**AULA 7 ELEM II 2020 (3ª AQUIS)**

**TEORIAS**

Empiristas: behaviorismo (Skinner) e conexionismo

Racionalistas: modular: inatismo (Chomsky)

construtivismo (não modular):

cognitivismo (Piaget) e interacionismo (Vygotsky)

**CONEXIONISMO**

* Fonte: <http://plato.stanford.edu/entries/connectionism/>
* O conexionismo é um movimento na ciência cognitiva que espera explicar as capacidades intelectuais humanas usando redes neurais artificiais (também conhecidas como 'redes neurais' ou 'redes neurais').
* As redes neurais são modelos simplificados do cérebro compostos por um grande número de unidades (os análogos dos neurônios) junto com pesos que medem a força das conexões entre as unidades.
* Esses pesos modelam os efeitos das sinapses que ligam um neurônio a outro.
* Experimentos com modelos desse tipo demonstraram capacidade de aprender habilidades como reconhecimento facial, leitura e detecção de estruturas gramaticais simples.
* Os filósofos se interessaram pelo conexionismo porque ele promete fornecer uma alternativa à teoria clássica da mente: a visão amplamente aceita de que a mente é algo semelhante a um computador digital que processa uma linguagem simbólica.

1. Uma descrição das redes neurais

* Uma rede neural consiste em um grande número de unidades unidas em um padrão de conexões.
* As unidades em uma rede são geralmente divididas em três classes: unidades de entrada, que recebem informações a serem processadas, unidades de saída onde os resultados do processamento são encontrados e unidades intermediárias chamadas unidades ocultas.
* Se uma rede neural modelasse todo o sistema nervoso humano, as unidades de entrada seriam análogas aos neurônios sensoriais, as unidades de saída aos neurônios motores e as unidades ocultas a todos os outros neurônios.
* Aqui está uma ilustração simples de uma rede neural simples:



* Cada unidade de entrada possui um valor de ativação que representa algum recurso externo à rede.
* Uma unidade de entrada envia seu valor de ativação para cada uma das unidades ocultas às quais está conectada.
* Cada uma dessas unidades ocultas calcula seu próprio valor de ativação, dependendo dos valores de ativação que recebe das unidades de entrada.
* Esse sinal é então passado para as unidades de saída ou para outra camada de unidades ocultas.
* Essas unidades ocultas calculam seus valores de ativação da mesma maneira e os enviam para seus vizinhos.
* No final, o sinal nas unidades de entrada se propaga por todo o caminho através da rede para determinar os valores de ativação em todas as unidades de saída.
* O padrão de ativação estabelecido por uma rede é determinado pelos pesos, ou força das conexões entre as unidades.
* Os pesos podem ser positivos ou negativos.
* Um peso negativo representa a inibição da unidade receptora pela atividade de uma unidade emissora.
* O valor de ativação para cada unidade receptora é calculado de acordo com uma função de ativação simples.
* As funções de ativação variam em detalhes, mas todas seguem o mesmo plano básico.
* A função soma as contribuições de todas as unidades emissoras, onde a contribuição de uma unidade é definida como [o peso da conexão entre as unidades emissoras e receptoras] vezes [o valor de ativação da unidade emissora].
* Esta soma é geralmente modificada posteriormente, por exemplo, ajustando a soma da ativação para um valor entre 0 e 1 e/ou definindo a ativação para zero, a menos que um nível limite para a soma seja alcançado.
* Os conexionistas presumem que o funcionamento cognitivo pode ser explicado por conjuntos de unidades que operam dessa maneira.
* Como se presume que todas as unidades calculam praticamente a mesma função de ativação simples, as realizações intelectuais humanas devem depender principalmente das configurações dos pesos entre as unidades.

A escola de modelagem computacional chamada processamento paralelo distribuído (PDP), conexionismo ou redes neurais artificiais introduziu um novo conjunto de mecanismos mentais possíveis para o campo da psicologia. Um deles, chamado de associador de padrão, projetado para memorizar pares específicos de itens em um conjunto de treinamento e generalizá-los para novas formas usando analogias parciais complexas. Em tais redes, uma entrada é geralmente representada como um padrão de ativação em uma grande coleção de "unidades", cada uma correspondendo a uma característica do item de entrada (uma "representação distribuída"). Uma saída é representada como um padrão de ativação em uma segunda coleção de unidades. Cada unidade de entrada é conectada a cada unidade de saída por um link cujo peso é modificável pelo treinamento; às vezes, uma ou mais camadas de unidades "ocultas" ficam entre as camadas de entrada e saída. Alguns teóricos tratam as unidades como modelos muito simples de neurônios e tratam as redes como modelos dos circuitos maciçamente paralelos do cérebro.

