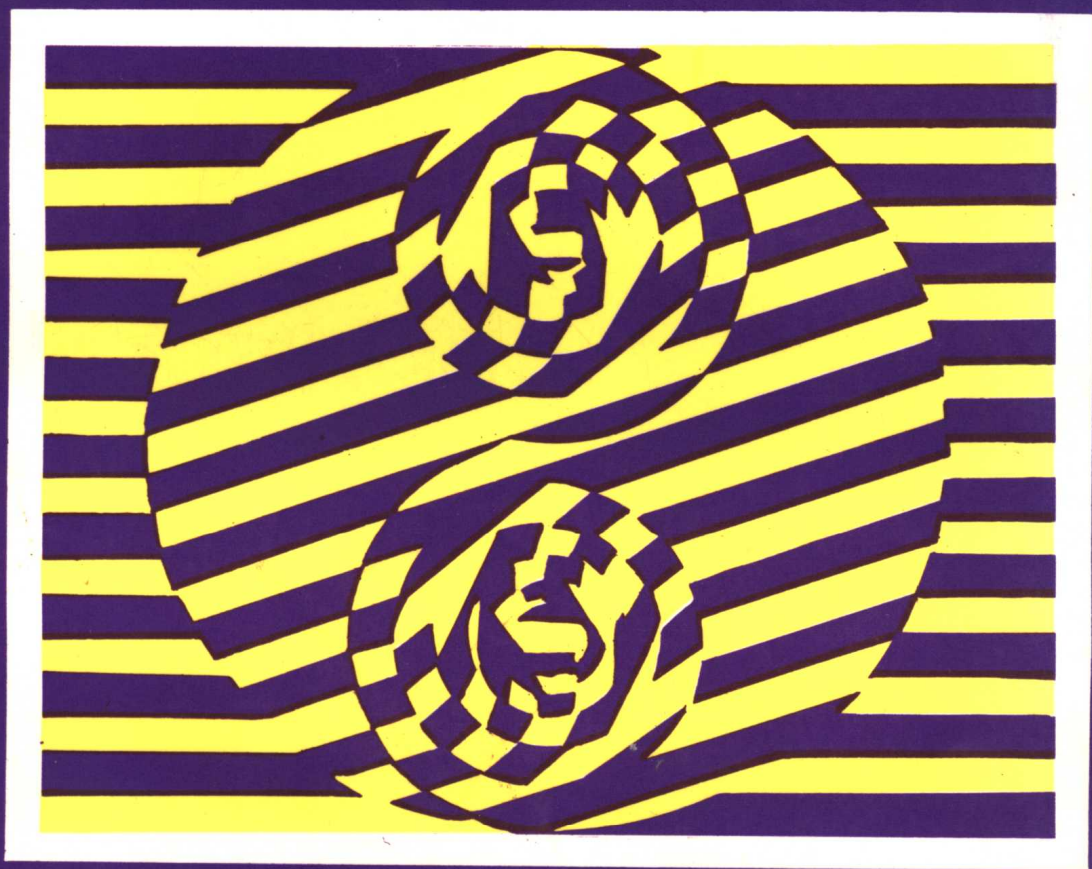


人工智能概论

侯广坤 邓卫安 编著



中山大学出版社

人工智能概论

侯广坤 邓卫安 编著

中山大学出版社

人工智能概论

侯广坤 邓卫安 编著

*

中山大学出版社出版发行

广东省新华书店经销

中山大学印刷厂印刷

787×1092毫米 16开本 18.5印张 45万字
1993年5月第1版 1993年5月第1次印刷

印数:1-1000册

*

登记证号(粤)第11号

ISBN7-306-00630-4

TP·13 定价:8.35元

前 言

随着我国国民经济的发展,对新科学技术的应用要求越来越高。人们对于人工智能的学科特点越来越了解,人工智能的研究和应用也越来越广泛。

艺术界关于人工智能已出现不少的作品,许多有关智能机器人的幻想小说和它的艺术作品已为人们所熟知。事实上,对于任何一门新兴的学科,首先是艺术家和文学家取得成就,其次才是科学家取得成就,最后是哲学家取得成就。对于人工智能的研究也不例外,目前正是自然科学家取得成就的阶段了。但是,由于受到文学家和艺术家的影响,人们一提起人工智能就联想起机器人和它那可怕的艺术形象。虽然机器人,尤其是智能机器人确是人工智能研究的一个重要课题,但机器人的研究不包括全部人工智能的研究,特别是一般机器人(非智能机器人)并没有人工智能的因素,就是智能机器人永远也达不到能对人类造成威胁的程度。因此有必要对人工智能研究的背景和前途,人工智能的最基本概念、原理、内容及主要应用范围作一些介绍。本书就是力图实现这一目的。

