

算法与模型

算法是解决问题的步骤序列,是计算机问题求解的“灵魂”;模型是算法应用后产生的用于预测或决策的抽象结果或数学表示,包含数据、参数和结构。模型与算法是人工智能的核心要素,它们不仅构成了智能系统的基础框架,更是连接理论创新与实际应用的桥梁。

本章将系统介绍算法与模型在问题求解、知识表示、知识发现中的核心概念与关键作用。首先探讨问题求解与搜索算法的基本原理,包括盲目搜索和启发式搜索等经典方法,分析这些算法如何为复杂问题提供解决方案。随后,深入剖析知识表示与推理模型,这是构建智能系统的基石,涵盖从一阶谓词逻辑到现代知识图谱的演进过程。最后,通过知识发现与学习模型的介绍,展示如何从数据中自动提取知识,并以基于 Apriori 的关联规则挖掘和支持向量机分类决策两个实践案例,呈现模型算法在实际问题中的应用价值。

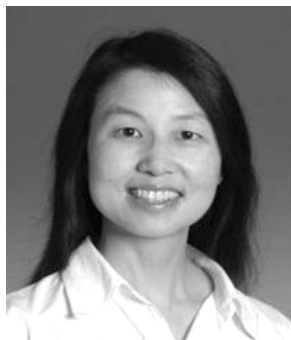
通过本章的学习,读者将能够:

- (1) 掌握问题求解的基本方法,理解不同搜索策略的特点和适用场景;
- (2) 熟悉知识表示的主要形式,了解符号推理的实现机制;
- (3) 认识机器学习的基本原理,了解统计学习和数据挖掘的核心思想;
- (4) 具备将模型与算法应用于实际问题的能力,培养计算思维和智能素养。

【人物档案】 周以真

周以真(Jeannette Marie Wing),1956年生于美国纽约,现任哥伦比亚大学计算机科学教授,曾任微软研究院副总裁(2013—2017年)。她本科、硕士、博士均毕业于麻省理工学院,师从图灵奖得主罗纳德·李维斯特(Ronald Rivest),早期研究形式化方法和高可信系统,在软件验证领域作出重要贡献。

2006年,她在《美国计算机学会通信》(CACM)上发表里程碑论文《计算思维》,首次系统阐述计算思维这一概念。她的思想引发全球教育变革:英国在2014年将计算思维纳入国家课程标准(参考UK教育部文件);中国2017年高中信息技术新课标将其列为学科核心素养。欧盟委员会评价:“周以真为数字公民的培养提供了理论基石。”作为卡内基-梅隆大学计算



机系首位女性系主任(2005—2007年),她推动成立全球首个“计算思维研究中心”,其理论被广泛应用于生物信息学(如癌症基因组模式识别);城市计算(基于时空数据的交通优化建模)。微软研究院评价:“她重新定义了计算机科学与其他学科的对话方式。”

2017年,她获得ACM杰出服务奖,评委会特别强调:“计算思维的概念比任何具体技术突破影响更为深远,它正在塑造未来社会的思维方式。”今天,从少儿编程到AI伦理,她的思维框架持续释放能量。正如她预言:“当每个人都学会像计算机科学家一样思考时,人类将获得解决问题的全新维度。”

在硅谷的算法黑箱与AI伦理争议日益凸显的今天,周以真始终强调计算思维的人文价值:“技术必须服务于人类尊严”。这位站在科学与人文十字路口的智者,仍在为科技文明的发展指引着理性而温暖的方向。

5.1 算法与模型概述

在人工智能尤其是机器学习领域,算法和模型是两个密切相关但又有所区别的核心概念。理解它们的定义和关系对于深入学习人工智能技术至关重要。

5.1.1 算法与模型的定义

在计算机领域,算法是对解决问题计算过程准确而完整的描述,是一系列清晰的指令。通过算法能够对规范的输入,在有限时间内获得所需的输出。算法中的指令描述了一个计算过程,从初始状态和输入开始,经过一系列有限且清晰定义的状态,最终产生输出并停止。例如,排序算法可以将一组无序数据排列成有序序列。

