4] J. L. McClelland and A. H. Kawamoto
Mechanisms of Sentence Processing: Assigning Roles to Constituents of Sentences.
in *Parallel Distributed Processing, Volume 2 - Psychological and Biological Models*, Chapter 19 (pages 272–325), J. L. McClelland and D. E. Rumelhart (Eds.), A Bradford Book, MIT Press, 1986.
- [5] J. L. G. Rosa
Fundamentos da Inteligência Artificial.
Editora LTC. Rio de Janeiro, 2011.

Bibliografia III

[6] J. L. G. Rosa

Um sistema híbrido simbólico-conexionista para o processamento de papéis temáticos.

Tese de Doutorado. Instituto de Estudos da Linguagem. Universidade Estadual de Campinas, 1999.

[7] J. L. G. Rosa and E. Françoço

Hybrid Thematic Role Processor: Symbolic Linguistic Relations Revised by Connectionist Learning.

in Proceedings of IJCAI'99 - Sixteenth International Joint Conference on Artificial Intelligence, Stockholm, Sweden, July 31-August 6, Volume 2, Morgan Kaufmann, 852–857, 1999.

Bibliografia IV

- [8] J. L. G. Rosa
A Thematic Connectionist Approach to Portuguese Language Processing.
Proceedings of the IASTED International Conference on Artificial Intelligence and Soft Computing, July 27-August 1, Banff, Canada, 240–243, 1997.
- [9] R. Setiono and H. Liu
Symbolic Representation of Neural Networks.
IEEE Computer, Vol. 29, No. 3, 71–77, 1996.
- [10] G. G. Towell and J. W. Shavlik
Extracting Refined Rules from Knowledge-based Neural Networks.
Machine Learning, 13, 71–101, 1993.