

图书在版编目 (CIP) 数据

人工智能导论/郭冰, 曾展挺, 陈挺华主编. —成都: 电子科技大学出版社, 2020. 9
ISBN 978-7-5647-8239-9

I. ①人… II. ①郭… ②曾… ③陈… III. ①人工智能 IV. ①TP18

中国版本图书馆 CIP 数据核字 (2020) 第 158516 号

人工智能导论

REN GONG ZHI NENG DAO LUN

郭 冰 曾展挺 陈挺华 主编

策划编辑 高小红 李述娜
责任编辑 李述娜

出版发行 电子科技大学出版社
成都市一环路东一段 159 号电子信息产业大厦 邮编: 610051
主 页 www.uestcp.com.cn
服务电话 028-83203399
邮购电话 028-83201495

印 刷 广州桐鑫印刷有限公司
成品尺寸 185mm×260mm
印 张 9.75
字 数 213 千字
版 次 2020 年 9 月第 1 版
印 次 2020 年 9 月第 1 次印刷
书 号 ISBN 978-7-5647-8239-9
定 价 32.00 元

版权所有 侵权必究

前言

一、编写理念

21 世纪，我国制造业快速发展，经济保持高增长，各行各业需要大量的专业人才，所以高职院校需根据职业活动所要求的职业能力结构体系来组织教育内容和教学过程，这是高等职业教育不得不完成的社会任务。

高职教育具有应时性的特点，只有适时调整高职教育结构和教育教学内容，才能适应不断变化的社会需要；高职教育也要有一定的前瞻性，才能有生存的土壤；高职教育是教育发展过程中理论和实践结合最密切的教育形式。上述三种情况最直接的体现就是高职教材的编写。从理论上来说，高职教材编写应该是以实训或实践操作为主，理论讲解为辅，而且教材内容应根据社会的变化而实时更新。实际上，到目前为止，我院乃至其他院校所用高职教材大多仍然按学科教育体系编制，这就出现了一个严重影响高职教育健康发展的矛盾。

可见，高职教材的编写及运用对高职学生的职业发展起到了很大的作用。因此，尽管从长远来说，学科教育是职业教育高层次发展之路，但是高职教材编写不能囿于学科教育理念的圈子，必须从高职学生的实际学业能力和社会需求出发，按高职教育各因素的实际状况编写符合校情、学情和师情的高职教材。

秉持这一编写理念，我们编写了《人工智能导论》一书。

二、教材结构

人工智能是一门非常复杂、庞大的科学，其中的机器学习、深度学习等细分学科是当下的研究热点。从 1956 年至今，无数科学家历经艰辛与坎坷，终于让人工智能迈出了属于全人类的一大步。人工智能在数据挖掘、计算机视觉、专家系统等方面的突破性进展，不但为企业带来了良好的经济效益，在智慧城市、智慧交通、智慧医疗等领域也为大众带来了更加舒适、便捷的生活。

在中国，人们对于真正的人工智能知之甚少，然而人工智能已经渗透到了我们生活的方方面面，给我们的生活带来了许多便利和好处。编者编写这本书就是为了能够让学生甚至老百姓了解人工智能的相关知识，让看起来“高大上”的人工智能更加“接地气”。

本书共分为六个章节，第一章讲解人工智能的概念性框架；第二章讲解人工智能的概述；第三章讲解人工智能涉及的相关技术；第四章讲解人工智能中的逻辑；第五章探讨了人工智能与机器人；第六章探讨了人工智能的奇点等人工智能哲学。书中内容尽量避免涉及晦涩难懂的专业知识，只有简单朴实的科普内容。

三、开发团队

本书的开发团队由职业教育的一线教师组成，主编为郭冰、曾展挺、陈挺华；副主编为杨立君、钟慈康、何映丽、甘华生；参与本书编写的还有彭锦强、文祝青、刘萍、曾俊义、简文胜、廖晓晖、张俐、罗忠亮、张意文、黄成葵、吴春忠、蓝祎天（排名不分先后）等。

四、教学建议

本书不仅可作为高职院校的学生公共课教材，还可以作为计算机相关专业和对人工智能基础技术感兴趣的读者的参考书。

本书的第三章涉及人工智能的专业技术，需要有一定的计算机基础，建议计算机相关专业研读，对于非计算机相关专业，建议仅学习第一、二、四、五、六章的内容即可。

为方便各职业院校的师生开展教学，我们组织老师为本书开发了配套的课件资源，如有需要的老师，请邮件联系：262362254@qq.com。由于编写时间仓促，编者水平有限，虽然我们对书稿做了反复校对，但难免存在疏漏的地方，恳请各位专家、老师批评指正，谢谢！

编写组
2020年9月

目录

第一章

○————— 人工智能概念性框架

1.1 人工与智能的概念	1
1.2 图灵测试	3
1.3 人工智能解决的问题	9

第二章

○————— 人工智能概述

2.1 人工智能的诞生	12
2.2 人工智能等级的划分	16
2.3 人工智能的三大学派	20
2.4 人工智能研究内容与应用领域 ...	22

第三章

○————— 人工智能相关技术

3.1 知识表示	27
3.2 搜索策略	34
3.3 确定性推理	38
3.4 不确定性推理	43
3.5 机器学习	49
3.6 数据挖掘	61
3.7 大数据	69
3.8 深度学习	89

第四章

○————— 人工智能中的逻辑

4.1 逻辑和表示	99
4.2 命题逻辑	101
4.3 谓词逻辑	109
4.4 其他一些逻辑	118



CONTENTS

第五章

○----- 人工智能与机器人

5.1 机器人的预言	121
5.2 机器人的历史	124

第六章

○----- 人工智能的奇点

6.1 人工智能的信息奇点	139
6.2 人工智能的摩尔定律	141
6.3 人工智能与人类的未来	142
6.4 人工智能造福人类社会	146

参考文献	148
------------	-----

早期，人类必须通过如轮子、火之类的工具和武器与自然做斗争。13—14 世纪，我国先祖发明的印刷术使人们的生活发生了广泛的变化。19 世纪，工业革命利用自然资源发展电力，这促进了制造、交通和通信的发展。20 世纪，人类通过对天空以及太空的探索，通过计算机的发明及其微型化，进而成为个人计算机、互联网、万维网和智能手机，并持续不断地向前进。过去的 60 年已经见证了一个世界的诞生，这个世界出现了海量的数据、事实和信息，这些数据、事实和信息必须转换为知识。本章主要介绍人工智能学科的概念性框架。

1.1 人工与智能的概念

对人工智能的理解因人而异。一些人认为人工智能是通过非生物系统实现的任何智能形式的同义词。他们坚持认为，智能的实现方式与人类智能实现的机制是否相同是无关紧要的。而另一些人则认为，人工智能系统必须能够模仿人类智能。没有人对是否要研究人工智能或实现人工智能系统进行争论，我们应首先理解人类如何获得智能行为（即我们必须从智力、科学、心理和技术意义上理解被视为智能的活动），这对我们才是大有裨益的。例如，如果我们想要开发一个能够像人类一样行走的机器人，那么首先必须从各个角度了解行走的过程，但是不能通过不断地声明和遵循一套规定的正式规则来完成运动。事实上，人们越要求人类专家解释他们如何在学科或事业中获得了如此表现，这些人类专家就越可能失败。例如，当人们要求某些战斗机飞行员解释他们的飞行能力时，他们的表现实际上会变差。专家的表现并不来自于不断的、有意识的分析，而是来自于大脑的潜意识层面。你能想象高峰时段在高速公路上开车并有意识地权衡控制车辆的每个决策吗？

想象一下力学教授和独轮脚踏车手的故事。当力学教授试图骑独轮车时，如果人们要求教授引用力学原理的同时，成功地骑在独轮车上，那么他极有可能会失败。同样，如果独轮脚踏车手试图学习这些力学知识，并在他展现车技时应用这些知识，那么他也极有可能会失败，也许还会发生悲剧性的事故。关键点是，许多学科的技能和专业知识的知识是在人类的潜意识中发展和存储的，而不是通过明确请求记忆或使用基本原

理来学会这些技能的。

1.1.1 人工的概念

在日常用语中，“人工”一词的意思是合成的（即人造的），这通常具有负面含义，即“人造物体的品质不如自然物体”。但是，人造物体通常优于真实或自然物体。例如，人造花是用丝和线等制成的类似芽或花的物体，它不需要以阳光或水分作为养料，却可以为家庭或公司提供实用的装饰功能。虽然人造花给人的感觉可能不如自然界的花朵，但它看起来和真实的花朵如出一辙。

另一个例子是由蜡烛、煤油灯或电灯泡产生的人造光。显然，只有当太阳出现在天空时，我们才可以获得阳光，但我们随时都可以获得人造光，从这一点来讲，人造光是优于自然光的。最后，思考一下，人造交通工具（如汽车、火车、飞机和自行车）与跑步、步行等相比，在速度和耐久性方面有很多优势。但是，人造交通工具也有一些显著的缺点——地球上无处不在的高速公路，充满了汽车的尾气；人们内心的宁静以及睡眠，常常被飞机的喧嚣打断。如同人造光、人造花和交通一样，人工智能不是自然的，而是人造的。要确定人工智能的优点和缺点，你必须首先理解和定义智能。

1.1.2 智能的概念

智能的定义可能比人工的定义更难以捉摸。斯腾伯格（R. Sternberg）就人类意识这个主题给出了以下有用的定义：

智能是个人从经验中学习、理性思考、记忆重要信息，以及应付日常生活需求的认知能力。

我们都很熟悉标准化测试的问题，比如，给定如下数列：

1, 3, 6, 10, 15, 21, ?, 要求提供下一个数字。

你也许会注意到连续数字之间的差值的间隔为1。例如，从1到3差值为2，从3到6差值为3，以此类推。因此问题正确的答案是28。这个问题旨在衡量我们在模式中识别突出特征方面的熟练程度。我们通过经验来发现模式。

不妨用下面的数列试试你的运气：

a. 1, 2, 2, 3, 3, 3, 4, 4, 4, 4, ?

b. 2, 3, 3, 5, 5, 5, 7, 7, 7, 7, ?

既然已经确定了智能的定义，那么你可能会有以下的疑问。

(1) 如何判定一些人（或事物）是否有智能？

(2) 动物是否有智能？

(3) 如果动物有智能，如何评估它们的智能？

大多数人可以很容易地回答出第一个问题。我们通过与其他人交流（如做出评论或提出问题）来观察他们的反应，每天多次重复这一过程，以此评估他们的智力。虽



然我们没有直接进入他们的思想，但是相信通过问答这种间接的方式，可以为我们提供内部大脑活动的准确评估。

如果坚持使用问答的方式来评估智力，那么如何评估动物智力呢？如果你养过宠物，那么你可能已经有了答案。小狗似乎记得一两个月没见到过的人，并且可以在迷路后找到回家的路。小猫在晚餐时间听到开罐头的声音时常常表现得很兴奋。这只是简单的巴甫洛夫反射的问题，还是小猫有意识地将罐头的声音与晚餐的快乐联系起来了？

脑的质量大小以及脑与身体的质量比通常被视为动物智能的指标。海豚在这两个指标上都与人类相当。海豚的呼吸是自主控制的，这可以说明其脑的质量过大，还可以说明一个有趣的事实，即海豚的两个半脑交替休眠。在动物自我意识测试中，例如镜子测试，海豚得到了很好的分数，它们认识到镜子中的图像实际上是它们自己的形象。海洋世界的游客可以看到，海豚可以玩复杂的戏法。这说明海豚具有记住序列和执行复杂身体运动的能力。使用工具是智能的另一个“试金石”，并且这常常用于将直立人与先前的人类祖先区别开来。海豚与人类都具备这个特质。例如，在觅食时，海豚使用深海海绵（一种多细胞动物）来保护它们的嘴。显而易见，智能不是人类独有的特性。在某种程度上，许多生命形式是具有智能的。

你应该问自己以下问题：“你认为有生命是拥有智能的必要先决条件吗？”或“无生命物体，例如计算机，可能拥有智能吗？”人工智能宣称的目标是创建可以与人类的思维媲美的计算机软件和（或）硬件系统，换句话说，即表现出与人类智能相关的特征。一个关键的问题：“机器能思考吗？”更简单来说，你可能会问人类、动物或机器拥有智能吗？

在这个方面，强调思考和智能之间的区别是明智的。思考是推理、分析、评估和形成思想和概念的工具。并不是所有能够思考的物体都有智能。智能也许就是高效以及有效的思维。许多人对待这个问题时怀有偏见，他们说：“计算机是由硅和电源组成的，因此不能思考。”或者走向另一个极端：“计算机表现得比人快，因此也有着比人更高的智商。”真相很可能存在于这两个极端之间。

正如我们所讨论的，不同的动物物种具有不同程度的智能。我们将阐述人工智能领域开发的软件和硬件系统，它们也具有不同程度的智能。我们对评估动物的智商不太关注，尚未发展出标准化的动物智商测试，但是对确定机器智能是否存在的测试非常感兴趣。

也许拉斐尔（Raphael）的说法最贴切：“人工智能是一门科学，这门科学让机器做人类需要智能才能完成的事。”

1.2 图灵测试

上一节中提出的“你如何确定智能”以及“动物有智能吗”这两个问题已经得到

了解决。第二个问题的答案不一定是简单的“是”或“不是”——一些人比另一些人聪明，一些动物比另一些动物聪明。机器智能也遇到了同样的问题。

1.2.1 人工智能之父——图灵

这个名字无论是在计算机领域、数学领域、人工智能领域还是哲学、逻辑学等领域，都可谓“掷地有声”。阿兰·图灵（Alan Turing）是计算机逻辑的奠基者，许多人工智能的重要方法也源自这位伟大的科学家。1912年6月23日生于伦敦的他，24岁时提出了图灵机理论，31岁参与了 Colossus（第二次世界大战时，英国破解德国通信密码的计算机）的研制，33岁时构思了仿真系统，35岁时提出自动程序设计概念，38岁时设计了“图灵测试”，后来还创造了一门新学科——非线性力学。虽然图灵去世时只有42岁，但在其短暂而离奇的生涯中的那些科技成就，已让后人享用不尽。人们仰望着这位伟大的英国科学家的同时，把“计算机之父”“人工智能之父”“破译之父”等等头衔都加冕在了他身上，甚至认为，他在技术上的贡献及对未来世界的影响几乎可与牛顿、爱因斯坦等巨人比肩。

