

连接机制：向传统计算机结构的挑战

黄玲 (武汉大学计算机科学系)

为了克服传统von Neumann结构符号思维处理的缺陷，研究者们提出了一种新的计算机结构——连接结构。本文论述了这一结构的产生、定义及特点，并着重讨论了几种典型的连接模型以及最近研制成功的连接机。连接机制，作为计算机结构上一种新的思想，为认知处理过程提供了强有力的工具，正日益引起人们的注意。

一、引言

人工智能技术的引入，使得当前计算机在模拟人的某些认知活动中取得了惊人的成功。例如，某一狭窄领域的专家系统可与人类最好的专家相匹敌；然而，尽理各种专家系统具有实在的能力，但它们解决的问题范围受到严格的限制。至今，还没有一个系统能够达到5岁儿童新具有的常识推理和感觉能力。显然我们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，这些因素主要有：

1. 人脑能够存贮大量繁多的知识，并且能够毫不费力地很快找到相关知识，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，视觉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干扰的不完全信息及缺省期望进行识别。而对当前计算机而言，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，却很少有人能精确描述其表面特征。同样，我们很难对运动、形状、声音、时空进行形式

符号描述，却很自如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。

如果我们追问致使上述情形出现的原因，那么可以归结为两个方面：一是我们对人脑思维过程的实质尚未彻底弄清(有待于其它学科的发展)，二是当前von Neumann结构本身的缺陷。从世界上第一台电子计算机诞生之日起，计算机几乎沿用了同样的结构模式——von Neumann模式，即采用单处理机/存贮器分离的二部分结构。这种结构的产生是当时的技术和条件的必然结果，并且在数字计算上取得了巨大的成功。然而随着科学技术的发展，VLSI的出现，这种结构愈发显出其不合理性，特别是AI的引入，符号思维处理的增多，von Neumann瓶颈问题变得越来越严重(von Neumann瓶颈：计算长度受处理机、存贮器间数据驱动时间控制)，即使采用了一些复杂而灵巧的AI技术(比如，Heuristic-search, cycle-intensive等)也不能从根本上加以解决。因此，从70年代末期开始，研究者们开始探索采用大量并行结构以克服von Neumann串行结构符号处理的局限，其中最具有挑战性的思想是，建造将存贮和处理相结合的更同构的并行计算机结构，从而拉开了对连接结构研究的序幕。

二、连接机制及连接结构

2.1 什么是连接机制 Jerry Feldman首先提出连接机制(Connectionism)一词用以言及人工智能中关于大量并行结构的某一类研究,即对连接结构的研究。连接结构是由大量可并行操作的小处理/存贮单元(Units)组成,单元之间可根据需要连成不同的模式,每个单元内部存贮信息很少,仅有一些标志位或简单标号“活动-电平”做为短期工作记忆,而长期记忆则是通过改变单元间的相连模式或修改相连“权”实现。“连接机制中的单个神经元(单元)不能处理大量符号信息,而是通过大量神经元的适当连接来实现”(Feldman, 1982)

2.2 连接结构的优点 通过对连接结构的研究,人们发现该结构具有许多令人吃惊的优点,这些优点主要有:

①并行协作-速度快。连接系统中的并行处理单元本身并不能执行独立的程序,仅能做一些很简单的操作(如接受,传送信号,对数据做布尔运算等),然而,这些并行处理单元却能通过反复的相互协同过程(调整相连“权”或改变相连模式),同时处理大量的知识,达到“最终、共同的决定”。这个过程被称为“松弛”(relaxation)。连接系统这种并行、协同处理问题的方法,使它比当前的串行计算机在符号思维处理中速度快得多。