我们认为,虽然人工智能的研究比起其它许多学科来都广泛,但有一轰而起的现象。诚然,人们还没有完全认识这个学科时,为了争得学科的地位,往往会在其是否可实现和是否有应用价值上多下一些功夫,因而出现了许多的人工智能的应用系统。特别是70年代,各种专家咨询系统应运而生。虽然国际上的一些专家咨询系统可达到中级专家的水平,但知识表示比较单调,推理机制比较低级,要进一步提高系统的智能水平存在着极大的困难。近年来,人们纷纷感到人工智能的研究缺乏自己的学科基础。那么,人工智能的学科基础是什么呢?有各种不同的看法。由于第一个算得上具有人工智能的课题是数学定理机器证明,而它的基础是一阶逻辑的理论,因而有的人认为人工智能的学科基础是数学,特别是离散数学。由于人的智能是同人脑有直接的联系,人工智能就是模拟人的智能,早期我国称人工智能为智能模拟就是这个原因,因此又有人认为人工智能的学科基础应是神经生理学。由于人工智能的许多应用系统都是建立在计算机上,而且用计算机字符表示知识时充分考虑人的心理因素,因此有的人认为人工智能的学科基础是心理学和计算机科学。不少人认为人工智能是模拟人脑的某些工作,人脑工作最具有同一性的是思维,不论人脑做什么工作,都是思维,因此,现在又有人认为人工智能的学科基础应是思维科学。人工智能是一个综合性的学科,它同前面所提到的学科有着十分密切的关系。一方面,必须对于思维模型和思维的表达有所了解,才能构造模拟人脑活动——思维的理论模型;另一方面,“计算”是人的一种智能活动,计算机是实现人的智能活动的一种机械装置,其中包括硬件和软件,在信息处理层面上,人脑的思维被看成是执行一系列程序。当然人工智能同其它许多学科有着密切的联系,比如,语言学、物理学、电子学、神经生理学、生物化学等。而它的应用更是向各个学科渗透。因此,本书以计算机科学为基础介绍人工智能的基本原理,在介绍人工智能的基本原理之前不介绍人们比较熟悉的学科,只简略地介绍思维科学。

本书根据被称为基于符号主义的人工智能一个方面研究的几个课题,分成若干章进行讨论,是一本入门的教材,其目的是比较广泛地将当前人工智能研究的课题介绍给读者,从而使读者了解人工智能的概况。有些问题的进一步研究将有专著论述。

人工智能研究的长远的目标是制造智能机,为了介绍国际上关于智能机的研制的竞争状况,本书的最后一章介绍日本对于新一代计算机研究的情况。但我们并不是认为日本的计划能够成功,事实会证明日本的计划是不会有什结果的,但它对国际人工智能研究具有一定的推动作用。大多数人认为目前人工智能的研究应当十分重视理论基础的研究,以多种途径达到使机器具有更多的智能的目标。人工智能是年轻的学科,更确切地讲是年轻人的学科,它的真正的突破是下几代甚至是下几十代的事。但事业的开创应从现在做起。

本书是我们和自己的研究生多年教学和研究成果的一些总结。由于我们水平有限,同时目前国际国内对于人工智能的看法各异,在综合国内外人工智能研究者的各种见解和自己教学实践的基础上提出自己的见解,其中有些问题会引起争论,有错误的地方请给予纠正。

作者

目 录

前 言

第一章 绪 论	1
第一节 人工智能的研究目标	1
一、计算理论和思维科学	1
二、现代计算机	4
三、人工智能研究的现实性	5
第二节 当前人工智能研究的状况	7
一、现阶段人工智能研究的几个课题	7
二、人工智能研究发展过程	10
三、人工智能的核心问题	13
第三节 关于思维科学	18
一、当前思维科学研究的概况	18
二、形式思维和推理思维	21
三、形象思维和直觉思维	25
四、灵感思维	32
第二章 一般问题求解	34
第一节 状态空间搜索法	34
一、状态空间的表示	34
二、状态空间搜索算法	36
三、搜索的控制	42
第二节 问题归约法	53
一、算符法	53
二、关键状态和关键操作法	55
第三节 与-或图法	61
一、与-或图解不定积分	61
二、与-或树的广度优先搜索法	65
三、与-或树的深度优先搜索法	67
四、与-或图的最好优先搜索法	70
第四节 博弈树	75
一、归约博弈图	75
二、最小最大值搜索法	76
三、 α - β 搜索法	79
四、A-B型 α - β 搜索法	82
五、向前修剪的搜索法	83