阿兰·图灵寻求可操作方法来回答智能的问题，欲将功能（智能能做的事情）与实现（如何实现智能）分离开来。

延伸阅读

抽象：

抽象是一种策略，这种策略忽略了对对象或概念的实现（例如内部的工作），这样，你就可以获得更清晰的人造物及其与外部世界关系的图像。换句话说，你可以将这个对象当作一个黑盒子，只关注对象的输入和输出（见图1-1）。

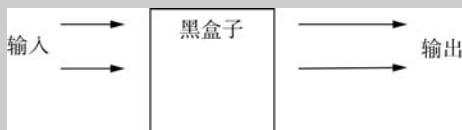


图 1-1 黑盒子的输入和输出

通常，抽象是一种有用而必要的工具。例如，如果你想学习如何驾驶，把车当作一个黑盒子可能是一个好主意。你不必一开始就努力学习自动变速器和动力传动系统，而是可以专注于系统输入，例如油门踏板、刹车、转向信号灯，以及输出，如前进、停车、左转和右转。

数据结构的课程也使用抽象，因此如果想了解栈的行为，你可以专注于基本的栈操作，比如 pop（弹出一项）和 push（插入一项），而不必陷入如何构造一个列表的细节（例如，使用线性链表还是循环链表，或使用链接链表还是连续分配空间）。

1.2.2 图灵测试的含义

阿兰·图灵提出了两个模拟游戏。在模拟游戏中，一个人或实体表现得仿佛是另一个人。在第一个模拟游戏中，一个人在一个中央装有帘子的房间中，帘子的两侧各有一人，其中一侧的人（称为询问者），必须确定另一侧的人是男人还是女人。询问者（其性别无关紧要）通过询问一系列的问题来完成这个任务。游戏假定男性可能会在他的回答中撒谎，而女性总是诚实的。为了使询问者无法从语音中确定性别，通过计算机而不是讲话的方式进行交流，如图 1-2 所示。



图 1-2 第一个图灵模拟游戏

如果在帘子的另一侧是男人，并且他成功地欺骗了询问者，那么他就赢了。图灵测试的原始形式是，一个男人和一个女人坐在窗帘后面，询问者必须正确地识别出其性别（图灵可能得到那个时代流行游戏的启发，发明了这个测试。这个游戏也促使了他进行机器智能测试）。

正如埃里希·弗罗姆（Erich Fromm）所写的：男女平等，但不一定要相同。例如，不同性别的人具有不同的关于颜色和花朵的知识，花在购物上的时间也不同。

区分男女与智能问题有什么关系？图灵认为，可能存在不同类型的思考，了解并容忍这些差异是很重要的。图 1-3 表示了图灵测试的第二个版本。



图 1-3 第二个图灵模拟游戏

第二个游戏更适合人工智能的研究。询问者还是在有帘子的房间里。这一次，帘子后面可能是一台计算机或一个人。这里的机器扮演男性的角色，偶尔会撒谎，但人是一直诚实的。询问者提问，然后评估答案，确定他是和人交流，还是和机器交流。如果计算机成功地欺骗了询问者，那么它就通过了图灵测试，因此也就被认为是具有智能的。

众所周知，在执行算术计算时，机器比人类快很多倍。如果帘子后面的“人”可以在几微秒内得到了三角函数的泰勒级数近似的结果，那么就可以不费吹灰之力辨别出在帘子后面的是计算机而不是人。自然，计算机可以在任意的图灵测试中成功欺骗询问者的机会非常小。为了得到有效的智能“晴雨表”，这个测试要执行许多次。同样，在这个图灵原始版本的测试中，人和计算机都在帘子后面，询问者必须正确地辨别它们。

延伸阅读

图灵测试

没有计算机系统通过了图灵测试。然而，1990年，慈善家 Hugh Gene Loebner 举办了一项比赛，这项比赛旨在实现图灵测试。第一台通过图灵测试的计算机将被授予金牌以及 100000 美元的奖金。同时，每年在比赛中表现最好的计算机将被授予铜牌以及大约 2000 美元的奖金。

在图灵测试中，你会提出什么问题？考虑以下示例。

- 1000017 乘以 725 等于多少？像这样的计算可能不是一个好主意。记住，计算机试图欺骗询问者。计算机可能不会在几分之一秒内做出响应，给出正确答案，它可能会有意地花费更长的时间，也许还会犯错误，因为它“知道”人类不熟悉这些计算。
- 当前的天气情况如何？假设计算机可能不会向窗外看一眼，因此你可能会试着问一下天气。但是，计算机通常连接着万维网，因此在回答之前，它也连接到了天气网站。
- 你害怕死亡吗？因为计算机难以伪装人的情绪，所以你可能会提出这个问题或其他的类似问题：“黑色给你的感觉如何？”或者“坠入爱河的感觉如何？”但是，记住，你是在试图判定智能，人类的情绪也许不是有效的智能“晴雨表”。

图灵预料到会有许多人反对他在最初论文中所提出的“机器智能”的想法。人们相信思考的能力使人变成万物之灵。承认计算机能够思考，这可能挑战了这个仅由人类享有的崇高的栖息地。图灵认为这种顾虑更多是带来安慰，而不是带来反驳。许多人认为，正是人的灵魂让人们可以思考，如果我们创造出拥有这种能力的机器，那么将会篡夺“上帝”的权威。

图灵反驳了这个观点，他提出人们仅仅是准备等待具有灵魂禀赋的容器来执行“上帝”的旨意。最后，我们提到洛甫雷斯伯爵夫人（Lady Lovelace）的反对意见（在文献中她经常被称为第一个计算机程序员）。在评论分析式引擎时，她无比轻松地说“单单这台机器不可能给我们惊喜”。她重申了许多人的信念：一台计算机不能执行任何未预编程的活动。

图灵反对这种意见，说机器一直都让他很惊喜。他坚持认为，这种反对意见的支持者认同人类的智慧可以即时推断给定事实或行动的所有后果。图灵的最初论文在收集上述异议以及其他的反对意见时提到了这些读者。

1.2.3 图灵测试的问题

1. 布洛克对图灵测试的批评

内德·布洛克 (Ned Block) 认为, 英语文本是以 ASCII 编码的, 换句话说, 是用计算机内一系列的 0 和 1 表示的。因此, 一个特定的图灵测试, 也就是一系列的问题和答案, 可以存储为一个非常大的数。例如, 假设图灵测试的长度有一个上限, 在测试中, “Are you afraid of dying? (你害怕死亡吗?)” 开始的前三个字符作为二进制数字存储, 如图 1-4 所示。

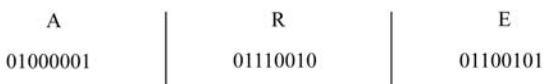


图 1-4 使用 ASCII 代码存储图灵测试的开始字符

假设典型的图灵测试持续一个小时, 在此期间, 测试者大约提出了 50 个问题, 并得到了 50 个答案, 那么对应于测试的二进制数应该非常长。现在, 假设有一个很大的数据库, 储存了所有的图灵测试, 这些图灵测试包含了 50 个或更少的已有合理答案的问题。然后, 计算机可以用查表的方法来通过测试。当然, 一个能够处理这么大量数据的计算机系统还未存在。但是, 如果计算机通过了图灵测试, Block 问: “你认为这样的机器有智能吗? 你感觉舒服吗?” 换句话说, Block 的批评意见是, 可以用机械的查表方法而不是智能来通过图灵测试。

2. 塞尔的批评: 中文室

约翰·塞尔 (John Searle) 对图灵测试的批评更为根本。想象一下, 询问者像人们预料的那样询问问题——但是, 这次用的是中文。另一个房间里的那个人不懂中文, 但是拥有一本详细的规则手册。虽然中文问题以潦草的笔迹呈现, 但是房间里的人会参考规则手册, 根据规则处理中文字符, 并使用中文写下答案, 如图 1-5 所示。



图 1-5 中文室的争论

询问者获得了语法上正确、语义上合理的问题的回答。这意味着房间里的人通晓中文吗? 如果你的回答是“不”, 那么人和中文规则手册的结合通晓中文吗? 答案依然是“不”——房间里的人不是在学习或理解中文, 而仅仅是在处理符号。同样, 计算机运行程序, 接收、处理以及使用符号回答, 而不必学习或理解符号本身的意思是

什么。

塞尔也要求我们设想，如果不是单个人持有规则手册这样的场景：在一个体育馆中，人们互相传递便条。当一个人接到这样的一张便条时，规则手册将确定这个人应该生成一个输出，还是仅仅传递信息给体育馆中的另一个人，如图 1-6 所示。

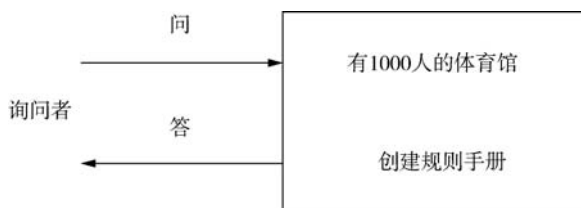


图 1-6 中文室争论的变体

现在，中文的知识存在于何处？属于全体人，还是属于体育馆？思考最后一个例子，描绘出一个确实通晓中文的人的大脑，如图 1-7 所示。这个人可以接收用中文提出的问题，并准确地用中文进行解释和回答。

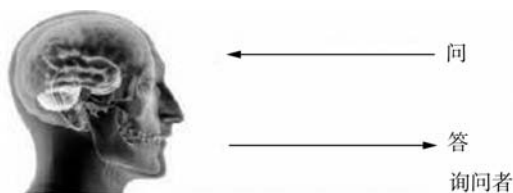


图 1-7 中文说话者用中文接收和回答问题

同样，中文的知识存在于何处？存在于单个神经元中，还是存在于这些神经元的集合中？（它必须存在于某个地方！）

Block 和 Searle 对图灵测试进行批评的关键点在于，图灵测试仅从外部观察，不能洞察某个实体的内部状态。也就是说，我们不应该期望通过将拥有智能的智能体（人或机器）视为黑盒来了解到一些关于智能的新东西。但是，这也并不总是正确的。19 世纪，物理学家欧内斯特·卢瑟福（Ernest Rutherford）通过用 α 粒子轰击金箔，正确地推断出物质的内部状态——它主要由空白空间组成。

他预测，这些高能粒子要么穿过了金箔要么稍微偏转。结果与他的原子轨道理论是一致的：原子由轨道电子包围着的致密核心组成。这是我们当前的原子模型，许多学过高中化学的人对此非常熟悉。Rutherford 通过外部观察成功地了解了原子的内部状态。

总之，定义智能很难。正是由于定义智能以及判定“智能体”是否拥有这一属性很困难，因此图灵开发了图灵测试。在他的论文中，他含蓄地指出，任何能够通过图灵测试的智能体必然拥有“脑能力”来应对任何合理的、相当于人们在普遍意义上接受的人类水平的智能挑战。

1.2.4 人工智能的两种学派

多年来，两种不同却很普遍的人工智能研究的分支得到了发展。一个学派与麻省理工学院相关，这个学派将任何表现出智能行为的系统都视为人工智能的例子。这个学派认为，人造物是否使用与人类相同的方式执行任务无关紧要，唯一的标准就是程序能够正确执行。在电子工程、机器人和相关的领域，人工智能工程主要关注的是得到令人满意的执行结果。这种方法称为弱人工智能。

另一种学派是以卡内基梅隆大学研究人工智能的方法为代表，他们主要关注生物可行性。也就是说，当人造物展现智能行为时，它的表现基于人类所使用的相同方法。例如，考虑一个具有听觉的系统，弱人工智能支持者仅仅关注系统的表现，而强人工智能支持者的目标在于，通过模拟人类听觉系统，使用等效的耳蜗、听力管、耳膜和耳朵其他部分的部件（每个部件都可以在系统中执行其所需执行的任务）来成功地获得听觉。弱人工智能的支持者单单基于系统的表现来衡量系统是否成功，而强人工智能的支持者关注他们所构建系统的结构。

弱人工智能的支持者认为，人工智能研究的存在理由是解决困难问题，而不必理会实际解决问题的方式；强人工智能支持者则坚持认为，单单凭借人工智能程序的启发法、算法和知识，计算机就可以获得意识和智能。

1.3 人工智能解决的问题

随着我们对人工智能了解的深入，理解了它与传统的计算机科学如何截然不同，我们必须回答这样一个问题：“人工智能能解决什么问题？”或者说“人工智能适用解决哪些问题？”大部分人工智能问题有 3 个主要的特征。

- (1) 人工智能问题往往是大型的问题。
- (2) 它们在计算上非常复杂，并且不能通过简单的算法来解决。
- (3) 人工智能问题及其领域倾向于收录大量的人类专门知识，特别是在用强人工智能方法解决问题的情况下。

采用人工智能的方法，一些类型的问题得到了更好的解决，然而涉及简单的决策或精确计算的另一些类型的问题更适合用传统计算机科学的方法来解决。让我们思考几个例子。

- 医疗诊断。
- 用具有条码扫描的收银机来购物。
- 自动柜员机。
- 二人博弈，如象棋和跳棋。

多年来，医疗诊断这个科学领域一直在采用人工智能的方法，并对于来自人工智

能的贡献，特别是利用专家系统（Expert System）的发展乐见其成。建立专家系统的领域一般具有大量的人类专门知识，并且其中存在着大量的规则（一些此类形式的规则：if 条件，then 动作。例如：如果头痛，那么你可以服用两片阿司匹林，并在早晨给我打电话）。这些规则比任何人类大脑能够记忆或希望记忆的规则都多。专家系统算得上一种最成功的人工智能技术，可能生成全面而有效的结果。