在这里,连接并行处理问题的方法与一般AI中的并行方法是不同的,其主要差别在于:前者假定,按问题需要(尽可能将问题分成足够小的数据元)系统有足够的处理元被分配(每一数据元分配一处理单元),这些处理元通过“松弛”过程,共同处理现有任务;而后者则假定:系统只有固定数目的处理机(通常2~1024台),遇到任务时,尽量将其分成许多能独立执行的子部分,同时分别在这些处理机上进行操作。前者的并行是一种“数据一级”上的并行,后者只是一种“控制一级”并行。

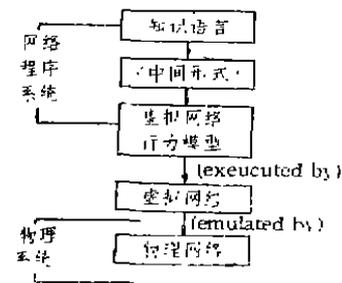
②强容错能力。在大多数连接网络中采用的是类似人脑结构的分布表示,即单个概念或特征不是由单个神经元而是大量神经元的激活模式表示,因而单个单元在整个网络行为中不起决定性作用。当

某些单元出错时,整个激活模式并不完全消失,仅仅有所退化,整个网络行为并不因此受到很大的影响。连接结构这种很强的容错能力,显示出它具有生物系统逐渐“退化”(degradation;记忆是由清晰—模糊—逐渐消失的)的特征。

③良好的“自由归结”(free-generalization)能力。由于连接系统中的知识是以相连模式形式存贮的,因此,相连“键”(strength)的微小变化使得许多记忆痕迹能同时存贮于一种相连模式中,这种记忆痕迹叠加存贮(superpositional)的方式,使得系统能以非常自然的方式归结出特殊的“经验”。

④构造灵活性。连接网络构造非常灵活,它不需要很多依赖于问题的信息(如规则等)静态地建立其并行组织结构,而是通过动态修改已有的处理机结构和通信路径来处理所面临的不同任务,显示出具有生物系统“自然选择”的特征。

2.3 一个基于连接机制的认知处理系统环境模型 连接结构所具有的上述优点,使之能够提供非常可靠的高性能的认知处理过程,一个基于连接机制的认知处理系统环境可用下图表示。



在该系统环境中,虚拟网络是通过高级语言(如KL-one等)根据不同的认知任务编程产生,开始该语言被编辑成网络行为描述的中间形式,然后加上适当的“键”和“权”产生高度一连接的虚拟网络,最后通过有效地模拟上述虚拟网络产生物理网络,以执行网络行为。下面,将以上图为基础,介绍几种连接模型。

三、几种典型的连接模型

到目前为止,研究者们已研制出许多连接模型,以证实该结构的行为和潜力,在众

多模型中,最有代表性的有以下三种。

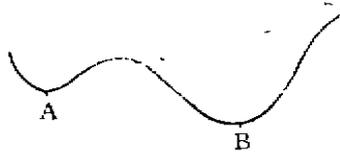
3.1 NETL—进行符号知识处理的连接系统 NETL系统是连接结构中采用局部表示的例子,它是由S.E.Fahiman于1979年提出的,用以存取大量的符号断言并进行简单的搜索和推理,其网络结构模拟语义网结构。语义网中的结点和链(link)都被表示成简单处理单元(能存储一些标志位或进行布尔运算),每个链单元与二个或二个以上结点单元相连,它通常用以将特殊标志从一个结点传送到另一结点单元。所有这些结点和链能同时对系统传送的命令(系统控制器或串行计算机)进行操作,完成其知识体内的搜索和推理。在NETL中“继承”是通过“标记-传播”实现的(沿is-a链的标记传播),例如,当查询一灰色、四腿大象时,可以通过所有特征结点的并行标记传播(用标记1标注所有灰色事物,标记2标注四脚事物,标记3标注所有大象),最后将所有同时含有这三个标记的结点单元送到控制器。NETL系统这种并行求“交集”的过程很快,仅需常量时间。另外,在NETL中,当一新的描述加入时,立即通过开关网络(增加结点和链)同已有的相关知识相连,成为网络识别过程的有效部分。因而NETL网络所能执行的搜索和推理远远超过了一个简单相连存储器所能提供的。但是,由于该模型采用的是局部表示,视每一特征为原子实体,要求网络执行完全匹配,因此,它只能在无干扰的理想环境中应用。