模型是一个概念外延非常广泛的词。它通过实体或虚拟表现,构成对客观事物形态结构的表达。模型可以是模型玩具、科学模型、经济学模型、逻辑模型等。例如,地球仪是地球的模型,理性人假设是经济学中关于人行为的模型,计算机网络中的7层协议是网络协议的模型。

科学模型是科学研究中对一类研究方法的通称,通过数学公式、计算机模拟或图来表示一个简化的自然界,帮助我们进一步理解科学现象。数学模型是关于部分现实世界的抽象和简化结构,用字母、数字及其他数学符号建立的等式或不等式,以及图表、图像等描述客观事物的特征及其内在联系。

模型是对问题域中各实体及其相互作用的抽象表示,两者的区别主要体现在定义、功能和应用层面。

算法是用于解决特定问题或执行计算任务的明确的指令序列,关注过程,强调步骤执行的效率和时间/空间复杂度。如梯度下降算法定义了如何优化模型参数,快速排序算法描述了排序步骤。模型是算法应用后产生的用于预测或决策的抽象结果或数学表示,是对现实系统的抽象或数据学习的产物,关注结果,用于预测或决策。如神经网络模型通过参数捕捉数据特征,线性回归模型用数学公式描述变量关系,CNN的图像识别能力基于层结构和训练后的权重。

算法是构建模型的方法,如决策树算法生成分类模型。模型是算法的具体实现,包含数据、参数和结构,如用特定数据集训练后的逻辑回归模型。

5.1.2 机器学习中的模型与算法

在机器学习领域,算法和模型有着更为具体的定义。

机器学习算法是一种具体的计算方法,用于求解**全局最优解**,使过程高效且准确。本质上,它是计算机算法的一种。机器学习算法执行“模式识别”,从数据中“学习”,或者对数据集进行“拟合”。这些算法可以用伪代码描述,并用现代编程语言实现。例如,Python中的Scikit-learn库提供了许多分类、回归和聚类等机器学习算法的实现。

机器学习模型是机器学习算法在数据集上的输出,表示算法所学到的内容,即用于预测的规则、数字和特定算法的数据结构。模型是由机器学习算法自动创建的程序,用于解决具体问题。从某种意义上说,数据结构也可以看作一种模型。

机器学习的主要问题是学习什么样的模型。例如,在监督学习中,模型是所要学习的条件概率分布或决策函数。样本空间中可能包含多种条件概率分布或决策函数,每个都对应一个模型。

与模型相关的另一个概念是策略。策略是从样本空间中挑选出参数最优模型的准则。模型的分类或预测结果与实际情况的误差(损失函数)越小,模型越好。因此,策略通过引入损失函数来度量模型的好坏。

在机器学习中,算法与模型密不可分。算法是模型建立流程的一个环节,赋予模型“思考”能力。模型则是算法的数学表达,包含数据和任务的概念。因此,我们评估模型的好坏,而不是算法的好坏。算法只有其适用环境和场景。通俗来说,模型可以近似理解为目标函数,算法则是求解该目标函数的方法。所有问题都可以有多种建模方法。

5.2 问题求解与搜索算法

许多人工智能问题在广义上都可以看成问题求解。因此问题求解是人工智能的核心问题之一,它通常是通过在某个可能的解空间中寻找一个解来解决。在问题求解过程中,人们所面临的大多数实际问题往往没有确定性的算法,需要使用搜索技术解决。

5.2.1 问题求解概述

人工智能中每个研究领域都有其特点和规律,但很多任务的解决过程都可抽象为一个问题求解过程。问题求解过程实际上是一个搜索过程,广义地说,它包含了计算机科学的许多技术。

问题求解智能体(agent)是一种基于目标的智能体,在寻找到达目标的过程中,当智能体面对多个未知的选项时,首先检验各个不同的导致已知评价状态的可能行动序列,然后选择最佳序列,这个过程就是搜索,从另一个角度看,搜索就是依靠经验,利用已有知识,并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,使问题得以解决。

