

## 内隐学习的人工神经网络模型\*

郭秀艳 朱 磊 魏知超

(华东师范大学心理学系, 上海 200062)

**摘 要** 近年来, 人工神经网络模型常被用来模拟各种心理活动, 从而为心理学的一些相关理论提供丰富的证据, 内隐学习也不例外。基于权重调整来学习正确反应的人工神经网络模型和内隐学习的两大本质特征间有着极为相应的匹配, 因此, 人工神经网络模型特别适用于内隐学习研究。到目前为止, 针对两种较为普遍的内隐学习任务, 已经相应地出现了两种使用较为广泛的神经网络模型——自动联系者和简单循环神经网络。

**关键词** 人工神经网络模型, 人工语法学习, 自动联系者, 序列学习, 简单循环神经网络。

**分类号** B842

### 1 引言

人工神经网络模型 (Artificial Neural Network Model, 简称ANN), 顾名思义, 就是用人造的程序、机械或设备来模拟人脑神经网络的模型。人工神经网络模型的用途有二: (1) 发明基于神经网络的人工智能系统, 来模拟人的学习、记忆、推理等智能活动, 以服务于人类的现实生活; (2) 构建各种心理活动和心理过程的模型, 以为各种心理学理论提供支持。前者是自动化、通信、制造、经济领域关注的, 我们平时所见的语音识别、经济领域使用的股票走势预测等智能系统大多是基于人工神经网络模型研制出来的。而后者则是心理学家所关注的领域。至今, 人工神经网络模型已被用来模拟诸如知觉、记忆、学习、判断等各种心理活动, 以解释矛盾的实验数据, 为有关的心理学理论提供丰富的证据。

和其他领域的研究者们一样, 内隐学习领域的研究者们也注意到了这一行之有效的工具。Cleeremans (1993) 指出根据已有的内隐学习理论构造人工神经网络模型, 将模型的输出数据与人类被试的实验数据进行比较, 能为原有的理论观点提供证据<sup>[1]</sup>。Dienes和Perner (1996) 也有类似的想法

<sup>[2]</sup>。然而, 在发挥人工神经网络模型在内隐学习研究上的功效之前, 必须解决如下问题: 人工神经网络模型是否正如Cleeremans等所言适用于内隐学习研究? 如果是, 用哪类人工神经网络模型来模拟内隐学习?

### 2 人工神经网络模型的工作原理及其研究内隐学习的适用性

人工神经网络模型之所以适用于内隐学习, 是因为它的基本工作原理和内隐学习的两个本质特征有着惊人的相似。

#### 2.1 人工神经网络模型的工作原理

早在 20 世纪 40 年代, 便有研究者对人工神经网络模型的工作原理做了最初的尝试。1943 年, McCulloch和Pitts用类似“开关”的阈限逻辑单元 (Threshold Logic Unit) \*\*来模拟神经元, 并将多个这样的单元以相等的强度 (权重) 连接起来, 形成网络, 这就是著名的MP模型<sup>[3]</sup>。1949 年, Hebb在论述条件反射的形成时, 无意间提到了神经元间连接强度更新的重要法则, 即两个彼此相连的神经元同时激活或同时抑制, 都能增加神经元间的连接强度, 后人称此为Hebb法则<sup>[4]</sup>。然而, MP模型和Hebb法则都不能构成真正意义\*\*上的神经网络模型, 虽然, MP模型已经具备将多个神经元连接起

收稿日期: 2006-09-20

\* 本研究得到教育部新世纪优秀人才支持计划项目 (批准号为 41193002) 和教育部高等学校全国优秀博士学位论文作者专项资金资助项目 (批准号为 200309) 支持。

通讯作者: 郭秀艳, E-mail: xyguo@psy.ecnu.edu.cn

\*\* 阈限逻辑单元: 其基本功能为综合来自其他神经元 (阈限逻辑单元) 的激活, 并简单求和, 算出总激活量, 判断总激活量是否达到某一阈限, 如果是, 则激活本单元, 并将激活传递给与之相连的其他阈限逻辑单元 (神经元)。

