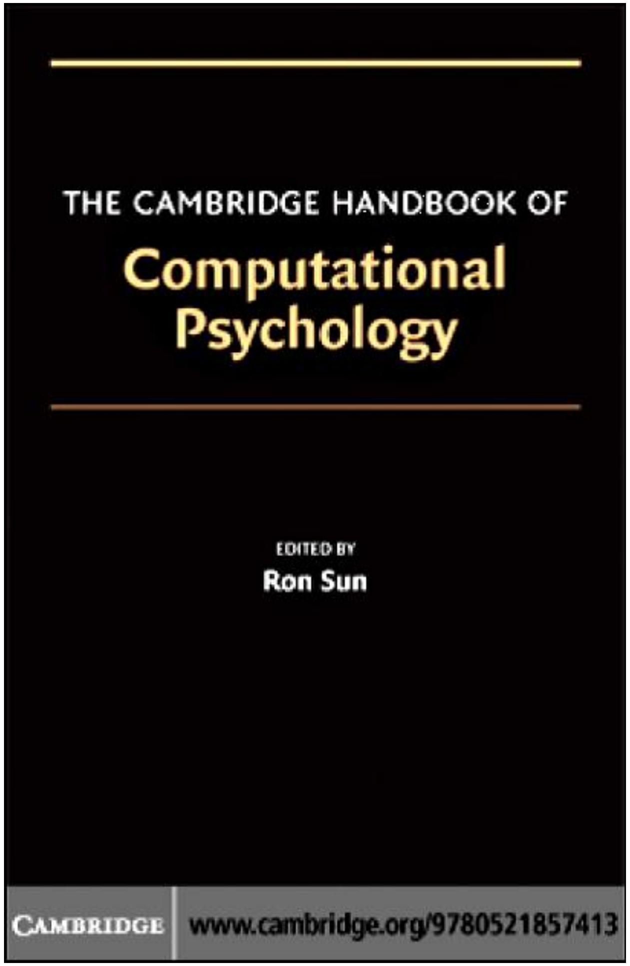
认知的连接主义模型

Q.L.

09-10-25



Chapter 2

Connectionist Models of Cognition

Michael S. C. Thomas and James L. McClelland

Michael S. C. Thomas

Centre for Brain and Cognitive Development, School of Psychology, Birkbeck University of London,

Position: Reader in Cognitive Neuropsychology
Postgraduate Tutor

My primary interests are in cognitive and language development, both in terms of developmental processes in children and in the final cognitive structures they produce in the adult.

James L. McClelland

Professor, Department of Psychology
Director, Center for Mind, Brain and Computation
Stanford University

James L. (Jay) McClelland (born December 1, 1948) is a Professor of Psychology at Stanford University. He is best known for his work concerning Parallel Distributed Processing, applying connectionist models (or neural networks) to explain cognitive phenomena such as spoken word recognition and visual word recognition. McClelland is to a large extent responsible for the "connectionist revolution" of the 1980s, which saw a large increase in scientific interest for connectionism..

----wikipedia

Over his career, McClelland has contributed to both the experimental and theoretical literatures in a number of areas, most notably in the application of connectionist/parallel distributed processing models to problems in perception, cognitive development, language learning, and the neurobiology of memory.

He was a co-founder with David E. Rumelhart of the Parallel Distributed Processing research group, and together with Rumelhart he led the effort leading to the publication in 1986 of the two-volume book, *Parallel Distributed Processing*, in which the parallel distributed processing framework was laid out and applied to a wide range of topics in cognitive psychology and cognitive neuroscience.

McClelland and Rumelhart jointly received the 1993 Howard Crosby Warren Medal from the Society of Experimental Psychologists, the 1996 Distinguished Scientific Contribution Award from the American Psychological Association, the 2001 Grawemeyer Prize in Psychology, and the 2002 IEEE Neural Networks Pioneer Award for this work.

McClelland has served as Senior Editor of *Cognitive Science*, as President of the Cognitive Science Society, and as a member of the National Advisory Mental Health Council, and he is currently president-elect of the Federtation of the Behavioral, Psychological, and Cognitive Sciences. He is a member of the National Academy of Sciences, and he has received the APS William James Fellow Award for lifetime contributions to the basic science of psychology.

一句话总结：这家伙很牛

Outline

- Background
- Some Illustrative Models
- Connectionist Influences on Cognitive Theory
- Conclusions