3.2 认知过程的分层“值-传送”网络系统 在现实世界中,人们无论处理视觉,语言理解还是其它认知问题都不可避免地面临着环境干扰,为此,产生了“值-传送”网络系统。在该网络中,每一连接有一“相连标量权”,每个处理单元计算所有输入“权”的总和,输出通过一非线性函数传送。基于这种值-传送单元,建立能够处理“不确定性观察或干扰”的分层认知系统。这种网络最简单的例子是简单的二层结构,如Perceptron

(Minsky and Papert 1969)。结构给定输入向量集及希望的输出,Perceptron收敛过程能找到一组适当的连接“权”以执行这种匹配。然而大多数认知任务都不是简单的二层网络所能完成的,往往在输入-输出之间还有一些隐藏因素的影响,仅靠输入特征简单地决定输出是不行的,需考虑在输入、输出之间引入一个或多个由隐蔽单元(hidden units)组成的干扰层,以便产生能更有效地决定输出的较高级特征。然而要确定适合的相连接权集合和隐蔽单元分配是非常困难的,为此,Hinton和Rumelhart 1985提出基于“反向-传播”过程(back-propagation)的学习算法。反向传播过程包括二次传送;正向传播:从输入单元开始,通过分层传播产生输出向量;反向传播:将上述结果输出与预期输出的偏差以相反方向通过同样的连接传送,网络对每一连接权计算输出偏差对应于该权的梯度,然后朝减少偏差的方向修改权值。因此,该网络中的学习是通过沿权空间中“错误表面”(error surface)的梯度降低实现,目前,反向-传播网络已证明能学习隐蔽单元间的许多不同的有趣表示。例如,它能学习从窄频通道排出的信息的最佳编码,还能将一组字母序列的编码输入向量转换成该序列语音特征的编码输出向量。然而由于梯度降低非常缓慢,还可能达到局部最小,因此该方式即使是对最简单的认知任务,学习过程都非常缓慢。

3.3 关于约束满足的Hopfield和Boltzmann网络 现实世界中的认知任务不仅需要同时识别多层特征而且需要获得的是一致的、相互约束的结构。例如,在一视觉场景中,不同的对象,它们的部分、子部分都彼此约束且以复杂的方式与系统接受的低层刺激物相连。为此,Hopfield提出用能量函数来满足这种约束,其函数意义为:单元*i*和单元*j*之间用“正权”(Positive weight)连接表示其约束关系为:其中一个点火,另一个也点火;用“负权”连接表示:一个点火,另

一个则不点火。网络中任何一个状态的能量表示为 $E = -\sum_i \sum_j S_i W_{ij} + \sum_i S_i Q_i$ (W_{ij} 是单元 i 、 j 间的连接权, 若单元 i 点火则为 1, 否则为 0, Q_i 是单元 i 的阈值)。给定 Hopfield 的能量二次方程及其它单元状态, 每个单元 i 能决定当它不点火时与点火时的网络能量差, 表示为 $E_{i,off} - E_{i,on} = \Delta E_i = \sum_j W_{ij}$, 若能量差为正, 则该单元应点火(或保持点火状态)以减少全局能量, 否则, 则关闭(或保持“不点火”)。通过网络中所有单元的这种朝减少能量差方向的运动, 最终可能达到全局能量最小。然而这种方法很容易陷入局部最小, 为此 Ackley, Hinton 等提出了 Boltzmann 机器结构, 该结构是在 Hopfield 网络基础上产生的, 采用模拟“退热”搜索的逃避局部最小。模拟“退热”(annealing)搜索是已应用在许多优化问题中的搜索技术。它通过在每一单元的决定过程中增加一随机成份以逃避局部最小, 即, 对每个单元 i 计算其值(如上), 假定单元 i 为 1 状态(点火状态)的可能性为: $P_i = 1/(1 + e^{-\Delta E_i/T})$ (其中 T 是控制干扰的比例因子, 将之比成温度 T), 当 $T=0$ 时, 随机成份消失, 其行为呈 Hopfield 网络。对任何给定 T , 一旦系统达到热平衡, 则状态 A 和状态 B 的相关可能性服从 Boltzmann 分布:



A: 局部最小
B: 全局最小

这种热平衡并不意味着系统已达到一稳定状态, 仅意味已“停下”(Settle down)的状态间的可能分布。由于使整个系统达到热平衡的最佳方法, 一般是从一较高温度 T 开始(便于越过能量障碍)然后减少 T 逐渐达到低能量状态。(当 T 减少得足够慢时, 系统达到全局最

佳状态或接近最佳的状态的可能性很大), 因此, Boltzmann 网络就采用上述类似于金属的“退热”(annealing)时程, 使系统达到全局能量最小, 目前, 该网络已证明能处理大量约束-满足问题, 如, 图型分离, 符号处理等, 特别是基于它而研制成功的 Boltzmann 机器, 能通过正向和反向二个阶段的学习循环, 学习许多复杂的模式, 目前已成功地应用于语言识别方面。但是, 由于能量原理本身的局限, 这种学习过程还是很慢, 尤其在大型网络上很难实用。

四、连接机

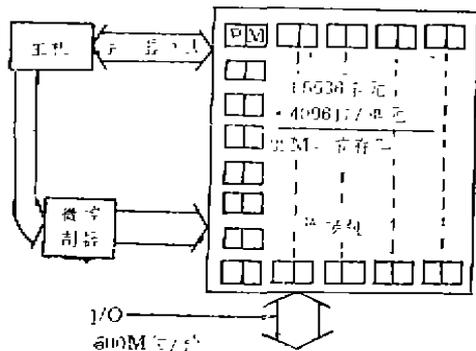
以上几种模型的实验成功, 使人们意识到建造真正具有智能的计算机的可能性, MIT 的 W. Daniel Hillis 等受连接结构思想的影响而研制成功的连接机 (Connection Machine) 标志着计算机工程上一场新的革命的到来!

Daniel 的连接机结构是基于以下两个需求提出的:

- 1、当需要时, 系统有充分的处理元可以分配(对应于问题的划分);
- 2、处理元能通过软件相连(编程实现)。

因此, 这是一种通用的全并行结构, 提供大量可并行操作的小处理/存贮单元(几千或几百万台带存贮器的小型处理机), 这些单元经可编程通讯网络(采用 CM-Lisp 编程)连接成任意所需模式, 程序设计策略要求尽可能按最接近问题的自然结构方式组织处理元。每个单元不能单独进行有意义的计算, 而是采用多个单元间连接成的数据-依赖模式(称活动-数据结构)表示和处理数据。其活动命令由连接外的主机(串行计算机)传送。主机将数据存入连接机, 如同存入到它的存贮器中, 不同的是, 这里不存在串行计算机中处理机/存贮器瓶颈问题。连接机中的计算是由活动-数据结构中单元间的交互协调完成;

相连是通过在二个处理/存贮单元内存贮指针表示,因此这种连接非常灵活,它或是由主机设置或通过I/O通道载入或动态地由通信网络改变。目前,根据上述结构,已建造出一个64K的样机CM-1,其断面图如下:由于建造CM-1的目的是用来评估连接机的行为,同时提供一种软件开发工具,因此,操作速度不是其主要设计目标,它强调的是并行结构组织的灵活性。



CM-1断面图

CM-1目前可提供从16K~64K单元的可用配置,每个存贮/处理单元带4K位存贮器和一个简单串行算术逻辑单元。处理机之间通过基于布尔n-立方拓扑的包-交换网络相连并采用适应式路径选择算法,所有处理机都执行来自主机指导下的小控制器的命令。该机器现已证明能执行大量的应用,应用系统主要有文档检索系统‘基于记忆的推理系统’、‘自然语言处理系统’、‘VLSI-电路模拟系统’等,这些应用系统的执行成功,显示出连接机的巨大潜力,越来越多的人开始从控制我们方法论的von Neumann结构中解放出来,投身于这场新的计算机革命中去!