来，形成网络的雏形，但是由于不同单元间的连接强度相等，且恒定不可变化，MP模型不具备人工神经网络的基本特征——学习性；而Hebb虽然提出了权重变化的一条有效法则，但却未将其应用到人工神经网络中来。真正将神经元连接成网络的思想与借助于权重更新使网络具有学习性的思想综合在一起的要属Rosenblatt。

Rosenblatt (1958) 提出了第一个真正意义上的人工神经网络模型——感知器 (Perceptron) [5]。其基本工作原理为：整个感知器由多个不同层次的加工单元组成，每个加工单元的功能类似于单个神经元或一组神经元，它能接收来自前一层的几个加工单元的激活，并综合这些激活，对此进行简单运算（例如：判断总激活量是否达到某一阈值），然后将运算的结果传递给下一层的加工单元。第一层单元的激活模式反映了外部刺激状态，即模型的输

入，而最后一层单元的激活模式则为模型的输出反应。某个单元对下一层的另一个单元的影响取决于两个单元之间连接的强度（权重）。为了在给定输入的情况下，使模型获得类似于人类被试的输出反应，必须不停地调节单元与单元间的连接权重。所以，在构建合适的人工神经网络模型时，研究者往往会先设置一系列初始权重，然后不断地给予模型不同的输入模式，在每个特定输入后，比较模型输出与正确输出间的差异，并据此调整单元间的连接权重，这一过程不断进行，直到模型输出和正确输出间的差异达到最小值，此时，模型便完成了整个学习过程。图1为一个典型感知器的例子，它旨在判断呈现于视网膜的光条是垂直的还是水平的。整个感知器由3个加工单元层组成，第一层为网膜层，即将整个视网膜分割为10×10的网格，用每个网膜单元对应于一个网格，共100个单元，当光条落

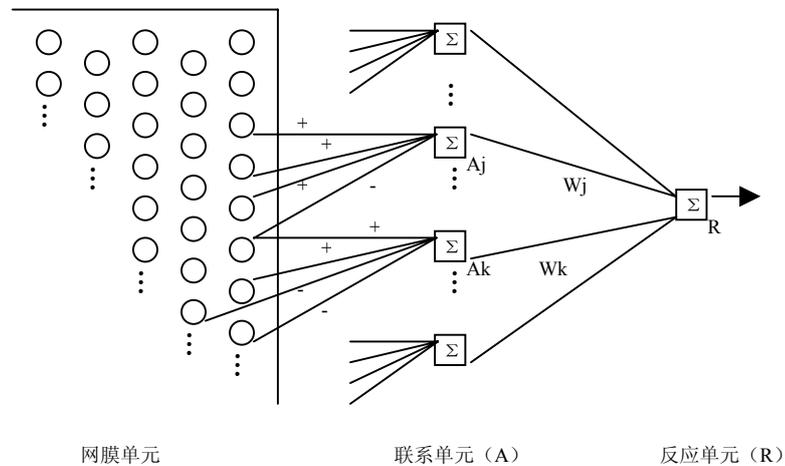


图1 感知器例子  
(资料来源：文献[3])

在视网膜的某几个网格上时，这些网格所对应的网膜单元被激活。第二层为联系层，其中的每个单元总是和某些网膜单元间存在兴奋或抑制连接，不论兴奋还是抑制连接，强度都是恒定的1或-1，当与联系单元连接的网膜单元的总激活量达到联系单元的激活阈值时，联系单元被激活，例如图1， $A_j$ 的接收到的总激活量为 $1+1+1-1=2$ ，如果 $A_j$ 的阈限为2，那么 $2=2$ ， $A_j$ 被激活。第三层为反应层，其中只包括一个反应单元，它与所有的联系单元连