人们可能会问：“为什么对专家系统而言，医疗诊断是个好的候选领域？”首先，医疗诊断是一个复杂的过程，有许多可能有效的方法。

诊断涉及基于患者症状和病史以及先例，确定疾病或治疗问题。在大多数情况下，不存在可以识别潜在疾病或病症的确定性算法。例如，MYCIN 是最著名的基于规则的专家系统，用于帮助诊断血液细菌感染。MYCIN 主要用作培训医学生的工具，其规则超过了 400 条。MYCIN 不提供确定的诊断，而是提供最可能存在的疾病的概率，以及诊断正确的程度。人们将开发这些规则的过程称为知识工程（Knowledge Engineer）。知识工程师与领域专家会面，在这个例子中，知识工程师在与医生或其他医疗专业人员的密集访谈过程中收集专家知识，使其变成离散规则的形式。专家系统的另一个特征是，它们可以得出让设计它们的领域专家吃惊的结论。这是由于专家规则可能的排列数量比任何人在他们大脑中记住的都多。用于构建专家系统的好的候选领域具有以下特征。

- 它包含大量的领域特定的知识（关于特定问题的领域知识，例如医疗诊断或人类努力发展的领域，如确保核电站安全操作的控制机制）。
- 它允许领域知识遵循某一种分层次序。
- 它可以开发成为存储了若干专家知识的知识库。

因此，专家系统不仅仅是构建该系统的专家知识的总和。在超市购物，通过扫描条形码将产品扫描到收银机中，这通常不被视为人工智能领域。然而，想象一下，杂货店购物体验发展到了与智能机器互动的阶段。机器可能会提醒购物者要购买什么产品：“你不需要一盒洗衣粉吗？”（因为机器已经知道从某日起，你还没购买过这些产品）系统可以提示消费者购买一些与已选择购买的食物很配的食物。这个系统可以作为平衡营养饮食的食物顾问，并且可以针对个人的年龄、体重、疾病和营养目标进行调整。由于它包含了关于饮食、营养、健康和各种产品的诸多知识，因此就是一种智能系统。此外，它可以做出智能决策，提供建议给消费者。

过去 30 年中使用的自动柜员机（ATM）不是人工智能系统。但是假设这台机器作为总财务顾问追踪了一个人的支出，以及所购买物品的类别和频率，可以解释娱乐、必需品、旅游和其他类别的支出，并就如何有力地改变支出模式提供建议（“你真的需要花那么多钱在高档餐馆吗？”）。那么，我们认为此处描述的自动柜员机就是一种智能系统。

智能系统的另一个例子是下国际象棋。虽然国际象棋的规则很容易学习，但是玩这个游戏达到专家级别可不是件容易的事情。关于国际象棋的书籍比所有其他游戏的



书籍的总和还多。人们普遍接受的是，国际象棋有 10^{42} 种可能合理的棋局（“合理”的棋局与之前给出的 10^{120} “可能”棋局的数目不同）。

这是一个巨大的数字，即使用全世界最快的计算机一起来解决国际象棋游戏（即开发一个程序进行一盘完美的博弈，总是做出最好的移动），这些计算机也不会在 50 年内完成这盘博弈。具有讽刺意味的是，尽管国际象棋是一个零和博弈（意味着最初没有一个玩家有优势），以及是一个具有完美信息的二人博弈（没有涉及机会，任何一方也没有未知的优势因素），但是它依然存在以下问题：完美博弈的结果如何？白方胜利，黑方胜利，还是不分胜负？大部分人认为这会是一个平局。

对于白方而言，最好的第一步是什么？大多数人相信是 1. e4 或 1. d4，这是国际象棋中的概念，即将白方国王前面的兵向前移动两个方格（1. e4），或将白方皇后前面的兵向前移动两个方格（1. d4）。统计数据支持这种观点，但是没有确凿的证据证明这种观点是正确的。

编写一个强大的国际象棋程序（大师级水平以上）目前都基于一个假设，即大师级水平的国际象棋程序需要并展示智能。

在过去的 20 年间，所开发的计算机国际象棋程序，可以击败除了顶尖棋手以外的所有棋手。但是，没有一款计算机程序是官方的世界国际象棋冠军。到目前为止，所有的比赛都相对较短，并利用了人类的弱点（人类会疲劳、焦虑等）。在人工智能领域，许多人强烈认为，程序还没有比所有人都下得好。此外，尽管最近国际象棋程序获得成功，但是这些程序不一定使用了“强人工智能的方法”。

一个真正智能的计算机国际象棋程序不仅会下世界冠军级的国际象棋，还能够解释走子背后的推理。这将需要大量关于国际象棋的知识（特定领域知识），并且程序能够将其作为决策过程的一部分，共享和呈现这些知识。

第二章

人工智能概述

自 1956 年人工智能的概念被第一次提及，人工智能发展至今的近 60 年时间里所取得的极大发展，不容忽视，它引起了众多学科和不同专业背景的学者们日益重视，逐步成为一门广泛的交叉和前沿科学。近些年来，随着现代计算机的不断发展及其在软硬件实现方面取得的长足进步，人工智能正在被应用到越来越广泛的领域中。目前来看，虽然人工智能在发展的过程中存在许多困难和挑战，但随着研究的不断深入，这些困难和挑战终将被战胜，并将推动人工智能继续向前发展。本书主要阐述传统人工智能的基本理论、原理、方法和应用，以及现代人工智能的新的技术与方法。

2.1 人工智能的诞生

人工智能是最近才有的吗？其实人工智能很早就有了，都快 60 多岁了呢。让我们进入历史的长河，一起沿着时间的足迹探究人工智能。

2.1.1 达特茅斯会议

人工智能（Artificial Intelligence）这个名词是 1956 年夏天，在美国东部的达特茅斯会议中初步登场。会议从 8 月开始，是一场为期两个月的会议。对于人工智能的研究者来说，这个会议是一场划时代的会议，会议将“像人一样思考的计算机”称为“人工智能”，于是“人工智能”这个词，诞生了！

这场会议中，有许多著名的计算机科学家参加，包括约翰·麦卡锡，马文·闵斯基，艾伦·纽维尔，如图 2-1 所示。许多研究者当时发布了最新的研究成果。其中马文·闵斯基曾在 1951 年，利用硬件实现了类神经网络。这是世界上第一个可以进行自我学习的网络。

在达特茅斯会议闭幕后的几年中，人工智能取得了令人瞩目的进步。人们制造出了可以解决学校中一般数学问题的机器；基于一个名为“伊丽莎”（Eliza）的程序，人们还研制出了世界上第一个聊天机器人，某些情况下用户甚至会觉得它是有意识的。

冷战期间，超级大国的军事研究机构有近乎无限的研发资金开支，这在一定程度

上使人工智能领域的这些成就成为可能。当时，全美国上下在科技的各个方面与苏联展开了全面竞争，作为竞争重点的计算机领域自然也不例外。



图 2-1 达特茅斯会议的与会者合影

2.1.2 第一次人工智能浪潮

达特茅斯会议之后，20 世纪 50 年代后期到 60 年代，整个人工智能领域流行的用计算机进行演算法，以解决特殊的问题。

以走迷宫为例，如图 2-2 所示，目标就是从迷宫的起点走到终点。人类走迷宫，碰到死路的时候稍微后退，再找其他路径，一步一步向终点靠近。

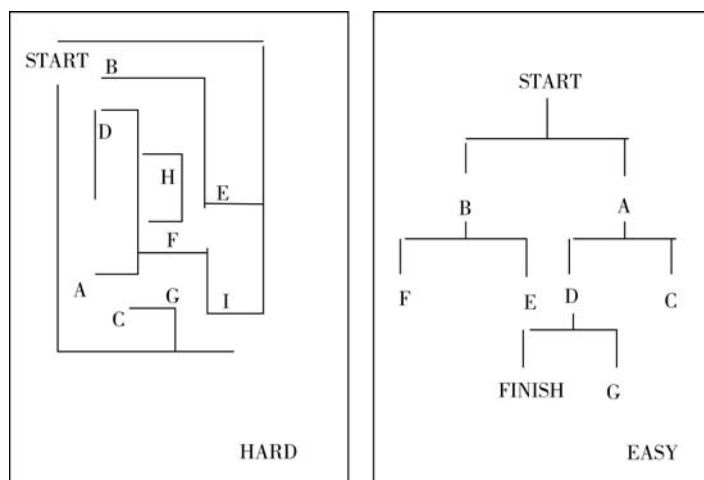


图 2-2 走迷宫示例图

如果让计算机去走迷宫，不会真的按照真实的道路前进，而是从起点开始分类，分成往 A 走的情况和往 B 走的情况等。接着往 A 走碰到的情况，以及往 B 走的情况，进行分类。在不断分类的情况下，最后能找到终点。这就是初期人工智能所使用的方法。

近些年，由于计算机的优异表现，广受媒体注意的国际象棋（IBM 深蓝）等棋类竞赛，用的都是这种演算法。

乍一看这种演算法处理问题似乎不是那么“聪明”，但是由于计算机处理的速度的提升，机器相对于人类慢慢体现出了压倒性优势（围棋除外）。

为什么我一直把围棋要除外呢，因为围棋的走法有 10 的 360 次方，数量要比天上的星星还要多，但是近期 AlphaGo 攻克的围棋，并不是第一次 AI 浪潮的产物。

那么，IBM 深蓝和 AlphaGo 有什么区别呢？深蓝用的演算法目的是证明 IBM 大型服务器的能力，而 AlphaGo 用的是深度强化学习，证明机器可以有智能，如图 2-3 所示。到了 20 世纪 60 年代，以演算法解决棋类竞赛的问题为核心的第一次 AI 浪潮，由于没有解决亟待解决的现实问题，美国政府切断了一切资金的供给，直接导致第一次 AI 浪潮的结束。20 世纪 70 年代末成了人工智能的寒冬。

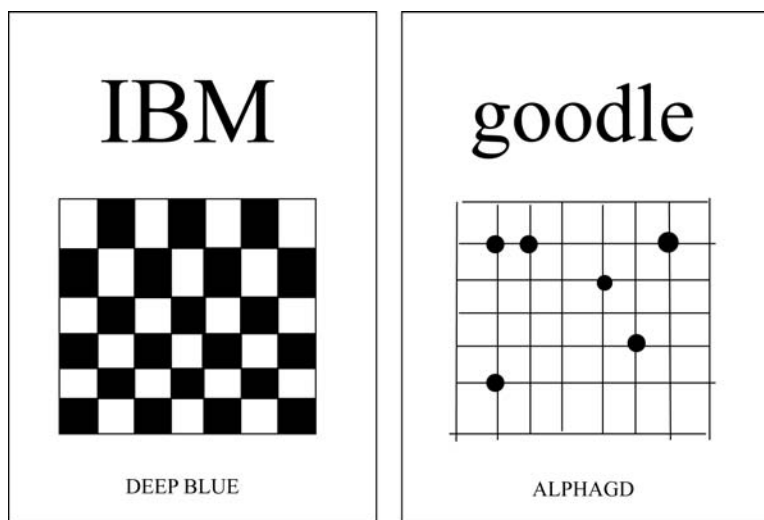


图 2-3 深蓝与阿尔法狗

2.1.3 第二次人工智能浪潮

在第一次 AI 浪潮中，人工智能无法为疾病治疗等人类实际问题做出贡献，使相关研究进入严冬。但是从 20 世纪 70 年代开始，研究人员利用计算机的存储功能，将“知识”存入电脑让它变得更加聪明。

斯坦福大学开发的 MYCIN 就是一个著名的例子。

MYCIN 可以将过去所有病人诊断为细菌感染的症状与其他情况等知识，记录在数

数据库中。当有新的患者出现时，输入患者症状和其他情况，就能够推测患者感染某种细菌的概率，如图 2-4 所示。

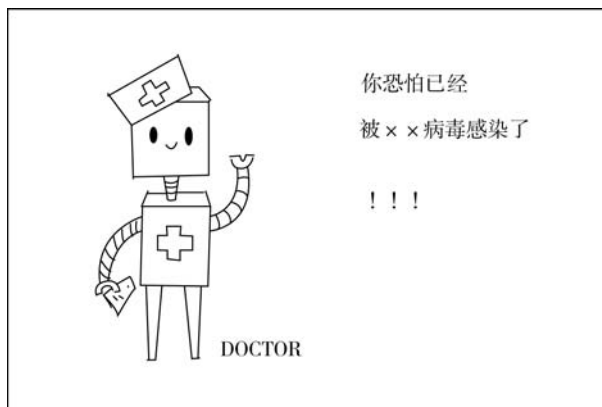


图 2-4 MYCIN 形象画

如“输入：发烧，头疼，口干；输出：这种症状的病人有 69% 的概率感染 * * 细菌”。然而，要让计算机具有这些知识，需要采纳许多专家的相关知识，并进行许多调查研究，以积累资料，相当耗费时间和精力。

为了让这项技术实用化，需要搜集和整理许多领域的资料，其中包括“觉得头重脚轻”这种含糊不清的症状描述。但是想让这些全部记录下来，实在是一件不容易的事情。

这时，有些计划将所有人类具有的知识全部输入计算机。比较著名的是 1984 年美国新创公司发起的 CYC 计划。但是世界的知识实在是太庞杂了。这个计划到目前为止，还在进行中……人工智能面临不知道该如何理解文字意义，以什么方式描述知识，才能让计算机容易处理的问题（我们会在后面详细说明）。于是，必须以人工的方式一条一条输入知识，才能构建人工智能的第二次 AI 浪潮，在 20 世纪 90 年代中期再次进入寒冬。

2.1.4 第三次人工智能浪潮

20 世纪 90 年代中期，互联网和搜索引擎相继诞生，到了 2000 年，随着网站的数量增加，人类的知识、资料在互联网呈现指数增长。到了 2008 年，随着智能手机的兴起和 4G 网络的普及，几乎全世界一半的人都成为了网民，为互联网贡献自己的数据。

能够让计算机自主学习，标志着进入了第三次 AI 浪潮。

从人工智能诞生到现在的历史，可以整理为图 2-5 中的内容。

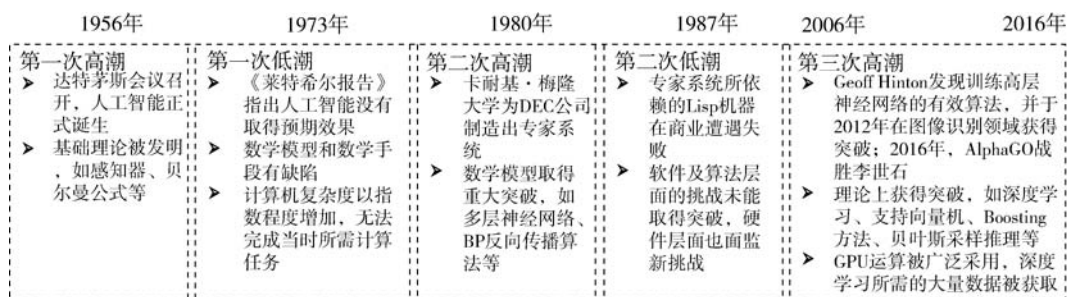


图 2-5 人工智能发展历程

当我们介绍人工智能浪潮的时候, 总会有人问“第三次人工智能浪潮”会很快结束吗? 这倒不一定, 至少从现在看, 人工智能已经是大势所趋, 在未来社会中, 是不可或缺的技术。

2.2 人工智能等级的划分

早在 2004 年, 好莱坞大片《我与机器人》的上映就已经让人们体会到了人工智能的强大 (见图 2-6)。到了 2016 年, 阿尔法狗与李世石的人机“世纪大战”再一次将人工智能推向了各大媒体的头条位置。在百度上搜索“人工智能”, 得到的相关链接有近 7000 万条, 人工智能已经从只有科学家才会讨论的话题, 变成了大街小巷饭后的谈资。那么, 人工智能究竟是什么呢?

