

神经网络的原理及应用

陈涌 王正志

(自动控制系)

摘要 人工神经网络是在现代神经科学研究成果的基础上提出的,反映了人脑功能的基本特征。但它只是人脑功能的某种抽象、模拟和简化。研究这一技术的目的在于,探索人脑加工、储存和搜索信息的机制,并将其原理应用于人工智能的可能性。本文介绍了一些神经网络模型及应用。

关键词 神经元,神经网络,模型,应用

分类号 T388.1 TP15

人工智能(AI)技术的引入,使得当前计算机在模拟人的某些认识活动中取得了惊人的成就。例如,利用专家的知识模拟专家问题求解的计算机程序——专家系统,作为人类智能的体现,使得某一狭窄领域的专家系统可以与人类最好的专家相匹敌。然而,尽管各种专家系统具有实在的能力,但它们解决问题的范围受到严格的限制。至今,还没有一个系统能够达到5岁儿童所具有的常识推理和感觉能力。阻碍专家系统发展的主要因素可概述为:知识获取的瓶颈问题;知识间的上、下文敏感问题和不确定性推理问题。同时,与当前的计算机在进行符号思维处理时,忽略了人类智能中一些重要的因素有关。这些因素主要有:

(1) 人脑能够存贮大量繁杂的知识,并且能够毫不费力地很快地找到相关知识。人类所固有的这种“常识”推理能力,是当前的串行计算机所望尘莫及的。虽然计算机也能存贮大量信息,但它却很难使这些信息成为一有机联系的整体。

(2) 在许多认知领域中,如语音识别,图象处理以及更高级的医疗诊断,关键都在于从大量代表中找出最佳的匹配。人类的这种识别能力是非常强的,甚至可以毫不费力地对含有干扰的不完全信息及缺少期望进行识别。而对当前的计算机而言,是非常困难的。

(3) 在多数情况下,人类处理信息采用的是非传统AI中符号断言的形式。比如,我们能很快识别一头象,却很少有人能精确地描述其表面特征。同样,我们很难对运动、形状、声音、时空进行符号描述,却能自由地生活在这些领域中。显然,人类有其自己内部的符号表示形式,不是AI中的符号断言所能代替的。

人工神经网络是在神经生理学、神经心理学基础上用来模拟人脑神经元组织的理论, 与传统专家系统的高层逻辑模型不同, 它们假设信息处理是通过大量称之为单元的简单处理元件之间相互作用进行的, 是一种低层数值模型。由于神经网络的并行性、分布性和可学习性, 为它们在人工智能的应用方面提供了良好的前景。

1982年, 美国加州理工学院的Hopfield提出了仿效人脑工作的神经网络模型^[7]。其重大意义在于, 与过去的类似模型相比, 它能够通过集成电路来实现。从此, 对人工神经网络(以下简称神经网络)的研究, 重又引起各界科技工作者的普遍关注。由于神经网络的重要特点是具有自我组织的能力, 在此基础上得以实现学习功能, 还可以通过网络来实现联想记忆。从本质上说, 它适合于进行并行信息处理, 所以, 非常适合于某些难以用文字表达的信息处理。在模式识别、语音处理、运动控制等领域, 最能发挥其长处。加之, 近年来在神经计算机方面的产品的出现, 如, Transputer, Mark III, Odyssey, Warp, AAP-2等, 使得这一理论得到了更为广泛的应用。

本文介绍一些神经网络模型及应用的例子。

1 神经元和神经网络模型

神经科学研究表明, 每个神经元由细胞体、树突(dendrite)、轴突(axon)组成。一神经元的轴突与另一神经元的树突的结合部称为突触(synapse), 它决定神经元之间连接强度及性质(兴奋或抑制)。在一定时间内, 神经元接到来自各方面的刺激总和(时空总和)达到一定程度(阈值)时, 就发生兴奋。图1为人工神经元模型, 其中 x_i 为输入, w_i 为突触联系强度, y 为输出, $f[\]$ 为非线性函数(如: 硬限函数, 阈值逻辑函数或S形函数等), θ 为阈值。图中 $y = f[\sum_{i=1}^n w_i x_i - \theta]$ 。

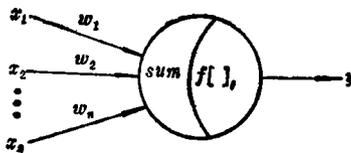


图1 人工神经元模型

人工神经网络是由大量处理元件, 如电子元件或光学元件等组成的非线性大规模自适应动力系统。它是在现代神经科学研究成果的基础上提出的, 反映了人脑功能的基本特性, 但它并不是人脑神经网络的真实描写, 而只是它的某种抽象、模拟和简化。研究这一系统的根本目的在于探索人脑加工、储存和搜索信息的机制, 进而将此原理应用于各种人工智能的可能