接，连接权重为 $W_j$ ，其中， $j$ 表示第 $j$ 个联系单元。反应单元将综合来自联系单元的激活信息，即将每个联系单元的激活量乘以它们之间的连接权重，然后简单求和，得出总激活量，并判断激活是否达到阈值，公式表示如下：

$$a_R = \text{sgn}(\sum_{j=1}^n a_j w_j - \theta)$$

$$\text{其中, } \text{sgn}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases},$$

$a_R$ 为反应单元的激活水平,  $a_i$ 为联系单元的激活水平,  $\theta$ 为阈值。如果, 总激活量达到反应单元的阈限, 反应单元被激活(激活量为1), 感知器决定光条为垂直, 否则, 反应单元不被激活(激活量为0), 感知器反应光条为水平。当然, 感知器必须经过一个漫长的学习阶段, 才能完成这一简单的判断任务。在学习阶段, 感知器接受各种不同的水平和垂直光条刺激, 并一一做出反应, 当反应正确时, 连接权重不做任何调整, 一旦反应错误, 感知器会自动调整联系单元与反应单元间的权重, 比如: 当反应单元的激活量为0时, 而实际光条为垂直, 说明反应单元所接收到的总激活量小于阈限, 此时, 则应增大那些被激活的联系单元与反应单元间的连接权重, 以提高总激活量, 使其更有可能达到阈限, 致使在下次刺激呈现时, 感知器更易做出正确反应。当然, Rosenblatt的感知器除了能调整权重外, 还会调整反应单元阈限。这种通过逐步调整连接权重和阈限, 以减少感知器反应和正确反应间的差距的方法就是著名的感知器收敛法则(perceptron convergence rule)。不过, 调整阈限的方法对于拥有多个反应单元的模型来说过于复杂, 所以未被以后的人工神经网络模型采纳。

可见, Rosenblatt的贡献是卓越的, 他给出了人工神经网络的基本工作原理, 基于感知器收敛法则发展而来的delta法则与斜率递减(gradient descent)法已成为如今人工神经网络最主要的算法, 本文第3点中将对对此做详细介绍。然而, 值得注意的一点是感知器仅在联系层和反应层间使用了权重概念, 学习过程也仅发生在这两层之间, 所以从本质上讲, 感知器只属于包含一个输入层和一个输出层的单层网络(single layer network), 这种单层网络在解决某些实际问题时, 遇到了障碍。Minsky等(1969)指出感知器甚至无法模拟诸如XOR(异或)等简单运算<sup>[3]</sup>。因此, 在接下来的将近20年中, 人工神经网络的发展一度进入低迷期。直到80年代中期, 逆向传导法(back propagation)<sup>[6]</sup>、自动联系者(atuoassociator)<sup>[7]</sup>、循环模型(recurrent model)<sup>[8,9]</sup>——提出, 多层网络广泛应用智能模拟任务中, 人工神经网络才得以迅速发展。然而, 这些算法和模型的基本工作原理与最初的感知器却并无两样。