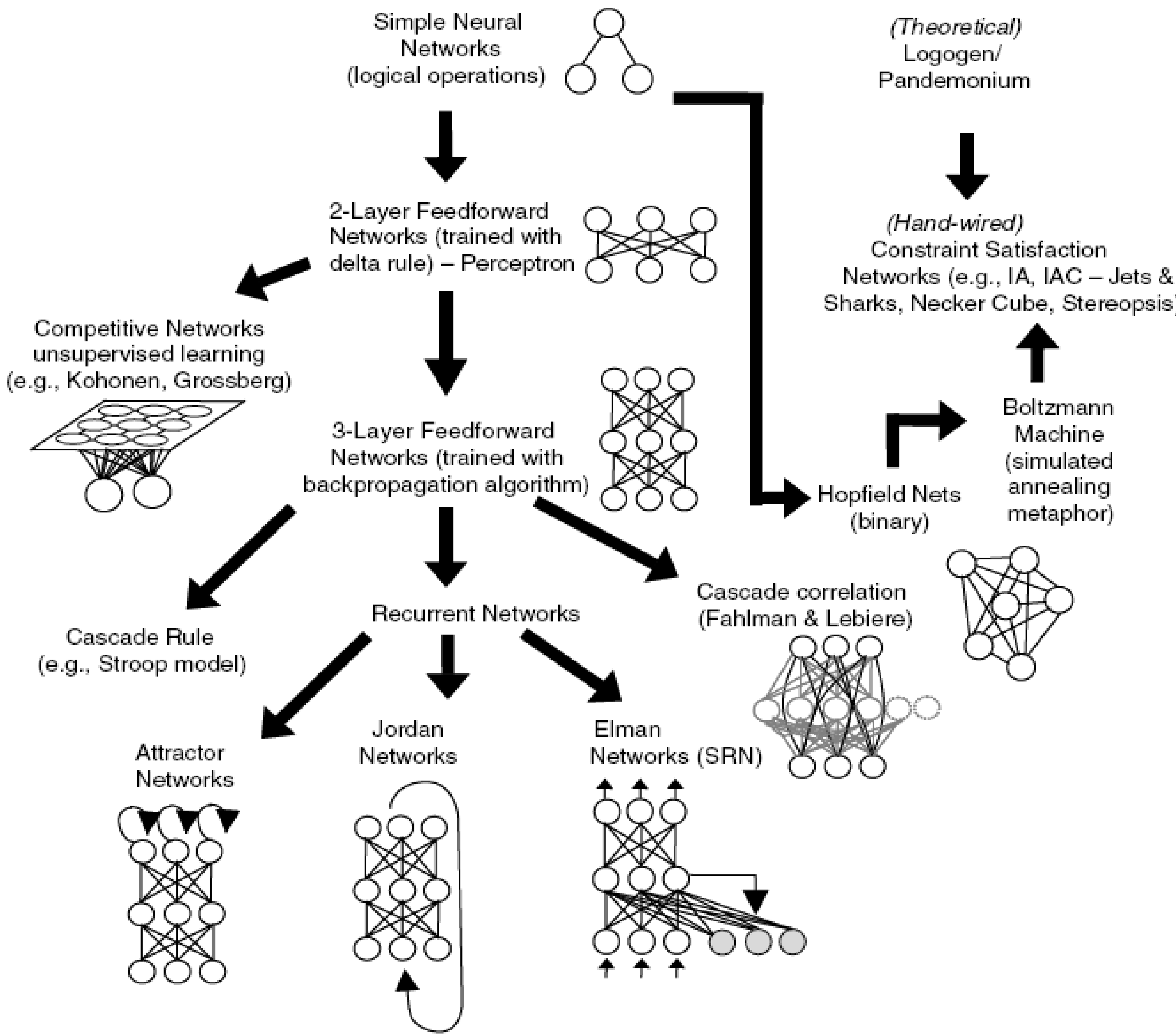
Background

What is Connectionist?

- Also known as Artificial neural network (ANN) or Parallel distributed processing (PDP) models

- has been applied to a diverse range of cognitive abilities, including models of memory, attention, perception, action, language, concept formation, and reasoning Although many of these models seek to capture adult function, connectionism places an emphasis on learning internal representations.

Histration context



- **萌芽期（20世纪40年代）**

☛1943年，心理学家McCulloch和数学家Pitts建立起了著名的阈值加权和模型，简称为**M-P模型**。发表于数学生物物理学会刊《Bulletin of Methematical Biophysics》

☛ 1949年，心理学家D. O. Hebb提出神经元之间突触联系是可变的假说——**Hebb学习律**。

- **第一高潮期（1950~1968）**

☛以Marvin Minsky，Frank Rosenblatt，Bernard Widrow等为代表人物，代表作是**单级感知器（Perceptron）。**

☛可用电子线路模拟。

☛人们乐观地认为几乎已经找到了智能的关键。许多部门都开始大批地投入此项研究，希望尽快占领制高点。

- **反思期（1969~1982）**

☛ M. L. Minsky和S. Papert，《Perceptron》，MIT Press，1969年

☛异或”运算不可表示

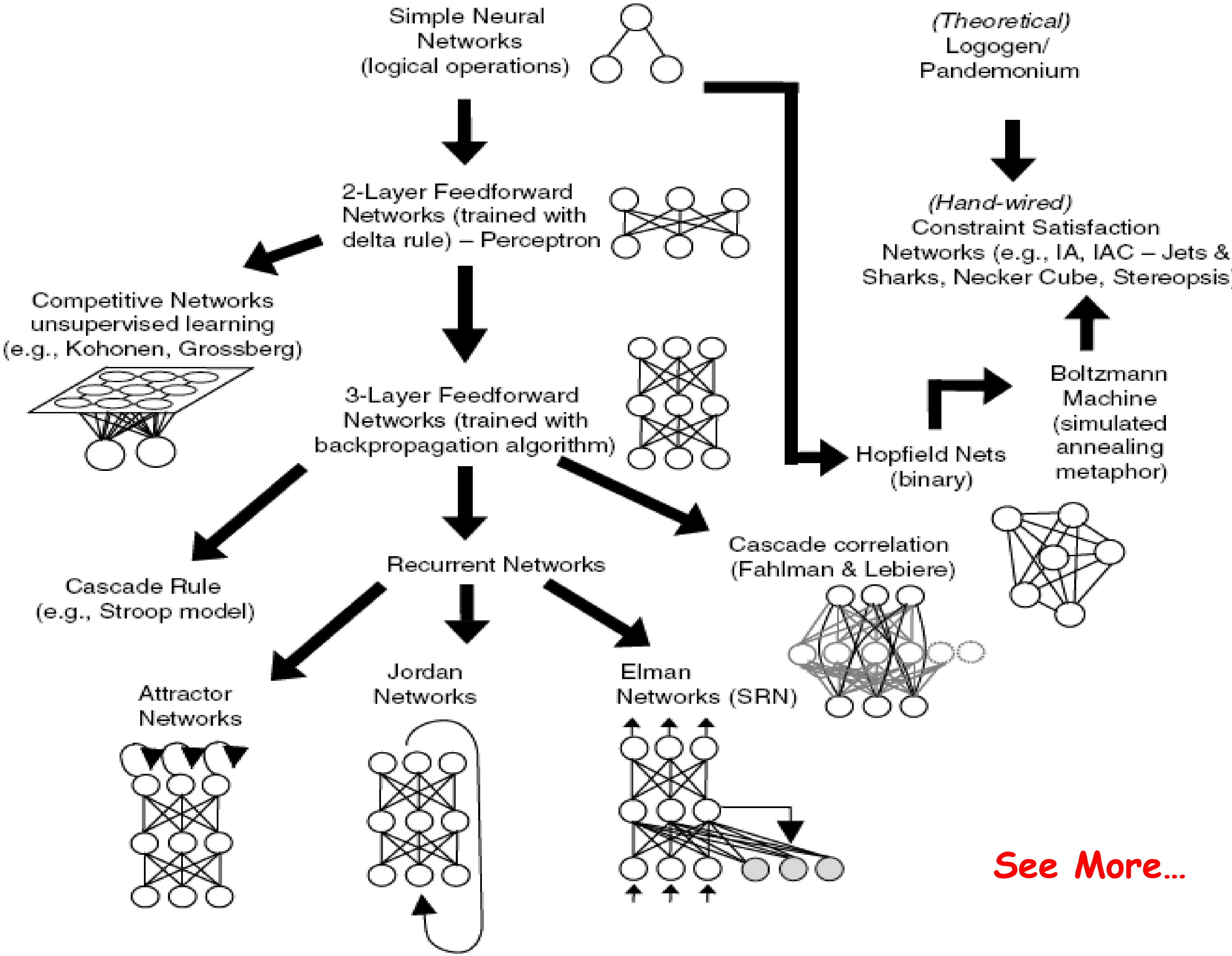
- **第二高潮期（1983~1990）**
  - ☛ 1982年，J. Hopfield提出**Hopfield网络**
    - 用Lyapunov函数作为网络性能判定的能量函数，建立ANN稳定性的判别依据
    - 阐明了ANN与动力学的关系
    - 用非线性动力学的方法来研究ANN的特性
    - **指出信息被存放在网络中神经元的联接上**
  - ☛ 1984年， J. Hopfield设计研制了后来被人们称为Hopfield网-Tank 电路。较好地解决了著名的TSP问题，找到了最佳解的近似解，引起了较大的轰动。
  - ☛ 1985年，Hinton、Sejnowsky、Rumelhart等人所在的并行分布处理（PDP）小组的研究者在Hopfield网络中引入了随机机制，提出所谓的**Boltzmann机**。
  - ☛ 1986年，并行分布处理小组的Rumelhart等研究者重新独立地提出多层网络的学习算法——**BP算法**，较好地解决了多层网络的学习问题。（Paker1982和Werbos1974年）