人工智能 (Artificial Intelligence) 简称 AI, 从科学的角度来说, 人工智能是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术应用系统的一门科学。这门学科的研究范围包括语言识别、图像识别、专家系统和我们最熟悉的机器人等。

人工智能是一种对人的意识、思维的模拟, 虽然它不是人类的智能, 但却能像人类那样思考甚至超过人类的智能。从字面上, 我们可以把人工智能分为“人工”和“智能”两个部分, 其中“人工”其实并不难理解。在科学界, 虽然表达的方式有所不同, 但是“人工”的意思却相差无几: 系统内的个体根据人为的、预先编排好的规则或计划好的方向运作, 以实现或完成系统内各个体不能单独实现的功能、性能与结果。简单来说, 就是由人工安排好了一切。虽然“人工”好解释, 但是“智能”这个问题却激起了无数科学家激烈的辩论, 什么是智能? 有没有

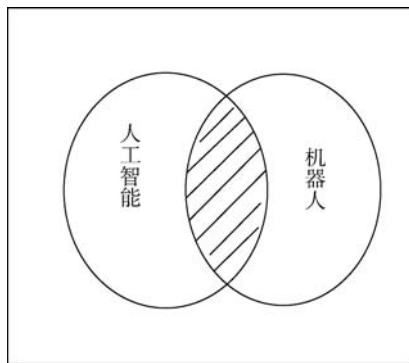


图 2-6 人工智能与机器人

超越人类的智能存在？由于我们对自身智能的理解也非常有限，因此就非常难定义什么是人工制造的“智能”。所以，通常认为：人工智能除了研究人类本身的智能，也研究其他人造系统或者动物的智能。

很多人工智能的专家都希望能够对“人工智能”这一概念做出构建性的阐释，但是以目前的技术水平还很难做到这一点。我们并不是科学家，也没有超越现代技术的智慧，因此不如暂且用一种坊间的说法来理解人工智能：“人工智能就是看起来好像拥有自己的思维。”比如内置人工智能的洗衣机，我们将衣物放进洗衣机内，洗衣机就能“自主”判断出应该加多少水、加多少洗衣液、洗多久等，而我们只需要坐着看一会电视，等着衣服洗好就可以了。

虽然我们暂时没有办法对整个“智能”做出系统性的阐述，但是我们却可以在已知的范围内对“智能”进行概括。在已知的人工智能技术领域，人们普遍认为人工智能可以被分为五个等级。我们目前处于第四级人工智能。

2.2.1 第一级人工智能

第一级人工智能，常见于搭载相对单纯的控制程序，输入与输出一一对应的家电产品（吸尘器、空调、洗机器、冰箱……）。近年来许多家电声称“搭载人工智能”，就是这个等级（不过有些产品具有第二级的人工智能）。

例如，有些人工智能空调主打温度在适当区间，有些人工智能洗衣机可以根据衣服的重量自动调整水量，如图 2-7 所示。不过，他们终究没有超出第一级人工智能的范围。

近期，微软与一家德国家电厂商联合推出一款“自动识别内部物品的冰箱”，引起大家关注。这样的功能已经超越了第一级人工智能，但是说白了，这只是利用冰箱内的摄像头进行图片识别的过程。虽然方便，但是仍然称不上人工智能的概念。



图 2-7 第一级人工智能

2.2.2 第二级人工智能

同样是家电产品，美国麻省理工学院 MIT 人工智能实验室成立的人形机器人公司，最近推出的扫地机器人 Roomba 已经是第二级人工智能，如图 2-8 所示。

有研究者认为，Roomba 的感应行动已和蟑螂有等级的智慧，它可以运用数十个感应器搜集房间的信息，并以每秒 60 次以上的频率判断状况，再从 40 种以上的行动模式汇总选择最适合的行动。

以这类扫地机器人为代表，能够判断、选择行动并执行的系统，即称为第二级人工智能。

让输入和输出变得更加精密的人工智能，就是第二级人工智能。一般应用程序中的棋类游戏，就属于这个等级的人工智能。

第一级和第二级人工智能就在这里结束。接下来就要说明第三级、第四级人工智能，将会涉及“机器学习”领域。机器学习顾名思义，就是让机器（计算机）去学习事物的特征和规则，然后计算机变得越来越聪明。

和第一级人工智能相比，更加复杂



图 2-8 第二级人工智能

2.2.3 第三级人工智能

这个等级的人工智能，可以自主学习，会变得越来越聪明。我们将在以后具体解释第三级人工智能，它具有机器学习功能。

那么什么是机器学习呢？以数据为基础，为各种输入与输出赋予关联性，以这种方法进行学习的算法，就是机器学习。第三级人工智能从 20 世纪 90 年代中期开始普及，一直进入 21 世纪前期。

我们耳熟能详的，藏在搜索引擎背后，会自动在网站上获取大量数据并自动分析判断的人工智能。就是第三级人工智能的典型代表，如图 2-9 所示。

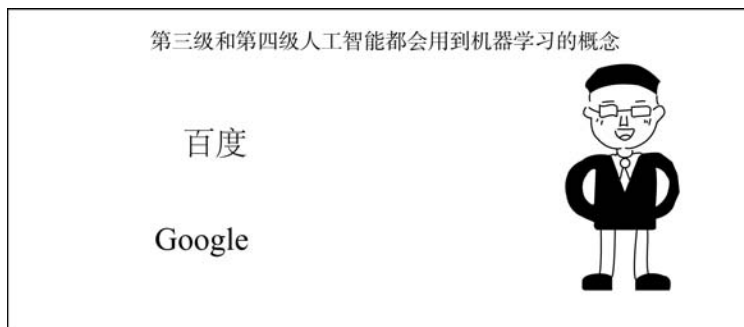


图 2-9 第三级人工智能

在第二级人工智能加入机器学习的方法后，进化成第三级人工智能。而在众多机器学习方法中，深度学习（Deep Learning）能够让计算机自行提取特征（见图 2-10），以进行学习。深度学习可以称得上是第四级人工智能。

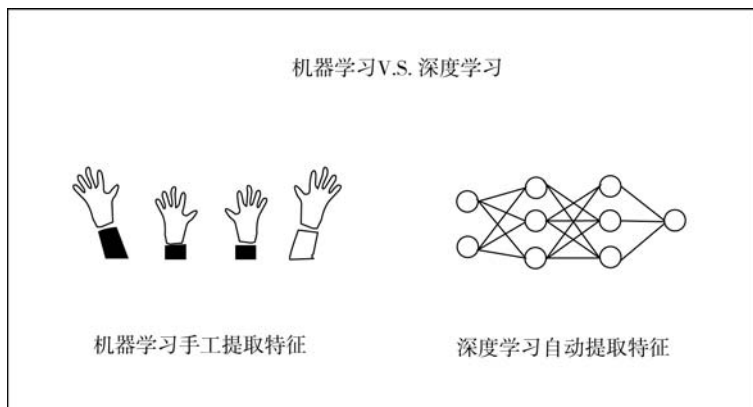


图 2-10 机器学习与深度学习

2.2.4 第四级人工智能

从第一级人工智能进化到第四级，有些 AI 甚至在某些领域超过人类（ImageNet 图片分类、AlphaGo）。这种人工智能能在特定领域发挥自己的作用，也称作“特型化人工智能”。

为了某个目的而特型化的 AI=特型化人工智能。

我们平时听到的，可以下围棋、辨别声音、参加益智问答、自动驾驶（见图 2-11）、对话机器人都是属于第四级特型化人工智能。

人类和第四级人工智能的区别：人类在学习过围棋



图 2-11 自动驾驶

后，或许可以举一反三，将经验运用到其他领域中。但是 AlphaGo 无论有多厉害，它也不具备围棋以外的功能。

2.2.5 第五级人工智能

第五级人工智能就是“泛人工智能”，指的就是类似于哆啦 A 梦等和人类相似的行为（见图 2-12），甚至能够发挥比人类更加优秀的能力。这样的人工智能可以了解人的喜怒哀乐，懂得物体的质感，能够懂得人的情感。

达到这种程度，可以称之为第五级人工智能。然而，一般认为，从第四级进化到第五级，非常困难，我们需要找到另一种特殊方法。



图 2-12 第五级人工智能

人工智能在第四级以前都是人类方便的工具，不过第五级“泛人工智能”不只是具有人类同等的智慧，还有第四级特型化的人工智能的能力，在特定领域超过人类，第五级人工智能可能不再是人类的工具，有可能对人类产生威胁，但是它还远远没有实现。

2.3 人工智能的三大学派

2.3.1 符号主义

符号主义又称逻辑主义、心理学派或计算机学派，其基本原理基于两点：物理符号系统假设和有限合理性原理。

物理符号系统假设是 Newell 和 Simon 在 1976 年提出的。该假设认为：物理符合



系统具有必要且足够的方法来实现普通的智能行为。Newell 和 Simon 把人类的一切精神活动、智能问题归结为计算问题。一个物理符号系统主要有 6 种功能：输入、输出、存储、复制符号、建立符号结构和条件性迁移。其中，建立符号结构即确定符号间的关系，条件性迁移需要依赖已经掌握的符号继续完成后续行为。根据 Newell 和 Simon 提出的这个假设可知，一个物理符号系统若能够完成上述 6 种符号操作就是具备智能的。计算机和人脑都是能操作符号的物理符号系统，因此，计算机和人脑可以进行功能类比。据此，我们可以用计算机的符号操作来模拟人的认知过程，相当于以此来模拟人的智能行为。

有限合理性原理是 Simon 提出的观点。他的观点强调：人类之所以能在大量不确定、不完全信息的复杂环境下解决难题，原因在于人类采用了启发式搜索的试探性方法来求得问题的有限合理解。

符号主义人工智能研究在自动推理、定理证明、机器博弈、自然语言处理、知识工程、专家系统等方面取得了显著成果。符号主义主张从功能上对人脑进行模拟，将问题和知识以逻辑形式表示，采用符号推演的方式来实现推理、学习、搜索等功能。然而，由于符号主义的核心是知识表示、知识推理和知识应用，对于“常识”问题以及不确定事物的表示和处理问题成为符号主义需要解决的巨大难题。

2.3.2 连接主义

连接主义又被称为仿生学派或生理学派，是基于神经元及神经元之间的网络连接机制来模拟和实现人工智能。人类智能的物质基础是神经系统，其基本单位是神经元。这也就是说，连接主义用人工神经网络来研究人工智能。

1943 年，McCulloch 和 Pitts 提出第一个神经元数学模型——MP 模型。该模型从结构上对人脑进行模拟。人工神经网络方法一般先通过神经网络的学习获得知识，再利用知识解决问题。神经网络以分布式方式存储信息，以并行方式处理信息，具有实现自组织、自学习的能力。人工神经网络为人工智能研究提供了一种新思路，在模式识别、机器学习、图像处理等方面很有优势。

局限于研究者对人脑的生理结构和工作机理的认识，目前的人工神经网络仅能近似模拟人脑的局部功能，且不适合模拟人类的逻辑思维过程。因此，单靠连接机制无法解决人工智能的所有问题。

2.3.3 行为主义

行为主义又被称为进化主义或控制论学派，是基于控制论和“感知—动作”控制系统的人工智能学派。其观点是：智能取决于感知和行为，取决于对外界复杂环境的适应。人类智能是经历漫长的演化形成的，真正的智能机器也应该沿着进化的步骤走。

行为主义的基本观点可以概括为：

- 知识的形式化表达和模型化方法是人工智能的重要障碍之一。
- 智能取决于感知和行为，在直接利用机器对环境作用后，以环境对作用的响应为原型。

- 智能行为体现在现实世界中，通过与周围环境的交互表现出来。
- 人工智能可以像人类智能一样逐步进化，分阶段发展和增强。

1991 年，麻省理工学院的 Brooks 教授提出了无须知识表示的智能和无须推理的智能。他认为智能只能在与环境交互过程中表现出来，不应该采用集中式的模式，应该需要具有不同的行为模式与环境交互，并以此产生复杂行为。他成功研制出了一种由相互独立的功能单元组成的 6 足机器虫，该机器虫由避让、前进和平衡等基本功能模块组成分层异步分布式网络。

行为主义实际上是从行为上进行智能模拟，使得机器具有自寻优、自适应、自学习和自组织能力。目前，行为主义在智能控制、机器人领域取得了巨大成就，在未来，控制论、系统工程的思想将进一步影响人工智能的发展。