问题求解有三个基本特征:目的性、操作序列和认知操作。问题求解必须有明确的目的,否则就不是问题求解。问题求解包括一系列操作过程,可以分为发现问题、分析问题、提出假设和检验假设4个阶段。

问题求解过程是搜索答案的过程,是指利用某些方法和策略,使问题从初始状态出发到达目标状态的过程,因此问题求解技术也叫搜索技术,是通过对状态空间的搜索而求解问题的技术,包括规则、过程和算法等。

搜索技术适用的情况主要有三种,第一种是不良结构或非结构化问题,第二种是难以获得求解所需要的全部信息。第三种是没有现成的算法可供求解使用。

下面介绍搜索技术的应用场景,在即时战略类游戏(例如著名的 *Red Alert*)的界面里,

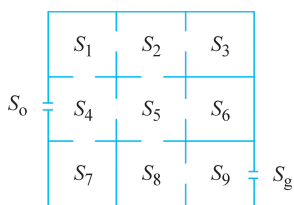


图 5-1 迷宫图示例

当坦克或间谍在地图上行走时,如何找到一条最优的路径呢?这被称为路径规划。搜索技术是解决路径规划的技术手段之一。图 5-1 是一个迷宫,走迷宫是人们熟悉的一种游戏,需要找到一条从入口开始到出口的路径。图中的边代表通道,实线表示走不通的通道,行进时需要避开,虚线表示可以通过,这个迷宫问题的求解需要使用搜索技术。

如果问题的规模很大,则会产生组合爆炸问题,也就是会产生数量特别多的中间过程。从初始状态到目标状态的可能性数量巨大,即使利用计算机解决仍然很费时间。这些问题包括魔方问题、博弈问题(包括围棋、中国象棋、国际象棋等)、八皇后问题、旅行商问题、排课问题、背包问题等。

搜索技术有不同的类型,根据搜索中是否使用启发式信息,搜索可以分为盲目搜索和启发式搜索,盲目搜索也称为无信息搜索,即只按预订的控制策略进行搜索,在盲目搜索过程中获得的中间信息不用于改进控制策略。启发式搜索是指在搜索中加入了与问题有关的启发性信息,用于指导搜索朝着最有希望的方向前进,加速问题的求解过程,并找到最优解。按问题的表示方式,搜索还可以分为状态空间搜索和与或树搜索。状态空间搜索是指用状态空间法求解问题时进行的搜索。与或树搜索是指用问题归约法求解问题时进行的搜索。本节对盲目搜索和启发式搜索进行介绍。

5.2.2 盲目搜索

盲目搜索也称为无信息搜索、无向导搜索,一般只适用于求解比较简单的问题。下面以八数码问题为例进行说明。

八数码问题: 在一个 3×3 的方格棋盘上放置 1,2,3,4,5,6,7,8 八个数码,每个数码占一格,有一个空格。这些数码可在棋盘上移动,其移动规则是:与空格相邻的数码方可移入空格。现在的问题是:对于指定的初始棋局和目标棋局,给出数码的移动序列。该问题被称为八数码问题或重排九宫问题。

可以想象,如果把整个棋盘所表示的一个棋局作为一个节点,则相邻的节点就可以通过移动数码一个一个地产生出来,这样,所有节点按相邻关系就可以连成一个有向图。可以看出,图中的一条边(即相邻两个节点的连线)对应一次数码移动;反之,一次数码移动也对应着图中的一条边。数码移动是按移动规则进行的,所以图中的一条边也就代表一个移动规则或移动规则的一次执行,于是,这个数码问题也就是从该有向图的初始节点(初始棋局)出发寻找目标节点(目标棋局)的问题,或者是在该有向图中寻找一条从初始节点到目标节点的路径问题。在人工智能技术中,把这种描述问题的有效图称为状态空间图,简称**状态图**。状态图实际上是一类问题的抽象表示。事实上有许多智力问题(如旅行商问题、八皇后问