## 2.2 人工神经网络模型研究内隐学习的适用性

如上述, 人工神经网络的工作原理为通过调整权重逐步学会正确反应。那么, 这些基于调整权重来学习正确反应的人工神经网络模型是否适于研究内隐学习呢? 内隐学习这一概念强调的是两个特征: 第一, 它是学习的一种形式; 第二, 这种学习是内隐的、无需意识努力的。人工神经网络的工作原理极好地匹配了这两个特征。首先, 要建立针对某一心理活动的合适的网络模型, 必须要经过长期的训练和学习过程, 即向模型呈现许多刺激, 要求其做出反应, 并将模型反应与要求的正确反应进行比较, 据此逐步调整内部结构, 使模型在以后的刺激情景中, 更易做出正确反应。这种学习过程和人类的学习有着惊人的相似, 人类在幼年时, 虽然脑神经已发展完全, 但是他们仍然不能执行各种认知任务, 这时候的人脑就好比初始状态的神经网络模型, 虽然, 已对模型基本结构做了设定(比如, 模型由几个加工单元层组成, 每层有几个单元, 哪些单元和哪些单元间存在联系), 但是由于所有的连接权重都是随机设置的, 模型不知道刺激与反应间的联系, 常常会给出错误反应, 而当经过几年的学习之后, 人类婴孩不断地从外界接收刺激, 做出反应, 并接收来自外界的反馈, 将之与自己的反应做比较, 逐步通过内部结构的改变, 来调整自己的行为, 从而能够知觉事物、理解言语、控制自身的机体运动, 这就好比初始状态的网络模型, 能够通过调整权重, 达到稳定状态, 对刺激模式做出正确反应。可见, 人工神经网络模型很好地匹配了内隐学习地第一个特征——两者都是学习过程。其次, 内隐学习强调学习是内隐地进行的。由于无法直接探索学习过程是否是内隐的, 研究者往往通过验证学习中所获得的知识是内隐的来间接推断学习过程是内隐的, 当被试在分类或选择任务中的表现高于随机, 却不知道自己是依据何种知识做出正确判断时, 习得的知识被认为是内隐的, 从而进一步推断学习过程也是内隐的<sup>[10-12]</sup>。人工神经网络模型也一样, 从根据输出结构来调整权重的过程来看, 根本无法判断学习过程是否是内隐的, 而类似的, 人工神经网络的知识表征形式显示神经网络模型所获得的知识很有可能是内隐的。当人工神经网络能做到正确反应时, 指导模型该如何反应的知识似乎是存在于单元的组织结构和他们之间的连接权重上, 根本无法直接从这种无序而杂乱的结构和权重

上看出什么规则，例如：上述判断光条垂直和水平的感知器，我们并无法从直观上判断哪些单元表征垂直光条，哪些单元表征水平光条，这种知识表征形式很有可能是内隐的。这种知识的表征形式被称为分布式表征（distributed representation），它和传统的局部表征（local representation）很不一样，局部表征中，每个用不同的单元表征不同的事物，例如：单元 1 表征事物 1，单元 2 表征事物 2……，在知识提取时，就好比查字典一样，根据不同的事物，找到表征的位置，便可以获悉有关事物的各种知识。而分布表征可以将不同的事物表征在同一组单元中，比如垂直光条和水平光条，在知识提取时，往往只需输入新刺激的内容，比如光条激活了视网膜的哪些区域，便能根据内容找到合适的反应。

### 3 广泛应用于内隐学习领域的两种人工神经网络模型

人工神经网络模型很好地匹配内隐学习的两大基本特征。因此，许多研究者针对不同的内隐学习任务，选择不同的人工神经网络模型来拟合内隐学习过程，以探讨内隐学习的本质和机制<sup>[13-17]</sup>。

纵观近 40 年来的内隐学习研究，虽然不同的研究者开发了许多不同的任务形式，获得了许多可喜的研究成果，但是这些任务都离不开两种根本的任务形式——人工语法学习和序列学习。相应地，根据两个任务的不同特点，研究者选择了两种不同的模型——自动联系者和简单循环网络（simple recurrent network, 简称 SRN）——分别对之加以模拟。

#### 3.1 人工语法学习与自动联系者

人工语法学习任务通常是：在学习阶段，要求被试在不知道内在规则的情况下记忆一大堆合法字符串，然后，在测试阶段，向被试呈现一批新的合法和非法字符串，并告知前面学习的字符串是一类符合语法规则的合法字符串，要求被试根据前面的学习，判断新的字符串是否合法<sup>[10]</sup>。从测试阶段的要求、学习阶段的要求和所学习的知识这三个不同的角度来看，人工语法学习具有三个主要特征：