第五节 非确定程序	84
一、状态空间法	85
二、与-或图法	86
第三章 数学定理机器证明	89
第一节 一阶逻辑理论	89
一、一阶逻辑公式的解释	90
二、数学定理的形式表达	93
三、一阶逻辑公式的前束范式	94
第二节 海布兰得定理	96
一、史柯伦(Skolem)标准型	97
二、子句集的海布兰得解释	100
三、语义树	103
四、海布兰得定理	104
第三节 归结原理	105
一、命题逻辑的归结原理	106
二、一阶逻辑的归结原理	107
三、归结原理的完全性	110
四、归结原理应用的例	111
第四节 归结原理的改进	115
一、消去策略	115
二、语义归结	115
三、有序归结	118
四、锁归结	120
五、线性归结	121
第五节 自动推理系统简介	122
一、关于自动推理的基本技术	122
二、自动推理系统的特点	123
第四章 自动程序设计	125
第一节 AP系统的外部表达式	126
一、输入/输出对	126
二、程序跟踪	127
三、形式的表达方式	128
四、自然语言表达方式	129
第二节 AP的基本方法	130
一、定理机器证明的方法	130
二、应用程序变换的方法	131

三、知识工程的方法	132
四、自动数据结构选择	133
五、传统的问题求解方法	133
六、归纳方法	134
第五章 机器人问题求解	139
第一节 关于机器人问题求解的描述	139
一、状态描述及目标描述	139
二、机器人的动作描述	140
三、画面(框架)问题	141
第二节 机器人问题求解方法	142
一、正向搜索系统	142
二、逆向搜索系统	146
第三节 STRIPS 系统	149
一、STRIPS 系统的工作方式	149
二、STRIPS 系统的控制策略	150
三、一个 STRIPS 系统不能求解的问题	153
第四节 用演绎系统生成机器人规则	153
一、GREEN 表示法	154
二、KOWAISKI 表示法	156
第六章 人工智能常用的语言	159
第一节 LISP 语言的特性	159
一、基本概念和基本 LISP 函数	159
二、常用的 LISP 系统函数	164
三、LISP 程序设计	166
四、关于 LISP 程序设计的例	168
第二节 逻辑程序设计语言——PROLOG 语言	173
一、PROLOG 语言基本语法	174
二、基本数据结构	175
三、PROLOG 系统内部的推理机制	177
四、PROLOG 的内部谓词	179
五、PROLOG 输入输出	186
六、PROLOG 程序设计	189
第七章 知识工程概论	195
第一节 知识工程研究的概况	195
一、知识和知识工程	195

二、知识工程主要研究的问题	196
第二节 知识工程中的知识获取的一些问题.....	199
一、知识获取的过程	199
二、事件记忆	200
三、学习	201
四、计算机学习系统	202
五、规则的获取	207
第三节 知识的表示.....	209
一、知识表示的一般要求	209
二、知识表示方法	211
三、知识表示的层次化	219
第四节 知识工程中的产生式推理.....	221
一、推理的产生式规则	221
二、产生式推理的可信度问题	223
三、产生式推理的理论模型	225
四、推理的最优化问题	228
第五节 模糊推理.....	230
一、模糊推理规则	230
二、知识图	233
三、真值问题	236
四、变量问题	239
第八章 自然语言的处理.....	242
第一节 问题的提出.....	242
一、关于“理解”的理解	242
二、为什么语言难于理解	243
三、如何进行语言理解	245
第二节 语法分析.....	246
一、关键字匹配法	248
二、增强式转换网络	249
三、确定性分析	252
第三节 语义分析.....	254
一、语义文法	254
二、格文法	255
三、概念从属	257
第四节 言语理解技术.....	260
一、言语理解中焦点的应用	260
二、故事文法	261

三、理解中手稿的使用	262
四、理解中规划的应用	263
附 录 新一代计算机的研究概况.....	266
一、信息化的社会需要新一代计算机	266
二、各国的五代机研究计划	269
三、日本的第五代计算机系统计划	272
四、对日本 FCGS 计划的分析	277
五、FCGS 计划的进展情况——第一阶段的成果	281
参考文献.....	285