题、农夫过河问题等)和实际问题(如路径规划、定理证明、演绎推理、机器人行动规划等)都可以归结为在某一状态图中寻找目标或路径的问题,在状态图中寻找目标或者路径的基本方法就是搜索。搜索过程中经过的节点和边,按连接关系便会构成一个树状的有向图,这种树状有向图,称为**搜索树**。随着搜索的进行,树会不断地生长,直到当搜索树中出现目标节点时,搜索便停止。

1. 广度优先搜索

广度优先搜索(Breadth-First Search, BFS)是一种典型的盲目搜索策略。广度优先搜索就是始终先在同一级节点中考查,只有当同一级节点考查完之后,才考查下一级节点。或者说,是以初始节点为根节点,向下逐级扩展搜索树。所以,广度优先策略的搜索树是自顶向下一层一层逐渐生成的。广度优先搜索策略的特点是具有完备性,但占用空间大。

广度优先搜索中使用的数据结构包括 OPEN 表和 CLOSED 表,如图 5-2 所示。广度优先搜索中的 OPEN 表是一个先进先出队列,用于存放待扩展的节点。CLOSED 表用于存放已被扩展过的节点。在所有的搜索策略中, CLOSED 表的作用和用法基本不变。

节点	父节点编号

(a) OPEN表

编号	节点	父节点编号

(b) CLOSED表

图 5-2 OPEN 表与 CLOSED 表示例

下面来看广度优先搜索的具体过程,如表 5-1 所示。

表 5-1 广度优先搜索的过程

广度优先搜索算法

- (1) 把初始节点 S_0 放入 OPEN 表中。
- (2) 若 OPEN 表为空,则搜索失败,退出。
- (3) 取 OPEN 表中前面第一个节点 N 放在 CLOSED 表中,并冠以顺序编号 n 。
- (4) 若目标节点 $S_g = N$,则搜索成功,结束。
- (5) 若 N 不可扩展,则转步骤(2)。
- (6) 扩展 N ,将其所有子节点配上指向 N 的指针依次放入 OPEN 表尾部,转步骤(2)

最终,八数码问题用广度优先搜索策略,则可得到如图 5-3 所示的搜索树。

2. 深度优先搜索

除了广度优先搜索,盲目搜索中还有一种常用的搜索策略,那就是深度优先搜索(Depth-First Search, DFS)。深度优先搜索的过程是,新节点优先扩展,直到达到一定的深度限制。若找不到目标或无法再进行扩展时,回溯到另一节点继续扩展。深度优先搜索策略的特点是需要有深度限制,需要有回溯控制。但与广度优先搜索相比,深度优先搜索更节省空间。

深度优先搜索中使用的数据结构依然包括 OPEN 表和 CLOSED 表。深度优先搜索的 OPEN 表是一个后进先出的线性结构(栈结构),存放待扩展的节点,而广度优先搜索中,使