第一，测试阶段要求被试进行的是一项典型的分类任务。相应地，检索 1943 年以来，人工神经网络模型发展的这段历史，便可以发现只有模式联系者（pattern associator）是专门针对模式识别和分类而开发出来的模型。模式联系者的诞生源自于条

件反射的思想，即建立刺激和反应间的联系。用模式联系者的话说，就是对于特定的输入刺激模式，给出相应的反应模式。有趣的是，当用一组单元表示输入模式，另一组单元表示输出模式，它们两两连接在一起（如图 2），并使用 Hebb 法则  $\Delta w_{ij} = \epsilon a_i a_j$  更新权重时（其中， $\Delta w_{ij}$  为输入单元  $j$  和输出单元  $i$  之间权重的变化量， $\epsilon$  为学习速率，是常数， $a_i$  为输出单元的激活量， $a_j$  为输入单元的激活量），模式联系者便能在同一个模型中存储不同的输入模式，将反应模式相同的输入模式归于同一类，并且当模式中的某些量缺失时，比如某个输入单元损坏，模型仍能正常运行<sup>[3]</sup>。可见，模式联系者能够有效地模拟分类任务，甚至是某些脑神经受损病人（相当于某些输入或输出单元受损）的分类活动。然而，人工语法学习并不仅是单纯的分类任务，它还具有其他特征，因此，只模拟测试阶段的分类任务是远远不够的。

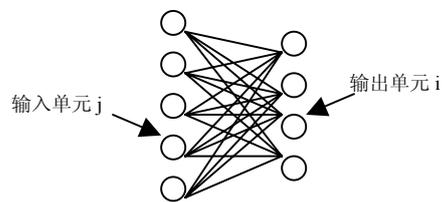


图 2 模式联系者示例

第二，人工语法学习的另一个特征是，学习阶段要求被试在没有任何指导（反馈）的情况下学习字符串，即并没有在被试学习某一字符串的同时，告诉他们这一字符串属于哪一类，合法还是非法。投射到模式联系者中，就是在学习对特定刺激模式的分类时，不伴随相应的反应模式，即没有输出单元。事实上，人类的许多学习过程都是在没有外部指导的情况下进行的。比方说，当孩子看到麻雀和燕子（刺激模式）时，家长告诉他这是鸟（反应模式），但很多时候，家长并不在身边，比如，当孩子看到鸽子时，没有外在反应模式指导，孩子仍能反应为鸟，人类似乎能够在没有外在的反应模式引导教育的情况下，学会分类。人工语法学习也是一样，虽然在学习时，没有被告知哪些合法，哪些不合法，但是被试仍能学会分类规则。这些事实提示研究者：必须开发一种没有外在引导者的自适应模型。于是，McClelland 和 Rumelhart（1985）在对

模式联系者稍做调整的基础上,提出了模式联系者的一个自适应特例——自动联系者<sup>[7]</sup>。由于没有外部引导,任何自适应模型必须具备自我反馈的功能,自动联系者也一样,如图3所示,每个单元除

了接收外部输入(e)以外,还必须接收来自其他单元的内部输入(i),这种存在内部反馈的模型被称为循环模型。为了达到分类的目的,自动联系者旨在再现单元的外部输入模式,就好像人工语法学