- 徐雷提出的 Ying-Yang 机理论模型
- 甘利俊一（ S. Amari） 开创和发展的基于统计流形的方法应用于人工神经网络的研究，

- 国内首届神经网络大会是1990年12月在北京举行的。

注：以上3页引自中科院计算所史忠植老师课件：

《神经信息学——平行分布式理论框架》


Simple Neural Networks (logical operations)
(Theoretical) Logogen/ Pandemonium
2-Layer Feedforward Networks (trained with delta rule) – Perceptron
(Hand-wired) Constraint Satisfaction Networks (e.g., IA, IAC – Jets & Sharks, Necker Cube, Stereopsis)
Competitive Networks unsupervised learning (e.g., Kohonen, Grossberg)
3-Layer Feedforward Networks (trained with backpropagation algorithm)
Boltzmann Machine (simulated annealing metaphor)
Hopfield Nets (binary)
Cascade correlation (Fahlman & Lebiere)
Recurrent Networks
Cascade Rule (e.g., Stroop model)
Attractor Networks
Jordan Networks
Elman Networks (SRN)
See More…

Key properties of Connectionist Models

1） a set of processing **units** ---u_i

2） a state of **activation** at a given time ---a(t)

3） a pattern of connectivity ---w_{ij}

4） a rule for **propagating** activation states throught the network

$$net_i = W \times a(t) = \sum_j w_{ij} a_j. \qquad (2.1)$$

5） an activation rule to specify how the net inputs to a given unit are combined to produce its new activation state ---F

$$a_i(t+1) = F(net_i(t)). \qquad (2.2)$$

6） the algorithm for modifying the patterns of connectivity as a function of experience

Hebbian learning rule $\Delta w_{ij} = \eta\, a_i a_j$ (2.3)

delta rule $\Delta w_{ij} = \eta\,(t_i(t) - a_i(t))\, a_j.$ (2.4)

backpropagation

7） a representation of the **environment** with respect to the systerm.

Neural plausibility

- **Whether these “brain-like” systems are indeed neurally plausible?**
- If they are not, should they instead be viewed as a class of statistical function approximators?
- And if so, shouldn’t the ability of these models to simulate patterns of human behavior be assessed in the context of the large number of free parameters they contain (e.g., in the weight matrix; Green, 1998)?

The advantage of connectionism, according to its proponents, is that it provides *better theories of cognition*.

Two sorts of criticism

- Many connectionist models either include properties that are not neurally plausible or omit other properties that neural systems appear to have.
- Endeavoring to show how features of connectionist systems might in fact be realized in the neural machinery of the brain.
- Stressing the cognitive nature of current connectionist models.

- Why connectionist models should be reckoned any more plausible as putative descriptions of cognitive processes just because they are "brain-like"?
- Connection vs symbolic

The relationship between Connectionist Models & Bayesian Inference

- There are strong links between the calculations carried out in connectionist models and key elements of Bayesian calculations (see chapter 3)

- First of all, that units can be viewed as playing the role of probabilistic hypotheses.

- Second, in stochastic neural networks, a network’s state over all of its units can represent a constellation of hypotheses about an input, and (if the weights and the biases are set correctly) that the probability of finding the network in a particular state is monotonically related to the probability that the state is the correct interpretation of the input.

连接主义模型与概率统计模型（如贝叶斯模型）

连接主义模型：

向底层发展——〉生物神经网络的构建。Blue Brain.模拟真实的神经结构，可用于模拟药物测试等。

向高层发展——〉建模，为认知心理学中的某些现象提供解释。模型的结构不一定符合生物的神经网络结构，可以看作是基于统计数据提出的数学模型（概率统计模型）。

概率统计模型：

本人了解的不多。不过根据miner的说法，差不多随便什么和认知相关的事都能用贝叶斯模型来解释，很强大。这部分将由jake在下下次活动时详细介绍。

Some Illustrative Models

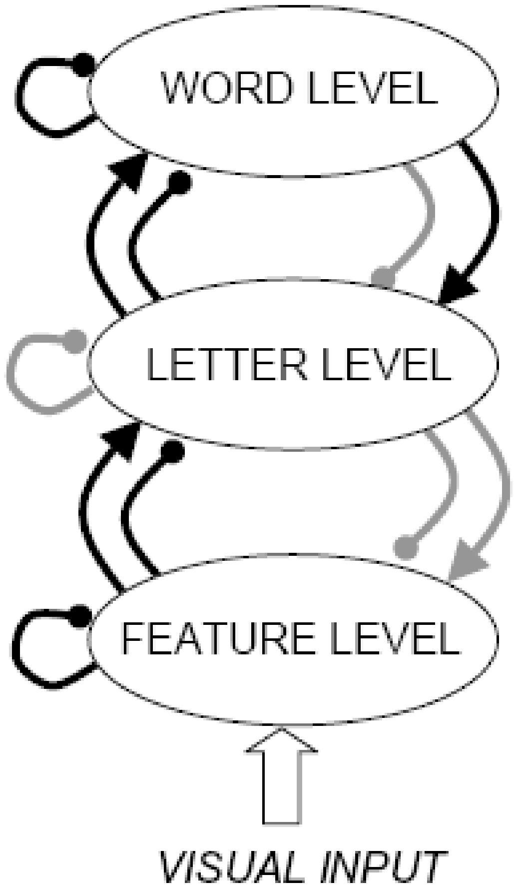
An Interactive Activation Model of Context Effects in Letter Perception