符号主义、连接主义和行为主义三大思想反映了人工智能的复杂性。这三个学派从不同的角度出发阐释了智能的特性，都取得了显著成就，但同时也存在发展的局限性。目前，对于人工智能的研究仍没有一个统一的理论体系，这种不统一的存在促进了新思潮、新方法的不断涌现，极大地丰富了人工智能的研究。三个学派各有优点也各有局限，可以融合这三个学派，取长补短，为未来的人工智能研究提供新的思路。

2.4 人工智能研究内容与应用领域

2.4.1 机器学习

机器学习研究如何使计算机能够模拟人类的学习功能，从大量的数据中发现规律提取知识，并在实践中不断完善和增强自我。机器学习是机器获取知识的重要途径，它是人工智能研究的核心问题之一，同时也是人工智能理论研究和实际应用中的一个主要瓶颈。Simon 认为，机器学习是系统在不断重复的工作中对本身能力的一种增强和改进，这样系统就能在下一次执行同样任务时拥有更高的执行效率。

传统的机器学习倾向于使用符号而不是数值表示，使用启发式方法而不是算法，传统的机器学习依赖使用归纳法，而不是演绎。按照系统对导师的依赖程度可以将学习方法分类为机械式学习、讲授式学习、类比学习、归纳学习、观察式学习等。近年来又发展了增强学习以及数据挖掘、知识发现等学习方法。

2.4.2 模式识别

模式识别一般指应用电子计算机及外部设备对给定事物进行鉴别和分类，将其归



入与之相同或者相似的模式中。模式识别的一般过程包括对待识别事物采集样本、数字化样本信息、提取数字特征、学习和识别。其核心是特征提取和学习过程。针对不同的识别对象,模式识别可以分为以下几种类型:

- 语音识别,主要研究各种语音信号的识别、翻译以及语音人机交互界面等。语音识别技术发展得比较成熟,计算机可以识别人类语音并将其转化为文本。
- 图形、图像识别,主要研究各种图形、图像的处理和识别技术,包括指纹识别、人脸识别、车牌识别等系统早已经进入实际应用领域。
- 信号识别,主要研究各种传感器信号,例如对雷达、声呐、地震波等信号的识别在军事、地震预测和医学上早有重要应用。
- 染色体识别,主要研究识别染色体以用于遗传因子研究等。

2.4.3 问题求解

人工智能的最早应用实践就是问题求解。该领域最有名的例子就是下棋程序。问题求解是指通过搜索的方法寻找目标解的一个合适的操作序列,并同时满足问题的各种约束。在解决良结构问题时,通常有巨大的搜索空间,理论上可以用穷举法找到最优解,但是实践时却由于时空约束而无法得到最优解。因此,问题求解的核心就是搜索技术。

一般搜索系统包括全局数据库、算子集和控制策略 3 部分。

全局数据库中含有与具体任务相关的信息,这些信息可以用来反映问题的当前状态、约束条件和预期目标。可以根据具体问题采用逻辑公式、数组、矩阵等不同的数据结构。状态分量的选择应该满足必要性、独立性和充分性。各分量不同的取值组合对应不同的状态,只有一些状态是求解问题所需要的。问题本身所具有的约束条件可以排除非法的状态和不可能出现的状态,数据库中只保留问题的初始状态、目标状态和中间状态。

算子集也称操作规则集,用来对数据库进行操作运算。算子一般包括条件和动作两部分。条件给出了适用于算子的先决条件,动作表述了适用算子之后的结果,即引起状态中某些分量的变化。

控制策略可以决定下一步选用哪一个算子以及在何处应用。控制策略通常选择算子集中最有可能导致目标状态或者最优解的算子,运用到当前状态上,不然可能会引起组合爆炸等问题。

搜索技术的最大难点在于寻找合理有效的启发式规则。

2.4.4 神经网络

人工神经网络简称神经网络,其研究始于 1943 年 McCulloch 和 Pitts 提出的神经元的数学模型(MP 模型)。人工神经网络是以对人脑和自然神经网络的生理研究成果

为基础，抽象和模拟人脑的某些机理、机制实现某方面的功能。Nielsen 给出的人工神经网络定义为：“人工神经网络是由人工建立的以有向图为拓扑结构的动态系统，它通过对连续或断续的输入作状态响应而进行信息处理。”

目前，一般的人工神经网络适用于难以应用严格解析方法的场合，包括特征提取、联想记忆、低层次感知、模式分类以及自适应控制等。主要的研究集中在以下几个方面：

- 对人工神经网络的软件模拟和硬件实现的研究。
- 人工神经网络在模式识别、信号处理、知识工程、专家系统、优化组合以及机器人控制等领域的应用。
- 利用神经生理与认知科学研究人类思维和智能推理。
- 深入研究神经网络算法和性能。
- 开发新的神经网络数理理论，如神经网络动力学和非线性神经场等。

2.4.5 专家系统

专家系统是一个具有大量专门知识与经验的程序系统，它采用人工智能技术，根据特定领域中一个或者多个人类专家提供的知识和经验进行推理和判断，模拟人类专家的决策过程，用来解决一些需要专家决定的复杂问题。目前专家系统在医疗诊断、故障诊断、资源勘探、贷款损失评估和教学等领域中得到了广泛应用。

专家系统是一种基于知识的系统，这种系统的设计方法以知识库和推理机为中心展开。知识库包含大量的领域专家提供的知识，推理机能够使用知识库里的知识进行推理并解决实际问题。

专家系统通常由知识库、推理机、综合数据库、人机交互界面、解释器和知识获取等部分构成。其中，知识库用来存放专家提供的知识，知识库中知识的质量和数量决定了专家系统的质量水平。推理机根据当前的已知信息，反复匹配知识库中的规则产生新结论，最后得到问题的求解结果。综合数据库用来存储推理过程中的原始数据、中间结果和最终结论。解释器对结论和求解过程做出解释。人机交互界面用于与用户进行交流。知识获取可以完成对知识的采集和输入到知识库的过程。

专家系统的工作流程如下：用户通过人机界面提交问题和已知条件。推理机将用户输入的信息与知识库中的规则进行匹配，并将中间结论存放在综合数据库中。若得到最终结论则推理结束并输出结果，否则输出推理失败。最后，解释器对结论做出解释。

由于专家系统知识获取需要依赖领域专家，需要大量的人工处理。当这些信息呈现爆炸式增长时，如何自动获取知识成为专家系统的瓶颈问题。这一问题制约了专家系统的发展。

2.4.6 智能控制

智能控制是一类无须或者尽可能少的人的干预就能够独立完成任务的自动控制。许多复杂系统无法建立有效的数学模型和用常规控制理论进行定量计算和分析，而必须采用定量的数学解析法与基于知识的定性方法的混合控制方式。随着人工智能和计算机技术的快速发展，已有可能把控制和人工智能以及系统科学的某些分支结合起来，建立一种适用于复杂系统的控制论和技术。

智能控制涉及的领域很多，目前主要有以下 6 个方面：智能机器人规划与控制、智能过程规划、智能过程控制、专家控制系统、语音控制以及智能仪器。

1998 年 10 月 24 日，宇宙飞船深空一号发射升空，其目的是为了测试 12 项高风险技术。飞行的成功使其使命延长，深空一号最终在 2001 年 12 月 18 日退役。该软件就是称为远程代理的一个人工智能系统，它能够规划和控制宇宙飞船的活动。

2.4.7 计算智能

计算智能主要借鉴仿生学的思想，应属于智能学习部分，目前还没有一个统一的定义，美国科学家贝兹德克（Bezdek）认为不具有计算适应性、计算容错力、接近人的计算速度和近似人的误差率这几个特性，就不能称之为计算智能，其涉及的主要领域有模糊计算、神经计算、进化计算、免疫计算等。就目前而言，在前 3 个领域发展相对比较成熟。

模糊计算是一种对人类智能的逻辑模拟方法，是用模糊逻辑去模拟人类的智能行为。其理论研究是基于 1965 年扎德（Zadeh）教授发表的《模糊集合》开始的。神经计算又叫作神经网络，是计算智能的基础和核心，其主要包括人工神经元的结构和模型，人工神经网络的互联网络结构和系统模型，基于神经网络的联结学习机制等。进化计算是一种模拟自然生物进化过程与机制进行问题求解的自组织、自适应的随机搜索技术。

2.4.8 计算机视觉

计算机视觉是一门用计算机实现或模拟人类视觉功能的新兴学科，其主要研究目标是使计算机具有通过二维图像认知三维环境信息的能力，这种能力不仅包括对三维环境中物体形状、位置、姿态、运动等几何信息的感知，而且还包括对这些信息的描述、存储、识别与理解。目前，计算机视觉已在人类社会的许多领域得到成功应用。例如，在图像、图形识别方面有指纹识别、染色体识别等；在航天与军事方面有卫星图像处理、飞行器跟踪、成像精确制导、景物识别、目标检测等；在医学方面有脏器的图像重建、医学图像分析等；在工业方面有各种监测系统和生产过程监控系统等。

2.4.9 数据挖掘和知识发现

数据挖掘和知识发现起源于 20 世纪 90 年代初期，在早期，它侧重于构建数据仓库，然后运用切片、下钻、上卷等操作从海量数据中发现有用信息。如今的数据挖掘对象包括无结构和半结构文本数据，音频、视频、图像等多媒体数据均可以成为数据挖掘的源数据来源。在数据库的基础上，实现知识发现系统，通过综合运用统计学、粗糙集、模糊数学、机器学习和专家系统等多种学习手段和方法，从大量的数据中提炼出抽象的知识，从而揭示出蕴涵在这些数据背后的客观世界的内在联系和本质规律，实现知识的自动获取。

2.4.10 其他

除了以上研究内容外，人工智能领域的研究内容还包括自然语言处理、分布式人工智能、人工生命、机器人学、智能检索以及最近风生水起的大数据、深度学习等。

3.1 知识表示

知识是人们在改造客观世界的实践中积累起来的认识和经验。这些经验的描述又需要涉及数据和信息的概念。数据是记录信息的符号，是信息的载体和表示。信息是对数据的解释，是数据在特定场合下的具体含义。信息仅是对客观事物的一种简单描述，只有经过加工、整理和改造等工序，并形成对客观世界的规律性认识后才能成为知识。

3.1.1 知识及知识的分类

从不同角度，可以将知识分成不同的类型，如下所示：

- 按知识性质分：概念、命题、公理、定理、规则、方法等。
- 按知识适应范围分：常识性知识，即通用知识；领域性知识，即专业性知识。
- 按知识的作用效果分：事实性知识（又称叙述性知识）；过程性知识；控制性知识（又称元知识或超知识）
- 按知识的确定性分：确定性知识；不确定性知识。
- 按知识的等级分：零级知识、一级知识、二级知识等。
- 按知识的结构分：逻辑性知识；形象性知识。

常识性知识是指通用通识的知识，即人们普遍知道的、适用于所有领域的知识。

领域性知识是指面向某个具体专业的专业性知识，这些知识只有相应专业领域的人员才能掌握并用来求解领域内的有关问题，如领域专家的经验等。

事实性知识也称叙述性知识，是用来描述问题或事物的概念、属性、状态、环境及条件等情况的知识，常以“……是……”的形式出现。事实性知识主要反映事物的静态特征，是知识库中低层的知识。例如，“我们是中国公民”“张三是一名人民教师”“小李和小张是好朋友”等都是事实性知识。

过程性知识是用来描述问题求解过程所需要的操作、演算或行为等的规律性知识，一般由规则、定律、定理及经验构成。

控制性知识是指有关如何选择相应的操作、演算和行动的比较、判断、管理和决策的知识。控制性知识常与元知识有所重叠，所谓元知识是指有关知识的知识，是知识库中的高层知识。对一个大的程序来说，以元知识或元规则形式体现控制知识更为方便，因为元知识存于知识库中，而控制知识常与程序结合在一起出现，从而不容易修改。

确定性知识是指可以给出其真值为“真”或“假”的知识，是可以精确表示的知识。

不确定性知识是指具有不确定特性（不精确、模糊、不完备）的知识。不精确是指知识本身有真假，但由于认识水平等限制却不能肯定知识的真假，可以用可信度、概率等描述。模糊是指知识本身的边界就是不清楚的，例如大、小等，可以用可能性、隶属度来描述。不完备是指解决问题时不具备解决该问题的全部知识，例如医生看病。

零级知识即陈述性知识或事实性知识，用于描述事物的概念、定义、属性，或状态、环境、条件等，回答“是什么”“为什么”。

一级知识即过程性知识或程序性知识，用于问题求解过程的操作、演算和行为的知识，即如何使用事实性知识的知识，回答“怎么做”。

二级知识即控制性知识或策略性知识，是关于如何使用过程性知识的知识，例如推理策略、搜索策略、不确定性的传播策略等。通常把零级知识和一级知识称为领域知识，把二级知识称为元知识（也称超知识）。

逻辑性知识是指反映人类逻辑思维过程的知识，一般具有因果关系或难以精确描述的特点，是人类的经验性知识和直观感觉，例如人的为人处事的经验与风格。

形象性知识是指通过事物的形象建立起来的知识，例如一个人的相貌。

3.1.2 知识的标识

人工智能问题的求解是以知识表示为基础的，如何将已获得的有关知识表示成计算机能够描述、存储、有效地利用的知识是必须解决的问题。知识表示实际上就是对知识的描述，即用一些约定的符号把知识编码成一组能被计算机接受并便于系统使用的数据结构。常用的知识表示方法有一阶谓词表示法、产生式表示法、语义网络表示法、框架表示法、过程表示法、脚本表示法、本体表示法等。

不同的知识表示方法有各自的优缺点，在考虑使用哪一种知识表示方法时应遵循以下相关原则。其一，是否能够充分表示领域性知识，即对症下药。不过有的时候也要根据具体情况，为了克服某种方法的不足，而采用多种知识表示方法结合，如为了弥补框架表示法不能表示过程性知识的不足，通常将其与产生式表示法结合使用。其二，是否具备可利用性。可利用性是指通过使用知识进行推理，以求解现实问题。如果不可利用，则会影响系统的推理效率。其三，是否可以对知识进行组织管理。在一个智能系统初步使用时，可能会发现知识表示方面存在某些问题，此时需要对知识进行增添或删减。长期使用后也需要进行维护。其四，在知识表示方