第一章 绪 论

本章我们就什么是人工智能？人工智能研究的背景是什么？人工智能研究的可能性问题怎样？人工智能的课题分类和人工智能的相关学科等问题加以简单的论述。

什么是人工智能？它在当前整个科学界的地位如何？这是一个很难回答的问题。著名的人工智能学者尼尔逊(Nilon)只是在建立智能科学的基础时曾对此学科建立的目的任务提出了看法。他在1974年提出：人工智能是研究智能的科学，即从其它学科的发展过程中，能否在研究和探索人类感觉和思维的最本质活动的基础上创立起智能科学，从而为设计出各种各样的自动机(或智能机—编著者)奠定基础。上述的研究任务是极其艰巨的。他指出：人工智能的研究同宇航，生命奥秘的探索构成当今三个最辉煌的科学事业。据国内外一些专家估计，人工智能的研究在21世纪及今后的若干世纪标志着一个国家的科学技术发展水平。就是现在也是衡量一个国家，一个民族在科学技术上是否处于世界领先地位的主要标志之一。人工智能的研究是一个长远的目标，沿着这个目标有着漫长的道路要走。

但是，对人工智能的研究往往被一些人看得过于简单，提出了一些不切合实际的大目标：可以在较短的时间里达到出神入化的程度，可以对人类智能活动器官结构进行直接的模拟，甚至可以对人脑进行完全的人工复制，有朝一日会给人类造成威胁。这很容易使人们认为人工智能是一种玄学，是20世纪的点金术。还有另一部分人，出于商业的需要，将智能作为标签使用，从而故意提高产品的价值和工作的价值，其实与人工智能毫不相干。一时导致了对人工智能的庸俗化的理解，实质上降低了人工智能研究的意义。

第一节 人工智能的研究目标

人工智能研究的目标是制造智能机，作为智能机器人的“大脑”，以代替人类去完成复杂的工作。那么智能机是否与现代计算机一样呢？在机器(计算机)能否代替人类思维的问题上始终都是有着两种截然不同的意见。一种认为，既然大脑是物质，为什么不能造出机器来模拟它呢？今天不能造出来并不是说明天不能造出来，更不等于永远造不出来。另一种人则认为，机器总是人造的，怎么能达到或超过人的水平呢？计算机的程序也都是人编的，它怎么能超过编程的人呢？机器只能做机械的事情，根本不可能做比如说辩证逻辑之类的思维。

我们很难在这两派意见中偏重于某种意见，但我们可以先讨论一些同这两种意见有关的问题。

一、计算理论和思维科学

在西方有关思维科学的各个领域，一个占统治地位的核心概念是“计算”，一个中心

议题是计算和符号操作在认知思维过程中的作用。思维科学在西方有各种学派,最有影响的是信息处理学派和表达学派,它们有这样那样的分歧,然而在计算的基本观点上却是一致的,他们都把智能和认知看成计算。西方思维科学的两位开创者,尼维尔(Newell)和西蒙(Simon)在1981年就明确声称,无论计算机和人脑在硬件水平上可能是如何地不同,但在计算理论水平上它们都具有产生、操作和处理抽象符号的能力;它们作为认知系统,无论是人脑还是计算机都是物理符号操作处理系统。在这样的计算理论看来,任何一种复杂的精神智力活动,本质上都是以离散符号的处理为核心的计算过程,解释一种心理现象被看成是为它提出某种计算机程序;通用图灵机给出“计算”的最一般的精确定义,精神、智力的任何一种状态都不外乎是图灵机的一种状态,精神、智力的任何活动都是由图灵机定义的离散符号的,可一步一步地实现的计算。

著者在1982年曾提出,计算概念之所以严格是因为它客观地构造一个计算模型——图灵机,客观性在于在计算的意义上所有的计算模型是等价的,模型的建立以及模型的工作同人脑没有直接的联系,因此计算可以用机器实现。

能否像构造计算模型一样来构造思维模型呢?既然计算机和人脑在计算理论上是一致的,那么,我们对计算理论的理解不能停留在目前的水平上,可以用构造图灵机的方法来构造思维模型。

有人曾经用构造计算的理论模型的方法构造了一个思维的数学模型。如果我们暂时只考虑思维的数学模型构造的可能性,而不考虑模型的功能和可实现性,这个模型是很有启发的。

思维被看成是生命的一种信息活动,它实际上是生物个体的各细胞之间在环境影响下的一种相互作用。这样可以先不局限于大脑。从生物学知道,生物个体的所有细胞都是由一个单细胞(受精卵)分裂而成的,不同细胞的基因都是相同的,同一的细胞处于许多不同的状态,而且状态的不同可以表现在形态上,还可以有另外的表现形式。不同的状态之间有许多不同的转换。这些状态在数学模型中用一个状态的有穷集合 Q 来表示,每一个时刻每一个细胞都处在 Q 中的一个状态 q 。从神经生理学可以知道,每个细胞又伸出许多支叉,用突触联向别的细胞。假定每个细胞有 k 个突触,用集合 $K=\{1,2,\dots,k\}$ 来表示每个细胞的 k 个突触,而且它的第 i 个突触所联系的细胞称为它的第 i 个邻居。这样,在每一个时刻,每一个细胞都可以根据自己当前的状态以及它的第 i 个邻居的状态(用数学语言说是根据 Q^{t+1} 中的一个元素)来做如下的几件事情:

(1)确定自己下一个时刻应转换到什么状态。这可以用一个映射(函数) φ 来表示:

$$\varphi: Q^{t+1} \rightarrow Q$$

(2)移动突触。把自己的第 i 个突触(用 K 中的一个元素 i 表示)从自己的第 i 个邻居 B 移到第 i 个邻居的第 j 个邻居 C (用 K 中的一个元素 j 表示)。移动前 C 是它的第 i 个邻居的第 j 个邻居,移动后 C 是它的第 i 个邻居,而 B 可能不是它的邻居了。这可以用下面的映射 ϕ_i 表示:

$$\phi_i: Q^{t+1} \times K \rightarrow K$$

(3)可以把第 i 个突触(用 K 中的一个元素 i 表示)从原来的第 i 个邻居分离,生成一个新细胞 S (称为它的儿子),把自己的第 i 个突触指向 S ,令 S 的突触的全部指向自己的

父亲,并令儿子 S 处于某状态 $q \in Q$ 。这可以用映射 ψ_2 表示:

$$\psi_2: Q^{t+1} \times K \rightarrow Q$$

但因为无论在任何一个突触上,(2)(3)两个动作只能有选择地执行一个,故可以把映射 ψ_1 和 ψ_2 并为一个映射 ψ :

$$\psi: Q^{t+1} \times K \rightarrow K \cup Q$$

Q 中的某个状态称为死亡状态。若细胞 B 进入死亡状态,则 B 和它的突触将消失掉。如果细胞 A 有突触指向 B ,则该突触将自动指向自己。

从以上讨论可以看出,一个生命在开始的时候是一个单细胞,处在 Q 中的某个状态 q_0 。然后一步步分裂,发展成一个互相用突触连在一起的细胞团。在发展的过程中,通过各细胞状态的转换交换信息。这一切都是由基因决定的,具体地说,是由两个函数 φ 和 ψ (称为基因函数)决定的。

生命(包括人脑)从环境感知信息,所以环境影响生命的发展。环境不仅有自己的发展变化规律,而且受生命对它的进行改造的影响。因此对于环境作合适的数学抽象是不容易的。但是,假如能将它抽象为一个不断变化的窗口,用一个函数 $g(i, j, t, \dots, z)$ 来表示。生命的发展过程是基因和环境的综合的结果,也即是先天和后天的综合的结果。在这个模型中,基因函数 φ 和 ψ 以及环境函数 g 完全地确定了生命(包括人脑)的发展过程。

以上的模型是一个抽象的模型,我们先不去考虑这个模型是否合理,也还不知道基因函数和环境函数的性质,就从模型的复杂性来看,应该用并行计算机模拟思维。并且指出:假定生命活动的时间为 T (或者说我们只讨论从时刻 0 到时刻 T 这一段的情况感兴趣),这段时间内活着的细胞最大数为 S 。假定并行计算机的运行时间为 T_1 ,硬件的规模(基本元件的个数)为 S_1 。再设基因的复杂度为 C ,并行计算机的程序复杂度为 C_1 ,那么有如下的定理。

定理 并行计算机可以模拟生命和思维的过程使得以下三式同时满足:

$$T_1 = O(T * \log S^*)$$

$$S_1 = O(S^*)$$

$$C_1 = O(C)$$

从而得到:思维模型和并行计算机模型是相似的。所以,用并行计算机模拟人的思维具有潜在的现实可能性。

从上面的论述说明这样的道理,要计算机实现较高级的思维活动的关键问题,是建立能解释思维的“计算”理论。虽然当前人们在论证这个理论的建立可能性时应用构造图灵机的方法来构造思维的数学模型,但是可以想象,思维的理论模型与图灵机是绝然不同的。

这里所说的计算机可以模拟人,并没有说目前的这种数字式的电子计算机可以模拟人,将来的智能计算机应当是模拟的和数字的相结合。我们不仅要研究逻辑思维,还要研究形象思维;不仅要研究离散系统,还要研究连续系统。作为人脑和智能机共同的“计算”理论也将是极为先进的理论。