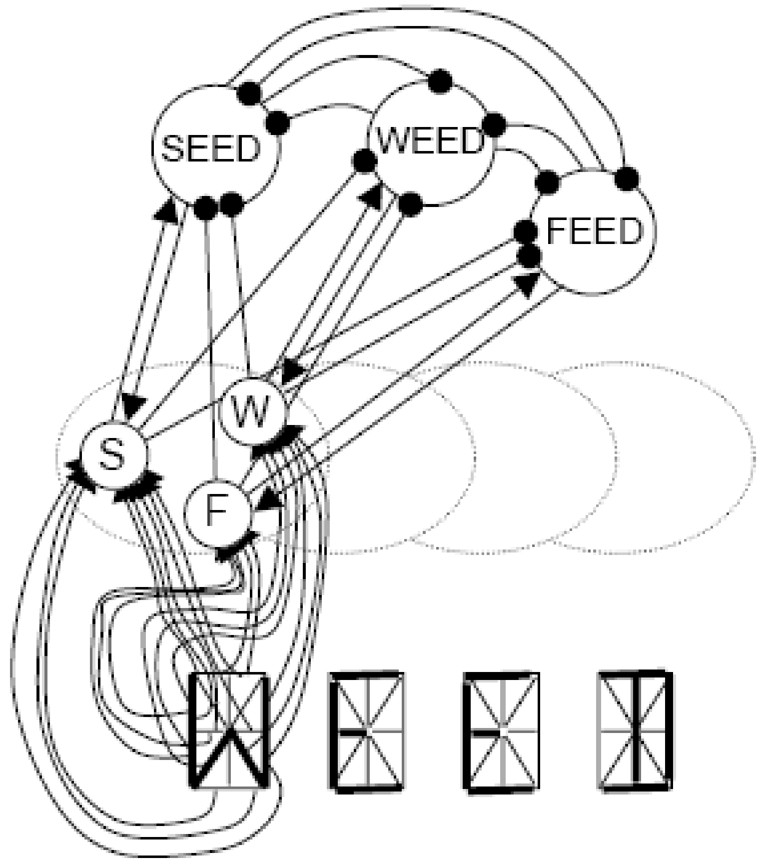
(McClelland & Rumelhart, 1981,Rumelhart & McClelland, 1982)

用于解释“词优越效应”

The protocol

He(the scientist) presented target letters in words, in unpronounceable nonwords, or on their own. The stimuli were then followed by a pattern mask, after which participants were presented with a forced choice between two letters in a given position. Importantly, both alternatives were equally plausible. Thus, the participant might be presented with WOOD and asked whether the third letter was O or R. As expected, forced-choice performance was more accurate for letters in words than for letters in nonwords or letters presented on their own.





识别在词中的字母时，自上而下（词—〉字母）与自下而上（形状—〉字母）相结合，最容易识别

three main assumptions of the IA model:

(1) Perceptual processing takes place in a system in which there are several levels of processing, each of which forms a representation of the input at a different level of abstraction;

(2) visual perception involves parallel processing, both of the four letters in each word and of all levels of abstraction simultaneously;

(3) perception is an interactive process in which conceptually driven and data driven processing provide multiple, simultaneously acting constraints that combine to determine what is perceived.

But

(1) not adaptive – connectivity was set by hand.

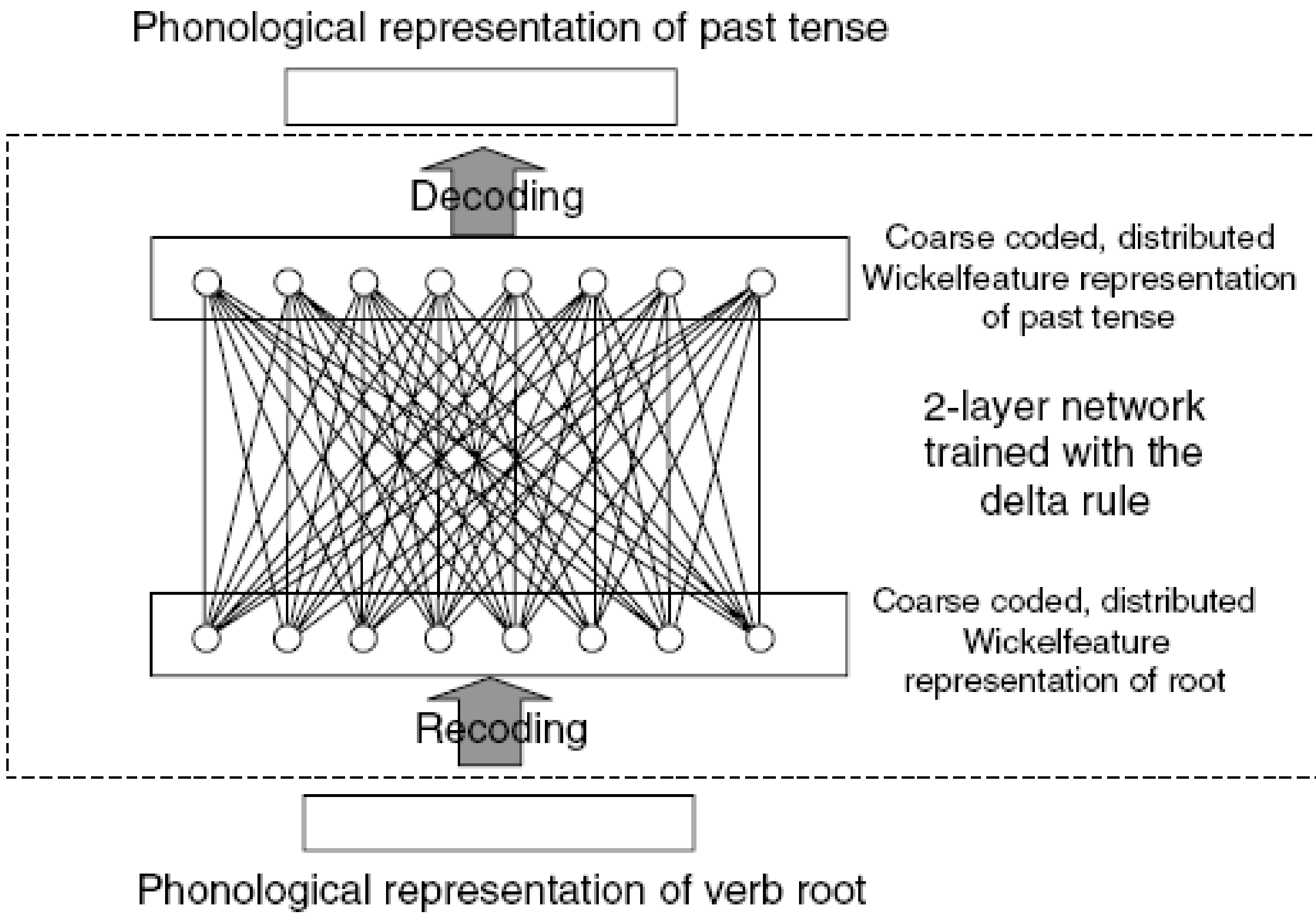
(2) the idea of bottom-up excitation followed by competition among mutually exclusive possibilities is a strategy familiar in Bayesian approaches to cognition.

