

## **Conexionismo e Redes Neurais**

A psicologia conexionista é um campo que se apóia nas chamadas redes neurais artificiais como veículo para a expressão matemática dos seus conceitos e teorias. Ela não é o único campo do conhecimento que faz isso, o que gera certa confusão entre estudantes iniciantes e leigos em geral que têm dificuldade em entender que o termo “redes neurais artificiais” pode significar diferentes coisas em contextos diferentes. Por esse motivo, a primeira coisa que será feita neste curso será uma tentativa de mostrar os diferentes significados do termo “redes neurais artificiais”.

O termo “redes neurais” se sobrepõe a alguns campos distintos:

- **Redes Neurais Biológicas** (Neurociência Computacional):
  - Ênfase na biologia: neurociência;
  - Objetivo: modelar o processamento de informação nos sistemas neurais biológicos, com base nas evidências anatômicas, fisiológicas e comportamentais;
  - Principais ferramentas: modelos biofísicos para a membrana neuronal, canais iônicos e contatos sinápticos, implementados por equações diferenciais resolvidas numericamente com o auxílio de neuro-simuladores (e.g., NEURON, NEST, BRIAN) ou outros programas; modelos “reduzidos”, tratáveis analítica ou numericamente com o uso de técnicas de sistemas dinâmicos e mecânica estatística.

- **Redes Neurais Artificiais** (Subdomínio da Estatística, da Inteligência Artificial e do Aprendizado de Máquina):
  - Ênfase nas aplicações: aprendizado de máquina, engenharia, processamento de dados e imagens, medicina, finanças, etc;
  - Objetivo: Resolver problemas de reconhecimento, classificação e agrupamento de padrões, controle adaptativo e previsão de séries temporais;
  - Principais ferramentas: modelos compostos por elementos (nós ou unidades) inspirados nos neurônios biológicos e interligados em rede cujos pesos das conexões são modificados por algoritmos de aprendizado supervisionado ou não-supervisionado (backpropagation, aprendizado hebbiano, etc), implementados via pacotes comerciais ou outros programas; técnicas estatísticas de processamento e tratamento de dados.
  - Recentemente, a área tem sido denominada de “Aprendizagem profunda” (*Deep learning*).
- **Modelos Conexionistas** (Psicologia e Ciência Cognitiva)
  - Ênfase nos processos mentais: psicologia, ciência cognitiva, neurociência cognitiva;
  - Objetivo: modelar os processos cognitivos e psicológicos usando modelos inspirados pela arquitetura do cérebro, representando a informação e os “estados mentais” por padrões de atividade distribuídos por uma rede de unidades simples ao invés de por estruturas simbólicas discretas;

- Principais ferramentas: modelos similares aos dos dois tópicos anteriores; simulações computacionais desses modelos.

### **Definição de Modelo Conexionista**

Segundo McClelland, Rumelhart e Hinton (1986), no capítulo 1 do clássico livro *PDP* (ver Bibliografia no fim desta aula):

“Esses modelos (conexionistas) assumem que o processamento de informação ocorre pela interação de um grande número de elementos processadores simples chamados de unidades, cada um enviando sinais excitatórios e inibitórios para os outros”.

Um modelo conexionista é, portanto, um sistema que possui uma forte inspiração na estrutura do cérebro, formado por redes de unidades interconectadas entre si por sinapses excitatórias e inibitórias.

Em geral, as propriedades do cérebro são apresentadas como o grande argumento para o estudo de redes neurais artificiais e modelos conexionistas.

O cérebro é superior a um computador digital em uma enorme variedade de tarefas. Aquelas tarefas que para nós são realizadas da maneira mais simples e direta, sem nem nos darmos conta de que as realizamos, são as mais difíceis de ser implementadas em um computador digital.

Algumas das características do cérebro que motivam o estudo de redes neurais artificiais são:

- O cérebro é robusto e tolerante a falhas: neurônios podem morrer sem que isso afete significativamente o desempenho do cérebro;
- O cérebro é flexível, adaptável e plástico. Ele pode se ajustar a uma nova situação ou condição aprendendo a partir dos novos exemplos;
- O cérebro pode lidar com informações corrompidas ou perturbadas por ruído;
- O cérebro processa informação de forma altamente paralela;
- O cérebro é pequeno, compacto e dissipa muito pouca potência.

