

本章探讨计算机科学的一个分支：人工智能。尽管该领域仍然处于发展的初期阶段，但它已经产生了一些令人惊讶的结果，例如，象棋大师，用来学习和推理的计算机，协调一致来完成一个共同的目标（如赢得一场足球比赛）的机器。在人工智能领域，今天的科学想象也许在明天就会变成现实。

人工智能是计算机科学的一个领域，旨在寻求建造自主的机器——无需人为干预就能完成复杂任务的机器。这个目标要求计算机能够感知和推理，尽管这对于人脑来说是天生的，但属于常识行为范畴的这两种能力对于机器来说证明是有困难的。结果是该领域的工作持续面临着挑战。本章就来探讨这个广阔研究领域中的的一些主题。

10.1 智能与机器

人工智能领域十分广阔，并且与其他学科相融合，如心理学、神经学、数学、语言学以及电子与机械工程等学科。为了让思考更集中，我们从考虑智能体的概念以及智能体可能呈现的智能行为类型着手。实际上，人工智能的许多研究都可归类于一种智能体的行为。

10.1.1 智能体

智能体（agent）是对环境的刺激做出响应的一种“装置”。很自然地会把一个智能体想象为一个像机器人这样的单个机器，尽管它可以来用别的形式，如一架自动飞机、计算机视频游戏里的一个角色或是通过因特网通信的一个程序（可能作为客户机或服务器）。大多数智能体具有传感器和效应器，前者接收来自环境的数据，后者对环境做出反应。传感器包括麦克风、摄像机、距离传感器以及空气或土壤采样设备等。效应器的例子有车轮、腿、翅膀、夹子以及语音合成器。

人工智能的目的是建造能够智能运转的智能体。这意味着智能体的效应器的行为必须对其传感器接收的数据做出合理的响应。最简单的情况下，这种响应可能就是反射动作，每个动作都是事先规定的对输入数据的响应。

要获得更“智能”的行为需要更高级别的响应。例如，我们可以把所处环境的知识赋予一个智能体，并要求该智能体相应地调整它的动作。投掷一个棒球的过程很大程度上就是一个反射动作，但是决定如何投、往哪儿投需要当前环境的知识。（一球已出，跑垒手在第一垒和第三垒。）这样的现实世界的知识如何存储、更新、访问乃至在决策过程中最终应用，一直是人工智能领域中具有挑战性的问题。

如果我们希望智能体寻求的目标是赢得一场棋赛或是蜿蜒通过一条拥挤的通道，那么就需另一种层次的响应。这种有目标性的行为需要智能体的响应（或是一系列的响应）应当是周密考虑的结果，构成一个行为计划，或是在当前的各种选项中选取最好的行为。

智能体作为电影角色

一个称为Massive的计算机程序用来为《指环王》三部曲产生兽类和人类的虚幻的军队。每一个屏幕上的战士在程序中都是作为一个不同的“智能的”智能体来表示的，智能体有自己的身体特征、随机指派的人物个性以及进攻或逃跑的意向。在测试第二部电影中Helms Deep之战的仿真中，兽族逃跑的意向值设置太高，当它们面对人类的战士时全部逃跑了。（这或许就是考虑到工作太危险而用虚拟临时演员的最初案例。）

在有些情况，随着智能体不断学习，它的响应能够不断地得到改进。这可以采取不断发展的过程性知识（procedural knowledge）（学习“怎样”）的方式，或者储备陈述性知识（declarative knowledge）（学习“什么”）的方式。学习过程性知识涉及一个反复试验过程，在这个过程中，智能体从出错受罚、正确受奖的过程中学习适当的反应。根据这个方法开发了一些智能体，它们能够随着时间改进在诸如跳棋和国际象棋这种竞赛性游戏中的能力。学习陈述性知识通常采取的形式是扩充或变更智能体的知识存储器里的“事实”。例如，一个棒球运动员必须不断地重复调整其知识数据库（虽然还是一球已出，但现在跑垒手在第一垒和第二垒），从中对将来的事件做出合理响应。

一个智能体要对刺激做出合理的响应，它必须“理解”由其传感器接收的刺激。也就是说，智能体必须能够从其传感器产生的数据里提取信息，换句话说，智能体必须能够感知。有些情况下，这是一个直接的过程。从一个陀螺仪获取的信号很容易就能编码成适合计算的形式，以确定响应。但是有些情况下，从输入数据提取信息并不容易，例如理解说话和理解图像就很难。同样，智能体也必须能够以与效应器的方式表达它们的响应。这可以是一个直接的过程，也可能要求智能体把其响应表达为一句完整的口语句子——这意味着智能体必须生成语音。所以，像图像处理和分析、自然语言理解以及语音发生这些主题都是重要的研究领域。

我们这里识别的智能体的属性既表示以前的研究范畴，又表示当今的研究范畴。当然，它们彼此之间并不是完全无关的。我们希望最终能够开发出处理所有这些属性的智能体，产生出能够理解来自环境的数据，并通过学习过程发展新的响应模式的智能体，学习的目的是最大限度地提高智能体的能力。然而，通过孤立各种推理行为并独立研究它们，研究人员获得了一个立足点，该立足点今后可以与其他领域的发展相结合，产生更加智能的智能体。

本节的最后我们介绍一个智能体，也将为10.2节和10.3节的讨论提供一个背景。该智能体是为解决8数码游戏（eight-puzzle）而设计的，该游戏由8个小方块组成，标号为1到8，放置在一个3行3列总共可容纳9片小方块的框架内（见图10-1）。这样，框内的方块间有个空位，任何邻接的方块可以推移，允许框内的方块随意排布。问题是要把杂乱布放的方块移回到它们初始的位置（见图10-1）。

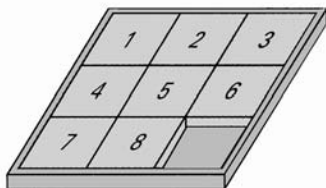


图10-1 8块拼图的最终布局

我们的智能体采用如图10-2所示的这样一种装置，该装置配备有一个夹具、一个摄像机和一个带橡皮头的指杆，橡皮是为了推移东西时不会打滑（见图10-2）。当首次开启该智能体时，

夹具会张合，好像要这个拼图一样。当我们把一个随意排布的拼图放进夹具时，夹具就会把它夹住。不一会，机器的指杆降低，并开始向框架内推移方块，直到所有的方块回复到原始位置。这时，机器就放开拼板并自己关掉电源。

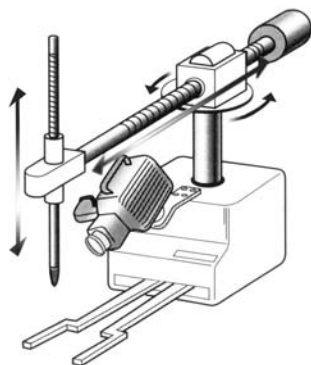


图10-2 解决8数码游戏的机器

这个解决8数码游戏的机器展现了已经提到过的两个智能体的属性。第一，它必须能够感知，就是必须从摄像机拍摄的图像中获取当前拼图的状态。我们将在10.2节阐述理解图像的问题。第二，它必须开发和实现一个达到目标的计划。这些问题将在10.3节中阐述。

10.1.2 性能与模拟

在最后的分析中，通过观察智能体的输入—响应模式来判定智能体的行为是智能的（或推理的）。这引出了人工智能的两种研究方法。一种是**面向性能的**（performance oriented）方法，在这种方法中，研究人员力求达到使智能体的性能最大化；另一种方法是**面向模拟的**（simulation oriented）方法，在这种方法中，研究人员对智能体如何产生响应更感兴趣，而不是它的效率。当直接目标是要产出一个产品时，面向性能的研究更为常用。面向模拟的研究更适合于那些长期研究与开发的项目，它们的目标在于扩展我们对一个对象的理解。这两种方法在人工智能的发展中都起着重要的作用。

作为一个例子，考虑自然语言处理和语言学领域。这些领域关系密切，研究成果相得益彰，但它们的深层目标却不同。语言学家的兴趣在于弄明白人类如何处理语言，而自然语言处理领域的研究者的兴趣在于开发能处理自然语言的机器。所以，语言学家以面向模拟的模式运作——建造用来检验理论的系统。相反，自然语言处理的研究者以面向性能的模式运作——建造执行任务的系统。后一种模式产生的系统（如文档翻译机和响应口头命令的机器系统）在很大程度上依赖于语言学家获取的知识，但对于特定系统的限定环境常加以“抄近路”的简化。

作为一个基本的例子，考虑为一个操作系统开发一个外壳（shell）的任务，它通过口述英语命令接收来自外部世界的指令。在这种情况下，shell（一个智能体）不需要考虑完整的英语语言。更准确地说，外壳不需要区分copy这个词的不同意思（是名词还是动词？是否有抄袭的含义？）。相反，外壳仅仅需要把copy这个词与其他命令（如rename、delete）区分开来就行。所以外壳只要将输入与预先确定的声音模式对比，就能够执行它的任务。这样一个系统的性能可能会令人满意，但在审美意义上也许不能令语言学家满意。

10.1.3 图灵测试

过去，**图灵测试**（Turing test）（1950年由阿兰·图灵提出）一直作为衡量人工智能领域的

进展时的一种基准。图灵的提议是，允许一个人（我们称他为询问者）与一个测试对象通过一个打字机系统进行通信，而没有告知询问者测试对象究竟是一个人还是一台机器。在这个环境中，如果询问者没能够把一台机器与一个人区分开来，那么可以宣称这台机器的行为是智能的。图灵预测，到2000年，机器将会有30%的机会通过一个5分钟的图灵测试——这个预见惊人的准确。

人工智能的起源

寻求建造能够模仿人类行为的机器有很长的历史，不过很多人会认同现代人工智能领域起源于1950年。就在这一年，阿兰·图灵发表了论文“Computing Machinery and Intelligence”，提出机器能够通过编程来展现智能的行为。这个领域的名字——人工智能——就在几年后由约翰·麦卡锡(John McCarthy)在一个现在看来非常具有传奇色彩的建议里提出，他建议“1956年夏天在达特默斯学院(Dartmouth College)开展人工智能的研究”，以探究“这种推测，即认为认知的每个方面或智能的任何其他特征原则上都能够被精确地描述，从而能够制造出模拟它的机器”。

图灵测试场景的一个著名示例是20世纪60年代中期由Joseph Weizenbaum开发的程序DOCTOR（更通用的一个版本称系统为ELIZA）。这个交互程序被设计用来反映一种指导心理医疗的罗杰斯心理分析的景象；计算机扮演了分析师的角色，而用户扮演病人。在内部，DOCTOR所做的一切是根据某些明确的规则将病人的陈述重新构造并反馈给病人。例如，回应病人的陈述“我今天觉得很累”，DOCTOR可能回答“为什么你今天觉得很累？”如果DOCTOR不能识别句子结构，它仅仅作出像“继续”或者“这很有趣”这样的响应。

