

思维模拟—人工智能的根本途径^{*)}

洪家荣 (哈尔滨工业大学计算机科学与工程系)

摘 要

This article discusses the major issues of the possibility of artificial intelligence and outlines several integrated learning systems for simulating human thought processes, including intuition, association, discovery, and insight. These results show that the simulation of thought processes is a fundamental way to advance the development of artificial intelligence.

一、引 言

人工智能一般被认为是研究如何使计算机去做需要人类智能的工作的一门科学。因此,人工智能旨在模拟人类行为和认知过程。

计算机能不能思维,或计算机能不能代替人脑?这个问题一直是心理学和人工智能界中的乐观派同悲观派长期争论的焦点。乐观派的立论根据是纽厄尔和西蒙的物理符号系统假设^[1],把人和计算机都看成信息加工系统,具有对符号信息的输入、输出、存储、复制、结构化和条件转换等六种功能,并认为这些功能是物理符号系统具有智能的充分必要条件。由此得出结论,在符号加工这一层次上,计算机和人脑等价。西蒙宣称:“现在计算机正在向又一个根深蒂固的观念:只有人才具有思维能力提出挑战”^[2]。费根鲍姆还预言日本研制的“第五代计算机甚至能在思维方面超过人类。”^[3]

悲观派对人工智能持否定态度^[4],认为在对人脑的思维机制尚不清楚的情况下,思维能否用计算机模拟只能是经验性的,而物理符号系统假设不过是一种杜撰。他们甚

至认为人工智能研究已经达到了它的极限,任何进一步的发展都不大可能。此外,最近出现的“神经达尔文主义”^[5]也反对物理符号系统假设。他们指出,在符号加工前,首先要对被加工的范畴(概念)命名(labelling),这就有很强的主观先验因素。另外,被命名的概念可以很低级(如视网膜上的一点座标);也可以很高级(如“椅子”这样的概念),因此系统在将环境转换为输入信息时,到底使用了多少智能是很难掌握的。他们还把物理符号系统假设和大脑皮层的定位假说(“侏儒小人”说)看成同样是先验的杜撰。他们主张人类范畴的形成是由神经元群体选择原理决定的。

作者认为上述两派观点既各有其合理性,也有其片面性。物理符号系统假设揭示了计算机和人脑在信息加工方面某些初等功能的相似性,为计算机模拟人类思维奠定了理论基础。但是,西蒙等人根据计算机和人脑某些功能上的相似性,得出计算机能够思维的结论不过是唯心主义的遐想。他们抹杀了人同物、精神同物质的根本区别。他们认为认识是结构化符号表示的加工,而符号从

^{*)} 国家自然科学基金和国家863资助项目。

同外部世界的对应中获得意义,所以思维是抽象的和客观的,它的意义需要从它所表达的外部世界去寻找。换句话说,思维可以脱离肉体而存在,这就导致了二元论。事实上,心理学、语言学、人类学和生物学的研究结果都同物理符号系统假设相矛盾。理论上,任何计算机都可以看成复杂程度不同的形式系统,它们的信息加工能力是极其有限的。根据数理逻辑中著名的哥德尔不完全性定理,即使在只包含初等数论和一阶逻辑的简单的形式系统中,也有不能证明其真伪的命题,即不可知的事物;并且一个形式系统的一致性(无矛盾性)不能在本系统内部得到证明,需要在一个新的更高级的系统内部得到解决,而这个新系统又有不能在其自身中解决的问题,等等。何况人脑比任何形式系统都复杂得多,它的机理一直是一个谜。即使在大脑皮层的所谓侏儒小人中,也有大片未知的区域。哺乳动物的进化约有数亿年,而人从哺乳动物又进化了约一百万年,这么漫长的自然选择与进化过程已经把哺乳动物和人类塑造成极为精美的“艺术品”了。谁能搞清楚人类的大脑机制呢?人类自己是决不可能的,人类至今不仅对哺乳动物、鱼类、甚至对与人类的进化过程相差数亿年的最低级的腔肠动物(如水螅)的神经机制都搞不清楚。换句话说,没有比人类先进数亿年的外星“超人”,是搞不清楚人类的大脑机制的。日本第五代计算机的失败宣告了“计算机能够思维”预言的破产。

悲观派对人工智能的某些批评是有益的,有助于帮助人们清除乐观派的不切实际的遐想。但悲观派认为计算机在智能模拟方面将无所作为,这种对人工智能全盘否定是错误的。作者相信,虽然原则上我们不能造出像人一样思维的计算机,但是我们有可能用计算机模拟人的某些思维过程,解决一些复杂的实际问题。例如,虽然对鸟类飞行的神经机制至今也没有搞清楚,但是人类根据鸟飞行的空气动力学原理造出的飞机却远远

超出了鸟类飞行的能力。类似地,虽然我们不能精确地了解人类专家思维的神经系统活动机制,但我们根据专家解决问题的外部启发式特征造出的专家系统却可以在一定(狭小)范围内和一定程度上解决只有专家才能解决的困难问题。由于人脑进化的历史比感官与运动系统短得多,对思维的模拟肯定比对其它器官功能的模拟容易得多,这使思维模拟不仅必要而且可能。人工智能出现四十多年来,在自然语言理解、模式识别、计算机视觉、定理证明、专家系统、机器学习和机器人学等方面取得的成果是令人瞩目的。如今,人工智能技术已经发展为一门科学。可以预言,人工智能的发展前景是不可估量的。