性。其中, 除了 $f[\]$ 以外, 神经元的连接方式(网络拓扑结构)和突触连接强度都对网络的特性和行为有很大影响。这些因素上的差别, 导致了许多功能和形态都有较大差异, 具有不同适应范围的模型。

(1) MP模型(McCuHoch-Pitts)与Hebb学习规则

神经网络的MP模型假设有 N 个神经元互连, 每个神经元的状态 s_i 只取0或1(抑制与兴奋)。

$$s_i = f[\sum_j w_{ij} s_j - \theta_i], \quad i=1, 2, \dots, N$$

Hebb规则就是在学习过程中调节 w_{ij} ($w_{ii}=0$)的原则, 若第 i 与第 j 个神经元同时

兴奋, 则其连接应加强, 即

$$\Delta w_{ij} = \alpha s_i s_j \quad (\alpha > 0)$$

这一规则与“条件反射”学说一致, 并已得到神经细胞学说证实。

至今, MP模型与Hebb规则仍在各种网络模型中起着重要作用。

(2) Hopfield模型

Hopfield网络模型有离散和连续两种。亦设网络由 N 个神经元组成, 神经元的状态随机地异步变化, 且为对称连接。通过引入能量函数 E , 且已证明 E 单调下降, 从而该网络必趋于定态吸引子。[注: 对任一初态, 当 $t \rightarrow \infty$, 系统所能达到的状态称为动力学系统的吸引子。当系统到达某一状态不再运行, 该状态称为稳定吸引子]。

两种Hopfield模型等价, 由此模型研制的神经计算机, 其计算时间就是网络趋于定态的时间。

(3) 多层网络^[1]

所谓多层网络就是在输入层与输出层之间引入隐单元, 它们与外界没有直接联系, 但其状态影响I/O关系。

(a) Rumelhart等提出的误差传播网络考虑的是前向式多层网络, 即每一层神经元的状态只影响下一层神经元的状态。能量函数 E 代表希望的输出与实际输出之差的平方和(误差), 学习算法是使 E 取极小值, 用梯度下降法修改权 w_{ij} 。若网络中某个 w_{ij} 变化时, 按前向传递法则, 它将逐级影响下一层神经元的状态, 从而可改变I/O特性。

(b) Lapedes和Farber提出的主/从网络则是对Hopfield模型的发展。Hopfield模型有两个局限性: 一是要求连接的对称性; 二是不能求解高阶问题。主/从网络模型的目标正是为了克服这两个限制。其从网络与Hopfield模型一致, 但其权 w_{ij} 由主网络决定, 且并不要求具有对称性。主网络方程就是 w 和网络输入的演变方程。当主网络适当选取参数时, 就可保证网络具有定态。

(c) 多层感知机(Perceptron)模型, 是在输入输出结点间具有一个以上结点层的前馈监督训练网络。从理论上说, 三层感知机可以创造分类器中所需的任何连续决策函数。经验表明, 利用误差回传(BP)与其它构造技术可以实现许多有用的非线性映射。

(d) CMAC模型, 即小脑模型连接控制器, 它最初是作为小脑神经生物模型研制的。CMAC模型是一种类似于多层感知机的前馈结构, 适当地形成连续值输入与输出间的非线性映射, 可以迅速地自适应。因为仅有一层连接体需要训练, 在机器人应用中, 当它执行预定的活动时, 加权就进行自适应。最初已证明CMAC模型能控制模拟机器人的手臂活动。最近借助于自适应模拟手臂的逆动力学模型已成功地改进了工业机器人手臂的动态行为, 在较慢的串行计算机中使用查表技术, 用不多的加权就能实时地实现。目前, 正在进一步研究用于机器人的这类算法及由此派生的更高级算法。

(e) 特征图模型, 使得非监督的簇化技术可以用于诸如语音与图像识别这类应用, 以执行数据压缩与减少区分的计算量。Kohonen提出了一种顺序簇化算法, 能产生一种所谓的自组织特征图。他描述了许多有关此算法的例子, 也说明此算法做矢量量化器用于语音识别器的方法。