On Learning the Past Tense of English Verbs

(Rumelhart & McClelland, 1986)

用于解释儿童语言学习的**U**形现象

During the acquisition of the English past tense, children show a characteristic **U-shaped developmental profile** at different times for individual irregular verbs. Initially, they use the correct past tense of a small number of high-frequency regular and irregular verbs. Later, they sometimes produce “overregularized” past tense forms for a small fraction of their irregular verbs (e.g., thinked; Marcus et al., 1992), along with other, less frequent errors (Xu & Pinker, 1995). They are also able to extend the past tense “rule” to novel verbs (e.g., wugwugged). Finally, in older children, performance approaches ceiling on both regular and irregular verbs (Berko, 1958; Ervin, 1964; Kuczaj, 1977)



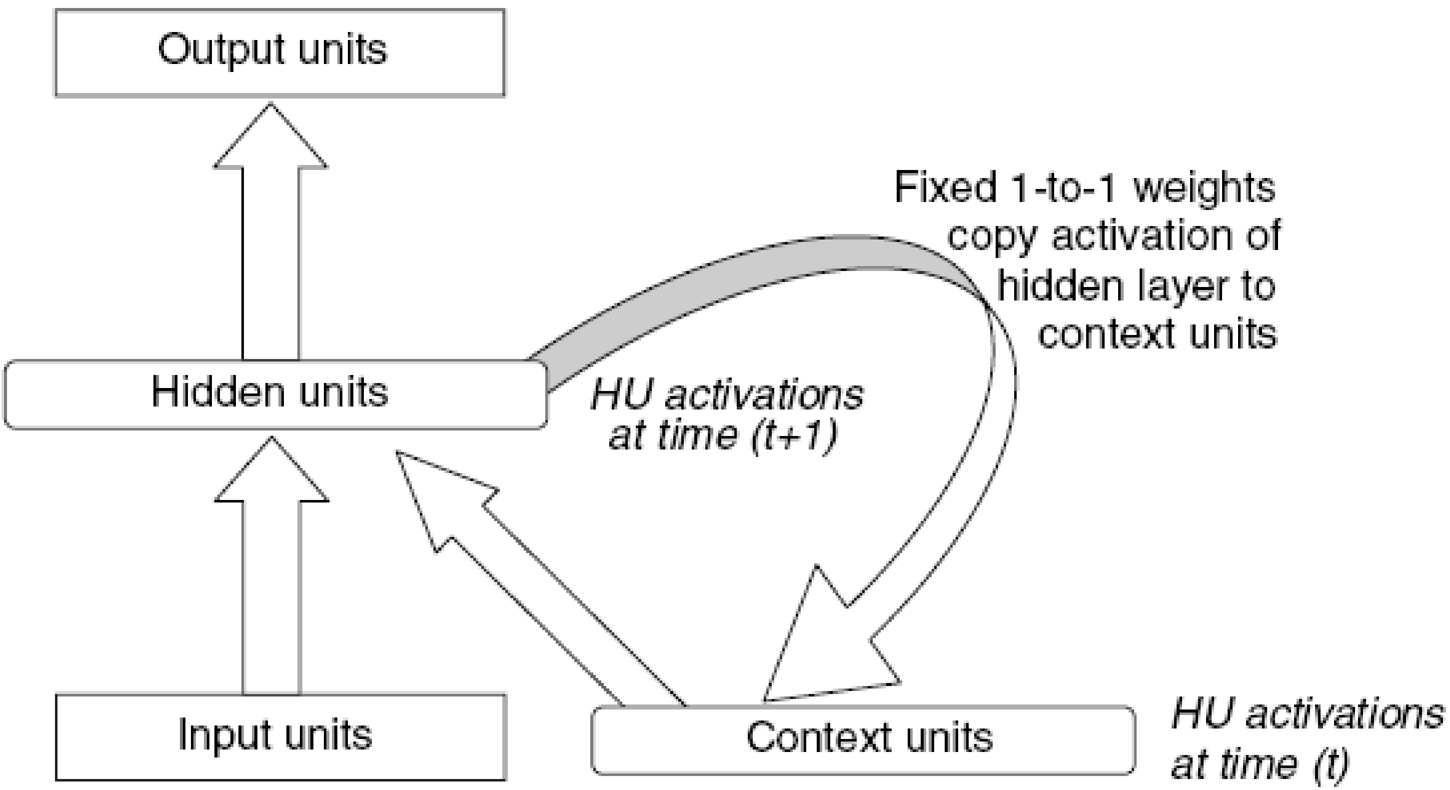
能通过学习自发“掌握”英语单词过去时态的变化规则

Finding Structure in Time

(Elman, 1990)