本文其余部分将扼要介绍作者在思维模拟方面取得的进展。我们的工作同当前国际上相关的工作的主要区别在于:现有工作一般只模拟较低级的思维活动,如刺激-反应(产生式系统)、手段-目的分析、组块(chunking)等。我们的工作模拟较高级的思维活动,如直觉、联想、发现、灵感等;现有工作用以模拟的手段是单一的学习方法(如示例学习、解释学习等)简单的搜索策略或几个简单搜索策略的组合(如SOAR系统^[6])。我们的工作使用多个较复杂学习方法和智能系统的集成。文中所介绍的集成化智能系统包括:(1)THOUGHT^[7],它能将处于新手水平的专家系统改造为具有专家直觉与联想能力的新型专家系统;(2)KD₁^[8],自动规则发现系统;(3)KDEBL^[9],适合于不完善领域理论的解释学习系统;(4)SWT₃^[10],基于问题求解的布尔函数化简系统;(5)CHWR^[11],手写体汉字识别系统;(6)LNES^[12],传统专家系统同神经网络相结合的专家系统合理模型;(7)ABAS^[13],具有联想能力的类比学习系统;(8)灵感的逻辑实现。

二、基本思维活动的模拟

人认识事物就是认识事物的本质,即事物的概念以及概念间的相互关系。事物的概念是抽象的,它一般用概念的外延(该概念代表的事物的集合)或概念的内涵(该概念的描述)来表示。当人们面对一大堆杂乱无章的事物(如数据或样本)时,他们往往首先根据事物的相似性加以分类,然后再抽取出每一类区别于其他类的基本特性。此即首先对事物进行分析与分类以形成某些有意义概念的外延,然后对形成的外延进行抽象与概括以产生这些概念的内涵。因此对事物的分析与分类、抽象与概括是人类基本的思维活动,也是思维模拟的基础。

我们在SUN₃工作站上实现了一个叫做COGNITRON的分类抽象系统,它由概念聚类系统LEOBS和示例学习系统GS^[14]组成。LEOBS是著名的概念聚类系统CLUSTER/2^[15]的新版本,其原理类似于一般动态聚类分析,但主要区别是,它以概念描述指导分类,因此所分的类具有更有意义的内涵。概念描述的抽象一般使用示例学习来实现。当前国际上流行两个有名的示例学习系统,一是J. R. Quinlan的ID₃^[16],二是作者与R. S. Michalski的AQ₁₅^[17]。ID₃速度快、效率高,但生成的描述难于显式表达;AQ₁₅有很强的表达能力,但速度较慢。GS吸取了两者的长处并且算法简单。

2.1 基本算法

一个例子 e 或事件是一个 n 维向量(属性值对),即 $e = \langle v_1, v_2, \dots, v_n \rangle$, $v_i \in D_i$, D_i 是离散有穷符号集。一个概念描述(简称描述)是公式的析取式,公式是选择子的合取式,选择子是形为 $[x_j = V_j]$ 的关系语句,其中 $V_j \subseteq D_j$, x_j 是第 j 个属性。例子 e 满足描述 F ,是指它满足 F 中的一个公式, e 满足公式 cpx ,是指它满足 cpx 中的所有选择子, e 满

足选择子 $[x_j = V_j]$,是指 $v_j \in V_j$ 。 e 满足描述(公式或选择子)也称该描述(公式或选择子)覆盖 e 。设 S_1 和 S_2 是两个例子的集合,一描述(公式或选择子)覆盖 S_1 并排除 S_2 ,是指它覆盖 S_1 的每一个元素但不覆盖 S_2 中的任何元素。求覆盖正例集 S_1 并排除反例集 S_2 的过程叫做示例学习,其结果是一个描述Cover(S_1, S_2, A, C)。其中 A 是学习算法, C 是学习规则(如较优覆盖,较简公式等要求*)。本文中,参数 A 和 C 一般省去不写。

2.1.1 LEOBS算法 假设 S 是训练例子的集合, k 是要求分的类数。

(1) 从 S 中任取 k 个例子作为“种子”, $e_1, \dots, e_k$ 。

(2) 用示例学习算法AE₁^[18]生成覆盖每一种子而排除其余种子的描述 $F_i = \text{Cover}(e_i, \{e_j | j=1, 2, \dots, k, i \neq j\})$, $i=1, 2, \dots, k$ 。

(3) 求出 S 中被 F_i 覆盖的所有例子的集合 S_i , $i=1, 2, \dots, k$ 。

(4) 将 S 中余下的例子逐个放入距离最近的集合中去。这里例子到集合的距离是该例子到该集合中所有元素距离的平均值,而两个例子间的距离定义为两例子不同的对应值数目的 n 分之一。

(5) 找出每一类的几何中心,仍记为 $e_1, \dots, e_k$ 。一类 S_i 的几何中心是该类中到所有其余元素距离和最小的例子。这些几何中心将做为下次迭代的种子。

(6) 如果某聚类条件满足^[16],终止;否则转向(2)。

不难看出,概念聚类LEOBS同动态聚类分析在算法上的主要不同是在第(2)步和第(3)步,即概念的抽取和概念的导向。

2.1.2 GS算法 设 S_1 和 S_2 是正例集和反例集,学习规则是较优公式问题, S_1' 和 S_2' 是临时正、反例集。

(1) $S_1' \leftarrow S_1, S_2' \leftarrow S_2, cpx \leftarrow \text{true}$ (公

*) 作者在国际上率先证明了最优覆盖和最简公式问题是NP困难的。^[18,19]

式赋初值)

(2) 从 S_1' 找出出现最多的属性值 v_0 , 设为第 j_0 个属性取值, 建立选择子 $[x_{j_0}=v_0]$, 并加入公式中, $cpx \leftarrow cpx \wedge [x_{j_0}=v_0]$ 。

(3) 检查 $[x_{j_0}=v_0]$ 是否还覆盖 S_2' 中的任何反例。若是, 则准备对 cpx 特化, 即转向(4); 否则新的公式 cpx 建成, 将其加入到生成的描述中去。并准备产生新公式, 即转向(5)。