另外, 还有一些网络模型, 如Darwin^[2]模型是一种自监督的复杂网络, 它模拟由

许多网络组成的自动机。C. Mead的视网膜片模型是用VLSI设计技术在硅片中模拟生物系统。迄今已制造出了一个人工视网膜片和一个人工耳蜗片。

虽然目前已有多种网络模型和学习算法, 但要找到功能强、易于学习训练和便于实现的神经网络模型, 仍有许多工作要做。

2 一些应用

2.1 文字识别、语音处理

1986年, 美国的T. Sejnowski和C. Rosenberg研制了NET talk程序。这个程序是一个网络模拟程序, 它能识别英文单词和朗读英文文章。该网络由安装在三层芯片上的309个神经元组成, 共18629个连接权, 用BP法修改网络的权。实验时使用1000个字的字典, Ridge32微机学习16个小时后, 程序对范文中的“发音”正确率为95%, 而对一篇新的文章, 正确率为90%。

T. Kohonen的最新研究结果表明^[2], 采用他称之为“第二类的联想记忆”, 使得神经网络系统通过学习巴赫作的曲子, 能够自己产生具有一定风格的新旋律段及旋律间的配合。用这种方法产生的音乐, 听起来是平稳的, 连续的, 而且非常悦耳, 同时还非常忠实于巴赫原来的风格。

2.2 优化计算

最优化的大多数算法或步骤的基本思路是在可能的组合空间中移动代表解, 并在目标函数递减至最小值的方向上移动, 而这里, 最优化问题被转化为神经网络问题, 此处的网络组合相应于问题的可能解, 能量函数 E 的选择应使组合表示可能的解。当 E 中相应于局部最小值的稳态组合实现时, 问题的解即从组合中得到解码。

Hopfield和Tank^[6]最早把神经网络用于求解TSP。现在, 很多学者已经把Boltzman机, 模拟退火, 高斯机, 柯西机等用于求解这一问题^[2], 得到了比Hopfield网络更为满意的结果。如: 对于10个城市的TSP, Hopfield网络求解得到最优路径 $D=2.691$ 的概率为12%; 而采用高斯机时, 找到这一最优路径的概率提高到22%。

2.3 智能控制、自适应控制^{[8][4]}

传统的控制理论由于受人的不确定因素的影响而不能用于许多实时问题。当然, 典型的应用有机器人的实时控制。但对大规模动力系统则存在如下问题: (1) 传感器中由于数据过载而引起的数据混乱问题; (2) 鲁棒性差; (3) 不能处理模糊信息; (4) 不能用于高速实时控制, 因为这需要智能计算; (5) 传感器的选择、设计问题仍未解决。此外, 有学习和自适应问题。

在神经智能控制系统中, 专家系统作为一自适应元件; 模糊分析用作决策元件; 而神经网络则用作补偿元件。这样可得到如图2所示的神经智能控制系统结构图, 它开拓了系统控制的新途。

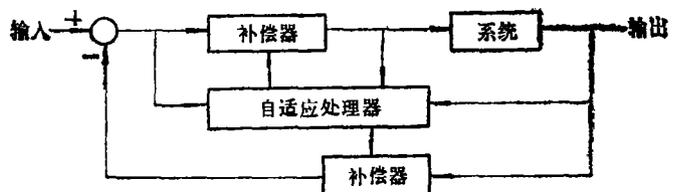


图2 神经智能控制系统

由于神经网络具有下述两个重要的计算特性：(1) 对于知识的联想记忆和恢复功能；(2) 与网络维数无关的网络动力学的一致收敛速度。把这两个特点用于自

适应控制，使得系统具有：(1) 更快的自适应速度；(2) 更简单的控制器结构；(3) 在离散域和连续域中都可自适应。与传统方法相比，把神经网络用于大系统的自适应控制，在计算速度上有一定的优势。神经自适应控制系统的结构如图3所示，其中神经估计器用于参数的辨识和从环境中分类以选择最合适的控制律。

目前，在控制上，已把神经网络用于机械手的控制，逆动力学的求解，眼手协调控制和结构不稳定系统的控制等等^{[2][8]}。

2.4 模式识别，图象处理^[2]

目前，神经网络在这方面得到了大量的应用。在图象处理中重要的一点就是图象的理解。用于图象理解的大多数方法都需要许多不同的步骤：噪声消除，边界增强，分段，特征抽取，模式识别等等。其中，分类器是至关重要的。现在，已把神经网络用于图象处理中的分类器。与传统的分类器相比，用神经网络作为分类器带来的好处是：它允许在特征空间中的更一般的决策边界。传统的分类器的决策边界，大多采用的是超平面或超二次曲面，而神经分类器采用的是分段线性的决策边界或分段二次的决策边界。由于任何平稳的决策边界都可由分段线性平面来近似，并达到任意的精度，这就使得神经分类器可达到更低的误差率。