Weizenbaum开发DOCTOR的目的是为了研究自然语言沟通。心理疗法的目标仅仅提供了一个程序可以“沟通”的环境。然而，Weizenbaum没有想到的是，个别的心理学家建议把这个程序用在实际的心理治疗中。（罗杰斯的论点是，在治疗期间，应该是病人来主导对话，而不是分析师。所以他们认为计算机也能像治疗师那样引导对话。）而且，DOCTOR表现出的理解的景象如此强以致许多与它“沟通”过的人变得依赖于这种与机器的问答式对话。在某种意义上，DOCTOR通过了图灵测试。其结果是，在伦理及技术上都产生了争议，Weizenbaum成为一位在这个技术进步的世界中维护人类尊严的倡导者。

较新的图灵测试“成功”的例子包括因特网病毒，病毒为了诱骗人类放弃对恶意软件的防护，它用人类受害者来传输病毒的“智能”变体。此外，与图灵测试类似的现象发生在像下棋这样的计算机游戏的场景中。尽管这些程序选择棋路仅仅通过应用蛮力技术（这与我们将在10.3节讨论的内容相似），但人类在同计算机的竞赛过程中常感觉机器拥有创造力甚至个性。相似的感觉发生在机器人技术领域，根据自然属性建造的机器表现出了智能的特征。例如，玩具机器狗，仅仅通过点头或者是竖耳朵来响应声音，呈现了可爱的个性。

问题 与练习

1. 指出一个智能体可能会做的几种“智能”动作。
2. 一棵放在只有一束光源的暗室里的植物，它就朝着这束光源方向生长。这是一种智能响应吗？植物拥有智能吗？那么，你对智能的定义是什么？
3. 假定一台售货机根据所按的按钮发售不同的物品。你认为这样的一台机器是否“知道”被按了哪个按钮？你对“知道”的定义是什么？
4. 如果一台机器通过了图灵测试，你会认同这台机器是智能的吗？如果不是，你是否否认该机器看上去是智能的？

10.2 感知

一个智能体要智能地响应从它的传感器接收的输入，它必须能够理解输入。也就是说，智能体必须能够感知。本节我们来探讨感知的两个已经证明是特别具有挑战性的研究领域——理解图像和理解语言。

10.2.1 理解图像

让我们考虑10.1节介绍的解决8数码游戏的机器所提出的问题。机器上夹具的张合没有表现出严重的障碍，在这个张合的过程中，因为这里的应用要求的精度不高，所以检测夹具中是否有拼图的功能也很简单，即使摄像机对拼图的对焦问题也不难通过设计夹具把拼图安置在一个事先确定的位置来加以解决。所以，机器需要的第一个智能行为是从一个视觉媒介中提取信息。

必须认识到，机器在看拼图时所面对的问题不单是产生和存储一张图像，这方面的技术好些年前在传统的摄影及电视系统中就能做到。相反，这里的问题是为了提取拼图当前的状态，要理解这个图像（可能随后还要监控方块的运动）。

在解决8数码游戏的机器的案例中，对拼图图像的可能解释是相对有限的。我们可以假定，在一个排列好的模式里所呈现的总是一幅包含数字1到8的图像。问题只是去提取这些数字的排列。为此，我们想象拼图的图像已经在计算机存储器中，按照位进行编码。编码中的每一位表示具体像素的亮度。假定图像的大小统一（机器把拼图放在摄像机前预定的位置），通过把图像的不同部分与用在拼图中的单个数字产生的位模式构成的预定模板相比较，机器就能够检测出哪个方块在哪个位置。如果发现匹配，则说明拼图达到了要求的状况。

这种图像识别技术是光学字符阅读器中使用的一种方法。但它是有缺点的，它对被读的符号在类型、大小以及方位上要求一定程度上的一致性。特别是，即使对于相同的符号，外形也相同，但字体较大的字符产生的位模式与较小字体的模板也不匹配。此外，可以想象的是，当试图处理手写材料时，问题会变得非常困难。

解决字符识别问题的另一个方法是基于匹配几何特征，而不是符号的精确外形。在这种情况下，数字1表示为一条单竖线，数字2可能代表一条不封闭的曲线，底部与一条水平直线相连，等等。这种符号识别的方法分两步：第一步是从要处理的图像中提取图像特征，第二步是把这些特征与已知符号的特征进行比较。与模式匹配方法一样，这种符号识别技术并不可靠。例如，图像的少许错误会产生一组完全不同的几何特征，比如区分字母O和C的情况，以及8数码游戏里的数字3和8。

在8数码游戏中我们比较幸运的是不需要理解一般三维图像。例如，我们拥有的便利条件是，保证识别的形状（数字1到8）相互孤立地处在图像上的不同部分，而不是一般常见重叠的图像。例如，在一张普通的照片中，面临的不仅是从不同的角度识别一个对象的问题，还包括隐藏在视线背后的对象的某些部分。

理解一般图像的任务通常采取两步进行处理：（1）**图像处理**（image processing），指标识图像的特征；（2）**图像分析**（image analysis），指对这些特征代表什么意思的理解过程。在利用符号的几何特征进行识别的描述中，我们已经提出了二分法的处理方法。在这个过程中，图像处理代表标识在图像中发现的几何特征的过程，图像分析代表标识那些特征的意义的过程。

图像处理带来大量研究主题。一个是轮廓增强，使用数学技术使图像中区域间的边界线变得更清晰。在某种意义上，轮廓增强试图将照片转换成线条画。图像分析的另一个活动是区域查找。这是标识图像中区域的过程，这些区域拥有共同的属性，比如亮度、颜色或者纹理。这

样的一个区域很可能代表属于一个对象的一部分。（这种识别区域的能力使得计算机可以给老式的黑白电影着上彩色。）图像处理领域还包含另一个活动——滤波，即去除图像中的缺陷的过程。滤波使图像中存在的错误不会混淆其他图像处理步骤，但是太多的滤波会导致重要信息的丢失。

滤波、轮廓增强以及区域发现都是用来标识图像中各种成分的步骤。图像分析是确定这些成分代表什么，以及最终确定这个图像代表什么的过程。这里我们还要面对从不同视角识别被部分遮挡对象这样的问题。图像分析的一个方法是开始假定一个图像大概是什么，然后尝试把图像中的成分与那些猜测的对象相联系。这看起来是人类所使用的方法。例如，有时我们会发现，在我们视觉模糊的情况下识别一个未预料的对象是很困难的，但是一旦我们有一个该对象可能是什么的线索，我们就能容易地识别它。

与一般图像分析相关的问题特别多，在这个领域还有许多研究要做。实际上，图像分析是证明人类能够很快又非常容易完成的任务是如何挑战机器能力的领域之一。

强人工智能与弱人工智能

能够通过对机器进行编程来展现其智能行为的那种推测能力被认为是**弱人工智能**（weak AI），今天在不同程度上已经被大众所接受。但是，机器能够通过编程而获得智力，亦即意识的那种推测能力，则被认为是**强人工智能**（strong AI）。这种人工智能受到了广泛的质疑。强人工智能的反对者认为，机器在本质上与人类不同，它永远不能像人类那样感受爱、区分对错，以及考虑自我。然而，强人工智能的支持者认为，人类的头脑是由许多小的成分构成，每个成分都不是人，没有意识，但是当它们结合在一起就成了人。所以他们辩称，为什么同样的现象就不可能出现在机器身上呢？

解决强人工智能争辩的难题在于，智能和意识这样的属性是内在特性，不能够直接加以确认。正如阿兰·图灵指出的一样，我们认为其他人属于有智能的是因为他们的行为表现出有智能——即使我们不能观察到它们内部的智力状态。那么，如果机器也呈现外在的意识特性，我们是否准备认可机器具备同样的水准呢？为什么是？为什么不是？

10.2.2 语言处理

理解语言是感知问题的另一个已证明的具有挑战性的问题。把形式化的高级程序设计语言翻译成机器语言（见6.4节）获得的成功使早期的研究人员认为，设计程序使计算机具有理解自然语言的能力不久将会实现。实际上，翻译程序的这种能力给我们一种错觉——机器真能理解被翻译的语言。（回忆6.1节中Grace Hopper讲的故事，那个经理以为她正在教计算机理解德语。）

这些研究人员没有明白形式化的程序设计语言与英语、德语以及拉丁语这些自然语言之间的差异。程序设计语言由精心设计的原语组成，每个语句只有一种语法结构，只有一种意思。相反，自然语言的一个语句会因为上下文的不同，甚至交流方式不同而有多种意思。因此，人类理解自然语言很大程度上依靠额外的知识。

例如，句子

Norman Rockwell painted people.

以及

Cinderella had a ball.

都有多种意思，但通过语法分析或单独翻译每个词并不能区分这些意思。实际上，要理解这些句子需要有理解句子上下文的能力。在其他场合，一个句子的真实意思与它的字面意思完全不

同。例如，

你知道几点了吗？

通常的意思是“请告诉我现在几点了。”或者如果说话者已经等候了很长时间，那么这句话意思可能是“你来得太迟了。”

要弄明白一种自然语言中的一个句子的意思需要几个层次的分析。第一层是**语法分析** (syntactic analysis)，其主要成分是语法分析。在这里，句子

Mary gave John a birthday card. (玛丽给约翰一张生日贺卡。)

的主语是Mary，而句子

John got a birthday card from Mary. (约翰收到玛丽的一张生日贺卡。)

的主语是John。

分析的第二层次称为**语义分析** (semantic analysis)。语法分析仅标识每个词语法上的作用，与语法分析不同，语义分析担负的任务是标识句子中的每个词在语义上的作用。语义分析试图标识这样的内容，如描述的动作、动作的主体（可能是句子的主语，也可能不是）以及动作的对象。正是通过语义分析，句子“玛丽给约翰一张生日贺卡”和“约翰收到玛丽的一张生日贺卡”被认为是在说同一件事情。

分析的第三层是**上下文分析** (contextual analysis)。在这个层次里，句子的上下文被引入到理解过程中。例如，很容易分辨出句子

That bat fell to the ground.

中的每一个单词的语法上的作用。我们甚至能通过识别动作“falling”和动作主体“bat”等等来实现语义分析。但只有当我们考虑到其上下文后，句子的意思才能变得明确。尤其是，这句话在棒球比赛这样的背景下和在洞穴中探险这样的背景下有着不同的意思¹。而且，在上下文这个层次，问题“你知道几点了？”的真正意思才能被最终揭晓。

我们应当注意，各个层次分析（语法、语义及上下文）并不一定相互独立。对于句子

Stampeding cattle can be dangerous.