五、连接机制之争

连接结构的提出,特别是连接机的研究成功,对传统计算机结构提出了猛烈的挑战,引起了各方面的巨大震动,特别是关于连接机制和一般AI方法的关系问题引起了极大的争论。目前主要有以下三种观点:

1)认为连接机制仅仅是标准理论的执行模型,与心理学无关;(D.Broadbent 1985)

2)认为连接机制同一般AI相比,显示出认知过程的微-结构细节,这种微-结构细节是心理学相关的。(McClelland and Rumelhart, 1985)

3)认为连接机制仅能解释有限范围的精神现象,而非认知现象,只有一般AI方法能模拟真正的认知结果。

我们认为第三种观点是对连接机制持不乐观的态度,前二种则是基于连接机制的二个极端的认识,目前,似乎还不能对连接方法和一般AI的关系作一明确的定义。连接机制有时仅是标准理论的执行模型,有时又显示出认知结构的微-结构细节,它不是对一般AI方法的取代,而是对一般AI所不能模拟的认知过程的补充。所有这些,都不是上述任何一种观点所能简单概述的。

六、结论

连接结构的提出为认知处理过程提供了强有力的工具,并为AI长期忽略的开发问题提供了一种新的方法。连接机的研制成功,证实了连接结构的合理性和可行性。目前该结构遇到的主要挑战在于:如何对难以划分数据元的认知任务分配处理单元?或许,这将取决于我们对人脑结构的进一步了解。“至今为止,我们都是将计算机作为理解人脑的模型,或许现在是应该颠倒过来的时候了——从人脑结构中获取一些建造计算机的线索”。[Bob Noyce, 1984]如果说人脑是由许多不同结构的虚拟计算机组成的,有的呈现连接结构,有的则更象von Neumann结构,那么我们当前最重要的任务则是怎样将这些结构有机巧妙地溶为一体,建造出具有真正智能的计算机。“也许,这一目标的真正实现需要几代人的努力,但是,Daniel及同事们已经为我们做了一个良好的开端! [Minsky, 1985]。

参考文献

1. "connectionist Architecture for Artifi-

(下转9页)

其中, P和Q代表被|隔开的赋值语句表, 因此, P∪Q恰好是这两张表的并。条件⟨c⟩⟨b⟩false等价于(c∧b)。由于运算对象是标准形式, 这样到处有定义且蕴涵着P∪Q中所有表达式也有定义。

$$\begin{aligned} & (5) \text{ 条件} \\ & (P \langle c \rangle \perp) \langle b \rangle (Q \langle d \rangle \perp) \\ & = (P \langle b \rangle Q) \langle c \rangle \langle b \rangle d \rangle \perp \quad (\text{根据} \\ & \text{条件法则(12)})。 \end{aligned}$$

如果 $P = \bigcup_{i=1}^n x := E_i$, 且 $Q = \bigcup_{j=1}^m x := F_j$, 则 $P \langle b \rangle Q = \bigcup_{i=1}^n \bigcup_{j=1}^m (x := E_i \langle b \rangle x := F_j)$ (根据条件法则(9))

$$= \bigcup_{i=1}^n \bigcup_{j=1}^m (x := (E_i \langle b \rangle F_j)) \quad (\text{根据赋值法则(6)})$$

因为 $c \Rightarrow \bigcup_{i=1}^n E_i$, 且 $d \Rightarrow \bigcup_{j=1}^m F_j$, 所以, 对所有的 i, j , 有 $c \langle b \rangle d = \bigcup_{i=1}^n \bigcup_{j=1}^m (E_i \langle b \rangle F_j)$ 。