法选择过程中,最重要的还是要便于理解和实现。如果不能被我们自己理解和实现,那同样还是徒劳的。

3.1.3 谓词逻辑表示法

掌握每一种语言都需要先学习一些单词、语法等,在理解与掌握各种知识表示方法前,也要掌握一些基本的概念、运算等。首先介绍论域的概念,所有讨论对象的全体构成的非空集合称为论域。论域中的元素称为个体。在谓词逻辑中,命题是用谓词来表示的,故在此给出谓词的定义。

定义 1 设 D 是论域, $P: D^n \rightarrow \{T, F\}$ 是一个映射,其中

$$D^n = \{ (x_1, x_2, \dots, x_n) \mid (x_1, x_2, \dots, x_n \in D) \}$$

则称 P 是一个 n 元谓词 ($n=1, 2, 3, \dots$), 记为 $P(x_1, x_2, \dots, x_n)$ 。其中, $x_1, x_2, \dots, x_n$ 个体变元。

定义 2 设 D 是论域, $f: D^n \rightarrow D$ 是一个映射, 则称 f 是 D 上的一个 n 元函数, 记为 $f(x_1, x_2, \dots, x_n)$ 。

其中, $x_1, x_2, \dots, x_n$ 是个体变元。

每一个谓词由谓词名和个体组成。其中, 个体是命题的主语, 用来表示某个独立存在的事物或者某个抽象的概念; 谓词名是命题的谓语, 用来表示个体的性质、状态或个体之间的关系等。

命题可以分为原子命题和复合命题, 后者由前者通过连接词复合而成。下面介绍常用连接词。

设 P, Q 是命题, 常用的连接词如下:

- (1) $\neg$ (否定或非)。对任一命题 P , $\neg P$ 则表示对命题 P 的否定。
- (2) $\vee$ (析取)。复合命题 $P \vee Q$ 表示 P 或 Q 的析取, 即 P 或 Q 。
- (3) $\wedge$ (合取)。复合命题 $P \wedge Q$ 表示 P 和 Q 的合取, 即 P 与 Q 。
- (4) $\rightarrow$ (条件或蕴涵)。它表示“若……则……”的语义。
- (5) $\leftrightarrow$ (双条件)。它表示“当且仅当”的语义。

我们已经知道谓词逻辑表示法, 主要依赖形式逻辑, 利用条件和结论之间的蕴涵关系。谓词逻辑表示法在使用过程中显得接近于自然语言系统且比较灵活, 接近人们对问题的理解, 易于被人们接受。该表示方法还具有模块化特点, 每个知识都是相对独立的, 它们之间不直接发生某种联系, 而是通过添加、删除、修改知识进行。但是, 谓词逻辑表示法也有自己的不足之处, 例如该方法只能表示确定性的知识, 对于模糊概念, 即不确定性知识, 是无法表示的, 然而, 人们生活中大多数实际问题都具有某种不确定性。

由于专业性太强, 具体的谓词逻辑表示法在此不再继续讨论。

3.1.4 产生式表示法

产生式表示法在人工智能中的应用非常广泛，因为它的求解过程和人类求解问题的思维过程很相像，可以用来模拟人类求解问题的思维过程。产生式系统是由美国数学家 E. Post 于 1943 年作为组合问题的形式化变换理论首先提出来的。产生式表示法也常称为产生式规则表示法，许多成功的专家系统都采用了这种知识表示方法。例如，1965 年斯坦福大学设计的第一个专家系统 DENDRAL 就采用了这种知识表示方式。1972 年，Newell 和 Simon 在研究人类的认知模型中开发了基于规则的产生式系统。

产生式系统的知识表示方法主要包括事实和规则两种表示。

1. 事实的表示

事实可以看作一个语言变量的值或断言或者多个语言变量间的关系的陈述句。语言变量的值或语言变量间的关系可以是一个词。对于确定性知识，事实通常用一个三元组来表示，即（对象，属性，值）或（关系，对象 1，对象 2），其中，对象就是语言变量。对于不确定性知识，事实通常用一个四元组来表示，即（对象，属性，值，可信度因子）或（关系，对象 1，对象 2，可信度因子），其中，可信度因子是指该事实为真的可信程度，类似于模糊数学中的隶属程度，可用一个 0~1 的数来表示。

比如：“雪是白的”可表示为（snow, color, white）或（雪，颜色，白）。

“老王和老张是朋友”可表示为（friendship, wang, zhang）或（朋友，老王，老张）。

“李二比王三年龄要大很多”可以用四元组来表示（large than, lier, wangsan, 0.9）。该表示可以理解为李二比王三大很多的可信度为 0.9，即当可信度因子越高，事实就越“真”。

2. 规则的表示

规则表示的是事物间的因果关系。其表现形式为

$$P \rightarrow Q$$

或

$$\text{IF } P \text{ THEN } Q$$

其中， P 表示前提条件， Q 表示所得到的结论或一组操作，这里需要注意的是，要得到结论，需要前提条件必须为真。该规则又称产生式，类似于谓词逻辑中的蕴涵式，但有所区别，区别在于蕴涵式只能表示确定性知识，而产生式不仅可以表示确定性知识，还能表示不确定性知识。例如，在 MYCIN 专家系统中，有这样的产生式：

P ：细菌革式染色阴性

形态杆状

生长需氧

Q ：该细菌是肠杆菌属，可信度为 0.8

在该规则中，所包含的事实就是不确定性知识，当前提条件满足时，就有结论“该细菌是肠杆菌属”，可以相信的程度是 0.8。这个 0.8 是规则强度，这是蕴涵式所不能做到的。

3.1.5 语义网络表示法

1968 年，奎廉 (J. R. Quilian) 在研究人类联想记忆时提出了一种心理学模型——语义网络，认为记忆是由概念间的联系实现的。随后奎廉又把它用作知识表示。1972 年，西蒙 (Simon) 在他的自然语言理解系统中也采用了语义网络表示法。1975 年，亨德里克 (G. G. Hendrix) 又对全称量词的表示提出了语义网络分区技术。

语义网络是一种用语义和语义关系来表示知识且带有方向的网络图。图中包含节点和弧，其中，节点代表语义，即各种概念、事物、属性、状态、动作等；弧代表语义关系，表示两个语义之间的某种联系，弧是有方向的，用来体现节点间的主次关系。为了区分不同语义以及语义关系，每个节点和弧都必须带有标识。

例如：用语义网络表示事实“小燕子是一种鸟”。

分析：事实描述中“小燕子”和“鸟”被定义为不同语义，可用两个节点进行表示，它们之间的语义关系应为“是一种”，则可将事实表示成如图 3-1 所示。

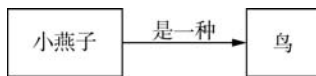


图 3-1 “小燕子是一种鸟”的语义网络

形如图 3-1 所示的结构称为语义基元，所谓语义基元是指语义网络中最基本的语义单元，可用三元组来表示，即

(节点 A，弧，节点 B)

如图 3-2 所示，A 和 B 分别代表节点，R 表示 A 和 B 之间的某种语义联系，该有向图称为基本网元。所谓基本网元是指一个语义基元对应的有向图。

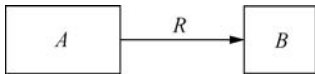


图 3-2 基本网元

当把多个基本网元用相应的语义联系关联在一起时，就可得到一个语义网络，如图 3-3 所示。

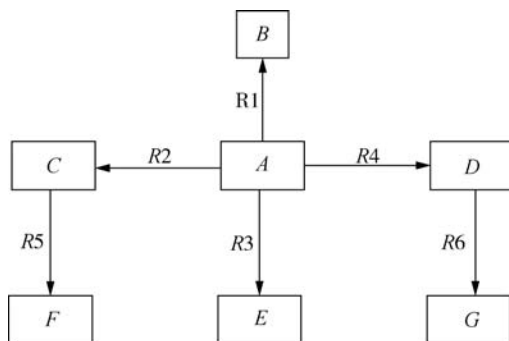


图 3-3 语义网络结构

在语义网络中，每个节点还可以是一个语义子网络，所以，语义网络实质上可以是一种多层次的嵌套结构。

谓词逻辑表示法也可以用语义网络来表示，因为三元组（节点 1，弧，节点 2）可写成 P （个体 1，个体 2），其中个体 1、个体 2 分别对应节点 1、节点 2，而弧及其上标注的节点之间的关系由谓词 P 来体现。

例如，对“小明与小旺是同学”可以表示为三元组（小明，同学，小旺），对应的语义网络如图 3-4 所示。



图 3-4 谓词逻辑表示法的语义网络示例

如果用一阶谓词表示法表示，则可写成 P （小明，小旺），谓词 P 表示小明和小旺为同学关系。

产生式表示法也可以用语义网络来表示。 R_{AB} 表示 A 与 B 之间的语义关系，即“如果……那么……”，如图 3-5 所示。

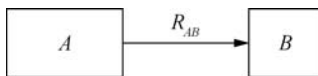


图 3-5 产生式表示法的语义网络

语义网络表示具有以下 5 个特点：

（1）结构性。语义网络是一种结构化的知识表示方法，易于被询问和学习，它能够将事物属性以及事物间的各种语义联系显式地表示出来，下层节点可以继承、新增和变异上层节点的属性，从而实现信息的共享。

（2）自然性。语义网络是一个带有标识的有向图，提供了自然的架构，可直观地把事物的属性及事物间的语义联系表示出来，便于理解，易于转换，符合人们表达事物间关系的习惯，自然语言与语义网络之间的转换也比较容易实现。



(3) 联想性。语义网络本来是作为人类联想记忆模型提出来的,它着重强调事物间的语义联系,体现了人类的联想思维过程。

(4) 非严格性。语义网络没有公认的形式表示体系,它没有给其节点和弧赋予确切的含义。推理过程中有时不能区分物体的“类”和“个体”的特点,因此通过推理网络而实现的推理不能保证其正确性。

(5) 复杂性。语义网络表示知识的手段是多种多样的,这虽然给其表示带来了灵活性,但同时也由于表示形式的不一致使处理变得复杂。

3.1.6 框架表示法

框架表示法是以框架理论为基础的一种结构化知识表示方法。这种表示方法可以表达结构性的知识,能够把知识的内部结构关系以及知识间的联系表示出来,能够体现知识间的继承属性,符合人们观察事物时的思维方式。

框架理论是明斯基于1975年作为理解视觉、自然语言对话及其他复杂行为的一种基础提出来的。他认为,人们对现实世界中各种事物的认识都是以一种类似于框架的结构存储在记忆中的,当遇到一个新事物时,就从记忆中找出一个合适的框架,并根据新的情况对其细节加以修改、补充,从而形成对这个新事物的认识。

框架表示下的知识推理方法与语义网络表示下的知识推理方法类似,即遵循匹配和继承的原则。与语义网类似,框架表示的问题求解系统由两部分构成,一是由框架及其相互关联构成的知识库,二是用于求解问题的解释程序,即推理机。前者的作用是提供求解问题所需要的知识;后者则是针对用户提出的具体问题,运用知识库中的相关知识,通过推理对问题进行求解。

求解问题的匹配推理步骤如下:

(1) 把待求解问题用框架表示出来,其中有的槽是空的,表示待求解的问题,称为未知处。

(2) 与知识库中已有的框架进行匹配。这种匹配通过对相应槽的槽名及槽值逐个进行比较来实现。

(3) 使用一种评价方法对预选框架进行评价,以便决定是否接受它。

(4) 若可接受,则与问题框架的未知处相匹配的事实就是问题的解。

由于框架间存在继承关系,一个框架所描述的某些属性及值可能是从它的上层框架继承而来的,因此两个框架的比较往往牵涉到它们的上层、上上层框架,从而增加了匹配的复杂性。

框架表示法有以下优点:

(1) 结构性。框架表示法最突出的特点是善于表示结构性知识,它能够把知识的内部结构关系以及知识间的特殊联系表示出来。

(2) 深层性。框架表示法不仅可以从多个方面、多重属性表示知识,而且还可以通过ISA、AKO等槽以嵌套结构分层地对知识进行表示,因此能用来表达事物间复杂

的深层联系。

(3) 继承性。在框架系统中，下层框架可以继承上层框架的槽值，也可以进行补充和修改，这样既减少知识冗余，又较好地保证了知识的一致性。

(4) 自然性。框架能把与某个实体或实体集的相关特性都集中在一起，从而高度模拟人脑对实体多方面、多层次的存储结构，直观自然，易于理解。

框架表示法有以下缺点：

(1) 缺乏框架的形式理论。至今还没有建立框架的形式理论，其推理和一致性检查机制并非基于良好定义的语义。

(2) 缺乏过程性知识表示。框架系统不便于表示过程性知识，缺乏如何使用框架中的知识的描述能力。框架推理过程需要用到一些与领域无关的推理规则，而这些规则在框架系统中又很难表达。

(3) 清晰性难以保证。由于各框架本身的数据结构不一定相同，从而使框架系统的清晰性很难保证。

知识表示的方法另外还有过程表示法、脚本表示法、面向对象表示法等，不再一一赘述。

3.2 搜索策略

如何在大量的知识甚至结构不良或非结构化的问题中获取对自己有用的信息，是人工智能中非常重要的一部分。对于这些问题，一般很难获得其全部信息，更没有现成的算法可供使用。因此，根据问题的实际情况，不断寻找可利用知识，从而构造一条代价最小的推理路线，就显得尤为重要。搜索就是要寻找一个操作序列，使问题从初始状态转换到目标状态。这个操作序列就是目标的解。因此，所谓搜索，就是根据问题的实际情况，按照一定的策略或规则，从知识库中寻找可利用的知识，从而构造一条使问题获得解决的推理路线的过程。搜索包含两层含义：一是要找到从初始事实到问题最终答案的一条推理路线，二是找到的这条路线是时间和空间复杂度最小的求解路线。