* David Rumelhart e James McClelland (1986) desenvolveram um modelo associador de padrões para a aquisição de formas do passado de verbos do inglês que adquiriu centenas de verbos regulares e irregulares, e generalizou adequadamente para dezenas de novos verbos nos quais não havia sido treinado.
* O mais impressionante é que o modelo parecia passar por uma sequência de desenvolvimento em forma de U, primeiro produzindo formas verbais irregulares corretamente e, mais tarde, superregularizando-as.
* Mas o modelo não tinha representação explícita de palavras e regras, nem qualquer distinção entre sistemas regulares e irregulares.
* Era apenas uma rede associadora de padrão único consistindo em uma matriz de unidades de entrada, uma matriz de unidades de saída e uma matriz de ligações ponderadas modificáveis ​​entre cada entrada e cada saída.
* O radical verbal foi representado ativando um subconjunto de nós de entrada, cada um correspondendo a um padrão de som no radical.
* Por exemplo, uma unidade pode representar "vogal entre duas consoantes"; outro pode representar "consoante no final da palavra"; um terceiro pode representar "vogal alta entre uma consoante sibilante e uma consoante não-soante".
* Para representar o radical verbal *pass*, a primeira e a segunda unidades (e dezenas de outras) seriam ativadas; o terceiro (e dezenas de outros) seria omitido.
* Isso enviou um sinal através de cada um dos *links* para os nós de saída, que representam os sons da forma de pretérito.
* Cada nó de saída somaria seus sinais de entrada e ligaria se a soma excedesse um limite; a forma de saída seria a palavra mais compatível com o conjunto de nós de saída ativos.
* Durante a fase de aprendizagem, a forma de pretérito computada pela rede é justaposta à versão correta fornecida por um "professor", e os pontos fortes dos *links* e limites são ajustados de modo a reduzir a diferença.
* Ao gravar e sobrepor associações entre sons radicais e sons passados, o modelo melhora seu desempenho e pode generalizar para novas formas na medida em que seus sons se sobrepõem aos antigos.
* Este processo é qualitativamente o mesmo para verbos regulares e irregulares: *stopped* é produzido porque as unidades de entrada *op* foram vinculadas às unidades de saída *opped* por verbos anteriores; *clung* é produzido porque *ing* foi vinculado a *ung*.
* É importante haver uma grande quantidade de dados fornecidos à rede. A frequência tem um papel importante
* Exemplo: plural de substantivos no português
	+ livro(s), mãe(s)
	+ mar(es), vez(es)
	+ degr**au**: degraus (regular mas pouco frequente), mas degrais tb é comum
	+ jorn**al**: jornais (irregular mas frequente)
* CONEXIONISMO importante na TEORIA DA OTIMIDADE (Prince & Smolensky 1993), importante principalmente na fonologia.

**INATISMO**

* Noam Chomsky: gramática gerativa
* teoria modular da mente e do conhecimento linguístico
* a mente é constituída de módulos: o conhecimento linguístico é um deles, sendo distinto de outros tipos de conhecimento
* o conhecimento linguístico também é modular para Chomsky
* para Chomsky, grande parte do nosso conhecimento linguístico é inato (algo determinado geneticamente)

**PRINCÍPIOS E PARÂMETROS**

* Princípios são características universais, ou seja, de **todas** as línguas
* Parâmetros são características que variam entre as línguas (o mais comum é considerar que existem duas possibilidades para cada parâmetro)
* modelo de 1981 a 1993: princípios e parâmetros. Parâmetro do núcleo (à direita ou à esquerda).
* para mim vi o filme

**PARÂMETRO DO SUJEITO NULO (OMISSÃO DO SUJEITO)**

Eu cheguei. Cheguei.

Você já chegou? Cheguei.

**Omissão e inversão.**

Uma carta chegou. (SV)

Chegou uma carta. (VS)

It is raining. Está chovendo.

Inglês: não admite sujeito nulo.

Português: admite sujeito nulo.