如果我们想象是一群牛在惊跑，那么主语是名词cattle（用形容词stampeding加以修饰）。但是如果语境是哪个恶作剧者以惊吓牛群取乐，那么主语就是动名词stampeding（宾语是cattle）。因此，这个句子不只有一个语法结构——哪一个正确意思依赖于上下文。

自然语言处理的另一个研究领域关系到整个文档，而不是单个句子。这里涉及的问题可分成两类：**信息检索** (information retrieval) 和**信息提取** (information extraction)。信息检索的任务是标识与手头论题有关的文档。一个例子是万维网的用户试图找到与特定主题相关的站点时所面对的问题。该技术的当前状态是为关键字搜索站点，但这经常产生大量的错误链接，并且由于其处理“automobiles”而不是“cars”，它会忽略一个重要的站点。所需要的就是理解所考虑站点内容的搜索机制。获取这样的理解是很困难的，这也就是许多搜索机制都转向采用像在4.3节介绍的XML这样的技术来产生一个万维语义网的原因。

信息提取是指从文档中提取信息这样的任务，并采用一种形式以方便用于其他应用程序。这个意思可以理解为为一个特定的问题确定答案，或者是将信息以某种格式记录，以备日后解

1. 英文单词bat有蝙蝠和球棒两种意思，在棒球比赛中应指“球棒”，在洞穴中探险时则可能指“蝙蝠”。

回答问题时使用。有一种这样的格式称作模板，这种模板实质上是一种记录细节的调查表。例如，考虑一个读报系统。该系统可以利用各种各样的模板，报纸上的每一类文章用一个。如果系统鉴别一篇文章是关于入室盗窃的报道，它会继续试图把这填写到入室盗窃模板中，该模板可能要求填写这样一些条目，如失窃地点、失窃时间和日期以及失窃物品等。相反，如果系统鉴别一篇文章是关于自然灾害的报道，那么它就填写自然灾害模板，引导系统确定灾害类型、损失的大小等。

信息提取者记录信息的另一种形式称为**语义网** (semantic net)。这实质上是一个大的链接的数据结构，结构中的指针用来指出数据项之间的联系。图10-3显示了一个语义网的一部分，突出显示的部分是从句子Mary hit John.中得到的信息。

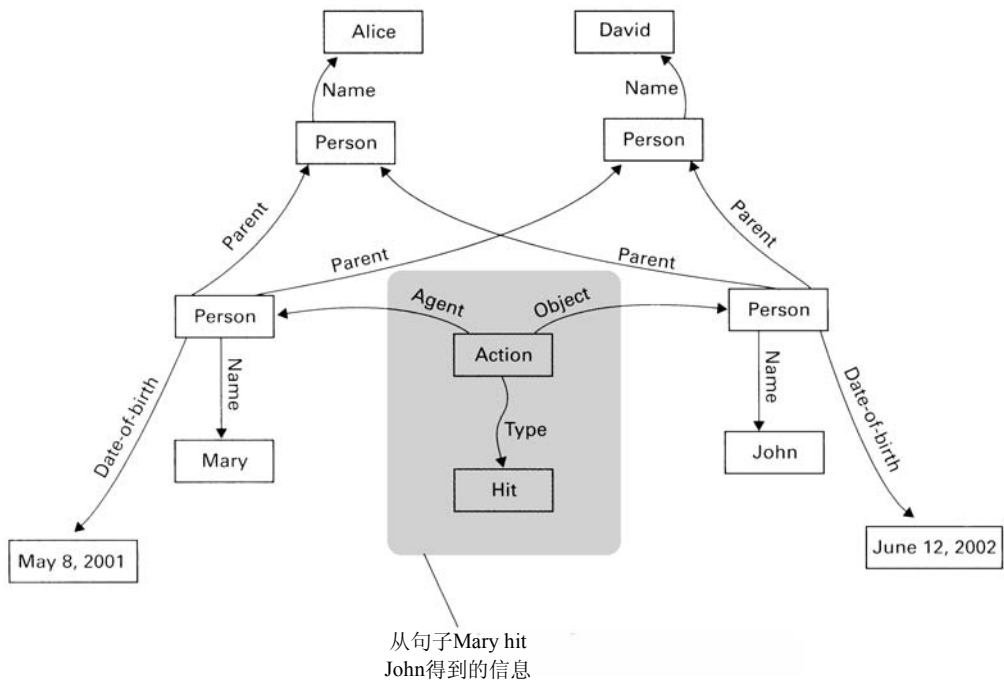
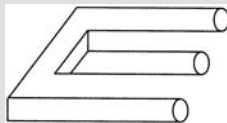


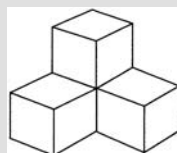
图10-3 一个语义网

问题与练习

1. 对于一个机器人视频系统，一种情况是机器人用来操控自己的活动，另一种情况是将其传递给遥控机器人的人。这两种情况对视频系统的要求有什么不同？
2. 是什么让你知道下图不合情理？这样的洞察怎样编程到机器中？



3. 下图中有几个方块？怎样对一个机器编程使其能正确回答这个问题？



4. 你是如何知道句子“Nothing is better than complete happiness”和句子“A bowl of cold soup is better than nothing”不是暗指“A bowl of cold soup is better than complete happiness”?你这种区分能力如何被移植到一台机器上?
5. 指出在翻译句子“*They are racing horses*”时的二义性。
6. 比较下列两个句子的语法分析结果。然后解释二者在语义上的不同。
The farmer built the fence in the field.
The farmer built the fence in the winter.
7. 基于图10-3中的语义网, Mary和John之间是什么家庭关系?

10.3 推理

现在,让我们来利用10.1节介绍的求解8数码游戏的机器来探讨开发具有基本推理能力的智能体的技术。

10.3.1 产生式系统

一旦解决8数码游戏的机器从看到的图像中解读出了方块的位置,它的任务就变成了决定需要哪些移动来求解难题。马上可以想到的一个方法是,把方块所有可能排列的解决方案都预先编制到机器中。然后,机器的任务就只是选择和执行合适的程序。然而,这个8数码游戏有几千种布局,所以对每一种提供一个直接的解决方法方案显然不可取。因此,我们的目标是对机器编程,让机器能够自己构建难题的解决方法。也就是说,必须对机器编程使其能够实现基本的推理活动。

开发机器的推理能力已经是一个研究多年的主题。有关这方面的研究已经形成一个共识,即有一大类推理问题具有共性,这些共性被孤立在一个抽象的实体中,该实体称为**产生式系统**(production system)。这种系统由三个主要部分组成。

(1) 状态集合。每个**状态**(state)是一个可能在应用环境中发生的情形。最初的状态称作**开始状态**(start state)(或者初始状态),期望的状态称作**目标状态**(goal state)。(在我们的案例中,一个状态就是指8数码游戏的一个布局,开始状态就是8数码游戏提交时的布局,目标状态就是图10-1所示的已经解决了难题的布局。)

(2) 产生式集合(又称规则或者移动)。**产生式**(production)是指能在应用环境中执行的一个操作,并使系统从一个状态转移到另一个状态。每个产生式可以与一些先决条件相关联,也就是说,在产生式被应用之前,环境中必定会出现一些可能存在的条件。(在我们的案例中,产生式就是方块的移动。一个方块每次移动的前提条件是其相邻的位置必须有空位。)

(3) 控制系统。**控制系统**(control system)是由解决问题使其从开始状态变换到目标状态的逻辑组成的。在处理过程的每一步,控制系统都要决定,在满足先决条件的那些产生式中,下一步该执行哪一个。(对于8数码游戏的例子,给定一个特定状态,在空位旁会有几个方块,因而存在几个可用的产生式。控制系统必须决定移动哪一个方块。)

注意,在一个产生式系统的上下文中,赋予解决8数码游戏的机器的任务可以被程式化。在这种情况下,控制系统采用程序的形式。该程序检查8数码游戏的当前状态,确定导向目标状态的一系列产生式,并执行这一系列产生式。因此,我们的任务就是为解决8数码游戏设计一个控制系统。

控制系统开发中的一个重要概念是**状态图**(state graph),它是一种方便的表示或至少概念化一个产生式系统中的所有状态、产生式以及先决条件的方法。这里,“图”这个词是指一种数

学家称为**有向图** (directed graph) 的结构, 是指一组由箭头连接起来的称为**结点** (node) 的位置。一个状态图由一组用箭头连接的结点组成, 结点表示系统中的状态, 箭头表示从一个状态转换到另一个状态的产生式。状态图中两个结点被一个箭头连接的条件是: 当且仅当有一个产生式, 它把系统从箭头起点处的状态变换到箭头终端处的状态。

我们应当强调的是, 正像在解决8数码游戏时拼图可能状态的数量使我们难以明确地提供预先编制好的解决方案一样, 数量太大的问题也使得我们不能明确地表示整个状态图。所以, 状态图是构思解决手头问题的一种方法, 但不能用来表示全部内容。虽然如此, 你会发现, 考虑图10-4显示的8数码游戏的一小部分状态图对于求解难题是有帮助的。

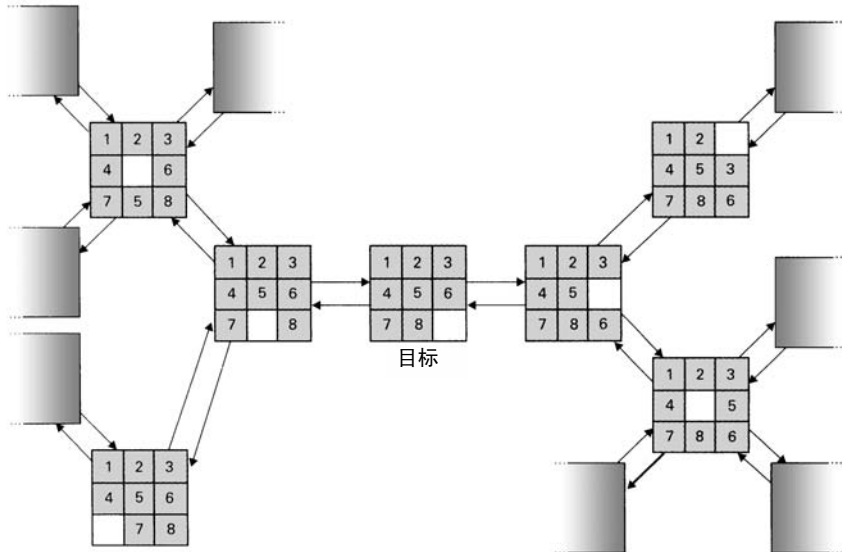


图10-4 8数码游戏状态图的一小部分

当根据状态图来考虑时, 控制系统面临的问题变成了寻找一连串从开始状态导向目标状态的箭头。毕竟, 这一连串的箭头代表了解决初始问题的一系列产生式。所以, 不管应用是什么, 控制系统的任务都可以看作是寻找一条贯穿状态图的路径。对控制系统的这种普遍观点是根据产生式系统对要求推理的问题进行分析的成果。如果一个问题能够根据产生式系统来描绘, 那么它的解决方法就能够根据搜索一条路径来明确表达。

为了强调这一点, 我们先来考虑其他任务是如何按照产生式系统设计, 然后在控制系统通过状态图发现路径的背景下完成的。人工智能的经典问题之一就是下棋这样的游戏, 这类游戏在一个明确规定的背景下属于中等复杂度, 因此, 它为理论测试提供了一个理想的环境。在下棋游戏中, 潜在产生式系统的状态是可能的棋盘布局, 产生式是棋子的移动, 控制系统就具体为棋手 (人或别的)。状态图的开始结点表示棋子在初始位置时的棋盘。从该结点出发的分支是一些箭头, 这些箭头指向游戏中在棋子的第一步移动之后会达到的那些棋盘布局, 而从这些结点出发的每一个分支又引向下一步移子会达到的那些布局, 依次类推。通过这种明确地表达, 我们可以把一个下棋游戏想象为有两个选手组成, 每个选手都试图通过在一个大的状态图中寻找一条通向自己选择的目标结点的路径。

或许从给定事实得出逻辑推论的问题是一个不太明显的产生式系统的例子。在这种情况下, 产生式是称为**推理法则** (inference rule) 的逻辑规则, 这些规则允许从旧命题中形成新命题。例如, 命题“所有超级英雄都是崇高的”和“超人是超级英雄”可以合并产生“超人是崇高的”。