(4) 用 S_1' 和 S_2' 中被选择子 $[x_{j_0}=v_0]$ 覆盖的例子组成新的 S_1' 和 S_2' 。准备建立新的选择子, 即转向(2)。

(5) 从 S_1 中去掉被新生成公式 cpx 覆盖的例子。如果 S_1 变空, 则终止, 描述生成; 否则转向(1)。

有了分类抽象系统COGNITRON, 就可以以它为核心, 构造更复杂的集成化智能系统, 模拟更复杂的思维过程。

三、专家直觉和联想的模拟

本节介绍如何把COGNITRON进一步同专家系统和经验学习结合起来形成一个新型的专家系统工具THOUGHT。THOUGHT不同于传统专家系统工具, 具有人类专家直觉、联想和启发式多种问题求解能力, 从而大大提高了现有专家系统工具的效率。

3.1 专家系统与专家思维

专家系统是模拟专家解决问题的启发式方法, 但一个实际专家系统的知识库通常很庞大, 用启发式方法求解时, 由于搜索空间太大, 效率一般都很低。另一方面, 每一组事实给出以后, 现有专家系统都必须使用规则从头一步一步推理, 尽管这组事实从前用到过。换句话说, 这些专家系统既不能根据问题性质采用不同的求解策略, 也不会根据以往的经验加快求解的速度, 永远只停留在见习生水平上。

然而人类专家求解问题的方法却不同, 他们能够根据问题的性质采用不同的方法求解问题。对于过去经历过的熟悉问题, 他们

能根据直觉, 直接跳跃到所求的解, 即一步解决; 对于过去虽未经历过但同所经历过的问题比较相似的问题, 他们能进行联想, 把问题空间缩小, 也能较快求解; 只有对于完全陌生的问题求解, 他们才象现存专家系统那样一步步推理。直觉和联想、特别是直觉是专家同新手的主要区别。一个专家系统是否具有直觉判断的能力是当前国际上评价专家系统是否达到专家水平的争论焦点。

3.2 THOUGHT的建造

THOUGHT系统能够模拟一个见习生通过实践和经验积累, 经验的分析与分类, 抽象与概括等思维活动, 使自己成为一个专家的过程, 这个系统还能将现存专家系统改造成具有专家直觉的名副其实的专家系统。

让我们先分析一下专家的成长过程。例如看一看一个见习生是如何成长为医生的。一个见习生在读完医科大学以后, 有了书本知识。然后他要经过临床实践, 这是他的感性认识阶段。在这一段时间, 他积累了丰富的经验。这些经验在他的头脑中被分析分类, 抽象概括, 他的感性认识上升为理性认识, 这是认识上的一个飞跃。同时, 他头脑中的知识结构也发生了相应的变化。数年以后, 当经验足够丰富, 理性认识达到一定高度, 知识结构日趋完善, 他就成为一个专家了。因此, 一个真正的专家系统应当具有经验积累、经验分类、经验抽象等功能。而且它的知识库结构应当不同于现存专家系统的知识库结构。

对专家系统, 一次问题求解的经验定义如下: 设 e 是所给的事实合取式(即一个例子), s 是所求的解, 及 $\langle i_1, \dots, i_m \rangle$ 是推理路径, 其中 $i_1, \dots, i_m$ 是在推理中所用规则 $R_{i_1}, \dots, R_{i_m}$ 的标号。 $e \rightarrow s$ 叫做经验规则。一个规则的联想是和该规则具有相同结论的规则。经验路径由推理路径中加入其中每个规则的联想标号组成。于是, 一次问题求解的经验由一条经验规则和一条经验路径组成。注意, 不同的经验规则可能对应同一

条经验路径。

在集成化系统THOUGHT中,经验分类与抽象由子系统COGNITRON完成,经验收集由经验学习系统SGESA^[20]执行,经验生成由普通专家系统运行实现。具体过程如下。

THOUGHT的工作过程,设NOVICE 是一个普通专家系统,其知识库是KB。

(1) 给定一问题,即一组事实 e , NOVICE用 KB中的规则进行推理。如果有解,则启动SGESA产生有关经验,即经验规则 $e \rightarrow s$ 和经验路径。给新的问题,重复上述过程,直到积累的经验足够丰富为止。

(2) 启动COGNITRON,对积累的经验进行分类与抽象,产生类似于决策树的结构。具体做法是:(i)用LEOBS对经验路径的集合 S 进行分类得子类 $S_1, S_2, \dots, S_k$ 。同时,相应的经验规则前提的集合 U 也被分为 k 个子类 $U_1, \dots, U_k$ 。(ii)然后,用GS产生每个经验规则前提集的局部描述 $F_1, \dots, F_k$,其中 $F_1 = \text{Cover}(U_1, U - U_1), \dots, F_k = \text{Cover}(U_k, U - U_k)$ 。这个分类与抽象过程再对

子类递归进行,则最终产生一棵决策树。其叶结点由一条或几条经验规则和一条经验路径组成。而中间结点由一组相似的经验路径与经验规则组成,其中经验路径的相似性由使用多少相同规则决定,而经验规则的相似性反映在有多少相同的事实。

THOUGHT的结构及其模拟的认知过程如图1所示。

定理1 假设THOUGHT从经验路径集合 S 和经验规则前提集合 U 构造成决策树状的 k 叉树。

(1) 如果所给问题是经历过的,则只需要大约 $\log_k |S|$ 次匹配即可求到解,其中 $|S|$ 是 S 的基数。

(2) 如果搜索达到某个结点(即从根结点到当前结点上的描述均被所给事实所满足),如果该问题在这个结点的规则集(叫经验网络)中无解,则在其他结点也不可能无解。

定理1断言,对熟悉的问题,被THOUGHT改造过的专家系统可以快速(对数级)