**DIREÇÃO/POSIÇÃO DO NÚCLEO**

Kanojo-wa [nihongo-o **hanasu**]. SV (sintagma verbal)

Ela-tóp [japonês-od falar].

Ela [**fala** japonês].

Kare-wa [hataraite **iru**]. SF (sintagma de flexão (verbal))

Ele-tóp [trabalhando estar].

Ele [**está** trabalhando].

Watashi-wa [Itaria-**e**] iku. SP (sintagma preposicional)

Eu-tóp [Itália para] ir.

Eu vou [**para** a Itália].

Kore-wa [are-to **onaji**] da. SA (sintagma adjetival)

 Isso-tóp [aquilo-com igual] ser.

 Isso é [**igual** àquilo].

* Chomsky não foi o primeiro a relacionar esse tipo de fatos. Greenberg indutivo. Chomsky dedutivo. Processo de aquisição seria apenas a fixação desses parâmetros.

**MOVIMENTO-QU**

A fita está na caixa. Onde está a fita? A fita está onde?

The tape is in the box. Where is the tape? \* The tape is where?

Teepu wa hako ni aru. Teepu wa doko ni aru ka? \*Doko ni teepu wa aru ka?

\* Teepu wa aru ka doko ni?

**MOVIMENTO-QU:**

Where is the tape? The tape is where? Is the tape where?

Onde está a fita? A fita está onde?

Japonês.

**TRÊS COMBINADOS: 8 possibilidades [na verdade, ainda há os casos intermediários]**

Inglês: não pro-drop núcleo inicial movto-qu obrigatório

Português: pro-drop núcleo inicial movto-qu facultativo

Japonês: pro-drop núcleo final sem movto-qu

|  |  |  |  |
| --- | --- | --- | --- |
| pro-drop | núcleo | movto-QU |  |
| sim | inicial | sim | espanhol |
| sim | inicial | não | chinês |
| sim | final | sim |  |
| sim | final | não | japonês |
| não | inicial | sim | inglês |
| não | inicial | não |  |
| não | final | sim |  |
| não | final | não |  |

* como funciona a aquisição de linguagem dentro desse modelo?
* os princípios são universais. Eles não precisam ser adquiridos (já vêm determinados geneticamente).
* quanto aos parâmetros, a criança não precisa adquirir quais eles são (isso já vem determinado geneticamente)
* ela só precisa fixar qual o valor de cada parâmetro:
	+ essa língua admite sujeito nulo?
	+ essa língua têm o núcleo sintático no início ou no final?

**COGNITIVISMO**

* Fonte: Field, John (2004). *Psycholinguistics. The Key Concepts*. Londres: Routledge.
* Abordagens para aquisição da linguagem que veem o processo como intimamente ligado à cognição geral e ao desenvolvimento cognitivo.
* Alguns relatos deixam em aberto a extensão em que certos aspectos da linguagem são inatos; mas todos consideram que a aquisição é impulsionada principalmente pela maneira como as capacidades cognitivas do bebê são exercidas sobre o *input* a que é exposto.
* Essas capacidades cognitivas podem refletir o desenvolvimento da consciência de objetos, relações espaciais, características definidoras, etc., ou podem assumir a forma de vieses perceptivos que inclinam a criança a reconhecer padrões no material linguístico.
* Entre as concepções sobre aquisição que podem ser caracterizadas como *cognitivistas* estão:
* Uma criança não pode expressar conceitos na linguagem, a menos que os tenha desenvolvido previamente.
* Exemplo: uma criança não pode usar a linguagem para se referir a objetos que não estão visíveis, sem ter aprendido a ideia de permanência do objeto.
* Para Piaget, a linguagem é o produto de processos cognitivos e perceptuais.
* Sua pesquisa com crianças o levou a concluir que havia **quatro estágios** de desenvolvimento cognitivo.
* Eles representam uma progressão gradual e não uma mudança repentina no comportamento; e a idade em que uma criança em particular passa por cada um varia consideravelmente.
* No entanto, todos os estágios estão intimamente ligados ao desenvolvimento linguístico.
* Tanto a linguagem quanto a cognição fazem parte de um programa de maturação em estágios, no qual operam em paralelo, apoiando-se mutuamente.
* Para o psicólogo infantil Jean Piaget (1896-1980), a linguagem era um fenômeno social e cognitivo. Não era uma faculdade modular independente, mas parte do processamento cognitivo e perceptivo geral. A aquisição da linguagem era, portanto, dependente do desenvolvimento cognitivo. O nível de linguagem da criança é determinado pelo fato de ela ter adquirido certos conceitos fundamentais e pela complexidade das operações de processamento de que era capaz.
* Piaget sugeriu que o desenvolvimento cognitivo se divide em quatro fases. Eles constituem uma progressão gradual em que os estágios anteriores são revisitados ciclicamente. A idade em que uma criança em particular passa por cada estágio varia consideravelmente. Cada estágio tem implicações para o desenvolvimento linguístico.