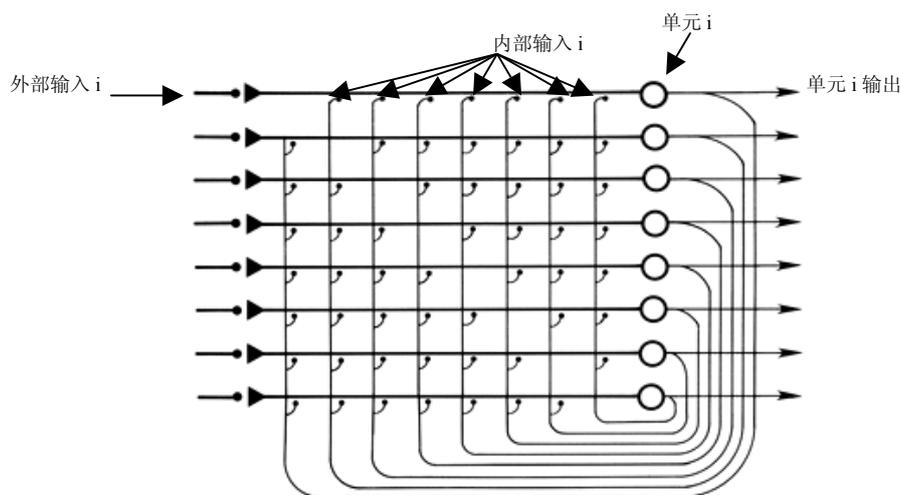


图3 含有8个单元的自动联系者

(资料来源:文献[7])

习中,记忆字符串一样。而为了达到这一点,必须调节连接权重,使单元的内部输入能匹配外部输入,用公式表示为:

$$\Delta w_{ij} = \varepsilon(e_i - i_i)a_j$$

其中,  $\Delta w_{ij}$  为单元j和i之间权重的变化量,  $\varepsilon$  为学习率,  $e_i$  为单元i的外部输入,  $i_i$  为单元i的内部输入,

$$i_i = \sum_j a_j w_{ij},$$

$a_j$  为单元j的激活量。虽然,自动联系者模型没有可比对的反应模式,但是,上述算法的核心仍是用内部状态去匹配外部状态,尽量减少两者间的差异,因此,上述公式仍可看成delta法则\*的变式。当然,除了使用delta法则来训练自动联系者外,也有研究者(例如,文献[14])使用类似模式联系者的Hebb法则训练该模型。和模式联系者一样,自动联系者也能习得不同的外部输入模式,并对此做出分类。可见,自动联系者能够很好地模拟人工语法学习任

务学习和测试阶段的表面属性——无外部引导的分类学习任务。然而,它是否能模拟语法学习这一内部属性呢?

第三,人工语法学习的本质特征在于所获得的是有关语法的知识。虽然,如前所述,自动联系者获得的是一种无法外显的分布性表征的知识,但是,有意思的是这种表征似乎代表了某一类别的原型。而原型或许就好比语法的抽象表示方法。McClelland 和Rumelhart(1985)曾构建了一个由24个单元组成的自动联系者,将一些有关狗的信息输入这24个单元<sup>[7]</sup>,其中前8个单元输入的是狗的名字信息,后16个单元输入的是狗的外貌信息,结果发现经过训练后,模型的权重矩阵呈现出特殊的构造,即前8个单元之间及前3个单元与后16个单元间的权重处于随机水平,而后16个单元间的权重却出现某种固定的模式,McClelland和Rumelhart认为这种固定的模式就是狗视觉表象的原型。虽然,无法从这种固定的权重模式中,看出典型的狗应该具有哪些特征,但是至少可以认为,如果两个单元间的连接权重比较大,那么当这两个

\* delta法则:是感知器收敛法则的便是,旨在减少模型输出和目标输出间的差异。

单元同时被激活时，模型有可能将目标判断为狗。也就是说，原型所表征的是每个视觉特征间的联系。这就好比，人工语法学习中的语法知识，语法或许是由字母间的联系所表征的。所以，从这点意义上讲，自动联系者习得的原型或许就是语法知识。

### 3.2 序列学习和简单循环网络

序列学习任务要求被试对一系列规则序列进行选择反应，其假设为：如果被试习得了序列间存在的固定规则，则他们可以依据前面呈现的序列来预测下一个项目是什么，成功的预测将会缩短被试的反应时<sup>[11]</sup>。序列学习和人工语法学习之间的本质区别在于：它是一项预测任务，而非分类任务，被试对某一项目的反应依赖于前面的项目。