英语语法学习：动词和主语数的一致

Recurrent网络（前两个是feed forward网络），前一时刻的输入在下一时刻仍有影响。



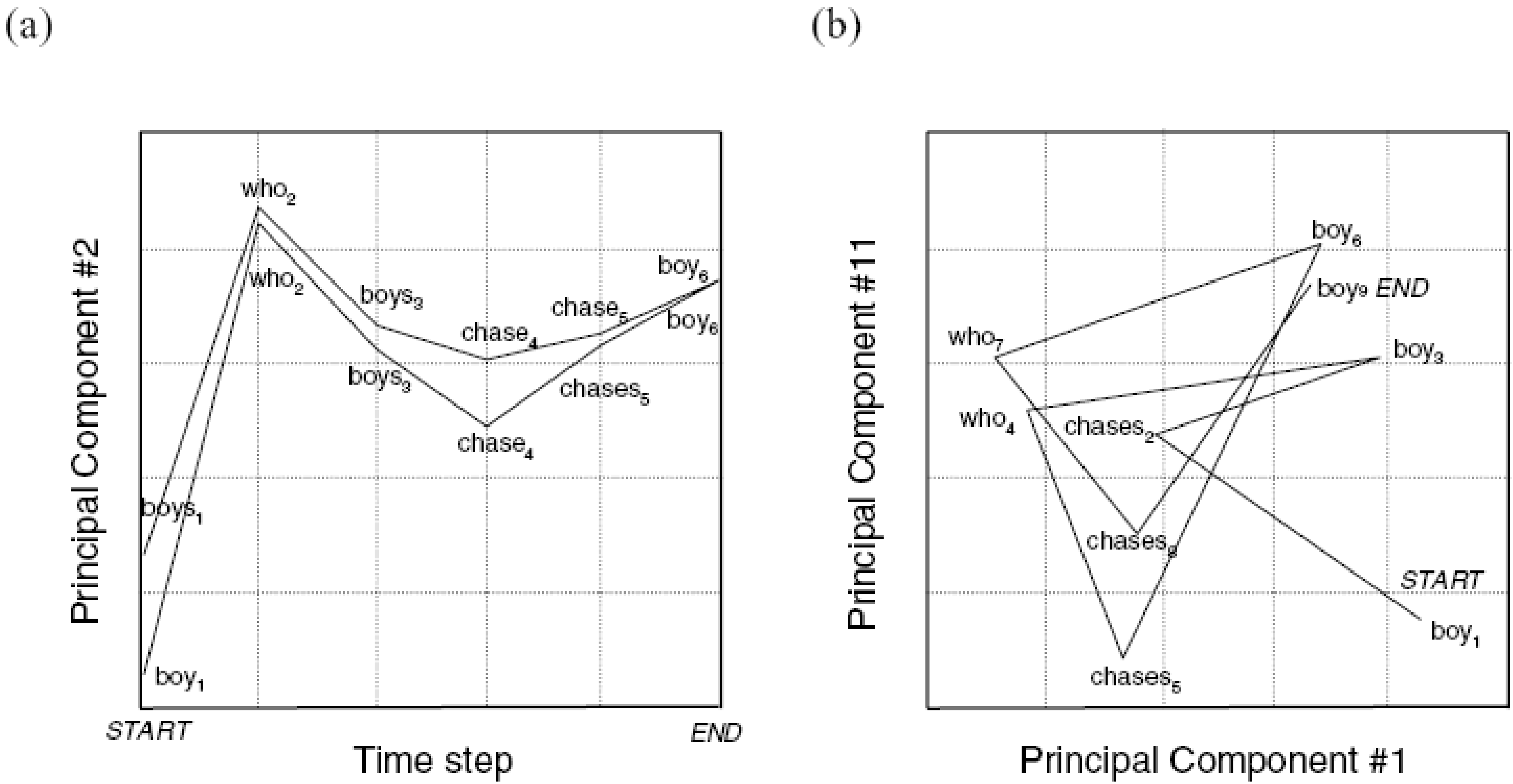


Figure 2.5. Trajectory of internal activation states as the simple recurrent network (SRN) processes sentences (Elman, 1993).

Related Models · · ·

- *Cascade-Correlation and Incremental Neural Network Algorithms*
- *Mixture-of-Experts-Models*
- *Hybrid Models*
- *Bayesian Graphical Models*

Connectionist Influences on Cognitive Theory

Knowledge vs Processing

Knowledge is hard to move about in connectionist networks because it is encoded in the weights.

- Knowledge encode: *active – recent – how things are now*
  latent – accumulated – how things will be

- If information does need to be moved around the system, this will require special structures and special processes. (see chapter 7 & 8)

- Information will be processed in the same substrate where it is stored. – control system has no content – placeholders
  efficiency

- Lastly, the connectionist perspective on memory alters how we conceive of *domain generality* in processing systems.

Cognitive Development

- Cascade-Correlation and Incremental Neural Network Algorithms

The Study of Acquired Disorders in Cognitive Neuropsychology

- the cognitive system comprises a set of independently functioning components. Patterns of selective cognitive impairment after acquired brain damage could then be used to construct models of normal cognitive function.

The Origins of Individual Variability & Developmental Disorders

- In addition to their role in studying acquired disorders, the fact that many connectionist models learn their cognitive abilities makes them an ideal framework within which to study *developmental disorders*, such as autism, dyslexia, and specific language impairment

Future Directions

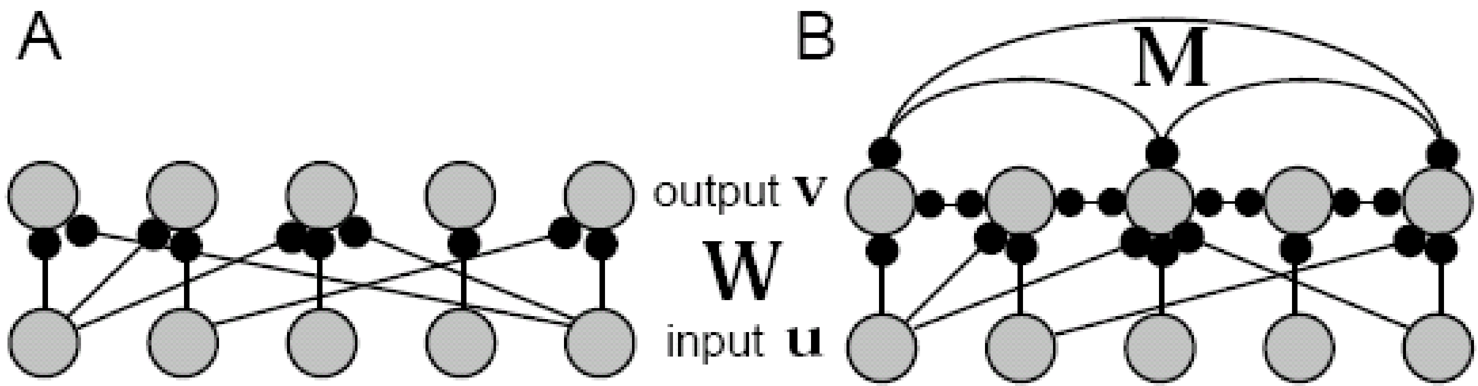
- Bespoke networks –> models fit together in the larger cognitive system
- Complexity vs Simplification & Understanding
- Connectionism will be affected by the increasing appeal to *Bayesian probability theory* in human reasoning.
- Connectionism will continue to have a close relation to *neuroscience*
- *Behavioral genetics -* build links between behavior (where variability is measured) and the substrate on which genetic effects act.

Conclusions

讲完了～～

谢谢大家！

Feedforward and recurrent networks



– **Feedforward connections** bring input to a given region from another region located at an earlier stage along a particular processing pathway

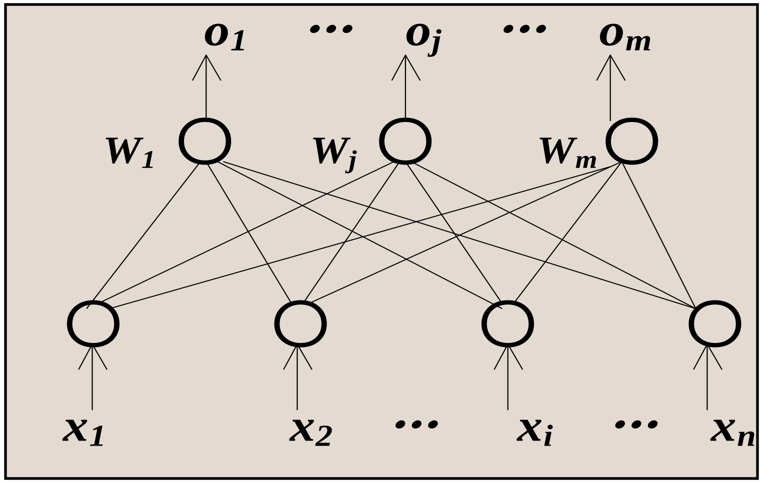
– **Recurrent synapses** interconnect neurons within a particular region that are considered to be at the same stage along the processing pathway

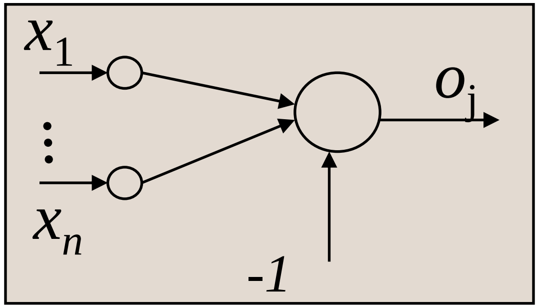
注：引自生物物理所武志华老师课件：《神经信息学——Network Models》