广度优先搜索



深度优先搜索

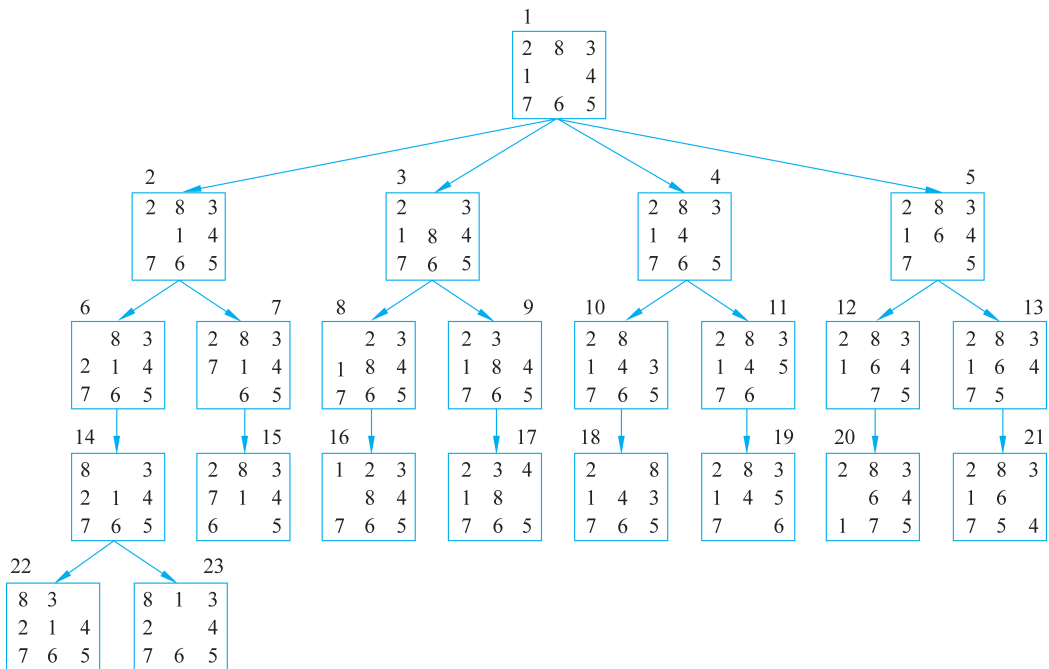


图 5-3 八数码问题的广度优先搜索

用的 OPEN 表则是一个先进先出的线性结构(队列结构),两者有显著的差异。CLOSED 表仍然存放已被扩展过的节点。在深度优先搜索中,除扩展后的子节点应放在 OPEN 表的最前面之外,其他与广度优先搜索算法一样。表 5-2 为使用深度优先搜索解决问题的详细过程。

表 5-2 深度优先搜索算法

深度优先搜索算法

- (1) 把初始节点 S_0 放入 OPEN 表中。
- (2) 若 OPEN 表为空,则搜索失败,退出。
- (3) 取 OPEN 表中前面第一个节点 N 放入 CLOSED 表中,并冠以顺序编号 n 。
- (4) 若目标节点 $S_g = N$,则搜索成功,结束。
- (5) 若 N 不可扩展,则转步骤(2)。
- (6) 扩展 N ,将其所有子节点配上指向 N 的返回指针依次放入 OPEN 表的首部,转步骤(2)

最终,八数码问题用深度优先搜索策略,则可得到如图 5-4 所示的搜索树。

5.2.3 启发式搜索

盲目搜索的缺点是可能会导致组合爆炸,中间状态的数量太多,导致盲目搜索的时间太长。因此人们又提出利用问题的某些控制信息(如解的特征)引导搜索,这种控制信息称为搜索的启发式信息,使用了启发式信息的搜索方法称为启发式搜索。