1990 年，Elman 开发了简单循环网络，专门用来模拟这类预测任务<sup>[9]</sup>。简单循环网络的目的是根据当前输入的项目来预测下一个项目，它的基本结构如图 4 所示，由 4 个加工单元层组成，它们分别是输入单元层、隐含单元层、上下文单元层和输出单元层。输入单元层用于表征当前输入的项目，输出单元层用于表征模型所预测的项目，上下文单元层用于表征在当前项目之前出现的项目序列，而隐含单元层负责在输入、输出和上下文单元层间的信息传递。简单循环网络的工作原理为：在接收第一

收到来自上下文层的激活，因此，它向输出层传递的激活则包含了第一和第二两个项目的信息，依此类推，输出层所做出的预测是基于包含当前项目之内的所有项目的信息。当然，和其他许多人工神经网络模型一样，简单循环模型依据 delta 法则来调整权重的，即将模型给出的预测与真实出现的后续项目进行比较，来调整权重。不过，简单循环网络所使用的 delta 法则和自动联系者有两点差异：（1）权重调整是通过斜率递减法进行的，即寻求预测反应和真实项目之间误差方差的最小值，用公式表示为：

$$E_p = (t_{out} - a_{out})^2$$

$$\Delta w = -\varepsilon \frac{dE}{dw}$$

其中， $t_{out}$ 为真实的正确输出， $a_{out}$ 为模型的预测输出；（2）不同于自动联系者，简单循环网络是多层模型，当调整输出层单元和隐含层单元间的权重时，计算预测和真实项目间的误差方差是可以直接计算获得的，然而当要调整隐含层单元与输入层单元或上下文层单元间的权重时，由于不存在真实的正确反应，误差方差则无从计算，所以Rumelhart和McClelland（1986）提出可以用输出层单元的误差来估计隐含层单元的误差，即某一隐含单元的误差为所有与之连接的输出层单元的误差与它们之间权重乘积求和的函数<sup>[6]</sup>

$$\delta_{hidden} = F(\sum_{out} w \delta_{out}),$$

然后，同样用斜率递减法求出权重的变化量，这种误差计算的法则和简单循环网络中原有的信息流方向正好相反，所以又被称为逆向推导。

简单循环网络能有效地模拟序列学习中的预测机制，因此，它刚提出不久，就被用于模拟内隐序列学习，用来研究序列学习的抽象性等问题<sup>[13,15,16]</sup>。

总之，针对不同的内隐学习任务，为了得到更佳的模拟效果，研究者往往倾向于选择不同人工神经网络模型加以模拟。然而，就像是任务之间的划分并不绝对一样（比如：Cleeremans等就曾将人工语法和序列学习结合在同一个任务中），神经网络模型和内隐学习任务间的匹配也并不绝对，比如：Boucher等（2003）就层用简单循环网络来模拟序列

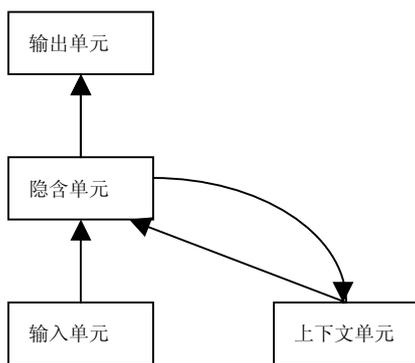


图 4 简单循环网络的基本结构

（资料来源：文献[16]）

个项目时，输入层中的某些单元被激活，并将激活传递给隐含层中的单元，隐含层除了将激活进一步传递给输出层，由输出层预测第二个项目外，还将自己的激活水平复制于上下文层，当接收第二个项目时，隐含层除了收到来自输入层的激活外，还将