这样一个系统中的状态由各种命题组成，这些命题在推导过程的一些特定点为真：开始状态是基本命题（常称为公理）的集合，从中可以得出结论；而目标状态则是包含了所提出结论的任何命题的集合。

作为一个例子，图10-5显示了以下推论所经历的状态图的一部分。从一组命题“苏格拉底是男人”、“所有男人都是人”及“所有人都终有一死”可以推论出“苏格拉底终有一死”。从中我们看到，随着推理过程应用合适的产生式来生成新的命题，知识的主体从一个状态转换到了另一个状态。

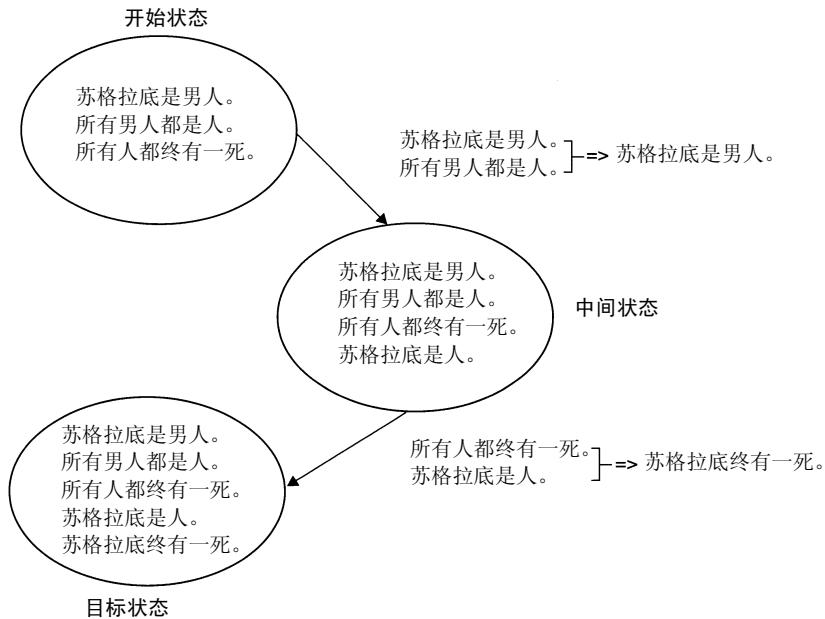


图10-5 产生式系统环境中的演绎推理

当今，这种推理系统经常应用在逻辑程序设计语言(6.7节)中，它是大多数专家系统(expert systems)的核心。专家系统是为模拟因果推理而设计的软件包，这些因果推理是人类的专家面对相同的情形所遵从的。例如，医疗专家系统用来协助疾病诊断或改良治疗。

小规模到大规模

一个在现论的发展和测试中不断出现的现象是，从小规模到大规模应用的转换。一个新理论的最初实验通常涉及一些小的、简单的例子。一旦取得成功，该实验环境就扩展到更现实的、大规模的系统。一些理论可以在这种转换中生存，有些则不能。有时小规模的成功足以鼓舞该理论的支持者在大规模中的失败已经让其他研究人员灰心之后坚持不懈。有些情况下，这种坚持最终得到回报，而另一些情况下则意味着白费力气。

这种事情在人工智能领域显而易见。有一个例子是发生在自然语言处理领域的，早期在有限集合中的成功曾导致很多人相信通用自然语言理解的曙光即将来临。遗憾的是，迈向大规模应用的进一步成功已经被证实极其艰难，巨大的付出却成效缓慢。另一个例子是人工神经网络学科，该学科刚刚兴起时呼声极高，这几年当其大规模应用能力受到质疑时则呼声渐弱，现在已经转为沉寂。正如文中所述，人工智能领域的不少学科，还有广义上的计算机科学，当前都面临着这种转换的考验。

10.3.2 搜索树

我们已经看到，在产生式系统的环境中，控制系统的工作涉及搜索状态图，找出从开始结点到目标结点的一条路径。完成这种搜索的一个简单的方法就是仔细考察每一个从初始状态发出的箭头，并记录下每一个目标状态，然后继续仔细考察从这些新状态发出的箭头，再记录下结果，依次类推。对目标的搜索像向水中滴入一滴墨水一样，从开始状态扩散开来。这个过程继续进行，直到一个新状态就是目标状态，在这里，解决方法就找到了，控制系统只需要沿着被发现的这条从开始状态到目标状态的路径应用产生式。

这种策略的结果实际上是建立一个树，称作**搜索树**（search tree），它由被控制系统分析后得到的部分状态图构成。搜索树的根结点是开始状态，每个结点的子结点是由那些应用一个产生式从父结点可到达的状态构成。搜索树中结点间的每条弧线代表一个产生式的应用，从根到叶子的每一条路径代表状态图中相应状态间的一条路径。

解决图10-6所示布局的8数码游戏将会产生的搜索树如图10-7所示。该搜索树最左边的分支代表试图通过最初向上移动方块6来解决问题，中间分支表示向右移动方块2的开局方法，而最右边的分支表示以向下移动方块5来开局。搜索树进一步显示，如果以向上移动方块6来开局，那么下一个允许的产生式只能是方块8右移。（实际上，这时还可以将方块6向下移动，但这仅仅是上一个产生式的倒置，因而是毫无关系的移动。）

1	3	5
4	2	
7	8	6

图10-6 一个尚未解决的8数码游戏

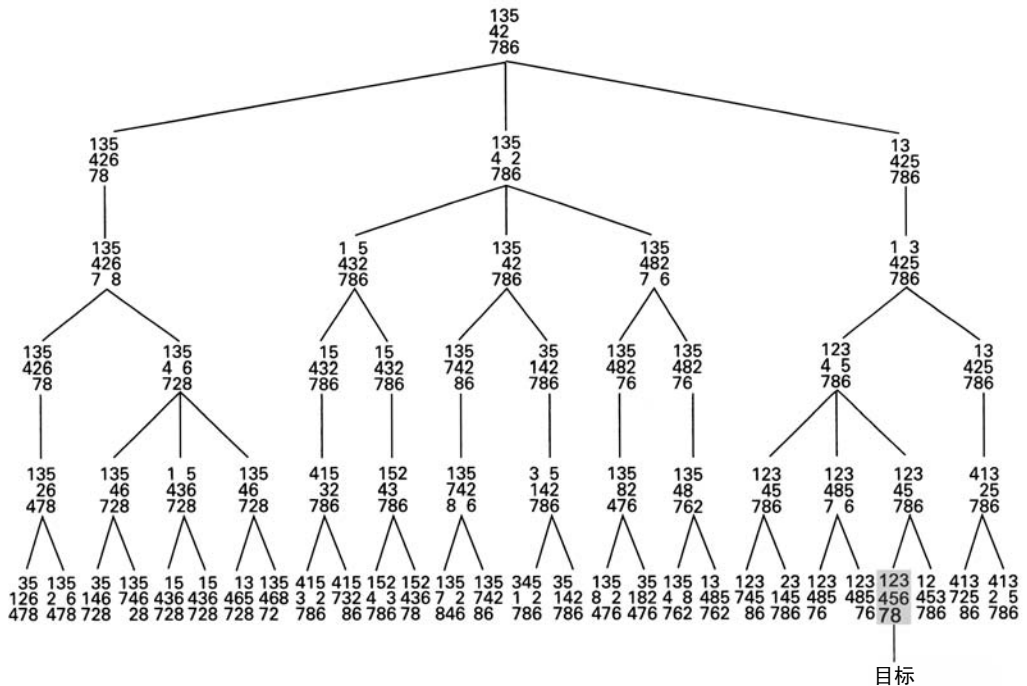


图10-7 搜索树的一个示例

目标状态出现在图10-7显示的搜索树的最下面一层。因为这预示了已经找到了一个解决方法，所以控制系统可以结束搜索过程并开始构建指令序列，该指令序列将用来解决外部环境中的拼图难题。说到底这只不过是一个简单的过程：从目标结点的位置上行，同时在遇到由树弧线表示的产生式时将它们压入栈。这种技术在图10-7所示的搜索树中的应用产生了如图10-8中所示的产生式栈。现在，只要把指令从该栈中弹出并执行，控制系统就能够解决外界的问题。

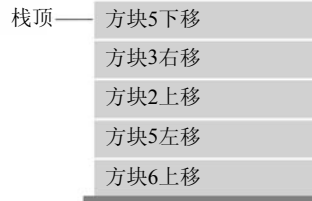


图10-8 压入栈的产生式以备后用

还有一点我们应当注意，回想第7章我们讨论的树，利用一个指针系统来指向下面的树，由此可以从一个父结点移动到它的子结点。但是在搜索树的情况下，控制系统必须能够从一个子结点移到其父结点，正如从目标状态向上移到开始状态。构建这种树的指针系统要向上指，而不是向下指。也就是说，每个子结点包含了一个指向其父结点的指针，而不是父结点包含了指向其子结点的指针。（有些应用中这两组指针都使用，允许在树中双向移动。）

10.3.3 启发

对于图10-7显示的例子，我们选择一个开始布局，产生了一个容易处理的搜索树。但是在试图解决一个比较复杂的问题时，产生的搜索树就会变得非常庞大。在国际象棋中，第一步就有20种可能，因此在这样的情况下，搜索树的根结点将会有20个子结点，而不是我们例子中的3个。而且，一局棋易手30至35次是常有的情况。即使是8数码游戏，若不能很快到达目标，搜索树也会变得非常大。结果，开发一个完整搜索树同表示出全部状态图一样都是不切实际的。

应对这种问题的一种策略是改变搜索树构建的次序。不用**广度优先**（breadth-first）的方式（这意味着树是一层一层地构建），我们可以沿着更有希望的路径往深度发展，只有在原来的选择不会导向成功时才考虑其他分支。结果就是以**深度优先**（depth-first）的方法构建搜索树，也就是说，树是以纵向路径的方式构建的，而不是以横向层次的方式。

深度优先的方法近似于人类面对8数码游戏时应用的策略。我们一般不会像广度优先方法那样，同时沿着几个可能的路径进行。相反，我们大概会选择看起来最有希望的路径并首先沿着这条路径走下去。注意，我们说的是“看上去”最有希望的。在一个特定点，我们很难确定哪个选择一定最佳。仅仅跟随知觉，当然可能误入歧途。但不管怎样，好于给与每种选择同等关注的蛮力方法，这种直觉信息似乎给了人类一个优势，所以在自动控制系统中应用直觉的方法似乎是明智的。

对于这样一个目标来说，我们需要一个方法来鉴别几个状态中哪一个看上去是最有希望的。我们的方法是采用**启发**（heuristic），这是给每个状态赋予的一个量化值，用来衡量这个状态与最近目标之间的“距离”。在某种意义上，一个启发值是衡量规划代价的一个尺度。给定的两个状态之间一个选择，那么从具有较小启发值的状态到达目标，显然花的代价小。因此，该状态代表了应遵循的方向。