检索达到所求的解,即类似于专家直觉求解;而对相似的问题,它的检索同样以对数级达到某个结点(因为它和该结点中的经验规则前提满足从根结点到该结点弧上的所有描述),然后在该结点内部的小规则集中推

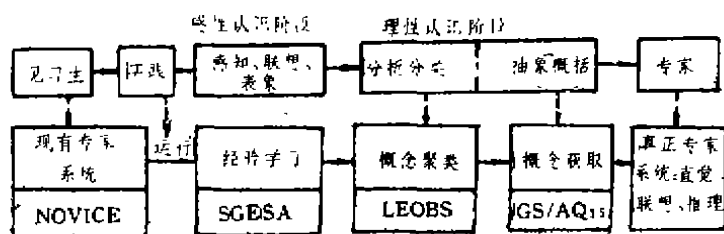


图1 THOUGHT的结构功能图

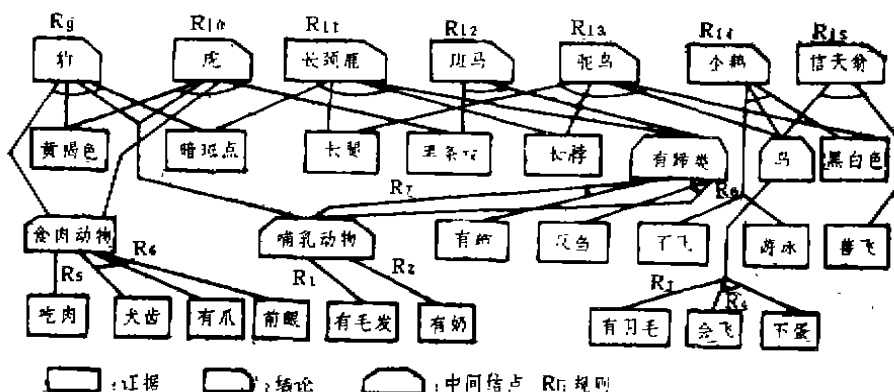


图2 动物识别专家系统的知识库和推理网络

理，其效率仍然很高。如果无解，就无需再回溯。换句话说，这个改造过的专家系统确实模拟了专家问题求解方式。

3.3 应用举例

表1

经验路径及其概念聚类

扩展 经验路径	规则号															概念聚类	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	一级	二级
EP (豹)	1	1			1	1			1							1	11
EP (虎)	1	1			1	1				1						1	11
EP (长颈鹿)	1	1					1	1			1					1	12
EP (斑马)	1	1					1	1				1				1	12
EP (鸵鸟)			1	1									1			2	21
EP (企鹅)			1	1										1		2	21
EP (信天翁)			1	1											1	2	21

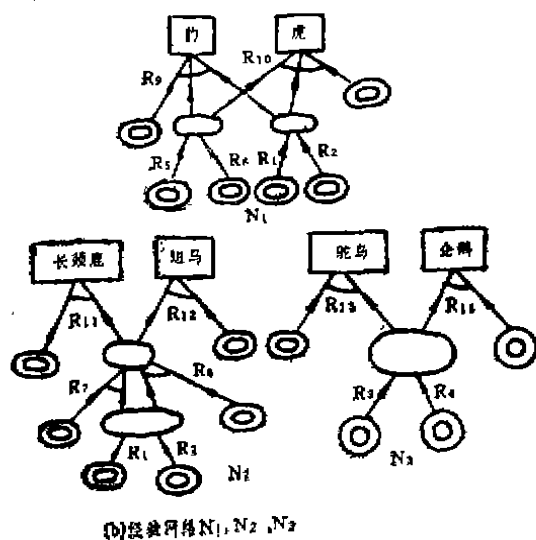
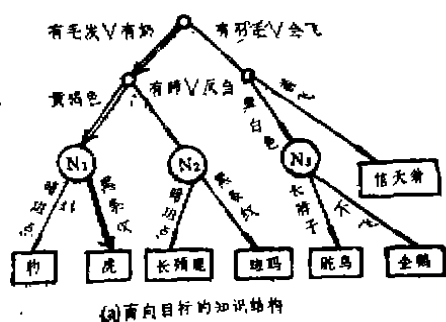


图3 动物世界专家系统面向解的新知识结构

假定已知动物识别专家系统ANIMAL，其知识库如图2所示，其中有向弧联结的结点构成从前提到结论的规则R_i。

表1是ANIMAL求解7次所产生的经验

路径。从表中可看到联想对聚类的重要作用。例如，豹和虎经联想后的经验路径就极为相似（规则1、2、5、6相同）。但如果豹由规则1、5、9推出，而虎由2、6、10推出，若不进行联想，则豹和虎很难分到一组。另外，我们发现不用经验路径，用LEOBS只对经验规则前提集聚类，便得不到理想的经验网络，所以，先对经验路径聚类，再对经验规则前提集抽象是必要的。

图3是ANIMAL被THOUGHT改造后所产生的知识库的决策树表示，这是知识库面向解的重组，它比图2中的原推理网络有更高的执行效率。我们可以将图2中的结构看成一个见习生的知识结构，而图3看成一位动物学家的知识结构。

四、知识发现

机器发现 (Machine Discovery) 是在没有外界帮助下从已知数据中自动地获取新知识，即蕴藏在数据中的规律和关系。随着科学技术的发展，积累的数据越来越多，所以机器发现被认为是最有应用前景的机器学习方法。由于专家系统发展的需要，规则发

现将起着更重要的作用。当前国际上已开始出现一些规则发现系统,但它们在理论和方法方面还很不成熟。作者根据人类知识发现的认知过程,实现了一个集成化知识发现系统KD₁^[8-21]。规则发现是从大量经验数据