单层感知器模型

$$\mathbf{X}=(x_1,x_2,\dots x_i,\dots,x_n)^T$$
$$\mathbf{O}=(o_1,o_2,\dots o_i,\dots,o_m)^T$$
$$\mathbf{W}_j=(w_{1j},w_{2j},\dots w_{ij},\dots,w_{nj})^T$$

$j=1,2,\dots,m$





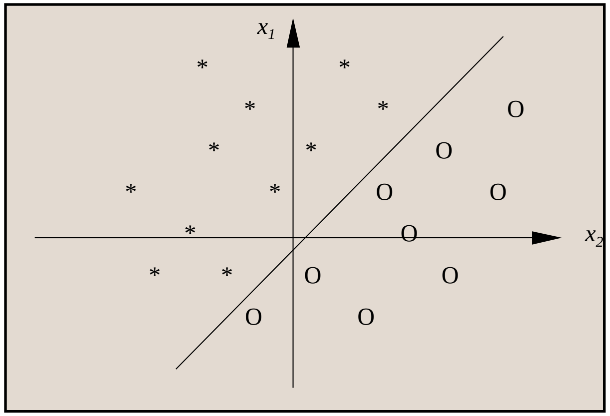
单计算节点感知器

净输入： $net_j=\sum_{i=1}^{n}w_{ij}x_i \qquad j=1,2,\dots,m$

输出： $o_j=sgn(net_j-T_j)=sgn(\sum_{i=0}^{n}w_{ij}x_i)=sgn(\mathbf{W}_j^T\mathbf{X})$

感知器的功能

（1）设输入向量$X=(x_1,x_2)^T$

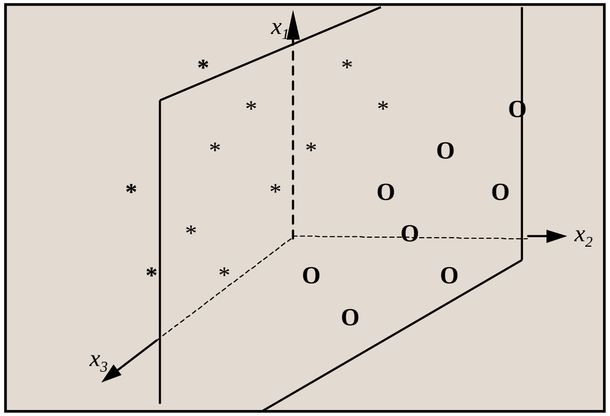


输出：$o_j=\begin{cases}1 & w_{1j}x_1+w_{2j}x_2-T_j>0\\ -1 & w_{1j}x_1+w_{2j}x_2-T_j<0\end{cases}$

则由方程 $w_{1j}x_1+w_{2j}x_2-T_j=0$

确定了二维平面上的一条分界线。

（2）设输入向量$X=(x_1,x_2,x_3)^T$



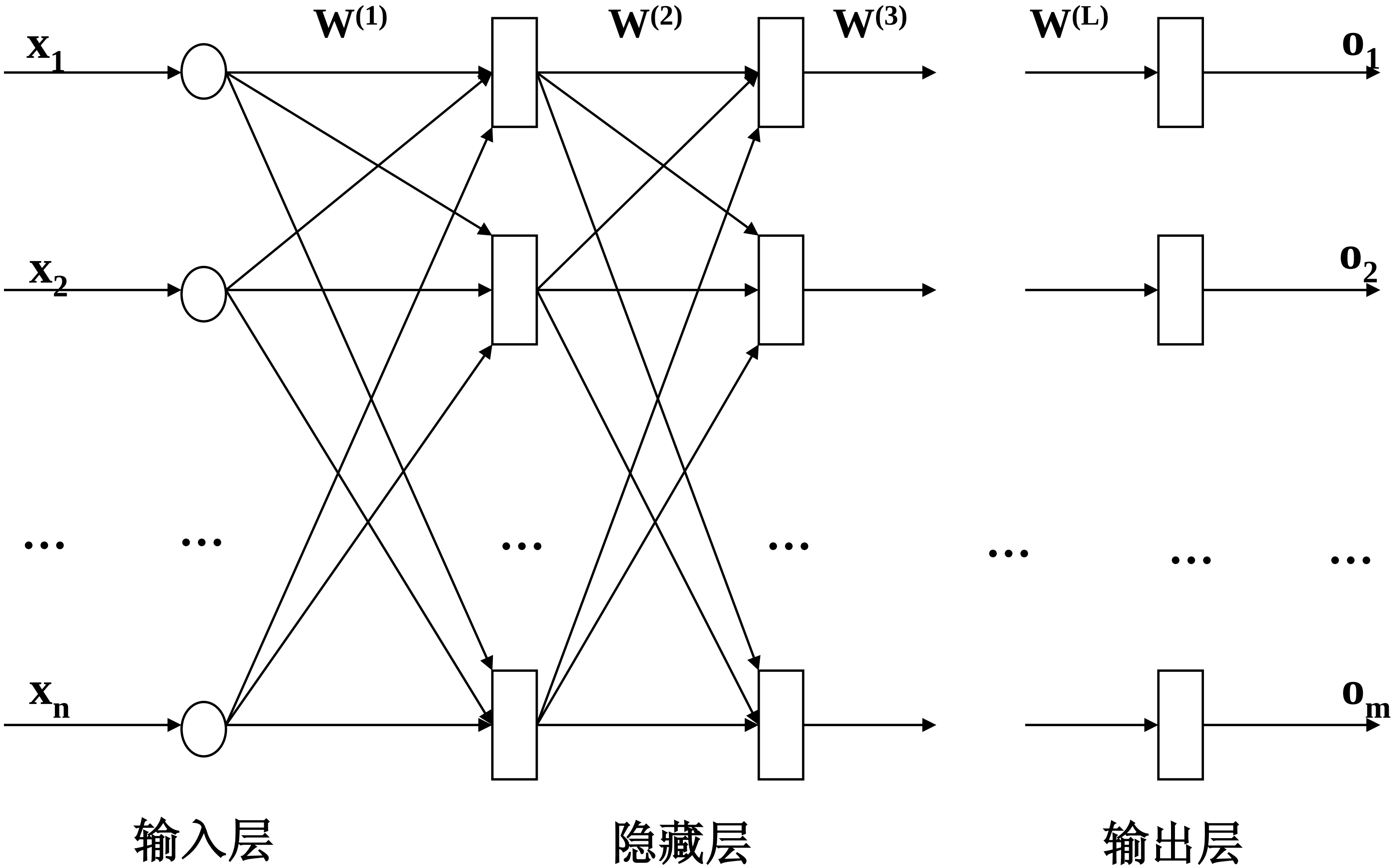
输出：$o_j=\begin{cases}1 & w_{1j}x_1+w_{2j}x_2+w_{3j}x_3-T_j>0\\ -1 & w_{1j}x_1+w_{2j}x_2+w_{3j}x_3-T_j<0\end{cases}$