2.5 专家系统

神经网络的并行性、分布性和可学习性为专家系统的知识表达和获取、不确定性推理提供了新的途径。它在核反应堆事故诊断专家系统中的实践^[9]表明了人工智能中的潜在前景。应用于专家系统的神经网络有两个基本要素：神经网络的知识表达体系和BP学习算法。基于此建立的Nesrad系统(一个基于神经网络核反应堆事故诊断专家系统)中，有关核反应堆的事故诊断知识不再仅仅是产生式、框架这样的逻辑形式，更主要的是知识在神经网络内部分布式隐含表达。BP学习算法的应用，使得网络的输出对一类特定的输入产生敏感性，从而达到象产生式那样的因果关系。与原来的专家系统RODES相比，Nesrad的论域知识从30个知识源变为两个神经网络，推理速度从原来的5秒或更长的时间(有时需人机交互)，缩短为3秒左右。

2.6 雷达^[2]

现在已把多层前馈神经网络用于脉冲雷达探测和脉冲压缩技术。采用的是BP法去训练网络。用这种方法，可以很容易地达到40dB峰值的信噪比。而且，与用传统的相关分析法和非匹配滤波法(mismatched filtering approaches)相比，在处理时间上更快。

3 结束语

在人工神经网络的研究中，1987年6月在美国召开了第一次神经网络国际会议，在

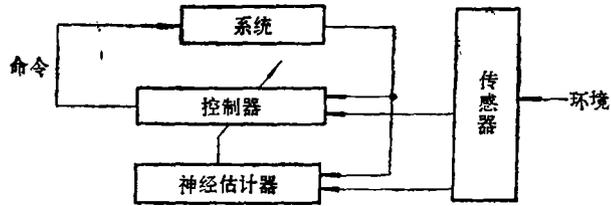


图3 神经自适应控制系统

此次会议上,有10家公司展出了神经计算机方面的软件产品。此外,由IBM、贝尔实验室等为代表,正在试制各种神经芯片,以便从硬件上实现神经计算机。日本从1988年起,许多大公司都投入巨大力量研制神经计算机,这一年在日本被称为神经计算机元年。富士通公司,日本电气公司已在软件方面作了不少工作。不久前,日本电气公司已使个人神经计算机实用化,具有超级小型机的速度,最大运算处理速度达每秒21.6万条指令,用它开发文字、语音识别系统时,开发时间可缩短到1/10。三菱电机公司在1988年11月宣布试制成功了世界上第一个光神经芯片。据估计,1988年美国仿神经系统的产品产值达2000万美元,到1990年将达1亿美元。日本通产省已拨款300亿日元,作为8年(1989年算起)的仿神经系统研究计划的费用。这些都说明,人工神经网络的研究已取得了初步令人满意的结果,前景是广阔的。

参 考 文 献

- [1] Lippmann R P. An Introduction to Computing with Neural Nets. IEEE ASSP Mag. 1987; 4~22
- [2] Washington D C. Proceedings of IJCNN, 1989
- [3] IEEE Cont. Sys. Mag., April, 1989
- [4] IEEE Cont. Sys. Mag., April, 1988
- [5] Denker J S. Neural Networks for Computing, 1986
- [6] Hopfield J J and Tank D W. Neural Computation of Decisions in Optimization Problems. Biol. Cybern. 52, 1985: 141-152
- [7] Hopfield J J. Neural Networks and Physical Systems with Emergent Collective Computational abilities. Natl. Acad. Sci. U.S.A. 1982, 79: 2554~2558
- [8] 黄玲. 连接机制: 向传统计算机结构的挑战. 计算机科学, 1989, (2): 24~28
- [9] 杨一平, 戴汝为. 从经验中获取知识: 人工神经网络专家系统及其在核反应堆事故诊断中的应用. 中科院自动化所

The Principles and Applications of Neural Networks

Chen Yong Wang Zhengzhi

(Department of Automatic Control)

Abstract

The developing of artificial neural networks is based on the research of modern neural science. Artificial neural networks reflect the basic feature of human brain's function, but it is not a real description of human neural networks, it is just some abstract, simulation and simplification of human neural networks. The fundamental goal to study this technique is to explore the mechanism of which human brain processes, memorizes and searches information, and the possibility of using the principle to artificial intelligence. This paper introduces some neural networks' models and their applications.

Key words: neuron, neural networks, model, application