(例子)中发现某些概念间的逻辑关系,如蕴涵和等价。一般来说,一个概念都有其外延(概念的例子)和内涵(概念的描述)。人们通常将大量杂乱无章的经验数据加以分类,并将这些类看作某些概念的外延,然后再从这些类中抽取概念的内涵。最后,通过寻找概念外延间的集合包含关系来发现其内涵间的逻辑蕴涵关系,这就是KD₁的基本思想。KD₁由COGNITRON和集合运算组成。COGNITRON负责数据分类和内涵抽取,集合运算用于构造规则。

范畴的概念构成人类知识和学习的核心部分。在规则发现中,类的命名将起到重要作用。KD₁中类U的名字或由用户给出或用函数Label(U)隐式表示。我们还定义,描述F₁逻辑蕴涵F₂即F₁→F₂,是指F₁的外延是F₂的外延的子集;F₁逻辑等价于F₂,F₁⇔F₂,是指F₁→F₂并且F₂→F₁。

4.1 KD₁中的规则发现方法

在KD₁中,子系统COGNITRON除了进行层次聚类,形成图3状的层次结构外,LE-

OBS还进行平行聚类,即依次取k=1, ..., r,将同一S分别聚类r次,产生1+2+...+r=r(r+1)/2个类的簇。在形成规则时,除了使用2.3节COGNITRON产生的描述和局部描述外,还使用类名。形式地,我们有如下定理:

定理2

(1) 任何一类的名字、它的所有覆盖和所有基本覆盖都彼此等价。

(2) 假定S_i和S_{ij}等是由层次聚类产生的集合。

(i) F_i和F_{ij}是相应的描述,即F_i=Cover(S_i, S-S_i), F_{ij}=Cover(S_{ij}, S-S_{ij}), 等, 则F_{ij}→F_i。

(ii) F_i和F_{ij}是相应的局部描述,即F_i=Cover(S_i, S-S_i), F_{ij}=Cover(S_{ij}, S_i-S_{ij})等, 则F_i∧F_{ij}∧...∧F_{ij...op}⇔Label(S_{ij...op}); Label(S_{ij...op})∧F_{ij...op}⇔Label(S_{ij...op})

(3) 假定S_i^(u)和S_j^(v)由平行聚类产生, F_i^(u)和F_j^(v)是相应的描述, 则如果S_i^(u)⊆S_j^(v), 则F_i^(u)→F_j^(v)。

定理2的(2)(i)叫层次规则发现,(2)(ii)叫做继承规则发现,(3)叫做平行规则发现。

4.2 KD₁用于自动建立专家系统的知识库

表2 动物的例子(*表示缺省值)

属性 编号	毛发	牙齿	眼睛	羽毛	脚	食物	奶	飞	产卵	会游泳
1 虎	有	犬齿	前方	无	有爪	肉	有	不	不	是
2 豹	有	犬齿	前方	无	有爪	肉	有	不	不	是
3 长颈鹿	有	钝	旁	无	蹄	草	有	不	不	是
4 斑马	有	钝	旁	无	蹄	草	有	不	不	是
5 鸵鸟	无	无	旁	有	有爪	谷	无	不	是	不
6 企鹅	无	无	旁	有	蹼	鱼	无	不	是	不
7 信天翁	无	无	旁	有	有爪	谷	无	是	是	不
8 鹰	无	无	前方	有	有爪	肉	无	是	是	不
9 毒蛇	无	犬齿	旁	无	无	肉	无	不	是	不
10 蜜蜂	无	无	旁	无	*	蜜	无	不	*	不

给出一些动物的例子,如表2所示, KD₁使用继承规则发现,自动构造了动物识别专家系统的整个知识库,如表3所示。图4是继承规则发现产生的结构树。用层次规则发现或平行规则发现还可以产生其它规则,例如[脚=蹄]→[类=哺乳]。

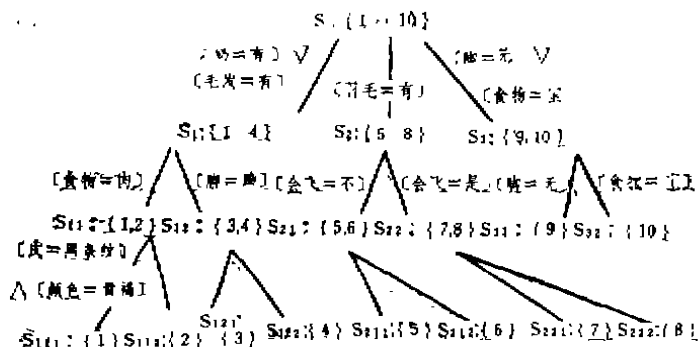


图4 继承法规则发现产生的结构树

表3 继承法规则发现产生的规则

- 1 [毛发=有]∨[奶=有]→[类=哺乳]
- 2 [羽毛=有]→[类=鸟]
- 3 [类=哺乳][食物=肉]→[类=食肉]
- 4 [类=哺乳][脚=蹄]→[类=有蹄]
- 5 [类=鸟][会飞=是]→[类=飞鸟]
- 6 [类=鸟][会飞=不]→[类=非飞鸟]
- 7 [类=食肉][皮=黑条纹][颜色=黄褐]→[动物=虎]
- 8 [类=食肉][皮=暗斑点][颜色=黄褐]→[动物=豹]
- 9 [类=有蹄][皮=暗斑点][颜色=黄褐][脖子=长][腿=长]
- 10 [类=有蹄][皮=黑条纹]→[动物=斑马]
- 11 [类=非飞鸟][颜色=黑白][脖子=长][腿=长]→[动物=鸵鸟]
- 12 [类=非飞鸟][颜色=黑白][会游泳=是]→[动物=企鹅]
- 13 [类=飞鸟][食物=谷]→[动物=信天翁]