二、现代计算机

我们有必要探讨图灵计算概念下所产生的计算机去完成人脑的一切工作的可能性问题。前面所讨论的是一个理想的抽象模型，人脑和计算机在“计算”理论的意义下是一致的说法是基于如下前提的：一是基于目前我们对信息处理的认识；二是基于“计算”理论不断发展。但在目前还是基于现代计算机去进行人工智能的研究的，因此我们根据现代计算机解决问题的能力，在现在的计算概念下，要计算机去解决某种问题，有三个基本前提：

(1) 必须把问题形式化。计算机，至少在它较低的层面上，只能进行有限符号集上的有限长度符号序列的决定性的形式变换。使用计算机时，常常从观念上以及实现手段上加上一些较高的层面。比如，可以不必事先假定符号集是有限的，可以认为符号串是有结构的，等等。但是无论如何放宽，计算总是作符号处理。因此，任何问题要交给计算机去解，必须先建立形式系统，规定所用的符号，规定符号连接成合法序列的有关规则（语法），以及合法符号串如何表示问题领域中的意义（语义解释），然后建立一些规则，说明对这些符号可以进行一些什么样的处理（演算）。于是问题便可以用符号表达出来，问题的解也表现为满足一定条件的符号序列。这样一来，计算机求解的过程就是从表示问题的符号序列出发，按规则进行加工直到得出符合要求的符号序列来（即解）为止。这就叫形式化。

然而，如果计算机能解决人脑能够解决的一切问题，那么，它也应能解决如何把某种问题形式化的问题。但这样一来，就出现一系列的困难。

首先，形式化过程是从非形式化领域向形式化领域转变的过程。如果要计算机来完成这一转变，就得把这个转变形式化，这又包含了如何表示转变的起点——一个非形式化的领域。这就造成一种回归现象（甚至悖论）。

要避免这种回归，必须假定我们有一种包罗万象的先验的形式化系统。但是形式化系统是属理性认识的范畴，不是先验地存在的。换言之，要计算机代替人脑所做的一切工作，就要预先把我们今天尚未认识而明天可能认识的领域预先形式化。这当然是办不到的（且不说大脑的能力也是不停地发展的）。即使能做到了这一点，由于客观世界是无穷无尽的，我们也会遇到一个无限多个符号、无穷多个规则的形式系统。这同计算机资源的有限性矛盾。

能否把问题领域逐个形式化，这又引起一个新的问题，即如何判断某个问题属于何种领域。这是比原来的问题更高层面上的新问题，这样，又会导致无穷的回归。

(2) 计算机要解决已形式化的问题，这类问题还必须是可计算的，即一定要有解题的算法，存在着算法和找出这个算法是两回事，前者是客观的，后者是人脑的功能。要用计算机代替人脑，就要找出一种算法来代替人脑找寻算法，这又遇到回归。

在人工智能系统中，常常使用搜索法。直接搜索某问题的解，在各种形式系统中搜索问题有解或无解的证明，从语义（或解释）的角度来看，这两者是一致的。但在形式系统中的符号处理，它们是不同的两个问题。这是在不同的层面上看待同一个问题。当我们为了从一个层面上向另一个层面过渡时（暂时不去考虑这些层面上的形式系统应如何建立），应在哪一个层面上停下来，是搜索法本身所不能解决的。

(3) 要用计算机实现一个算法以解决某问题，这种问题就要有一个合理的复杂度。这

常常被说成是避免指数爆炸。是否会发生指数爆炸的情况，是问题本身固有的性质所决定的，不是任何巧妙的技术可以绕过的。人工智能目前面对的多数难题都与指数爆炸有关。

由于计算机解题必须具备以上三个条件，所以计算机要完成人脑能做的一切工作，是毫无希望的。

计算机是处理信息的机器，要用它来模拟大脑，最自然的莫过于在信息处理的层面上进行模拟，即模拟大脑的信息处理的功能。但是还会出现双重困难：第一，大脑的功能是否都可以通过信息处理活动来描述；第二，大脑的信息处理活动和思维功能是否直接联系在一起。其实，困难尚不止如此。大脑的功能根本不是只由其生理机制决定性地发展出来的。许多实验材料都证明，人脑的先天能力远比我们设想的要复杂得多（在上小节中的单细胞分裂及有限状态的模型是理想的抽象化模型，同 50 年代有人将图灵机称为理想自动机一样），而这些能力后来在与周围环境的联系中得到了发挥，不适应的部分得到改造，没有用的部分则退化了。这又从另一方面引起双重困难：