Para os cientistas cognitivos, porém, o apelo dos modelos de redes neurais artificiais ou conexionistas não se deve exclusivamente à sua similaridade com o cérebro. Ainda no capítulo 1 do livro *PDP*, McClelland, Rumelhart e Hinton (1986) escrevem:

“Apesar de que o apelo dos modelos conexionistas é definitivamente fortalecido pela sua plausibilidade fisiológica e inspiração neural, essas não são as bases primárias do seu apelo para nós. Nós somos, acima de tudo, cientistas cognitivos, e os modelos conexionistas nos chamam a atenção por razões psicológicas e computacionais. Eles contêm a esperança de que se possam oferecer descrições mecanicistas computacionalmente suficientes e psicologicamente acuradas dos fenômenos da cognição humana que têm

escapado a uma explicação satisfatória via formalismos computacionais convencionais”.

Notem que, para os autores, o interesse pelo conexionismo não vem do seu apelo biológico, mas do seu apelo computacional.

O “formalismo computacional convencional” ao qual eles se referem é o que pode ser chamado de paradigma simbólico para a cognição.

Segundo o paradigma simbólico, a cognição envolve a manipulação de símbolos. No caso dos fenômenos cognitivos, os símbolos possuem um significado semântico. Eles são considerados como entidades duradouras que podem ser armazenadas e retiradas da memória e manipuladas ou transformadas de acordo com regras. As regras de manipulação de símbolos definem a sintaxe da cognição; elas determinam como os símbolos podem ser compostos e transformados gerando algum desempenho cognitivo.

Por outro lado, segundo o paradigma conexionista, a cognição resulta do processamento coletivo feito pelas várias unidades simples que compõem uma rede neural. Dado algum padrão de atividade inicial cobrindo todas ou algumas das unidades da rede, esse padrão vai se propagando pela rede via as conexões excitatórias e inibitórias interligando as unidades até que um estado estável seja atingido (o estado estável não precisa ser estático; ele pode ser dinâmico, como um padrão que se repete ciclicamente, por

exemplo). Esse estado final é interpretado como a resposta dada pelo sistema ao padrão inicial, determinando o seu desempenho cognitivo.

Apesar de colocar ênfase mais no aspecto computacional do que no biológico das redes neurais, os modelos conexionistas não têm como escapar de serem chamados de reducionistas.

O reducionismo pode ser considerado como a estratégia padrão utilizada pelas ciências (pelo menos no Ocidente) para entender o mundo. Por exemplo, a física e a química procuram reduzir o comportamento da matéria aos seus componentes moleculares e atômicos. Um programa similar existe em biologia, em que a vida e as propriedades dos organismos vivos seriam redutíveis à ação de genes e proteínas.

Da mesma forma, os modelos conexionistas procuram reduzir os complexos fenômenos cognitivos e mentais a redes neurais compostas por unidades mais simples interconectadas por sinapses.

Pode-se argumentar que toda forma de explicação implica em um tipo de reducionismo, por tentar explicar algo inicialmente inexplicável em termos de conceitos ou componentes mais conhecidos e familiares. Por exemplo, quando se procura pelo significado de uma palavra desconhecida num dicionário o que se obtém é uma definição dela em termos de palavras mais simples cujo significado é conhecido. De fato, ao longo da história, várias têm sido as tentativas de explicar a mente humana em termos de construtos

e conceitos (muitas vezes metafóricos) tomados como de natureza mais fundamental, ou seja, de símbolos. Pode-se concluir, portanto, que tanto o paradigma simbólico como o paradigma conexionista são reducionistas.

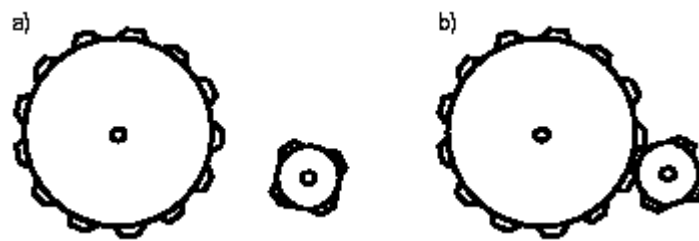
O que caracteriza a essência do conexionismo não é apenas o reducionismo, mas a exigência de que a redução seja feita em termos de componentes baseados no substrato físico da cognição humana, o cérebro. Segundo O'Reilly e Munakata (2000), no livro que pode ser considerado como a versão do clássico *PDP* para o início do Século XXI, isso é chamado de **reducionismo físico**.