启发式搜索是指利用启发式信息进行有引导的搜索。启发式搜索的优点是深度优先、效率高、无回溯。缺点是不能保证得到最优解。

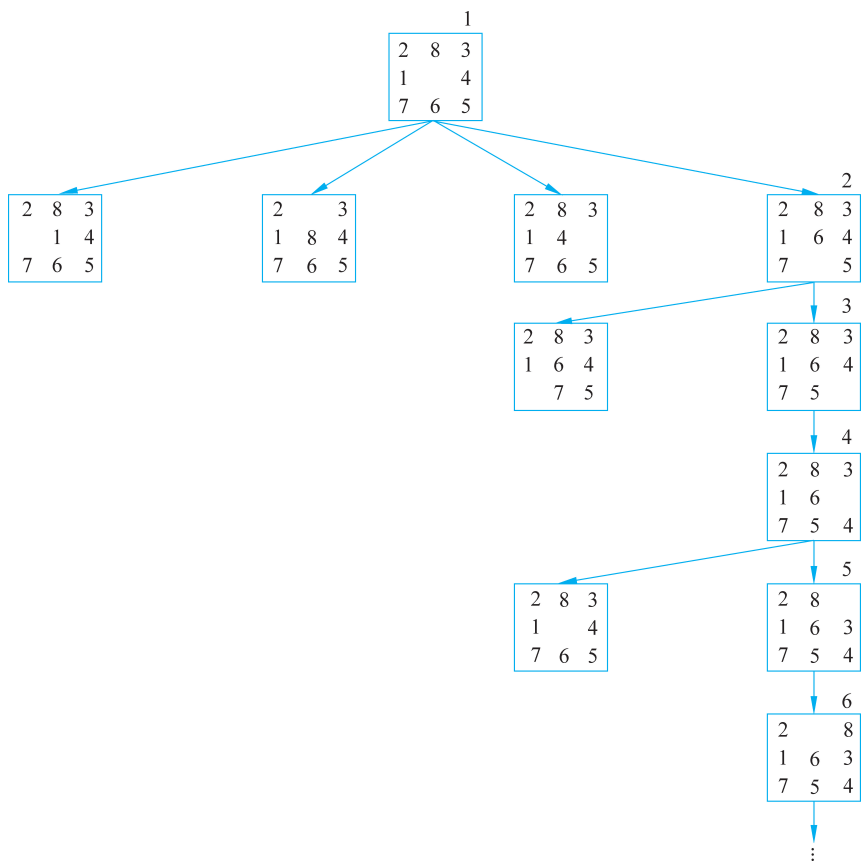


图 5-4 八数码问题的深度优先搜索

启发式信息是有利于尽快找到问题之解的信息。按其用途划分,启发式信息可分为以下三类。

- (1) 用于扩展节点的选择,即用于决定应先扩展哪一个节点,以免盲目扩展。
- (2) 用于生成节点的选择,即用于决定应生成哪些后续节点,以免盲目地生成过多无用节点。
- (3) 用于删除节点的选择,即用于决定应删除哪些无用节点,以免造成进一步的时空浪费。

例如,由八数码问题的部分状态图可以看出,从初始节点开始,在通向目标节点的路径上,各节点的数码格局同目标节点相比,其数码不同的位置个数在逐渐减小,最后为零。所以,这个数码不同的位置个数便是标志一个节点到目标节点距离远近的一个启发式信息,利用这个信息就可以指导搜索。可以看出这种启发式信息属于上面的一种类型。

需指出的是不存在能适合所有问题的万能启发式信息,或者说不不同的问题有不同的启发式信息。

在启发式搜索中,通常用启发函数来表示启发式信息。启发函数是用来估计搜索树上节点 x 与目标节点 S_g 接近程度的一种函数,通常记为 $h(x)$ 。

启发函数并无固定的模式,如何定义一个启发函数,需要具体问题具体分析。通常可以参考的思路有:一个节点到目标节点的某种距离或差异的度量;一个节点处在最佳路径上

的概率;根据主观经验打分等。例如,对于八数码问题, $h(x)$ 就可定义为节点的数码格局同目标节点相比数码不同的位置个数。

启发式搜索要用启发函数来导航,其搜索算法就要在状态图一般搜索算法基础上,再增加启发函数值的计算与传播过程,并且由启发函数值来确定节点的扩展顺序。启发式搜索包括两种搜索策略:全局择优搜索和局部择优搜索。

1. 全局择优搜索

全局择优搜索就是利用启发函数制导的一种启发式搜索方法。该方法亦称为最好优先搜索法。其基本思想是在 OPEN 表中保留所有已生成而未考查的节点,并用启发函数 $h(x)$ 对它们全部进行估价,从中选出最优节点进行扩展,而不管这个节点出现在搜索树的什么地方,如表 5-3 所示。

表 5-3 全局择优搜索算法

全局择优搜索算法

- (1) 把初始节点 S_0 放入 OPEN 表中,计算 $h(S_0)$ 。
- (2) 若 OPEN 表为空,则搜索失败,退出。
- (3) 移出 OPEN 表中第一个节点 N 放入 CLOSED 表中,并冠以序号 n 。
- (4) 若目标节点 $S_g = N$,则搜索成功,结束。
- (5) 若 N 不可扩展,则转步骤(2)。
- (6) 扩展 N ,计算每个子节点 x 的函数值 $h(x)$,并将所有子节点配以指向 N 的返回指针后放入 OPEN 表中,再对 OPEN 表中的所有子节点按其函数值大小以升序排序,转步骤(2)。