第一，大脑模拟程序的初始状态应该是什么？

第二，怎样把大脑模拟程序与周围环境联系起来？

此外，无论用什么方法克服上述困难，总是要有办法模拟感觉器官、运动器官的功能。于是会发现一种两难的处境，如果通过感觉，运动与大脑模拟程序联系的外部世界只是一形式化的模型，那么就无法最终避免把外部世界形式化的问题；如果要与客观世界本身真的联系起来，设法从外部世界提取抽象的信息，这又是一个难以完成的任务。

综上所述，计算机不应该是也不会是最终的智能机器。人工智能要向前迈进，就不应把自己局限于计算机的应用。应该研究思维活动的更深入、更具体的规律，提出新的概念，新的方法，新的机制，比信息处理和图灵机更广泛、更深刻地描述思维的某些功能，并把这些与某些理论的机器模型联系起来，以期最终在工程上得到实现。要实现这个目标，必须对人类的思维模型进行研究。

三、人工智能研究的现实性

上面两小节，一方面说明思维和计算机在“计算”理论是一致的，从而用智能机模拟思维是可能的；另一方面又说明用现代电子计算机模拟人脑是没有希望的，这不是引起矛盾吗？其实不是。这是一个问题的两个方面，一方面说明了人工智能研究的意义及从理论上论述智能机最终目标是可以实现的，另一方面说明目前的计算理论和以它作为基础的计算机，是不能实现这个最终目标的。要想克服这些困难就必须重视人工智能的基础理论的研究，加强思维科学等的研究。

那么，现代人工智能的研究是不是缺乏现实的意义呢？虽然，当前的人工智能的研究不是十分扎实，缺乏自己的理论基础，但从工程上和应用上是有现实意义的。即存在人工智能的系统，它能代替人类智能活动一部分以前只有人才能完成的那些琐碎的、重复性的工作，作为人的智能的一个延伸。并不一定是对人的智能机构器官的模拟，而是用机器模拟人类智能活动的外部行为和智能活动的效果。换言之，人能够用智能活动完成某项工作，达到某效果，那么可以制造这样的机器或编制计算机程序同样可以完成这项工作，得到与人完成这项工作相同或相似的效果。至于人是怎样完成的，机器又怎样去完成，可以

完全不相同。

虽然目前的计算机在对于人工智能的研究上有很大的局限,但是,我们还不得不承认,现代计算机仍然是现在人工智能研究(不论是理论研究和应用研究)的基础和有效的工具。通用的程序内存式计算机实际上就是这样的人工智能系统。众所周知,人类在从懂得使用和制造工具以减轻和代替体力劳动以来,就梦想能创造出能代替和减轻人的脑力劳动的工具,也正因为如此创造了文明。计算机的出现和它的发展无疑为实现这个梦想准备了必要的条件。正因为如此在我们后面的章节中是以计算机作为讨论的基础的。计算是人类思维的一个表现,计算机所作的计算就是完成以前只有人才能完成的智能活动。这是一个直观的理解。此外还有更深一层的意义,就是现代计算机使人们对于人工智能的研究从理想化逐步进入现实化,而且在人工智能研究的过程中促进计算理论和计算机的发展。因此,现代的人工智能研究是工程的研究。

在这里我们首先要回答两个问题:工程上智能的概念是什么?在怎样的情况下才认为机器(计算机)具有智能?

弗莱得(Flud)从工程上理解知识、智能和智能行为,认为知识和智能是两个不同的概念。知识从工程上来讲,是人类在自己发展过程中积累起来的有用的信息,而智能从工程上来讲是人获取、处理和应用这些有用信息的能力。知识的概念是对全人类而言的,而智能是对于具体人或具体一类人而言。那么在上述智能概念下,人的智能行为是人的智能活动通过其器官或工具表现出来的使别人能理解的行为以及为社会所能接受的社会效果。从这里可以看出,这是在信息处理的层次上理解智能行为的。

那么,智能行为怎样表达呢?现在任何一个学科都还未能提出一个人脑思维的内部表达式模型来。如果将人的自然语言理解为人类用自己的动作、语言、文字、图象、图形以及其它一切外部行为来表达自己的思维和传递信息,那么,自然语言就可认为是人类思维的一种外部表达式。我们必须从人类的自然语言中总结某些人类思维的外部表达式,建立思维的外部理论模型,作为完成有局部思维能力的智能机的基础。