一个启发值应具备两个特征。第一，如果到达相应的状态，它必须包含一个在所关联的状态到达后离最终解决还剩余工作量的合理估计。这意味着它在多个选项中做出选择时能提供有

意义的信息——启发提供的估计越精确，根据此信息所作的决定就越正确。第二，启发值应容易计算。这意味着它的利用应有益于搜索过程而非成为一种负担。如果计算启发值非常复杂，那么可能倒不如把时间花费在推导一个广度优先树。

在8数码游戏中，一个简单的启发要通过计算不在合适位置上的方块数目来估计到达目标的“距离”——这种推测指的就是一个有4个方块不在合适位置上的状态，相对于只有2个方块不在合适位置上的状态来说离目标更远（也就因此更缺少吸引力）。然而，启发并没有考虑方块离其位置有多远。假如这两个方块离它们相应的位置太远，就需要许多产生式来移动它们。

于是，有一个比较好的启发是测算每个方块离其终点的距离，并把这些值相加来得到一个量。一个直接邻近其终点的方块可伴有距离值1，而一个角与其终点方块相接触的方块可伴有距离值2（因为它至少要在垂直方向和水平方向各移动一次）。这种启发容易计算，并对拼图从最初状态到目的状态过程中需要移动的步数有一个粗略的估计。例如，图10-9所示布局相应的启发值为7（因为方块2、5和8每个离其终点的距离都为1，而方块3和6每个离终点的距离都为2）。实际上，它的确需要移动7步来完成拼图。

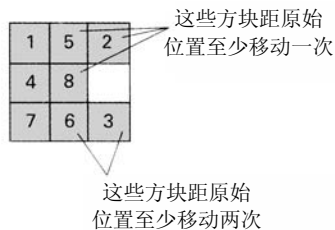


图10-9 一个未解的8数码游戏

既然我们有了8数码游戏的一个启发值，下一步就把它结合进决策过程。记得，一个人在做决定的时候倾向于选择看起来更接近目标的选项。所以我们的搜索过程应当考虑树中每个叶子结点的启发，并且从启发值最小的一个叶子结点进行搜索。这就是图10-10所采纳的搜索策略。图10-10中给出了开发一个搜索树并执行得到的解决方法的算法。

```

创建状态图的开始结点作为搜索树的根结点，并记录它的值
while (目标结点还没有到达) do
    [选择所有叶结点中有最小值的最左边叶结点
     将这个选定的结点作为子结点附加到通过单个产生式能到达的结点
     在搜索树中结点的旁边记录每一个新结点的值
    ]
从目标结点向上遍历搜索树，把与每个经过的弧相关联的产生式压入栈
通过执行从栈中弹出的产生式解决原始问题
    
```

图10-10 采用启发的控制系统的算法

让我们把这个算法应用到8数码游戏，从图10-6给出的初始布局开始。首先，我们建立初始状态并将它作为根结点，记录下它的启发值是5。然后，如图10-11所示，while循环体的第一次循环添加了3个从初始状态能达到的结点。注意，我们已经在结点下的括号里记录了每一个叶子结点的启发值。

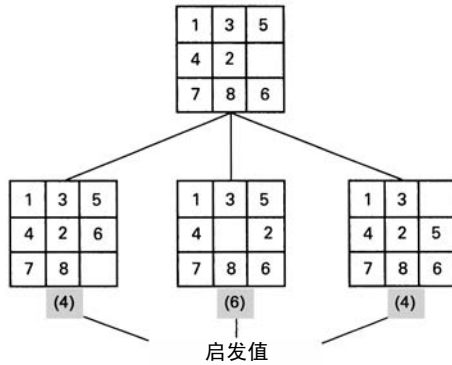


图10-11 启发的开始

目标还没有达到，所以我们再次执行while循环体，这次是从最左边的结点扩展搜索（“有最小启发值的最左边叶子结点”）。结果搜索树呈图10-12所示的形式。

现在，最左边叶子结点的启发值是5，说明这个分支也许最终不是一个好选择。算法注意到这一点，在下次经过时，循环指示从最右边结点扩展（它现在是“有最小启发值的最左边叶子结点”）。这样扩展后的搜索树如图10-13所示。

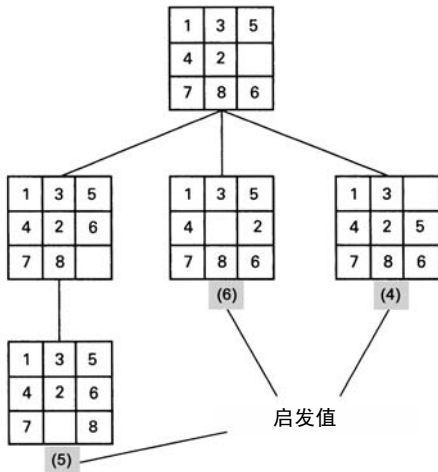


图10-12 两次搜索后的搜索树

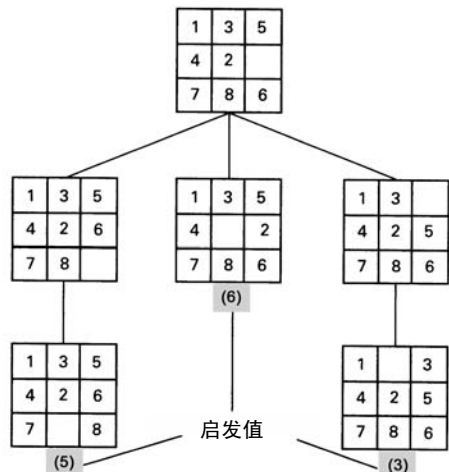


图10-13 三次搜索后的搜索树

这时，算法好像走上正轨。因为最右边结点的启发值只是3，while语句指示继续沿着这条路径进行，搜索直瞄目标，产生了如图10-14所示的搜索树。这个树同图10-7所示的搜索树相比较表明，新算法即使早期走了点弯路，但启发信息的利用已经大大减少了搜索树的大小，并且处理效率大大增加。

在到达目标状态之后，while语句终止，我们从目标结点反向向上移动到根结点，沿途把遇到的产生式压入一个栈。结果产生的栈如图10-8所示。

最后，当这些产生式从栈中弹出时，我们就被指示执行它们。这时，我们可以看到解决拼图难题的机器放下它的指杆，开始移动方块。

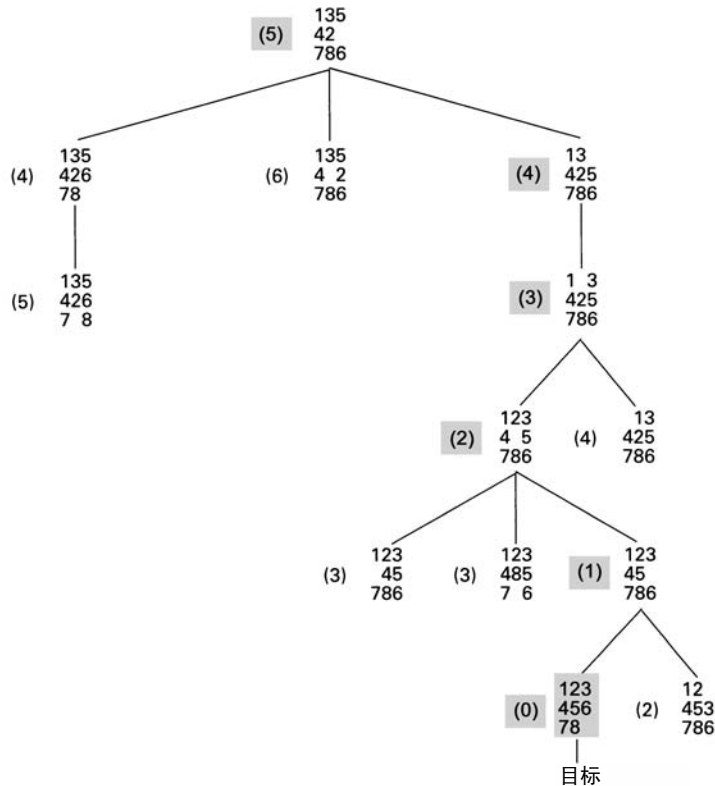


图10-14 用启发系统形成的完整搜索树

基于行为的智能

人工智能早期工作着手的课题涉及直接编写程序来模拟智能。但是，今天许多人认为，人类的智能并非基于复杂程序的执行，而是经历了世代进化而来的简单的刺激——反应功能。这种关于“智能”的理论称为基于行为的智能，因为“智能的”刺激——反应功能似乎是一些行为的结果，这些行为导致某些个体在其他个体遭难时得以幸免且繁殖。

基于行为的智能似乎能回答人工智能范畴的若干问题，例如，为什么基于冯·诺依曼体系结构的机器在计算能力上能轻易地胜过人类，却难以展现常识性的判断力。因此，基于行为的智能有希望成为人工智能研究中的一个重要的影响因素。正如文中描述的那样，基于行为的技术已经应用在人工神经网络领域，训练神经元如何按所期望的方式表现；应用在遗传算法领域，为更传统的编程过程提供一个可供选择的方法；应用在机器人学领域，通过反应策略来改进机器的性能。

问题与练习

1. 产生式系统在人工智能中有什么重要意义？
2. 画出8数码游戏中，围绕着代表下图状态的结点的那部分状态图：

4	1	3
	2	6
7	5	8

3. 使用广度优先搜索的方法，画出以下图为开始状态解决8数码游戏时控制系统构建的搜索树：

1	2	3
4	8	5
7	6	

4. 用笔、纸以及广度优先方法，构建出以下图为开始状态解决8数码游戏时所产生的搜索树（不必做完），你会碰到什么问题？

4	3	
2	1	8
7	6	5

5. 登山者为到达顶峰，只考虑当地地形，并总是沿着最陡峭的上坡行进，解决8数码游戏的启发式系统与登山者之间有什么相似之处？
6. 利用本节所讲的启发式方法，采用图10-10所示的控制系统算法，解决下面给出的8数码游戏：

1	2	3
4		8
7	6	5

7. 改进我们计算8数码问题中一个状态的启发值的方法，使图10-10所示的搜索算法不会做出错误的选择，就像在本节中的例子中那样。你能否举出一个例子，改进的启发仍然会导致搜索走入歧途？

10.4 其他研究领域

本节，我们来探究人工智能领域一直挑战研究人员的另外两个主题：知识处理和学习。这两个主题涉及的能力对人类大脑来说看上去很简单，但是对机器的能力来说却负担沉重。目前，本质上通过避免直接面对这些问题（或许通过应用聪明捷径或限制问题出现的范围）在开发“智能”体方面已经取得了很大的进展。

10.4.1 知识的表达和处理

在关于感知的讨论中，理解图像需要大量的关于图像细节的知识，理解一句话的意思可能要依赖其所处的上下文。这些都是知识仓库起作用的例子，这些称为**真实世界知识**（real-world knowledge），它们是由人脑维护的。人类以某种方式存储大量的信息并且以非凡的效率从信息中汲取有用的信息。赋予机器这种能力是人工智能面临的一个重要挑战。

潜在的目标是找到表示和存储知识的途径。这是很复杂的，就像我们已经看到的事实那样，知识以陈述性和过程性这两种形式出现。因此，知识表示不仅是事实的表示，而是包含了一个更广泛的领域。因此，能否最终找到一种用来表示所有形式的知识的单一方案是值得怀疑的。