用全局择优搜索算法求解八数码问题,设启发函数 $h(x)$ 为节点 x 的格局与目标格局相比数码不同的位置个数,以这个函数制导的搜索树如图 5-5 所示。此八数码问题的解为: S_0, S_1, S_2, S_3, S_g 。

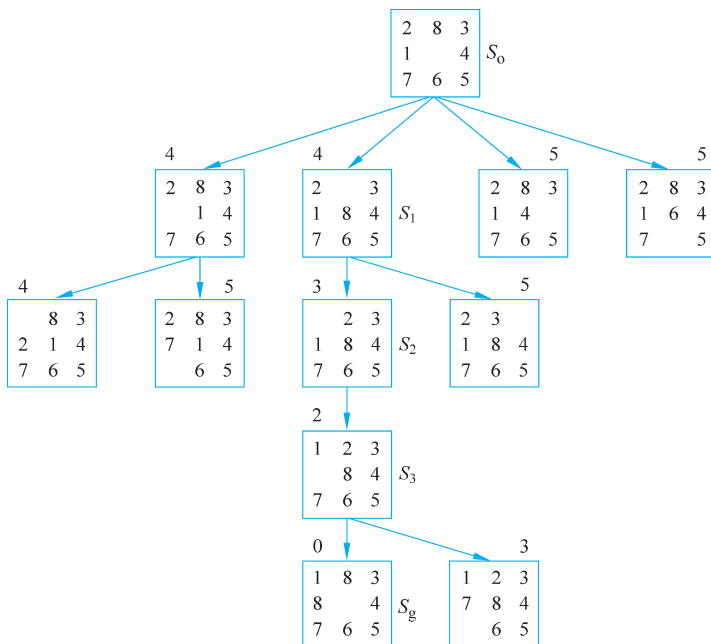


图 5-5 八数码问题的全局择优搜索

2. 局部择优搜索

局部择优搜索与全局择优搜索的区别是,扩展节点 N 后仅对 N 的子节点按启发函数值大小以升序排序,而将它们依次放入 OPEN 表的首部,故算法从略。

5.3 知识表示与推理模型

知识与智能之间有着密切的关系,人类的智能活动主要是获得并运用知识,因此知识是智能的基础。为了使计算机具有智能,能模拟人类的智能行为,就必须首先使它具有知识。但是,人类的知识需要用适当的形式表示,才能存储到计算机中并被运用。因此,知识的表示就成为人工智能中一个十分重要的研究课题。

一个智能系统不仅应该拥有知识,还应该能够很好地利用这些知识,即利用知识推理和求解问题。

5.3.1 知识概述

对于知识,很难给出一个明确的定义,只能从不同侧面加以理解,不同的人有不同的理解。1994年图灵奖获得者、知识工程之父费根鲍姆认为:知识是经过消解、塑造、解释和转换的信息。英国著名教育学家伯恩斯坦(Basil Bernstein)认为:知识是由特定领域的描述、关系和过程组成的。从知识库的观点来看,知识是某领域中所涉及的各有关方面的一种符号表示。

一般认为知识是人们在长期的生活及社会实践中、在科学研究及实验中积累起来的对客观事件的认识和经验。人们把在实践中获得的信息有机地结合在一起,就形成了知识。一般来说,把有关信息关联在一起所形成的信息结构称为知识。

信息之间有多种关联形式,其中用得最多的是一种用“如果……那么……”表示的关联形式。在人工智能中这种知识被称为规则,它反映了信息之间的某种因果关系。例如,我国北方的人们经过多年的观察发现,每到冬天即将来临,就会看到大雁一群一群地向南方飞去,于是把“大雁向南飞”和“冬天就要来了”这两个信息关联在一起,得到了如下知识:“如果大雁向南飞,那么冬天就要来了。”又如,“雪是白色的”也是一条知识,它反映了“雪”与“白色”之间的一种关系。在人工智能中,这种知识被称为事实。