计算是人类思维的一种表现,图灵根据人类在作计算的外部表现而建立了计算的理论模型。人在作计算时,用一支笔和一张纸。在人的外部表现出来的无非是如下三个动作:第一,人不断地用笔在纸上写上一些或清除一些符号,从而把计算的问题、中间结果和最后的结果写在纸上;第二,人在计算时根据原来写在纸上的符号和计算的要求不断地移动笔到相应的位置;第三,人在计算的过程中所依赖的法则的外部表现是将所得出来的结果在纸上改变一些符号。

根据以上人在计算时的外部表现,图灵构造了一个计算的理论模型——图灵机。如果一个计算是将符号变换(指令)的规律(按顺序)写在图灵机内,则是程序内存式的计算的理论模型。符号的变换实质上是谓词逻辑的演绎,所以计算机的内部表达式是谓词演算,不管计算机在做如何复杂的信息处理,在其内部所表现的是谓词演算。那么计算行为如何表现呢?我们称它为计算的外部表达式。计算机的外部表达式为形式语言。哥德尔(Godel)编码理论使计算机各层次的语言相互解释(编译系统),形式语言也只是基于谓词演算。计算机的外部表达式和内部表达式是一致的。

图灵成功地建立了计算理论模型后,进而想建立人思维的理论模型。他首先要解决的

是机器具有智能的客观标准,即怎样才算是机器具有智能?这里有著名的图灵试验。图灵试验是1950年由图灵提出来的。试验以如下的方式进行:首先有两个房子住进两个人分别是甲和乙(如图1.1(a)),然后通过传话机分别同外面的一批人对话,例如100人。这批人每人向房子尽量提出一系列问题让房子里的人回答,乙回答这些问题总是讲实话,而甲回答这些问题总是讲谎话。然后要求问话人判别哪个是甲,哪个是乙?经过一个回合这个过程后对判断错误的次数进行统计,通过许多批试验之后得到一个错误的判断的概率 p_0 。然后将甲换成机器,乙仍然是人(如图1.1(b)),重复前面的试验(当然可以是另外一些人来进行试验)。也得出一个判断错误的概率,用 p_1 来表示。如果经过多次的试验都有 $p_0 \geq p_1$,那么我们称机器具有智能。图灵预言,完成图灵试验的机器(智能机)在本世纪末将出现。

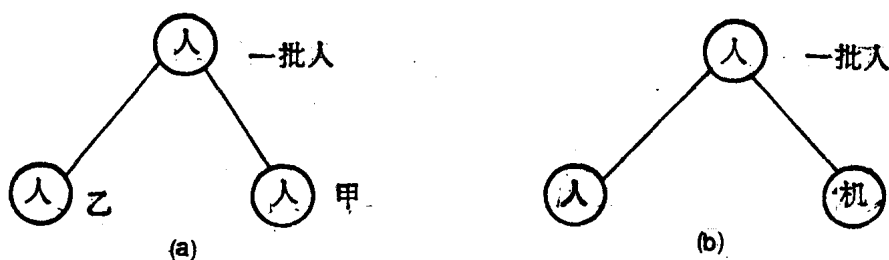


图 1.1 图灵试验

现在看来,图灵的预言是过于乐观。显然图灵试验是受其建立的计算概念影响的。

试验的提出先于智能的工程概念的建立。而且当时计算机还处于初级状态,还不能为人工智能的工程研究提供必要的工具。不过图灵对人工智能的研究提出了一个看法:“人工智能就是研究如何使计算机去做过去只有人才能做的具有智能的工作,人工智能的中心目标就是使计算机更有益于探讨智能的原理。”根据这样的人工智能研究目标,人工智能的研究有着非常现实的意义。要探讨清楚人类智能行为的本质及其内在规律,或许需要几十世纪甚至更长时间的努力,但创造按工程定义的智能机是可以实现的。人工智能研究的现实性还在于近20年来出现了许多人工智能系统,在各个学科和工程领域中得到广泛的应用。

第二节 当前人工智能研究的状况

人们对于人工智能研究的现实性认识,使得人们着手从许多方面开展对于人工智能的研究。

一、现阶段人工智能研究的几个课题

从大的方面,人们从两个方面对人工智能进行研究(为了叙述的方便我们便分别称它们为结构派和非结构派,但并没有学派的意思):

一个称为结构派,又称为联结主义派。在结构派中又有微结构派和宏结构派两种。微结构派又称为神经元派,是从研究神经元如何实现对信息的传输、处理、加工等来研究人的智能行为,提出各种各样的模型。其中最有代表性的是罗申布拉特(Rosenblatt)在1960