于是, (5) 的左边化归到标准形式。

(6) 顺序复合

$$((\bigcup_{i=1}^n x := E_i) \langle b \rangle \perp); ((\bigcup_{j=1}^m x := F_j(x)) \langle c(x) \rangle \perp)$$

根据对U的分配, 上式化归到

$$\begin{aligned} & \bigcup_{i=1}^n \bigcup_{j=1}^m ((x := E_i \langle b \rangle \perp); (x := F_j(x) \langle c(x) \rangle \perp)) \\ & = \bigcup_{i=1}^n \bigcup_{j=1}^m (x := F_j(E_i) \langle c(E_i) \rangle \langle b \rangle \text{false} \rangle \perp) \quad (\text{根据赋值法则(9)})。 \end{aligned}$$

(4) 中所述方法把并分配到条件, 从而得到

$$(\bigcup_{i=1}^n x := F_i(E_i)) \langle \bigwedge_{i=1}^n (c(E_i)) \rangle \langle b \rangle \text{false} \rangle \perp,$$

其中∧归纳定义为

$$\begin{aligned} & \bigwedge_{i=1}^n c_i = c_n \\ & \bigwedge_{i=1}^n c_i = (\bigwedge_{i=1}^{n-1} c_i) \langle c_n \rangle \text{false}。 \end{aligned}$$

这样证明了所有有限程序都是可化归的。

标准形式的重要性在于它们提供了一种完整的检测方法, 用以确定两个有限程序文本是否表示相同的程序。首先把两个程序化归到标准形式。若它们的标准形式相等, 则两个程序也相等; 否则, 两个程序不相等。

两个标准形式 $(\bigcup_{i=1}^n x := E_i) \langle b \rangle \perp$ 和 $(\bigcup_{j=1}^m x := F_j) \langle c \rangle \perp$ 相等, 当且仅当 $b=c$ 且 $\{v \mid \exists i \leq n \cdot v = E_i\} = \{w \mid \exists j \leq m \cdot w = F_j\}$, 其中, 对表达式b、c、E_i和F_j中所有的变量值, 这些方程必须都成立。

这些方程的真假(例如, 在整数算术中)可能是不可判定的。这里列出的结果仅建立了相对完备性。

(未完待续)

[宋国新译自Comm. ACM, vol 30, No8, 1987, 孙永强校]

(上接28页)

cial Intelligence" Scott E. Fahman and Geoffrey E. Hinton Carnegie-Mellon University, Computer 1987.1.

- "The Connection Machine" W. Daniel Hillis 1986, MIT Press
- "The Cognitive Architecture Project" Dan Hammerstrom, David Maier, Shreckant Thakkar, Computer Science & Engineering The Oregon Graduate Center, 1986. 2.7
- "Connectionism and Cognitive Science" Andy Clark. Cognitive Studies Programe, University of Sussex Brighton, U. K. "Advances in AI" 1987.4. Edinburgh.
- "Applications of the Connection Machine" David I. Waltz, Think Machine Corporation IEEE, 1987. 1.

- J. A. Feldman and D. H. Ballard, "Connectionist Models and their Properties", Cognitive Science, Vol. 6, no. 3, Ablex, Norwood, NJ, 1982, PP. 205-254.
- G. E. Hinton, J. L. McClelland, and D. E. Rumelhart, "Distributed Representation, "Parallel Distributed Processing, Explorations in the Microstructure of Cognitive, eds. D. E. Rumelhart, J. L. McClelland, and the PDP Research Group MIT Press, Cambridge, MA, 1986
- G. E. Hinton, "Learning Distributed Representations of Concepts, " Proc. 8th. An. Conf. Cognitive Science, Soc., Lawrence Erlbaum Assoc., Hillsdale, NJ, 1986.