然而，问题不仅仅是表示和存储知识。知识也必须是容易理解的，并且获取这种理解是一个挑战。10.2节介绍的语义网通常用来作为知识表示和存储的一种手段，但是，从中提取信息可能是有问题的。例如，句子“Mary hit John”的意思依赖于Mary和John的相对年龄（是2岁和

30岁或是相反?)。这种信息可能被存储在图10-3给出的完整的语义网中,但是在上下文分析的过程中,提取这样的信息需要对语义网进行大量的搜索。

处理知识存取还有另外一个问题是,知识的鉴别不是明确的,而是含蓄的,是与手头的工作相联系的。相对于直接用一个“没有”来回答问题“亚瑟赢了比赛吗?”,我们想要一个系统可以这样回答:“没有,他因为流感病倒了,没有完成比赛。”下一节我们将探究联想记忆的概念,这是试图解决有关系信息问题的一个研究领域。然而,任务不仅仅是找到有关系的信息,我们需要系统能够区分有关系的信息和相关联的信息。例如,“不,他是一月份出生,他妹妹的名字叫利萨。”这样的回答就不会被认为是对于前面问题的一个有意义的回答,即使该信息是通过某些相关的方式提交。

开发更好的知识提取系统的另一个方法是向提取过程插入各种各样的推理,结果产生了称为**元推理**(meta-reasoning)的方法,即关于推理的推理。举一个例子,数据库搜索最初是应用**封闭世界假设**(closed-world assumption)。这种假设是一个句子为假,除非它能够可能从可用的信息中明确得出。例如,封闭世界假设允许数据库做出Nicole Smith没有订阅特定的杂志这样的推断,即使数据库中没有包含任何关于Nicole的信息。处理过程就是观察Nicole Smith不在订阅列表中,然后应用封闭世界假设推断Nicole Smith没有订阅。

封闭世界假设在表面上看起来微不足道,但是结果证明,元推理并不是只发挥细微的、不合需要的作用。例如,假定我们仅有的知识是一句话。

Mickey is a mouse OR Donald is a duck.

孤立地看这个句子,我们不能推断出Mickey实际上是一只老鼠。因此,封闭世界假设强制断定句子

Mickey is a mouse.

为假。以相同的方式,封闭世界假设强制断定句子

Donald is a duck.

为假。这样,尽管两个句子中至少有一个为真,封闭世界假设已经引导我们得出了相矛盾的结论:两个句子都为假。理解这种元推理技术是人工智能和数据库这两个领域研究的一个目标,同时也强调了涉及智能系统开发的复杂性。

最后,有一个问题称为**框架问题**(frame problem),用来在变化的环境中使存储的知识保持最新。如果一个智能体打算使用它的知识来决定其行为,那么,该知识必须是当前的。但是,支持智能行为所需知识的数量是庞大的,在变化的环境中维护这些知识是一项繁重的工作。一个复杂的因素是,在一个环境中发生的改变经常会间接地改变信息中的其他细节,而且说明这种间接影响的结果是很困难的。例如,如果一个花瓶被敲碎了,尽管水洒了是打碎花瓶的唯一间接结果,但对于这种状况,你的知识不会再包括水在花瓶中这样的事实。因此,解决框架问题不仅需要以一种有效的方式存储和获取大量信息的能力,而且要求存储系统能够正确地反应间接的推论。

10.4.2 学习

除了表示和处理知识,我们还希望赋予智能体获取知识的能力。我们可以通过编写和安装一个新程序或者直接向数据库中添加数据来“教”基于计算机的智能体,但是我们更希望智能体能够自己学习。我们希望智能体能够适应环境的变化并执行任务,这些任务并不是通过事先简单地编写程序就能够完成的。一个为做家务而设计的机器人将面对新家具、新设备、新宠物甚至是新主人;一辆能自己驾驶的轿车必须能适应道路交通线的变化;博弈智能体应当能够开

发和应用新的策略。

一种把计算机学习的途径进行归类的方法是根据需要人类干涉的程度。学习的第一层是**模仿** (imitation)，人类直接演示一个任务的步骤（可能是通过执行一系列的计算机操作或是通过一系列动作将机器人移动），而计算机仅仅是记录这些步骤。这种形式的学习应用在像电子表格和字处理软件这样的应用程序中已经很多年了，在这些应用软件中，记录频繁发生的指令序列，然后通过一个请求就可以重放。注意，通过模仿学习智能体承担的任务很少。

学习的下一个层次是通过**监督学习** (supervised training)。在监督学习过程中，人对一连串的示例确定正确的反应，然后智能体对这些示例进行归纳，开发出一种适用于新案例的算法。这一连串的示例称为**训练集** (training set)。监督学习的典型应用包括学习识别一个人的笔迹或声音，学习区分垃圾邮件和受欢迎的邮件，以及学习如何从一组症状中验明疾病。

学习的第三个层次是**强化学习** (reinforcement)。在强化学习过程中，给予智能体一个一般规则，通过反复试验，使其在工作中当成功或失败时能自我判定。当胜负容易分辨时，强化学习对于学习如何玩国际象棋或跳棋这样的游戏是很有益的。相对于监督学习，当智能体学习改善其行为时，强化学习允许智能体独立行动。

因为还没有发现覆盖所有可能的学习行为的通用的学习规则，所以学习一直是一个有挑战性的研究领域。然而，有很多研究进展的例子。其中一个就是在卡内基-梅隆大学开发的ALVINN（基于神经网络的自动驾驶车辆）系统，该系统学习如何驾驶一辆配有一台车载电脑的大篷车，车载电脑使用一台摄像机作为输入。这里所采用的方法就是监督学习。ALVINN从人类驾驶员那里搜集数据并且利用这些数据调整自己的驾驶决策。像它所学习的，该系统可以预测向哪里驾驶，对照人类驾驶员的数据来检查自己的预测，然后修改自己的参数使其更接近人类的驾驶选择。ALVINN获得了很大的成功，它能够以每小时55英里的速度驾驶大篷车，同时引发了其他方面的研究，已经产生了可以成功在道路上驾驶的控制系统。

与开发单一智能体的学习技术相对应，另外一种对学习进行研究的方是开发一种智能体的后代能够通过一个进化过程来学习的技术。这个领域称为**遗传算法** (genetic algorithms)，它启发如何把自然进化理论用在智能体的开发上，其目标是通过应用“适者生存”法则设计问题的解决方法。建立和评估一组被提议的解决方案，然后选择和混合最好的方案，满怀希望地创造相对于原始集合具有改进的新一代解决方法。通过一遍一遍地重复该过程，目标是进化得到越来越好的解决方案，直到“学习”到一个成功的解决方案。

逻辑程序设计知识

在知识表示和存储中一个很重要的关注点就是采用何种方式与系统相匹配并使得系统能够存取知识。在这种背景下，逻辑程序设计（见6.7节）经常证明是有利的。在这样的系统中，知识用

Dumbo is an elephant.

和

X is an elephant implies X is gray.

这样的“逻辑”语句来表达。这样的语句能够用符号系统来表达，这对于推理规则的应用是易于实现的。正如在图10-5所见的演绎推理序列就可以采用直接的方式实现。因而在逻辑程序设计中，知识的表达和存储与知识的提取和应用过程很好地整合在一起。可以说，逻辑程序设计系统为知识的存储和应用之间提供了一个“无缝”边界。

把遗传算法应用到程序开发工作中的方法称为**进化程序设计**（evolutionary programming）。这里，目标是通过允许程序进化而不是直接编写程序来开发程序。研究人员已经使用函数编程语言把进化规则技术应用到程序开发过程中。这种方法以包含多种函数的一组程序开始，这个开始集合里的函数构成一个“基因池”，程序后代就是从这个基因池构建而来的。然后，希望通过上一代中最好的执行者所产生的后代经过很多代的进化最终得到目标问题的解决办法。

最后，我们应该认识一个与学习紧密相关的一个现象：发现。其区别是，学习是“基于目标的”而发现不是。名词“发现”含有意料之外的意思，不是现有的可以学习的。我们可以着手去学习一门外语或如何驾驶轿车，但是我们可能发现那些任务比我们想象得更加困难。一个探索者可能会发现一个大的湖，而目标仅仅是学习那儿有什么。

开发具有有效发现能力的智能体需要该智能体能够识别潜在的富有成效的“思考训练”。这里，发现在很大程度上依赖推理的能力以及启发的使用。此外，许多发现的潜在应用需要智能体能够区别有意义的结果和无意义的结果。例如，一个数据挖掘智能体不应当报告所发现的每一个微不足道的关系。

计算机发现系统中成功的例子包括以哲学家弗朗西斯·培根爵士命名的Bacon，它已经能发现（或许应该说是“重新发现”）电工学上的欧姆定律，行星运行的开普勒第三定律以及动量守恒原理。系统AUTOCLASS更有说服力，它采用红外光谱数据，已经发现了目前在天文学上未知的新型的恒星——计算机完成的一个真实的科学发现。

问题 与练习

1. 名词“真实世界知识”是什么意思？它在人工智能领域有何重要意义？
2. 一个关于杂志订阅者信息的数据库通常包含一个关于每一种杂志订阅者的列表，但是不包含没有订阅的人的列表。那么，这种数据库如何确定一个人没有订阅一种特定的杂志？
3. 概述框架问题。
4. 给出训练一台计算机的三种途径，哪一种没有涉及直接的人为干预？
5. 进化技术如何区别于更传统的计算机学习技术？

10.5 人工神经网络

伴随着人工智能所取得的所有进展，这个领域里的许多问题仍然使得基于冯·诺依曼体系的计算机的能力备受重负。执行指令序列的中央处理单元的感知和推理能力看来不能与人的大脑相匹敌。由于这个原因，许多研究者的目标转向了其他体系结构的机器。其中之一就是人工神经网络。

10.5.1 基本特性

如第2章介绍的那样，人工神经网络由许多单个的处理器以模仿活的生物体神经网络的方式构成，这些处理器称为**处理单元**（processing unit）。一个生物神经元是一个细胞，具有一些称为树突的输入触角和一个称作轴突的输出触角（图10-15）。经由一个细胞的轴突传递的信号反映了细胞是处于抑制状态还是兴奋状态。这种状态由细胞的树突接收到的信号的合成来决定。这些树突从其他细胞的轴突通过称为突触的小间隙采集信号。研究表明，一个突触的传导性是由突触的化学成分控制的。也就是说，具体的输入信号将对神经元起兴奋作用还是抑制作用是由突触的化学成分决定。所以可以认为，一个生物神经网络是通过调整神经元间的这些化学连接来学习的。