五、规则发现和解释学习相结合

解释学习(Explanation-based Learning)
是一种重要的学习方法,近年来,被广泛而深

入地研究。解释学习可概述为：已知目标概念，该概念的一个实例和领域理论。用领域理论产生该例子满足目标概念的一个解释（证明树），并将该解释推广以适合新的例子。虽然现在已有解释学习同其他学习方法（如类比学习或相似概括）相结合的报导，但由于仅仅是两个方法结合，都没有明显的

效果。尤其是解释学习的主要困难,领域理论不完善问题,还没有出现有效的解决方法。我们已实现一个发现式解释学习系统 **KDEBL**,它由规则发现系统 **KD₁**与解释学习系统 **EBLOG**集成而成。**KD₁**能够从所给事例中发现**EBLOG**所缺少的目标概念和领域知识,然后**EBLOG**用这些领域知识和所给实例构造满

足目标概念的解释及其概括。我们用例子说明KDEBL的能力。

5.1 解释学习的例子

我们引用[22]中自杀的例子。

目标概念: $Kill(x, y) \iff Hate(x, y) \wedge Weapon(c) \wedge Possess(x, c)$; x 杀死 y 等价于 x 恨 y 且 x 拥有一件武器 c 。

领域理论:

$$\text{Depressed}(w) \longrightarrow \text{Hate}(w, w);$$

如果w沮丧，则他恨自己。

$$\text{Buy}(u, v) \longrightarrow \text{Possess}(u, v),$$

如果 u 买了 v ，则 u 拥有 v

$\text{Gun}(z) \longrightarrow \text{Weapon}(z),$

如果 z 是枪，则它是武器。

训练实例（自杀）：John很沮丧，并且买了一支枪Obj₁₀。

Depressed (John), Buy(John, obj₁),
Gun(obj₁)

操作准则：目标概念必须用在例子中出现的谓词表示。

确定: 训练实例的一个概括, 它满足操作准则并成为目标概念的一个充分条件。

运行 EBLOG, 可产生自杀例子的一个

解释,如图5所示。这里解释的概括省略。

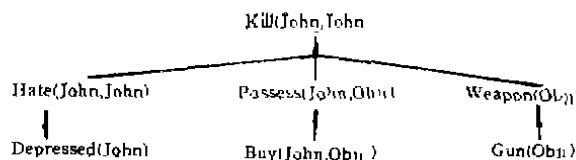


图5 自杀实例的解释

表4 训练例子 (y代表“是”, n代表“非”)

	Hate(x, y)	Weapon(z)	Possess(x, z)	Kill(x, y)	Buy(x, z)	Gun(z)	Depressed(x)
1	y	y	y	y	y	y	y
2	y	y	y	y	y	y	y
3	y	n	y	n	y	n	y
4	y	y	n	n	n	y	y
5	y	y	n	n	n	y	y
6	n	y	y	n	y	y	n
7	n	n	y	n	y	n	n

我们让KDEBL中的发现系统 KD_1 做平行规则发现(见4.1节)。对于 $k=2$, KD_1 可将表4的例子集 $S=\{1\cdots 7\}$ 分为两类 $S_1=\{1, 2\}$ 和 $S_2=\{3\cdots 7\}$ 。其中 S_1 的一个覆盖为 $[Kill(x, y)=y]$,一覆盖为 $Hate(x, y) \wedge Weapon(z) \wedge Possess(x, z)$ 。再根据定理2(1),它同 $[Kill(x, y)=y]$ 等价。我们用2.2.1节的GS算法具体地产生这一合取覆盖 $Cover(S_1, S_2)$ 。首先,GS在 S_1 (见表4头两行)中查找属性值出现最多的建立选择子 $[Hate(x, y)=y]$,但这个选择子还覆盖了三个反例 $\{3, 4, 5\}$ 。因此要进行特化,即要求排除反例 $\{3, 4, 5\}$ 。类似地选出 $[Weapon(z)=y]$,但它又覆盖了集合 $\{3, 4, 5\}$ 中的反例 $\{4, 5\}$ 。继续特化,排除集合 $\{4, 5\}$,因此得到合取覆盖 $[Hate(x, y)=y] \wedge [Weapon(z)=y] \wedge [Possess(x, z)=y]$ 。

六、布尔函数化简系统 SWT_3

作者在[23]中介绍了一个布尔函数化简系统 SWT 。它的核心算法是示例学习的扩张矩阵算法 AE_1 [24]。 SWT 基本思想是,已给 n 个变元的布尔函数 $f(x_1, \cdots, x_n)$ 。生成 2^n 个

5.2 发现式解释学习的例子

假定已知几个训练例子如表4所列。

使用发现式解释学习系统KDEBL,可立即发现出5.1节中的目标概念及所有领域理论,并使解释学习得以实现。我们仅以目标概念的自动发现为例说明KDEBL的工作原理。

指派,然后将满足 f 的所有指派做成正例集,其余的指派做成反例集。于是布尔函数 f 的化简问题变为求一个比 f 更优化的覆盖正例集并排除反例集的概念描述。 SWT 应用于一些实用布尔函数,均能获得满意的结果。但 SWT 的一个根本缺陷是输入变元数不能太大(例如,在SUN工作站上运行, n 不能超过16)。系统 SWT_3 正好弥补了这个缺陷。已给 n 元布尔代数 $f(x_1, x_2, \cdots, x_n)$,一般 $n > 16$ 。设 f 是析取范式形式,即概念描述形式。现在我们造成新的例子集,每个例子是中的析取项,即公式。例子中变元的取值是1,如果该变元在该析取式中出现;否则取值0。然后用LEOBS对例子集聚类。如果所生成的类中有的类含的不同变元数仍大于某个给定的数 r , $r < 16$,则对这些类递归聚类,直到所有类都含小于 r 的不同变元为止。尔后,再调用 SWT 对每一小类进行化简。化简后,再把各类结果用析取联结起来,给出最后的结果。