A adoção do reducionismo físico coloca em cena o assim chamado problema do **reconstrucionismo** (ver O'Reilly e Munakata, 2000): para explicar a cognição, não basta apenas reduzi-la a componentes mais simples e fundamentais (os neurônios e suas interconexões), mas explicar como a combinação desses componentes produz (reconstrói) a cognição.

Na área conhecida como **sistemas complexos**, que vem se desenvolvendo rapidamente nas últimas décadas, um tipo de explicação que se costuma dar para o aparecimento de padrões complexos a partir da combinação de elementos constitutivos mais simples é que esses padrões complexos são **fenômenos emergentes**, isto é, em que o todo é maior do que a soma das partes. Segundo essa explicação, embora cada neurônio individual não tenha capacidades cognitivas ou seja capaz de pensar, os bilhões de

neurônios do cérebro com suas trilhões de conexões sinápticas produzem a mente humana e todos os fenômenos associados a ela.

A figura abaixo, retirada do livro de O'Reilly e Munakata (2000), ilustra a importância do reconstrucionismo para se entender um sistema. Não basta apenas dizer que o sistema é composto por duas componentes (as duas engrenagens mostradas em a), mas deve-se especificar como elas interagem para produzir um fenômeno novo (mostrado em b). Por exemplo, se for a engrenagem menor que move a engrenagem maior, teremos uma baixa velocidade angular e um maior torque; mas se for a engrenagem maior que move a menor, teremos alta velocidade angular e baixo torque.



Como cada engrenagem sozinha não pode alterar seu padrão de movimento rotacional, necessitando da interação com a outra engrenagem para tal, podemos dizer que os “comportamentos” novos surgem da união das duas engrenagens e do tipo de interação entre elas.

O problema do reconstrucionismo torna ainda mais evidente a importância do uso de modelos computacionais para o conexionismo, pois a pretendida reconstrução da cognição humana a partir de neurônios e sinapses não tem como ser feita por meios puramente verbais.



## **Bibliografia:**

1. Ellis, R. and Humphreys, G., *Connectionist Psychology: a text with readings.*, Psychology Press, Publishers, Hove, UK, 1999.
2. Rumelhart, D. E. and McClelland, J. L., *Parallel Distributed Processing: explorations in the microstructure of cognition*, vols. 1 e 2, The MIT Press, Cambridge, MA, EUA, 1986.
3. O'Reilly, R.C. and Munakata, Y., *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain.* MIT Press, 2000. Disponível eletronicamente no site: <http://psych.colorado.edu/~oreilly/cecn/cecn.html>.
4. Anderson, J. A., *An Introduction to Neural Networks*, The MIT Press, Cambridge, MA, EUA, 1995.
5. Schmajuk, N. A., *Animal Learning and Cognition: a neural network approach*, Cambridge University Press, Cambridge, UK, 1997.
6. Bechtel, W. and Abrahamsen, A., *Connectionism and the Mind*, Second Edition, Blackwell, Oxford, UK, 2002.
7. Levine, D. S., *Introduction to Neural and Cognitive Modeling*, Second Edition, Lawrence Erlbaum Associates, Mahwah, NJ, USA.
8. Anderson, J. A. and Rosenfeld, E. (Eds), *Neurocomputing: Foundations of Research*, MIT Press, Cambridge, MA, USA, 1989.
9. Anderson, J. A., Pellionisz, A. and Rosenfeld, E. (Eds), *Neurocomputing 2: Directions for Research*, MIT Press, Cambridge, MA, USA, 1990.
10. Pfeifer, R. and Scheier, C., *Understanding Intelligence*, MIT Press, Cambridge, MA, EUA, 1999.

11. Schmajuk, N., Brain-behaviour relationship in latent inhibition: a computational model. *Neuroscience and Biobehavioral Reviews*, 29: 1001-1020, 2005.
12. Elman, J. L., Connectionists models of cognitive development: where next? *Trends in Cognitive Sciences*, 9: 111-117, 2005.
13. Cruse, H., *Neural Networks as Cybernetic Systems*, 2<sup>nd</sup> and revised edition. Livro publicamente disponível no site da revista eletrônica *Brains, Minds and Media*: <http://www.brains-minds-media.org/archive/615>.
14. Nielsen, M. A., *Neural Networks and Deep Learning*, Determination Press, 2015. Livro online disponível no sítio: <http://neuralnetworksanddeeplearning.com/index.html>