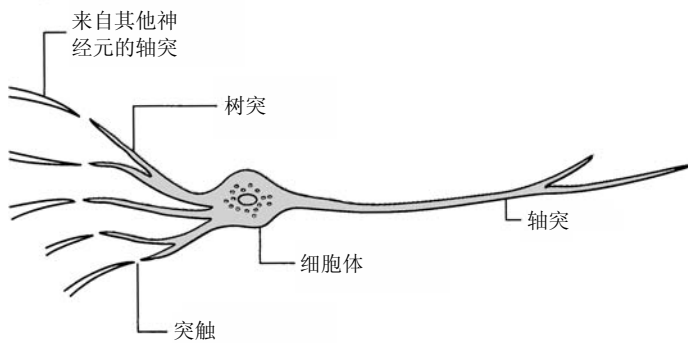


图10-15 活的生物体中的一个神经元

人工神经网络的一个处理单元是模仿对生物神经元这种基本了解的一个简单装置。根据其有效输入是否超过了一个给定的值（这个值称为处理单元的**阈值 (threshold value)**）产生0和1作为输出。如图10-16所示，这个有效输入是许多实际输入的一个加权和。图中，3个处理单元（记为 v_1 、 v_2 和 v_3 ）的输出用作另一个处理单元的输入。第4个单元的这些输入与称为**权 (weight)**的一些值（记为 w_1 、 w_2 和 w_3 ）相关联。接收单元把每个输入值与相应的权值相乘，再把这些乘积相加形成有效输入（ $v_1w_1+v_2w_2+v_3w_3$ ）。如果这个和超过该单元的阈值，那么该单元就产生一个输出值1（模拟神经元的兴奋状态）；否则就产生一个输出值0（模拟神经元的抑制状态）。

按照图10-16，我们采用矩形作为表示处理单元的约定符号，在单元的输入端为每个输入放置一个较小的矩形，并在矩形内写上与这个输入相关联的权值，最后在大矩形中央写上这个单元的阈值。图10-17的例子表示了一个有3个输入且阈值为1.5的处理单元。第1个输入的权值为-2，第2个输入的权值为3，第3个输入的权值为-1。因此，如果单元接收的输入分别为1、1、0，那么其有效输入为 $(1)(-2)+(1)(3)+(0)(-1)=1$ ，所以其输出为0。但是，如果单元接收的输入分别为0、1、1，那么其有效输入为 $(0)(-2)+(1)(3)+(1)(-1)=2$ ，超出了阈值，所以单元的输出为1。

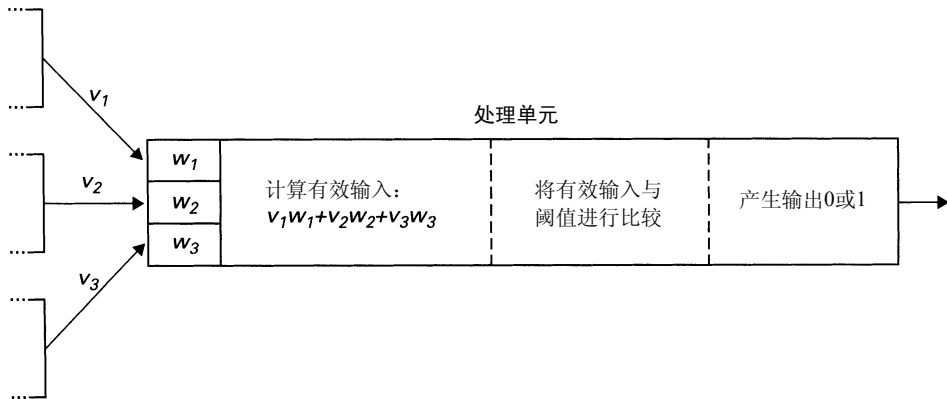


图10-16 一个处理单元中的活动

权值可以是正值，也可以是负值，说明相应的输入对接收单元的作用可以是兴奋或是抑制。（若权值为负，接收的输入为1就减少了加权和，故有效输入偏向低于阈值；相反，一个正的权值使相应输入对加权和起增加作用，故增加了加权和超过阈值的机会。）此外，权的实际大小控制了相应输入单元对接收单元起抑制作用还是兴奋作用的程度。因此，通过调节整个人工神经网络中的权值，就能够对网络编程，以预定的方式对不同的输入做出响应。

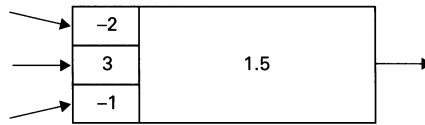


图10-17 一个处理单元的表达

图10-18给出了一个简单的神经网络的例子。图10-18a编程为：若两个输入不同，则产生输出1；否则输出0。但如果改变权值如图10-18b所示，那么这个网络无论其两个输入都是1，或者有一个是0，其响应都是1。

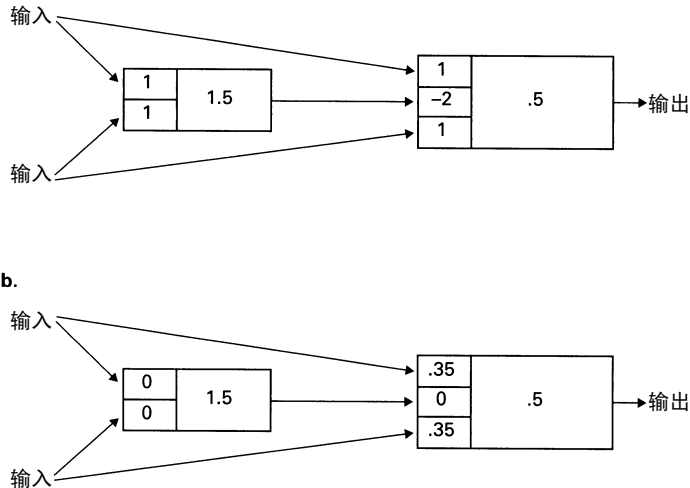


图10-18 有两个不同程序的神经网络

我们应当注意，图10-18所示的网络比实际的生物神经网络实在是太过于简单。一个人的大脑大约包含 10^{11} 个神经元，每个神经元约有 10^4 个突触。事实上，一个生物神经元的树突多得像一个纤维网，而不像图10-15中所表示的一个个触角。

10.5.2 训练人工神经网络

神经网络的一个重要特征是其不是传统意义上的被编程而是被训练。也就是说，一个程序员不再决定解决一个特定问题所需权的值，然后并把这些值“插”入网络中；相反，一个神经网络通过监督训练学习获得合适的权值，该训练是一个反复的过程，从训练装置而来的输入被应用到网络，然后用小的增量调整权值使网络的性能接近期望状态。如何调整权值是一个值得研究的课题。修改权值需要的一种策略是，每个新的调整都导向总目标而不是破坏前一步调整取得的进展。

为了说明这个问题，考虑培训如图10-19所示网络（其中所有权的值设为0）当其输入不同时正确地产生一个输出1的任务。也就是说，我们想要输入模式1,0和0,1产生输出1，而输入模式0,0和1,1产生输出0。（在图10-18a中，我们已经看到了对于这个问题的解决方法。）我们给两个输入都赋值1来开始这个训练过程。如图10-20a所示，我们观察到输出是期望的0，于是乎不管这个网络，尝试输入模式1,0（如图10-20b），继续训练过程。这次产生了输出0，而我们希望输出为1。我们通过把第二个过程单元上面的权值改为1（如图10-20c所示）来调整这个值。现在，网络正确地执行了输入模式1、0。这时，我们回去再次尝试输入模式1、1。但令人失望的是，网络不再正确地处理那个模式。实际上，网络现在产生了一个输出1（如图10-20d）。通过把第

二个过程单元上面的权值改回0（如图10-20c），我们调整这个值。我们又回到了起点，继续这个过程仅是带着我们经过一个无穷尽的训练循环，我们所做的每一个修改都与前一个修改抵消了。

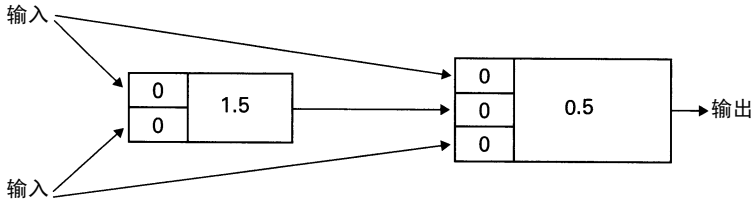
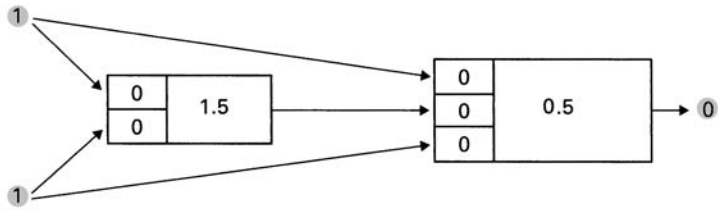
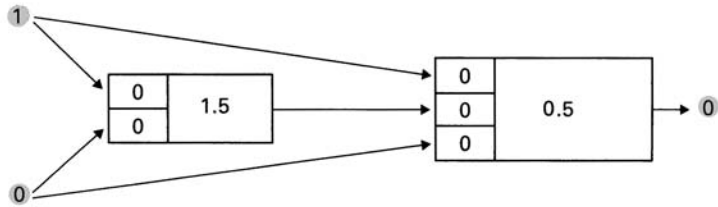


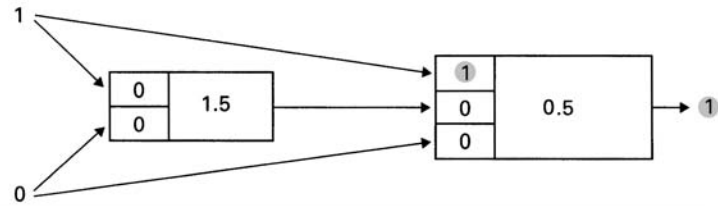
图10-19 一个神经网络



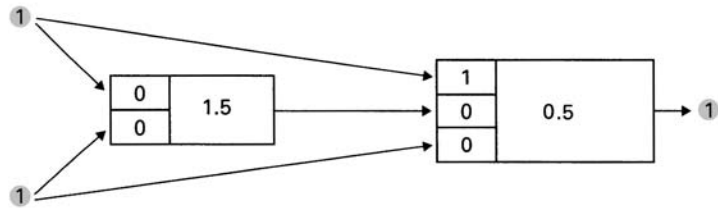
a. 网络正确地执行输入模式1,1。



b. 网络错误地执行输入模式1,0



c. 第二个处理单元上部的权被调整



d. 但是，网络不再正确执行输入模式1,1

图10-20 训练一个神经网络

幸运的是，如何开发成功的训练策略已经取得了重大的进展，前一节引用的ALVINN项目论证了这一点。实际上，ALVINN是一个人工神经网络，如图10-21所示，其构成出奇地简单。ALVINN从 30×32 阵列的传感器中获得输入，每个传感器负责观察前面道路视频图像的一个特定部分，并且向4个处理单元的每一个报告其发现。（从而，这4个单元的每一个都有960个输入。）4个单元的每一个的输出与30个输出单元的每一个相连接，其输出预示了驾驶的方向。这30个单元行一端处于兴奋的处理单元指示了一个向左急转弯，而另一端处于兴奋的单元则指示了一个向右急转弯。