通常搜索策略的主要任务是确定选取规则的方式。可根据是否使用启发式信息分为盲目搜索和启发式搜索，也可以根据问题的表示方法分为状态空间搜索和与/或树搜索。

盲目搜索是不考虑给定问题所具有的特定知识，系统根据事先确定好的某种固定排序，依次调用规则或随机调用规则，一般统称为无信息引导的搜索策略。由于搜索总是按照预定的控制策略进行搜索，因此这种搜索策略具有盲目性，效率不高，不便于复杂问题的求解。启发式搜索考虑问题领域可应用的知识，动态地确定规则的排序，优先调用较合适的规则，加速问题的求解过程，使搜索朝着最有希望的方向前进，找

到最优解。

状态空间搜索是指用状态空间法来求解问题所进行的搜索。与/或树搜索是指用问题归约法来求解问题时进行的搜索。下面分别介绍状态空间法与问题归约法。

3.2.1 状态空间搜索

状态空间的搜索策略分为盲目搜索和启发式搜索两大类。广度优先搜索、深度优先搜索和有界深度优先搜索都属于盲目搜索策略。

盲目搜索策略的一个共同特点是它们的搜索路线是已经决定好的,没有利用被求解问题的任何特征信息,在决定要被扩展的节点时,并没有考虑该节点到底是否可能出现在解的路径上,也没有考虑它是否有利于问题的求解以及所求的解是否为最优解,下面将大致的介绍一般图搜索与广度优先搜索两个算法。

1. 一般图搜索

一般图搜索是在状态空间中搜索从初始状态到目标状态解答路径的过程。由于问题的状态空间可以用一个有向图来表示,因此状态空间搜索实际上就是对有向图的搜索。从图搜索的角度来看,状态空间搜索的基本思想可以概括为:将问题的初始状态作为当前扩展节点对其进行扩展,生成一组子节点,然后检查目标状态是否出现在这些节点中。如果出现,表明搜索成功,即找到了该问题的解;如果没有出现,则再按照某一种搜索策略从已生成的子节点中选择一个节点作为当前的扩展节点。重复上述过程,直到目标状态出现在子节点中或者没有可供扩展的节点为止。所谓对一个节点进行“扩展”是指对该节点用某个可用操作施加作用,生成该节点的一组子节点。

在开始搜索过程之前,先定义两个数据结构 OPEN 表与 CLOSED 表。OPEN 表用于存放刚生成的节点,对于不同的搜索策略,节点在 OPEN 表中的排列顺序是不同的。例如,对广度优先搜索,节点按生成的顺序排列,先生成的节点排在前面后生成的节点排在后面。CLOSED 表用于存放将要扩展或已扩展的节点。

搜索步骤如下:

第 1 步,把初始节点放入 OPEN 表,并建立目前只包含 $S()$ 的图,记为 G 。

第 2 步,检查 OPEN 表是否为空,若为空则问题无解,失败退出。

第 3 步,把 OPEN 表的第一个节点取出放入 CLOSED 表,并记该节点为节点 n 。

第 4 步,判断节点 n 是否为目标节点。若是,则求得问题的解,成功退出。

第 5 步,考查节点 n ,生成一组子节点。把其中不是节点 n 先辈的那些子节点记作集合 N ,并把这些子节点作为节点 n 的子节点加入 G 中。

第 6 步,针对 M 中子节点的不同情况,分别进行如下处理:

- 对那些未曾在 G 中出现过的 M 成员设置一个指向父节点(即节点 n)的指针,并将它们放入 OPEN 表。

- 对那些先前已在 G 中出现过的 M 成员,确定是否需要修改它指向父节点的

指针。

- 对那些先前已经在 G 中出现并且已经扩展了的 M 成员，确定是否需要修改其后继节点指向父节点的指针。

第 7 步，按某种搜索策略对 OPEN 表中的节点进行排序。

第 8 步，转第 2 步。

2. 广度优先搜索

广度优先搜索又称为宽度优先搜索，是一种先生成的节点先扩展的简单策略。从初始节点 S_0 开始逐层向下扩展，只有当同一层的节点全部被搜索完以后，才能进入下一层继续搜索。

在搜索的过程中，要建立两个数据结构：OPEN 表和 CLOSED 表，其形式分别如表 3-1 和表 3-2 所示。

表 3-1 OPEN 表

节点	父节点编号

表 3-2 CLOSED 表

编号	节点	父节点编号

OPEN 表用于存放刚生成的节点，对于不同的策略，节点在此表中的排列顺序是不同的。CLOSED 表用于存放将要扩展或者已扩展的节点（节点 n 的子节点）。

所谓对一个节点进行扩展，是指用合适的算符对该节点进行操作，生成一组子节点。一个节点经一个算符操作后一般只生成一个子节点，但对一个可使用的节点可能有多组，故此时会生成一组子节点。需要注意的是，在这些子节点中，可能有些是当前扩展节点（节点 n ）的父节点或者祖父节点等，此时不能把这些先辈节点作为当前扩展节点的子节点。

在广度优先搜索策略中，OPEN 表中的节点是按进入的先后排序，先进入 OPEN 表的节点排在前，后进入的节点排在后。

因此，广度优先搜索的基本思想是：从初始节点 S_0 开始，逐层地对节点进行扩展并考察它是否为目标节点，在第 n 层的节点没有全部扩展并考察之前，不对第 $n+1$ 层的节点进行扩展。其搜索过程如下：

第 1 步，把初始节点 S_0 放在 OPEN 表中。

第 2 步，若 OPEN 表为空，则问题无解，退出。

第 3 步，把 OPEN 表中的第一个节点（记为节点 n ）取出放入 CLOSED 表中。

第 4 步，考察节点 n 是否为目标节点。若是，则得到问题的解，成功退出。

第 5 步, 若节点 n 不可扩展, 则转第 2 步。

第 6 步, 扩展节点 n , 将其子节点放入 OPEN 表的尾部, 并且为每一个子节点设置指向父节点的指针, 然后转第 2 步。

由于广度优先搜索总是在生成扩展完 n 层的节点后才转到 $n+1$ 层, 所以总能找到最优解。但是实用意义不大, 广度优先算法的主要缺点是盲目性较大, 尤其是当目标节点距初始节点较远时, 将产生许多无用节点, 最后导致组合爆炸, 尽管耗尽资源, 在可利用的空间中也找不到解。

3.2.2 博弈树的启发式搜索

博弈是一类富有智能行为的竞争活动, 如下棋、打牌、战争等。博弈可以分为双人完备信息博弈和机遇性博弈。所谓双人完备信息博弈, 就是两位选手对垒, 轮流走步, 每一方不仅知道对方已经走过的棋步, 而且还可以估计出对方未来的走步。对弈的结果是一方赢, 另一方输或者是双方和局。这类博弈的实例有象棋、围棋等。所谓机遇性博弈, 是指存在不可预测性的博弈, 如掷币等。对机遇性博弈, 由于不具备完备信息, 因此本节不讨论。本节主要讨论双人完备信息博弈问题。

在双人完备信息博弈的过程中, 双方都希望自己能获胜。因此, 当任何一方走步时, 都是选择对自己最为有利, 而对另一方最为不利的行动方案。假设博弈的一方为 MAX, 另一方为 MIN。在博弈过程中的每一步, 可供 MAX 和 MIN 选择的行动方案都可能有很多种。从 MAX 方的观点来看, 可供自己选择的那些行动方案之间是“或”的关系, 原因是主动权掌握在 MAX 手里, 选择哪个方案完全是由自己决定的; 而那些可供对方选择的行动方案之间是“与”的关系, 原因是主动权掌握在 MIN 的手里, 任何一个方案都可能被 MIN 选中, MAX 必须防止那种对自己最为不利情况的发生。

若把双人完备信息博弈过程用图表示出来, 就可以得到一棵与/或树, 这种与/或树称博弈树。在博弈树中, 那些下一步该 MAX 走步的节点称为 MAX 节点, 而下一步该 MIN 走步的节点称为 MIN 节点。博弈树具有如下特点:

- (1) 博弈的初始状态是初始节点。
- (2) 博弈树中的“或”节点和“与”节点逐层交替出现。
- (3) 整个博弈过程始终站在某一方的立场上。所有能使自己一方胜利的终局都是本原问题, 相应的节点是可解节点; 所有使对方获胜的终局都是不可解节点。例如, 站在 MAX 方, 所有能使 MAX 方获胜的节点都是可解节点, 所有能使 MIN 方获胜的节点都是不可解节点。

下面来看一下极大极小过程

对简单的博弈问题, 可以生成整个博弈树, 找到必胜的策略。但对于复杂的博弈, 如国际象棋, 大约有 10^{120} 个节点, 可见要生成整个搜索树是不可能的。一种可行的方法是使用当前正在考察的节点生成一棵部分博弈树, 由于该博弈树的叶节点一般不是哪一方的获胜节点, 因此, 需要利用估价函数 $f(n)$ 对叶节点进行静态估值。一般来

说, 那些对 MAX 有利的节点, 其估价函数取正值; 那些对 MIN 有利的节点, 其估价函数取负值; 那些使双方均等的节点, 其估价函数取接近于 0 的值。

为了计算非叶节点的值, 必须从叶节点向上倒推。对于 MAX 节点, 由于 MAX 方总是选择估值最大的走步, 因此, MAX 节点的倒推值应该取其后继节点估值的最大值。对于 MIN 节点, 由于 MIN 方总是选择使估值最小的走步, 因此 MIN 节点的倒推值应取其后继节点估值的最小值。这样一步一步地计算倒推值, 直至求出初始节点的倒推值为止。由于我们是站在 MAX 立场上, 因此应该选择具有最大倒推值的走步。这一过程称为极大极小过程。

3.3 确定性推理

在现实生活中, 人们对各种事物进行分析、综合并最后做出决策时, 通常是从已知的事实出发, 通过运用已掌握的知识, 找出其中蕴涵的事实或归纳出新的知识, 这一过程通常称为推理。因此, 从智能技术的角度来说, 所谓推理就是按照某种策略由已知判断推出另一种判断的思维过程。在人工智能系统中, 推理通常是由一组程序来实现的, 人们把这一组用来控制计算机实现推理的程序称为推理机。例如, 在医疗诊断专家系统中, 知识库存储专家经验及医学常识, 数据库存放病人的症状、化验结果等初始事实, 利用该专家系统为病人诊治疾病实际上就是一次推理过程, 即从病人的症状及化验结果等初始事实出发, 利用知识库中的知识及一定的控制策略, 对病情做出诊断, 并开出医疗处方。像这样从初始事实出发, 不断运用知识库中的已知知识逐步推出结论的过程就是推理。

3.3.1 推理的控制策略及其分类

推理过程不仅依赖于所用的推理方法, 也依赖于推理的控制策略。推理的控制策略是指如何使用领域知识使推理过程尽快达到目标的策略。由于智能系统的推理过程一般表现为一种搜索过程, 因此, 推理的控制策略又可分为推理策略和搜索策略。其中, 推理策略主要解决推理方向、冲突消解等问题, 如推理方向控制策略、求解策略、限制策略、冲突消解策略等; 搜索策略主要解决推理线路、推理效果、推理效率等问题。

推理方向用来确定推理的控制方式, 即推理过程是从初始证据开始到目标, 还是从目标开始到初始证据。按照对推理方向的控制, 推理可分为正向推理、逆向推理和混合推理等。无论哪一种推理方式, 系统都需要有一个存放知识的知识库, 一个存放初始证据和中间结果的综合数据库和一个用于推理的推理机。求解策略是指仅求一个解还是求所有解或最优解等。限制策略是指对推理的深度、宽度、时间、空间等进行限制。冲突消解策略是指当推理过程有多条知识可用时, 如何从这多条可用知识中



选出一条最佳知识用于推理的策略。常用的冲突为选择知识的依据。新鲜知识优先,是指把知识前提条件中事实的新鲜性作为选择知识的依据。例如,综合数据中后生成的事实比先生成的事实具有更大的新鲜性。

推理策略可分为正向推理、逆向推理、混合推理等,接下面简要的介绍这几个推理策略。

1. 正向推理

正向推理是一种从已知事实出发,正向使用推理规则的推理方法,也称为数据驱动推理或者前向链推理。其基本思想是:用户需要提供一组初始证据,并将其放入综合数据库。推理开始后,推理机根据综合数据库中的已有事实,到知识库中寻找当前可用知识,形成一个当前可用知识集,然后按照冲突消解策略,从该知识集中选择一条知识进行推理,并将新推出的事实加入综合数据库,作为后面继续推理时可用的已知事实。如此重复这一过程,直到求出所需要的解或者知识库中再无可用知识为止。

2. 逆向推理

逆向推理是一种以某个假设目标作为出发点的推理方法,也称为目标驱动推理或逆向链推理。其基本思想是:首先根据问题求解的要求,将要求证的目标(称为假设)构成一个假设集,然后从假设集中取出一个假设对其进行验证,检查该假设是否在综合数据库中,是否为用户认可的事实。当该假设在数据库中时,该假设成立,若此时假设集为空,则成功退出;若假设不在综合数据库中,但可被用户证实为原始证据时,将该假设放入综合数据库,此时若假设集为空,则成功退出;若假设可由知识库中的一个或多个知识导出,则将知识库中所有可以导出该假设的知识构成一个可用知识集,并根据冲突消解策略,从可用知识集中取出一个知识,将其前提中的所有子条件都作为新的假设放入假设集。重复上述过程,直到假设集为空时成功退出,或假设集非空但可用知识集为空时失败退出为止。

3. 混合推理

由以上讨论可知,正向推理和逆向推理都有各自的优缺点。当问题较为复杂时,单独使用其中的一种,都会影响到推理效率。为了更好地发挥这两种算法各自的长处,避免各自的短处,互相取长补短,可以将它们结合起来使用。这种把正向推理和逆向推理结合起来所进行的推理称为混合推理。

混合推理可以有很多具体的实现办法。例如,可以采用先正向推理,后逆向推理的方法;也可以采用先逆向推理,再正向推理的方法;还可以采用随机选择正向和逆向推理的方法。由于这些方法仅是正向推理和逆向推理的某种结合,因此对这3种情况不再进行讨论。