1. Sensório-motor (0–2 anos). A criança adquire o reconhecimento da permanência do objeto (o fato de que um objeto ainda existe mesmo quando não está à vista). Este é um pré-requisito para a formação de conceitos (incluindo conceitos lexicais). Pode ser uma consciência nascente da permanência do objeto que primeiro leva a criança a nomear as coisas e dá origem ao crescimento rápido do vocabulário por volta dos 18 meses. As primeiras palavras relacionais (‘NÃO’ ‘ACIMA’ ‘MAIS’ ‘SEI’) também refletem a permanência do objeto, com aquelas que indicam presença emergindo antes daquelas (‘TUDO SEJA’) relacionadas à ausência. A linguagem da criança tem suas origens em sinais simples (uma garrafa significa comer) e, em seguida, em relações indiciais (um cuidador com um casaco significa sair). As primeiras palavras são empregadas para referência simbólica (CACHORRO ou AU-AU se referindo a um cachorro específico que está presente), mas posteriormente adquirem sentido simbólico ("cachorro" se referindo à classe dos cães). As produções da criança podem mostrar uma consciência dos meios-fins (a palavra LEITE dá uma bebida à criança) e uma consciência espacial limitada.

2. Pré-operatório (2–6 anos). O comportamento da criança reflete o pensamento egocêntrico: ela é incapaz de se identificar com as opiniões dos outros. A linguagem da criança progride através de ecolalia (repetir as declarações dos outros) para monólogos (falar em voz alta o que normalmente seriam pensamentos privados). Pode envolver-se em monólogos coletivos com outras crianças, nos quais os participantes parecem se revezar, mas expressam suas próprias ideias sem responder às dos outros.

3. Operações concretas (7-11 anos). O vocabulário da criança mostra sinais de organização em categorias hierárquicas. Animal: pássaro: papagaio.

Ele desenvolve o conceito de conservação (o reconhecimento de que o tamanho ou a quantidade não depende do recipiente) e mostra sinais de descentralização, a capacidade de considerar vários aspectos de um problema físico.

Ele aprende a receber e responder a idéias externas.

4. Operações formais (11-15 anos). O adolescente torna-se capaz de raciocínio abstrato. Ele aprende a construir suas próprias estruturas de argumento, pode representar situações hipotéticas e se engaja mental e verbalmente na resolução de problemas.

* Tendências cognitivas inatas podem nos predispor:

a. A encontrar padrões em dados de linguagem (como em dados em geral). Uma teoria de *bootstrapping* sintático postula que os bebês chegam a conclusões sobre as palavras com base em suas flexões e outras propriedades gramaticais: assim, a criança aprende que a diferença entre é irmão e é irmão serve para distinguir entidades do mundo real que são massivas daquelas que são contados.

b. Adotar certas estratégias em resposta aos dados de linguagem. Slobin (1973) conclui que os bebês aplicam um conjunto de estratégias universais ou princípios operacionais para desconstruir o input ao qual são expostos. (Prestar atenção ao final das palavras. Prestar /atenção à ordem das palavras e morfemas.) Características cognitivamente mais complexas são adquiridas posteriormente.

c. Aplicar estilos de aprendizagem individuais aos dados linguísticos. Alguns bebês parecem dividir o *input* em palavras, enquanto outros adquirem pedaços de linguagem de uma maneira holística.

* A capacidade cognitiva limitada do bebê o torna mais sensível às características da linguagem do que poderia ser antes ou depois. O argumento "menos é mais" sustenta que podem ser as próprias limitações do estado cognitivo inicial do bebê que permitem que ele identifique a estrutura da linguagem e reconheça que a linguagem constitui um conjunto de símbolos inter-relacionados.