学习<sup>[17]</sup>。

#### 4 小结

基于权重调整来学习正确反应的人工神经网络模型和内隐学习的两大本质特征间有着极优的匹配,人工神经网络模型在内隐学习领域的适用性毋庸置疑。在这样的背景下,出于深入探讨内隐学习的目的,研究者纷纷根据不同的内隐学习任务,选用不同的人工神经网络模型对之加以模拟<sup>[18]</sup>。到目前为止,针对两种较为普遍的内隐学习任务,也相应地出现了两种使用较为广泛的神经网络模型——自动联系者和简单循环网络。在实际研究中,合理地使用这两个模型,必将为内隐学习的理论和人工模拟提供更有力的证据。

#### 参考文献

- [1] Cleeremans A. Mechanisms of Implicit Learning: Connectionist Models of Sequence Processing. In: Jeffrey L E ed. Neural Network Modeling and Connectionism. London: MIT Press, 1993
- [2] Dienes Z, Perner J. Implicit knowledge in people and connectionist networks. In: Underwood G. Implicit cognition. Oxford: Oxford University Press, 1996
- [3] McLeod P, Plunkett K, Rolls E T. Introduction to Connectionist Modelling of Cognitive Processes. Oxford: Oxford University Press, 1998
- [4] Milner P. A Brief History of the Hebbian Learning Rule. Canadian Psychology, 2003, 44(1): 5-9
- [5] Rosenblatt F. The perceptron: A probabilistic model for information storage and organization in the brain. Psychological Reviews, 1958, 65: 386-408
- [6] Rumelhart D, McClelland J. Parallel distributed processing (vol. 1). Cambridge: MIT press, 1986a
- [7] McClelland J, Rumelhart D. Distributed Memory and the Representation of General and Specific Information. Journal of Experimental Psychology: General, 1985, 114(2): 159-188
- [8] Jordan M. An introduction to linear algebra in parallel distributed processing. In: Rumelhart D, McClelland J ed. Parallel distributed processing (vol. 1). Cambridge: MIT press, 1986a. 365-422
- [9] Elman J. Finding structure in time. Cognitive Science, 1990, 14(2): 179-212
- [10] Reber A S. Implicit learning of artificial grammars. Journal of Verbal Learning and Verbal Behavior, 1967, 6(2): 317-327
- [11] Nissen M J, Bullemer P. Attentional requirement of learning: Evidence from performance measures. Cognitive Psychology, 1987, 19(1): 1-32
- [13] 郭秀艳, 邹玉梅, 李强等. 中学生颜色内隐学习特征的研究. 心理与行为研究, 2003, 1(2): 122-127
- [14] Cleeremans A, McClelland J L. Learning the Structure of Event Sequence. Journal of Experimental Psychology: General, 1991, 120(3): 235-253
- [15] Dienes Z. Connectionist and Memory-Array Models of Artificial Grammar Learning. Cognitive Science, 1992, 16(1): 41-79
- [16] Jiménez L, Méndez C, Cleeremans A. Comparing Direct and Indirect Measures of Sequence Learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 1996, 22(4): 948-969
- [17] Dienes Z, Altmann G, Gao S J. Mapping across Domains Without Feedback: A Neural Network Model of Transfer of Implicit Knowledge. Cognitive Science, 1999, 23(1): 53-82
- [18] Boucher L, Dienes Z. Two ways of learning associations. Cognitive Science, 2003, 27(6): 807-842
- [19] 郭秀艳, 朱磊. 神经网络模型对内隐学习的探索. 心理科学, 2006, 29(2): 410-415

## Artificial Neural Network Model about Implicit Learning

Guo Xiuyan, Zhu Lei, Wei Zhichao

(Department of Psychology, East China Normal University, Shanghai 200062, China)

**Abstract:** These years, Artificial Neural Networks have been used to simulate many kinds of psychological activities, including implicit learning. Neural networks has some traits which are good for simulating implicit learning, such as learning to response correctly through weight adjustment. There are two types of neural network models named autoassociators and simple recurrent networks which are used broadly in cognitive simulation.

**Key words:** artificial neural network model, artificial grammar learning, autoassociator, serial learning, simple recurrent network.