3.3.2 推理的逻辑基础

在前文中，讨论了知识表示的逻辑基础，已经引入了谓词逻辑的一些简单概念。在此主要讨论推理所需要的一些逻辑基础。

定义：设 D 是谓词公式 P 的非空个体域，若对 P 中的个体常量、函数和谓词按照如下规则赋值：

(1) 为每个个体变量指派 D 中的一个元素。

(2) 为每个 n 元函数指派一个从 D^n 到 D 的映射，其中

$$D^n = \{ (x_1, x_2, \dots, x_n) \mid (x_1, x_2, \dots, x_n \in D) \}$$

(3) 为每个 n 元谓词指派一个从 D^n 到 $\{F, T\}$ 的映射。则称这些指派为 P 在 D 上的一个解释。

为了以后推理的需要，下面先定义谓词公式的永真性、永假性、可满足性与不可满足性。

定义1：如果谓词公式 P 对非空个体域 D 上的任一解释都取得真值 T ，则称 P 在 D 上是永真的；如果 P 在任何非空个体域上均是永真的，则称 P 永真。

由此定义可以看出，要判定一个谓词公式为永真，必须对每个非空个体域上的每个解释逐一进行判断。当解释的个数有限时，尽管工作量大，公式的永真性毕竟还是可以判定的；但当解释个数无限时，其永真性就很难判定了。

定义2：对于谓词公式 P ，如果至少存在 D 上的一个解释，使公式 P 在此解释下的真值为 T ，则称公式 P 在 D 上是可满足的。

谓词公式的可满足性也称为相容性。

定义3：如果谓词公式 P 对非空个体域 D 上的任一解释都取真值 F ，则称 P 在 D 上是永假的；如果 P 在任何非空个体域上均是永假的，则称 P 永假。

谓词公式的永假性又称不可满足性或不相容性。

3.3.3 自然演绎推理

从一组已知为真的事实出发，直接运用经典逻辑中的推理规则推出结论的过程称为自然演绎推理。在这种推理中，最基本的推理规则为三段论推理，包括假言推理、拒取式推理、假言三段论等。

在自然演绎推理中，需要避免两类错误：肯定后件的错误和否定前件的错误。所谓肯定后件的错误是指，当 $P \rightarrow Q$ 为真时，希望通过肯定后件 Q 为真来推出前件 P 为真，这是不允许的。原因是当 $P \rightarrow Q$ 为真及 Q 为真时，前件 P 既可能为真，也可能为假。所谓否定前件的错误是指，当 $P \rightarrow Q$ 为真时，希望通过否定前件 P 来推出后件 Q 为假，这也是不允许的。原因是当 $P \rightarrow Q$ 及 P 为假时，后件 Q 既可能为真，也可能为假。

一般来说，自然演绎推理由已知事实推出的结论可能有多个，只要其中包含了需

要证明的结论，就认为问题得到解。

自然演绎推理的优点是定理证明过程自然，易于理解，并且有丰富的推理规则可用。其主要缺点是容易产生知识爆炸，推理过程中得到的中间结论一般按指数递增，对于复杂问题的推理不利，甚至难以实现。

3.3.4 归结演绎推理

归结演绎推理是一种基于鲁宾孙归结原理的机器推理技术，鲁宾孙归结原理也称为消解原理，是鲁宾孙于1965年在海伯伦（Herbrand）理论上提出的一种基于逻辑反证法的机械化定理证明方法。

在人工智能中，几乎所有的问题都可转化为一个定理证明问题，而定理证明的实质就是要对前提 P 和结论 Q 证明 $P \rightarrow Q$ 永真。由前面章节可知，要证明 $P \rightarrow Q$ 永真，就是要证明 $P \rightarrow Q$ 在任何一个非空的个体域上都是永真的。这将是非常困难的，甚至是不可能实现的。为此，人们进行了大量的探索，后来发现可以采用反证法的思想，把关于永真性的证明转化为关于不可满足性的证明，即要证明 $P \rightarrow Q$ 永真，只要能够证明 $P \wedge \neg Q$ 为不可满足即可，这正是归结演绎推理的基本出发点。

由此定理可知，要证明一个谓词公式是不可满足的，只要证明其相应的标准子句集是不可满足的就可以了。至于如何证明一个子句集的不可满足性，由鲁宾孙归结原理来解决。

鲁宾孙归结原理

鲁宾孙归结原理通过对子句集的子句做逐次归结来证明子句集的不可满足性，它是对定理自动证明的一个重大突破。

由谓词公式转化为子句集的方法可以知道，在子句集中子句之间是合取关系。其中，只要有一个子句为不可满足的，则整个子句集就是不可满足的。另外，前面已经指出空子句是不可满足的。因此，一个子句集中如果包含空子句，则此子句集就一定是不可满足的。

鲁宾孙归结原理就是基于上述认识提出来的。其基本思想是：首先把欲证明问题的结论否定，并加入子句集，得到一个扩充的子句集 S' 。然后设法检验子句集 S' 是否含有空子句，若含有空子句，则表明 S' 是不可满足的；若不含有空子句，则继续使用归结法，在子句集中选择合适的子句进行归结，直至导出空子句或不能继续归结为止。鲁宾孙归结原理分为命题逻辑归结原理和谓词逻辑归结原理。

命题逻辑的归结：归结推理的核心是求两个子句的归结式，因此需要先讨论归结式的定义和性质，然后讨论命题逻辑的归结过程。

谓词逻辑的归结：在谓词逻辑中，由于子句集中的谓词一般都含有变元，因此不能像命题逻辑那样直接消去互补文字，而需要先用一个最一般合一对变元进行代换，然后才能进行归结。可见，谓词逻辑的归结要比命题逻辑的归结复杂一些。

归结演绎推理实际上就是从子句集中不断寻找可进行归结的子句对，并通过对这

些子句对的归结，最终得出一个空子句的过程。由于事先并不知道哪些子句对可进行归结，更不知道通过对哪些子句对的归结能尽快得到空子句，因此就需要对子句集中的所有子句逐对进行比较，直到得出空子句为止。这种盲目地全面进行归结的方法不仅会产生许多无用的归结式，更严重的是会产生组合爆炸问题。因此，需要研究有效的归结策略来解决这些问题。

目前，常用的归结策略可分为两大类，一类是删除策略；另一类是限制策略。删除策略是通过删除某些无用的子句来缩小归结范围；限制策略是通过参加归结的子句进行某些限制来减少归结的盲目性，以尽快得到空子句。

3.3.5 基于规则的演绎推理

规则是一种比较接近于人们习惯的问题描述方式，按照这种问题描述方式进行求解的系统称为基于规则的系统，或者叫作基于规则的演绎系统。规则演绎系统的推理可分为正向演绎推理、逆向演绎推理和双向演绎推理 3 种形式，在此主要讨论前面的两种。

规则正向演绎推理

规则正向演绎推理和前面所讨论过的正向推理相对应。它是从已知事实出发，正向使用规则，直接进行演绎，直至达到目标位置的一种证明方法。一个直接演绎系统不一定比反演系统更有效，但其演绎过程容易被人们理解。

在一个基于规则的正向演绎系统中，其前提条件中的事实表达式是作为系统的初始综合数据库来描述的。对事实的化简，只需转换成不包含蕴涵符号 $\rightarrow$ 的与/或形式表示即可，而不必化为子句集。把事实表达式化为非蕴涵形式的与/或形的主要步骤如下：

- (1) 利用规则 $P \rightarrow Q \Leftrightarrow \neg P \vee Q$ 消去蕴涵符号。实际上，在事实表达式中很少有蕴涵符号 $\rightarrow$ 出现，因此可把蕴涵式表示成规则。
- (2) 利用摩根定律及量词转化律把 $\neg$ 移到紧靠谓词的位置，直到每个否定符号的辖域最多只含有一个谓词为止。
- (3) 对所得到的表达式进行前束化。
- (4) 对辖域内的变元进行改名和标准化，使不同量词约束的变元有不同的名字，并利用 Skolem 函数消去存在量词。
- (5) 消去全称量词，而余下的变元都被认为是全称量词量化的变元。
- (6) 对变元进行换名，使得各主合取式之间的变元名互不相同。

规则正向演绎推理过程是从已知事实出发，不断运用 F 规则，推出欲证明的目标公式的过程。即先用与/或树把已知事实表示出来，然后再用 F 规则的前件和与/或树的叶节点进行匹配，并通过一个匹配弧把匹配成功的 F 规则加入到与/或树中，依次使用 F 规则，直到产生一个含有以目标节点为终止节点的解图为止。

规则逆向演绎推理

基于规则的逆向演绎推理是从目标公式的与/或树出发，逆向使用规则（ B 规则）



对目标表达式的与/或树进行变换,直到得出含有事实节点的一致解图为止。

逆向系统中的目标公式采用与/或形表示,而不像正向系统那样需要化成子句形式。在把任意目标公式化为与/或形表示时,采用正向系统中对事实表达式处理的对偶形式。即要用存在量词约束变元的 Skolem 函数来替换由全称量词约束的相应变元,消去全称量词,再消去存在量词,并进行变元换名,使主析取元之间具有不同的变元名。

规则逆向演绎推理过程

规则逆向演绎推理过程是从目标公式的与/或树出发,不断用 B 规则的右部和与/或树的叶节点进行匹配,并将匹配成功的 B 规则用最一般合一匹配弧加入到与/或树中,直到产生某个终止在事实节点上的一致解图为止。所谓一致解图是指在推理过程中所用到的置换应该是一致的。

3.4 不确定性推理

在现实生活中遇到的问题通常都具有不确定性,能够进行精确描述的问题只占较少的一部分。对于这些不确定性的问题,若采用前述确定性推理方法显然是无法解决的。因此,为满足现实世界问题求解的需求,人工智能需要研究不确定性推理方法。可以将知识分成确定性知识和不确定性知识,不确定性是智能问题的一个本质特征,是建立在不确定性知识和证据的基础上的推理。而知识的不确定性主要体现在两个方面:随机性和模糊性。

随机性主要由概率知识来体现,因此研究随机性即通过概率和马尔可夫链等知识进行。模糊性与随机性一样,在生活中几乎无处不在,我们常说的“高、矮、长、短、大、小”都体现了一种模糊性现象,研究模糊性则通过模糊集理论,该理论以及模糊推理设计面叫专业,在此不再讨论。

采用不确定性推理是客观问题的需求,其原因主要包括以下几个方面。

(1) 所需知识不完备、不精确。在很多情况下,解决问题所需要的知识往往是不完备、不精确的。所谓知识的不完备是指在解决某一问题时,不具备解决该问题所需要的全部知识。例如,医生在看病时,一般是从病人的部分症状开始诊断的。所谓知识的不精确是指既不能完全确定知识为真,又不能确定知识为假。例如,专家系统中的知识多为专家经验,而专家经验又多为不确定性知识。

(2) 所需知识描述模糊。知识描述模糊是指知识的边界不明确。例如,平常人们所说的“很好”“好”“比较好”“不很好”“不好”“很不好”等概念,其边界都是比较模糊的。那么,当用这类概念来描述知识时,所描述的知识当然也是模糊的。例如“如果王刚这个人比较好,那么我就把他当作好朋友”所描述的就是比较模糊的知识。

(3) 多种原因导致同一个结论。在现实世界中,可由多种不同原因导出同一结论的情况有很多。例如,引起人体低烧的原因至少有几十种,医生在看病时只能根据病

人的症状，低烧的持续时间以及病人的体质、病史等做出猜测性的推断。

(4) 解决方案不唯一。现实生活中存在的问题一般都存在着多种不同的解决方案，并且这些方案之间又很难绝对地判断其优劣。对于这种情况，人们往往是优先选择主观上认为相对较优的方案，这也是一种不确定性推理。

总之，在人类的知识和思维行为中，确定性只是相对的，而不确定性才是绝对的。人工智能要解决这些不确定性问题，必须采用不确定性的知识表示和推理方法。

3.4.1 贝叶斯分析方法

贝叶斯分析方法 (Bayesian Analysis) 是贝叶斯学习的基础，它提供了一种计算假设概率的方法，这种方法是基于假设的先验概率、给定假设下观察到不同数据的概率以及观察到的数据本身而得出的。其方法为，将关于未知参数的先验信息与样本信息综合，再根据贝叶斯公式，得出后验信息，然后根据后验信息去推断未知参数的方法。

计算后验分布期望的传统数值计算方法是数值积分、拉普莱斯近似计算和蒙特卡罗 (Monte Carlo) 重要抽样。MCMC 方法，即马尔可夫链——蒙特卡罗 (Markov chain Monte Carlo) 方法已经变成了非常流行的贝叶斯计算方法。一方面是由于它处理非常复杂问题的效率，另一方面是因为它的编程方法相对容易。

贝叶斯分析方法 (Bayesian Analysis) 提供了一种计算假设概率的方法，这种方法是基于假设的先验概率、给定假设下观察到不同数据的概率以及观察到的数据本身而得出的。其方法为，将关于未知参数的先验信息与样本信息综合，再根据贝叶斯公式，得出后验信息，然后根据后验信息去推断未知参数的方法。

在贝叶斯统计理论中，统计推断中的相关量均作为随机量对待，而不考虑其是否产生随机值。概率被理解为基于给定信息下对相关量不完全了解的程度，对于具有相同可能性的随机事件认为具有相同的概率。在进行测量不确定度的贝叶斯评定时，与测量结果推断或不确是度评定相关的每一个物理量均被分配一个随机变量，分布宽度常用标准差表示，反映了对未知真值了解的程度。

按照贝叶斯理论，与测量或相关评定工作有关的每一个物理量均被分配一个随机变量，尽管每一个估计量和它所表示的相关被测量是不相同的，但它是用来估计被测量的待定真值的。为了简单起见，估计量、估计量的值和该被测量均用相同的符号表示，如用表示样本，同时也用它表示样本值，这可从上下文区别，不会发生混淆，因为样本是随机变量，而样本值是一些常量，这与经典统计理论是不同的。

贝叶斯理论基础

经典统计在对随机分布参数进行参数估计时，假定待估计参数是未知常数，并认定这些参数的信息仅由样本携带，于是通过对样本“毫无偏见”的加工来获得参数估计。由于估计量可能有不完善之处，估计误差在所难免，因此经典统计理论中用置信区间表示这些误差的大小。

在对概率的理解上，经典统计认为概率就是频率的稳定值。一旦离开了重复试验，