知识是有不同类型的,从不同的角度、不同的侧面对知识有着不同的分类方法。首先从内容上知识可以分为原理(客观)性知识和方法(主观)性知识,其中原理性知识具有抽象概括性,而方法性知识则具有通用性;从形式上知识可以分为显式知识和隐式知识;从逻辑思维角度,知识可以分为逻辑型知识和直觉型知识;从可靠性上知识可以分为理论知识和经验知识。

知识的要素是指构成知识的必需元素。在这里,重点关注的是一个人工智能系统所处理的知识的组成成分。

一般而言,人工智能系统的知识包含事实、规则、控制和元知识等几类。

(1) 事实。包括事物的分类、属性、事物间关系、科学事实、客观事实等,它是有关问题环境的一些事物的知识,常以“……是……”的形式出现,也是最低层的知识。例如,雪是白色的。

(2) 规则。指事物的行动、动作和联系的因果关系知识。这种知识是动态的,常以“如果……那么……”的形式出现,例如启发式规则“如果下雨,则出门带伞或穿雨衣”。

(3) 控制。指当有多个动作同时被激活时,选择哪一个动作执行的知识。控制是有关问题的求解步骤、规划、求解策略等技巧性知识。

(4) 元知识。是关于如何使用规则、解释规则、校验规则、解释程序结构的知识。元知识是有关知识的知识,是知识库中的高层知识。元知识与控制知识有时会有重叠。

知识的特性包括如下 4 方面。

(1) 知识的相对正确性。知识是相对正确的,也就是说,知识的正确性需要一定的前提条件。知识是人类对客观世界认识的结晶,并且受到长期实践的检验。因此在一定的条件及环境下,知识是正确的。“一定的条件及环境”是必不可少的,它是知识正确性的前提,因为任何知识都是在一定的条件及环境下产生的,因而也就只有在这种条件及环境下才是正确的。例如,牛顿力学定律在一定的条件下才是正确的;又如 $1+1=2$,是在十进制的条件下才是正确的,如果是二进制,它就不正确了。

(2) 知识的不确定性。由于现实事件的复杂性,信息可能是精确的,也有可能是不精确的、模糊的;关联可能是确定的,也可能是不确定的。这就使知识并不总是只有“真”和“假”两种状态,而是在“真”和“假”之间还存在许多中间状态,也就是说,存在为“真”的程度问题,也就是知识的不确定性。

(3) 知识的可表示性。知识的可表示性是指知识可以用适当形式(如语言、文字、图形、神经网络等)表示,这样才能被存储和传播。

(4) 知识的可利用性。知识的可利用性是指知识可以被利用,这是显而易见的,每个人每天都在利用自己掌握的知识解决各种问题。

在知识处理中总要问到以下问题:如何表示知识,怎样使机器能读懂这些知识,能对其进行处理,并能以一种人类能理解的方式将其处理结果告诉人们,知识表示是人工智能研究中最基本的问题之一。

知识表示(knowledge representation)就是对人类的知识进行形式化或者模型化,以便计算机能够存储管理和提供访问。知识表示的目的是能够让计算机存储和使用人类的知识。知识表示方法是面向计算机的知识描述或表达形式和方法,是一种数据结构或控制结构的统一体,既考虑知识的存储,又考虑知识的使用。知识表示可看成一组事物的约定,把人类知识表示成机器能处理的数据结构,是知识工程的核心领域。

同一知识可采用不同的表示方法,不同的表示方法可能产生不同的效果。知识表示的目的在于通过有效的知识表示,使人工智能程序能利用这些知识做出决策,获得结论。

5.3.2 推理概述

在现实生活中,人们对各种事物进行分析合并,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,利用专家系统为病人诊治疾病实