SWT_3 把用 SWT 不能化简的输入变元数大的布尔函数,在析取项联结最薄弱处切开,分解成小类,再用 SWT 化简。这就可能

3.2节所述。这些经验然后用来训练神经网络,如图7所示。例如,设已知事实为 $e=\{\text{黄褐色, 黑条纹, 有奶, 吃肉}\}$,则专家系统产生经验规则为 $e \rightarrow \{\text{虎}\}$,经验路径为: $\{1, 2, 5, 6, 10\}$ (见图2)。当训练网络时,在输入层代表黄褐色,黑条纹、有奶、吃肉的节点取值1,其他节点取值0;在输出层,在结论部分仅节点虎取1,其他取0。当专家系统问题求解足够多,经验足够丰富时,神经网络训练完毕。

8.1.2 LNES的问题求解 现在给定一个问题 e , LNES首先启动神经网络求解。这可出现下列三种情况:①对熟悉问题,即 e 和过去的一个训练实例相同,则结论层有一节点为1,即为所求解。这时经验路径的输出构成解释;②相似问题,即 e 不是用过的训练实例,但却与一些实例有些相似(如某些输入事实相同),则结论层可有几个节点的值大于某个事先规定的阈值 r 。这时取节点值最大者为解;或者为精确起见,用经验路径结点值大于另一阈值 s 的规则组成小知识库,并以候选结论为目标进行推理求解。③完全陌生问题,即神经网络获得的输出结论值均小于规定阈值 r 。这时神经网络求解失败。需再启动专家系统进行启发式推理。

不难看出, LNES的三种问题求解方式是对专家直觉、联想和启发式推理的模拟。有趣的是,虽然THOUGHT和LNES方法各异,但它们的主要功能,问题求解方式以及获得的结果却极为相似,这表明这两种思维模拟方法的合理性。

8.2 ABAS——类比中的联想实现

众所周知,类比是人类最重要、最富创造性的思维方式之一。当前类比学习模拟的主要困难是类比源的联想。即给定一个目标域,如何从无数个错综复杂的结构找出一个或数个候选的类比源。例如,当卢瑟福研究原子结构时,他是怎么想到同太阳系类比的;当阿基米德要测王冠金子的纯度时,是如何在洗澡盆中发现浮力定律的。这个问题实质

上可以归结为灵感思维问题。作者感到,灵感的实现是个理论问题,即必须搞清楚灵感的机理,在这方面已有较多的研究报导。但目前对灵感的模拟只停留在功能上。这个问题是一个实践问题,即用什么方法和工具来进行这种模拟,我们发现人工智能的现有方法中,神经网络比较适合进行类比的联想。我们实现了一个基于结构映射的类比学习系统ABAS^[13],这个系统就是用神经网络的BP模型来进行联想的。由于神经网络具有高度并行、分布式存储和很强的容错与概括能力,比较适合从大量的现象中联想识别一个近似的候选解。相关的BP网络的简单模型如图8所示。

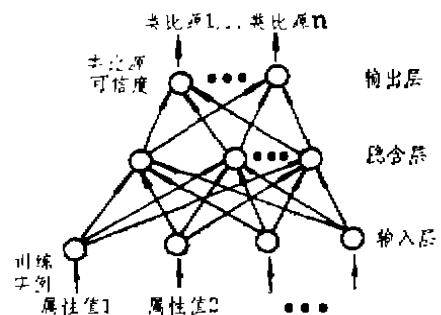


图8 一个BP网络的模型

具体地, ABAS把所有类比源做成一个个由属性/值组成的实例(编码表),这些属性包括谓词的名字,其在结构中的层次等等。当训练网络时,实例做为输入,类比源的名字做为相应的输出。当目标源给出时,它的编码表做为输入,网络的输出可信度最大的几个类比源做为类比候选集。

九、灵感的模拟实现

黑猩猩够香蕉是一个著名的灵感试验^[25]。将被试黑猩猩关在笼子里,它可以拿笼中的竹杆够外面的香蕉吃。当香蕉被移开一点,黑猩猩够不着就急得乱蹦乱跳,拿起一根竹杆够一下再丢下。有一次黑猩猩突然将两只手上的两根竹杆接起来去够香蕉,果然够着了。奇怪的是,被试以前并没有把

两根竹杆接起来的任何经验,因此这是典型的灵感实验。为了用计算机解决这个问题,首先需要将这个试验形式化,即表示成一组产生式规则的序列,然后通过推理,求得所求的解。相关的产生式规则有:

R_1 : 如果黑猩猩手中有竹杆而笼外有香蕉,则它就会用竹杆去够香蕉;

R_2 : 如果黑猩猩拿竹杆够着香蕉,则它可以吃到香蕉;

R_3 : 如果笼里有几根竹杆,则黑猩猩每只手都可以拿一根竹杆;

R_4 : 如果两根竹杆接在一起,则黑猩猩可以用它够着香蕉。

显然,从仅有的四条规则无法得出黑猩猩可以吃到香蕉的结论。原因是黑猩猩没有将两根竹杆接起来的经验,即推理在规则4出现了“断点”。这个断点是由黑猩猩的灵感来接通的。而灵感的诱发往往同过去的某些相关经验有关。那么这个相关的经验是什么呢?就是接竹杆的逆过程掰竹杆。试想,黑猩猩从小到大,一定有无数次将一根树枝或竹杆掰成两截的经验。这样在它两手各拿起一根竹杆时,可能会联想起它过去掰断一根树枝的情景,于是掰断的逆过程——接上的灵感就产生了。这样为了实现灵感思维,我们加上一条**可逆性原理**: 一个过程在一定条件下是可逆的。过去掰竹杆的经验可表示为:

R_5 : 如果黑猩猩手中拿一根竹杆,则它可能将这根竹杆掰成两截且每只手拿一截。

根据可逆性原理,我们从规则 R_5 产生一条新规则:

R_6 : 如果黑猩猩两手各拿一根竹杆,则它能够将这两根竹杆接到一起成为一根长竹杆。

现在由规则 R_1 — R_6 我们可以推出黑猩猩可以吃到香蕉了。推理使用规则的次序是 R_3 , R_6 , R_1 , R_4 和 R_2 (推理过程留给读者)。

十、结 论

本文讨论了当前国际上有关人工智能可

能性的争论,简要地介绍了作者近年来在思维模拟以及复杂问题求解方面一些探索性工作。这些工作同国际上现存智能系统有着明显的不同: 我们是在思维科学指导下,用多个复杂学习算法联合来模拟高级的思维活动,如直觉、灵感和发现。由于哲学、辩证法和思维科学在我国很普及,这有可能使我们在思维模拟方面保持一定的优势。因此,作者宁愿把本文介绍的工作看做具有中国特色的研究成果之一。

参 考 文 献

- [1] 司马贺, 人类的认识——思维的信息加工理论, 科学出版社, pp10-13
- [2] 赫伯特·A·西蒙, 普罗米修斯还是潘朵拉, 《世界科学》1983, 1
- [3] 爱德华·费根鲍姆, 帕梅拉·麦克达克, 第五代计算机, 人工智能和日本计算机对世界的挑战, 《人工智能对世界的挑战》, 新华通讯社参考消息编辑部, 1984年8月
- [4] 休伯特·德雪福斯, 人工智能的极限——计算机不能做什么, 生活, 读书, 系统的三联书店, pp172-180
- [5] Gerald M. Edelman, Neural Darwinism, The Theory of Neuronal Group Selection, Basic Books, New York, 1987
- [6] J. E. Laird, P. S. Rosenbloom and A. Newell, Chunking in SOAR, The anatomy of a general learning mechanism, Machine Learning 1 (1986) pp11-46
- [7] 洪家荣, THOUGHT——一个能将新手培养成专家的新一代专家系统工具, 全国第二届机器学习研讨会大会报告论文, 1989年7月, 桂林。
- [8] J. R. Hong and C. J. Mao, Thought/KD1: Theory Methodology and Implementation of Knowledge Discovery, Proc. UCAI WORKSHOP on Knowledge Discovery in databases Aug, 1989, Detroit, MI

- [9] 洪家荣、五岩、邓松, 解释学习中的知识发现, 全国第三届机器学习研讨会, 论文1991年7月
- [10] C. J. Mao and J. R. Hong, A Fast Minimization of Boolean Functions with High Optimization, International Conference on Switching Circuits and System录用, Shenzhen, 1991
- [11] J. R. Hong and X. Q. Xu, A Machine Learning Approach to Chinese Hand-Written Character Recognition, Int. Conference on Int. & Orient. Lang. Processing, Nov. 1989
- [12] 叶文, 洪家荣, 符号——神经网络型专家系统LENS的设计与实现, 《中国神经网络首届学术大会论文集》下册, pp933~936, 1990年12月, 北京
- [13] 赵德开, 洪家荣, 黄春湛, ABAS——一个基于联想的类比学习系统, 《第一届中国人工智能联合学术会议论文集》吉林人民出版社, 1990年6月
- [14] J. R. Hong and C. Unrik, A New Similarity-based Learning Algorithm GS and a Comparison with ID₃, Proc. of International Computer Science Conference, 88, Hong Kong, Dec. 1988
- [15] Michalski, R. S. & Stepp, R. E. (1983), Automated construction of classifications, Conceptual clustering versus numerical taxonomy, IEEE Transaction on Pattern Analysis and Machine Intelligence, 5:4, pp396—410
- [16] J. Quinlan, Induction of Decision Trees, Machine Learning, 1:1 (1986)
- [17] J. R. Hong, I. Mozetic, and R. Michalski, AQ₁₅: Incremental Learning of Attribute-Based Descriptions from Examples, The Method and User's Guide, Rept, ISG 86-5, UIUCDCS-F-86-949 Dept. of Comp. Sci. Univ. of Illinois, Urbana, 1986
- [18] J. R. Hong, AE₁: An Extension Matrix Approximate Method for The General Covering Problem, International Journal of Computer and Information Science, Vol. 14, No. 6, 1985
- [19] D. Haussler, Corrigendum, Artificial Intelligence 45 (1—2) 1990
- [20] J. R. Hong and X. Song, SGESA: A Second Generation Expert System Architecture Proc. of International Symposium on New Generation, Computer, Beinne Apr. 1989
- [21] J. R. Hong, Incremental and Constructive Knowledge Discovery, Chapter 10 in book, Knowledge Discovery in Databases, AAAI Press, USA, 1991
- [22] G. DeJong and R. Mooney, Explanation-Based Learning, An Alternative View, Machine Learning 1 (1986)
- [23] 洪家荣, SWT: 一个基于机器学习的布尔代数化简系统, 《电子学报》No. 6, 1989
- [24] J. R. Hong and Carl Unrik, the Extension Matrix Approach to Attribute-Based Learning, In book, Progress in Machine learning, I. Branko and N. Lavroc (EDS.) SIGMA Press, Wilmoslow, England, May 1987
- [25] J. R. 安得森, 《认知心理学》, 吉林出版社