则由方程 $w_{1j}x_1+w_{2j}x_2+w_{3j}x_3-T_j=0$

确定了三维空间上的一个分界平面。

多层感知器

网络的拓扑结构

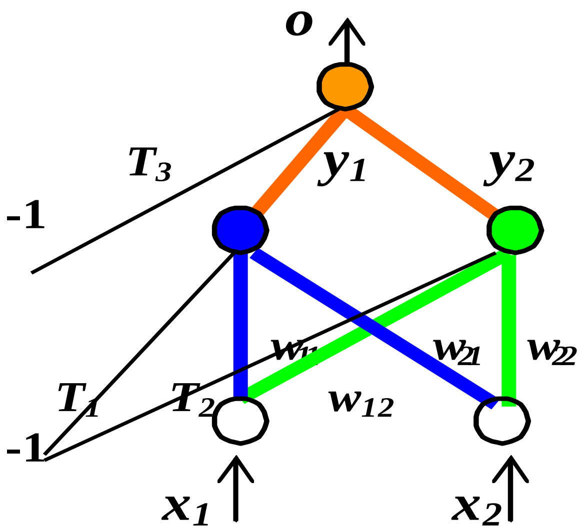


用两计算层感知器解决“异或”问题。

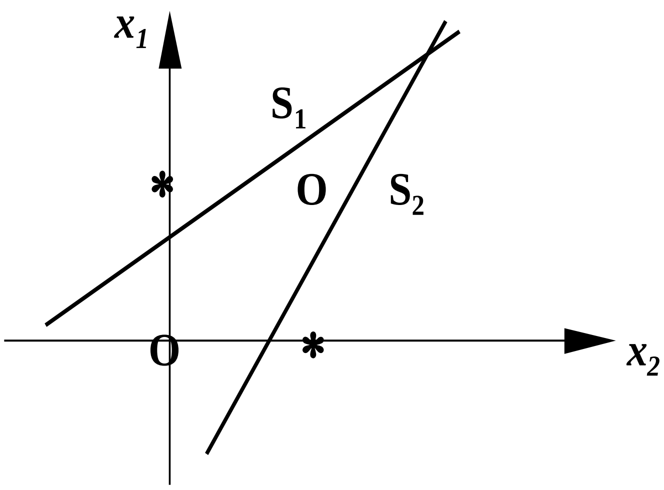
“异或”的真值表

x1	x2	y1	y2	o
0	0	1		
0	1	1		
1	0	0		
1	1	1		

双层感知器



“异或”问题分类



误差反向传播（BP）网路

输入向量： $X=(x_1,x_2,\ldots,x_i,\ldots,x_n)^T$

隐层输出向量： $\boldsymbol{Y}=(y_1,y_2,\ldots,y_j,\ldots,y_m)^T$

输出层输出向量： $\boldsymbol{O}=(o_1,o_2,\ldots,o_k,\ldots,o_l)^T$

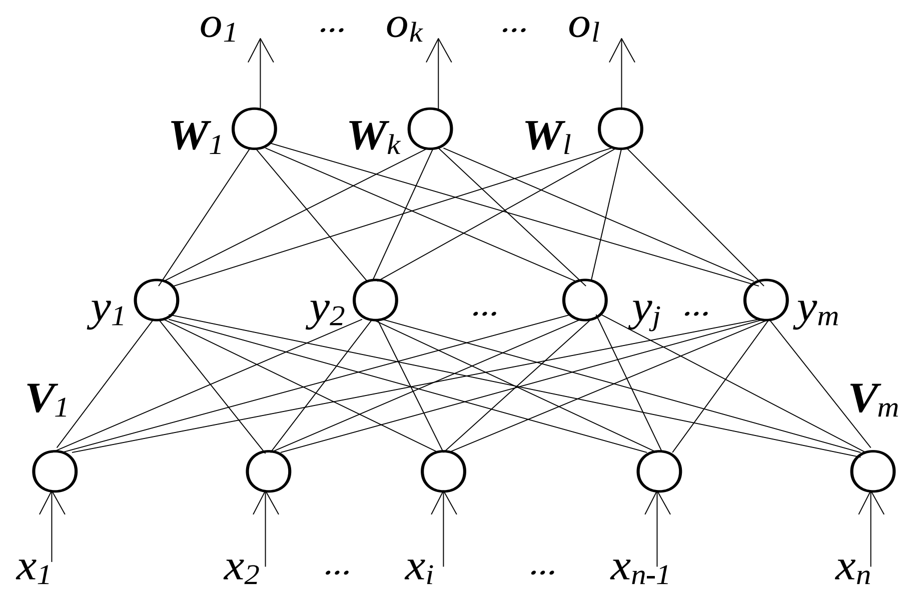
期望输出向量：$\boldsymbol{d}=(d_1, d_2,\ldots,d_k,\ldots,d_l)^T$

输入层到隐层之间的权值矩阵：

$\boldsymbol{V}=(\boldsymbol{V}_1,\boldsymbol{V}_2,\ldots,\boldsymbol{V}_j,\ldots,\boldsymbol{V}_m)$

隐层到输出层之间的权值矩阵：

$\boldsymbol{W}=(\boldsymbol{W}_1,\boldsymbol{W}_2,\ldots,\boldsymbol{W}_k,\ldots,\boldsymbol{W}_l)$



基于BP算法的多层前馈网络模型

对于输出层：$o_k = f(net_k)$ **k=1,2,…,l**

$$net_k = \sum_{j=0}^{m} w_{jk} y_j \quad \textbf{k=1,2,…,l}$$

对于隐层：$y_j = f(net_j)$ **j=1,2,…,m**

$$net_j = \sum_{i=0}^{n} v_{ij} x_i \quad \textbf{j=1,2,…,m}$$

网络误差 定义与权值调整思路

输出误差E定义： $E=\frac{1}{2}(\mathbf{d}-\mathbf{O})^2=\frac{1}{2}\sum_{k=1}^{l}(d_k-o_k)^2$

将以上误差定义式展开至隐层：

$$E=\frac{1}{2}\sum_{k=1}^{l}[d_k-f(net_k)]^2=\frac{1}{2}\sum_{k=1}^{l}[d_k-f(\sum_{j=0}^{m}w_{jk}y_j)]^2$$

进一步展开至输入层：

$$E=\frac{1}{2}\sum_{k=1}^{l}\{d_k-f[\sum_{j=0}^{m}w_{jk}f(net_j)]\}^2$$

$$=\frac{1}{2}\sum_{k=1}^{l}\{d_k-f[\sum_{j=0}^{m}\underline{w_{jk}}f(\sum_{i=0}^{n}\underline{v_{ij}}x_i)]\}^2$$

注：引自中科院计算所史忠植老师课件：《神经信息学——平行分布式理论框架》

Stroop task

红 绿 蓝

红 绿 蓝

红 绿 蓝