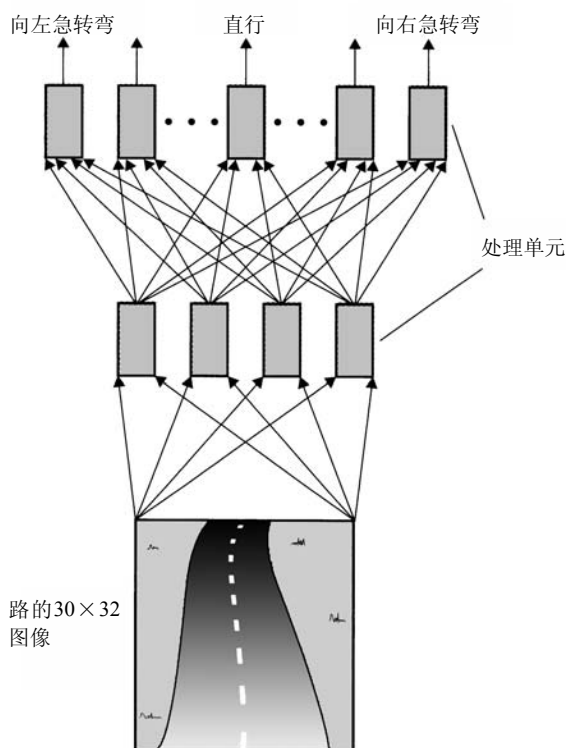


图10-21 ALVINN的结构

ALVINN的训练是通过“观察”一个人的驾驶进行的。当它要做出自己的驾驶决定时，它把自己的决定与人的决定相比较，并且稍微修改其权值使其更接近人的决定。然而，存在一个有趣的附加问题，尽管ALVINN通过它的简单技术学习驾驶，但是它没有学习如何从错误中恢复过来。因此，从人那里收集的数据也要人工强加以包含恢复状态。（这种恢复训练的一种方法是把它在最初就考虑进去，使人令交通工具偏离方向，以便ALVINN可以通过观察人来学习如何恢复。除非当人完成初始的偏离过程时ALVINN不可用，否则，ALVINN可以学会偏离以及恢复——显然，这并不是一个受欢迎的方法。）

10.5.3 联想记忆

人脑具有惊人的能力，能够从当前关心的情景中提取与之关联的信息来。当闻到特定气味时，我们很容易勾起对儿时的回忆；朋友的声音会唤起友人的身影和一段美好时光的回忆；特定的音乐可能会产生对某个假日的怀念。这些就是**联想记忆**（associative memory）的例子——提取与手头信息相关联的或相关的信息。

构建具有这种联想记忆能力的机器是许多年来研究的一个目标。途径之一是应用人工神经网络技术。例如，考虑一个由许多处理单元组成的网络，这些处理单元相互连接形成一个没有输入和输出的网。（有些设计中，一个单元的输出连到其他单元，作为每个单元的输入，这种设计称霍普菲尔德网络（Hopfield network）。在其他一些设计中，一个单元的输出可能只连到与其直接相邻的单元。）每个单元都可以处于兴奋状态，也可处于抑制状态。如果用1表示兴奋状态，用0表示抑制状态，那么整个网络的状况可以想象成一个1和0的配置。现在假定以这样的方式对该网络进行编程，也就是使得某些0和1配置是稳定的，也就是说，当网络发现自己处在这几种配置之一时，它就保持在那个配置状态。如果网络所处的配置不稳定，那么处理单元之间的互动将引起配置改变，并且一直改变，直到变成一个稳定配置为止。

如果网络一开始处于一个与某个稳定配置很近的非稳定配置，那么我们希望网络会向那个稳定配置行进。在某种意义上，当给定一个稳定配置的一部分时，网络能够完成这个配置。或者，换句话说，网络能够找到与给定部分模式相关联的位模式。所以，如果一些位用来编码成“气味”，另一些位用来编码成“儿时回忆”，那么，根据某个稳定配置初设的“气味”，能够导致其余的位找到关联的“儿时回忆”。

下面我们考虑图10-22所示的人工神经网络。图中每个圆圈代表一个处理单元，其阈值记于圆中。连接圆圈的线代表相应单元间的双向连接。也就是说，一条连接两个单元的线表示每个单元的输出连到另一个单元作为输入。因此，中央单元的输出连到其周边每个单元作为输入，而周边每个单元的输出也都连到中央的单元作为输入。两个相连的单元相互的输出都有相同的权值。这个共同的权值记在连接线旁。于是，图中顶部那个单元从中央单元接收的输入伴有权值-1，从其两个周边邻居接收到的输入伴有权值1。类似地，中央单元从周边各单元接收的输入伴有权值-1。

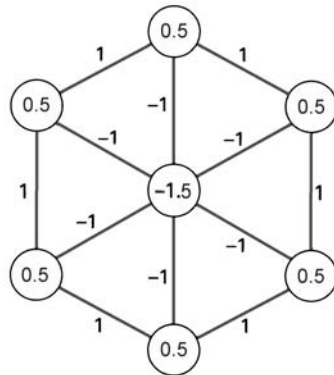


图10-22 实现联想记忆的一个人工神经网络

网络以离散的步骤运转，每一步，所有的处理单元都以同步方式对其输入作出响应。为了从网络的当前配置确定其下一步配置，我们要确定整个网络中每一个单元的有效输入，再让所有的单元同时响应其输入。结果，整个网络遵循一个协调的顺序运作：计算有效输入，响应输入，计算有效输入，响应输入，依次类推。

如果网络初始化为最右边两个单元为抑制状态，其他单元为兴奋状态（见图10-23a），我们来考虑会发生的一系列事件。最左边两个单元有效输入为1，所以保持兴奋；但它们周边的邻居有效输入为0，所以会变成抑制。类似地，中央单元有效输入为-4，所以变成抑制。于是，整个网络转变成图10-23b所示的配置，只有最左边两个单元兴奋。因为中央单元现在抑制，所以最左边两个单元的兴奋状态将导致顶部和底部两个单元再变成兴奋。同时，因为有效输入为-2，

中央单元继续保持抑制。于是网络转变成图10-23c所示的配置，然后它又导致了图10-23d所示的配置。（如果网络初始化为只有上面4个单元兴奋，那么会出现一种闪烁现象。顶部单元保持兴奋，而其2个周边邻居及中央单元会在兴奋与抑制两种状态间不断切换。）

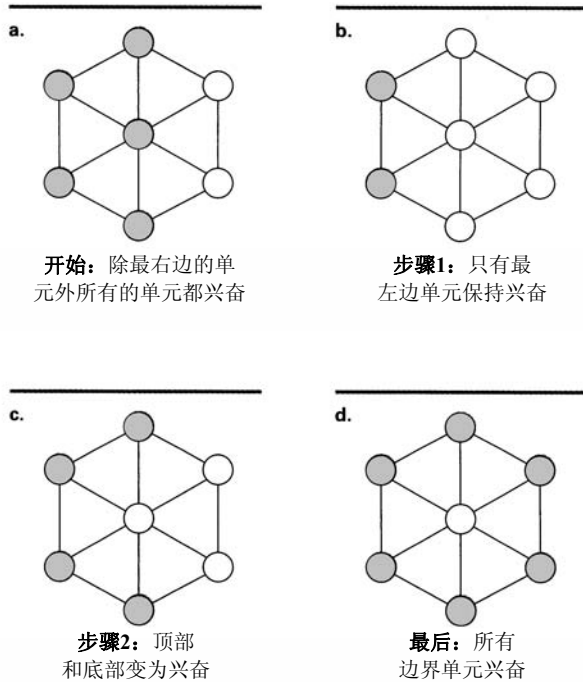
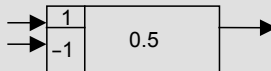


图10-23 导向稳定配置的步骤

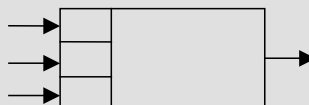
最后，我们观察到这个网络有两种稳定配置：一种是中央单元兴奋，而其他单元抑制；另一种是中央单元抑制，而其他单元兴奋。如果网络初始化为中央单元兴奋而其他单元不会有2个以上兴奋，那么网络会走向前一种稳定配置。如果网络初始化为至少4个相邻周边单元兴奋，那么网络会走向后一种稳定配置。所以，可以说，这种网络，如果初始模式为中央单元及少于3个周边单元处于兴奋状态，就与前一种稳定配置关联；如果初始模式为4个或4个以上周边单元处于兴奋状态，那么就与后一种稳定配置相关联。简单地说，这个网络呈现一种初步的联想记忆。

问题与练习

1. 对于下面的处理单元，当两个输入都是1时，输出是多少？如果输入模式是0,0、0,1以及1,0呢？



2. 调整下面的处理单元的权值和阈值，使得当且仅当至少有两个输入为1时输出为1。



3. 举出在训练一个神经网络时可能会发生的一个问题。
4. 图10-22中，如果初始化为所有处理单元抑制，那么网络会走到哪个稳定配置？

10.6 机器人学

机器人学（robotics）是研究具有智能行为的物理上的自主智能体的一门学科。对于所有的智能体，机器人在所处的环境中必须能够感知、推理和发生作用。因此，机器人学涵盖了人工智能的所有研究范围，并在机械和电子工程方面引起了巨大的反应。

机器人需要用机械装置来回移动和操作目标物体来与外界交互。在早期的机器人学中，该领域与操作器械的发展机密联系，这些操作器械通常是带有肘、腕及手或工具的机械臂。研究不仅涉及这样的装置如何操作，而且涉及如何维护和应用有关它们的定位和定向的知识。（你闭上眼睛也能够用手摸到你的鼻子，因为你的大脑保存有你的鼻子和手指在什么地方记录。）随着时间的推移，机械臂已经越来越灵巧，使用基于强反馈的触觉，机械臂能够成功地握住鸡蛋和纸杯。

最近，快速、轻便计算机的发展引发了移动机器人方面更重大的研究。这种灵活性的实现导致了大量的富有创意的设计。在机器人移动能力方面，研究人员已经开发出可以像鱼一样游、像蜻蜓一样飞、像蝗虫一样跳跃、像蛇一样蜿蜒爬行的机器人。

因为带有轮子的机器人相对容易设计和建造，所以非常受欢迎，但是它受到了可以穿过的地形的限制。使用轮子和导轨的联合体，克服这种限制爬楼梯或翻越岩石是当前的研究目标。例如，美国国家航空航天局的火星探路者号就是使用特殊设计的轮子在火星的岩石层上行走。

有腿的机器人提供了较大的可移动性，但是相当复杂。例如，设计能像人一样行走的两条腿机器人必须持续地监视和调整其姿态，否则它将会跌倒。但是，这种困难能够被克服，例如本田公司开发的两条腿的具有人的特征的机器人Asimo，能够上楼梯，甚至能够跑。

尽管在操作器械和移动能力方面取得了巨大的进步，但是大多数机器人仍然不是非常自主的。工业机械人手臂是为每个任务专门严格设计的，工作时不用传感器，它假设零件将会按照指定的位置被精确地传送给它们。其他的移动机器人（如美国国家航空航天局的火星探路者号和军用无人机）其智能依靠人的操作来实现。

克服这种对人的依赖是当前研究的一个主要目标。一个问题涉及一个自主的机器人需要知道关于其所处环境的哪些知识，以及需要预先计划其行为到什么程度。建造机器人的一个方法是维持所处环境的详细记录，该记录还包含目标物体的一个详细目录以及它们的方位，通过这些信息制定行动的详细计划。这个方向的研究很大程度上依靠知识表示和知识存储的进展以及推理和规划技术的改进。

另一个可选择的方法是开发反应型机器人，该方法不用保持复杂的记录以及在构建详细行动计划上耗费大量的精力，只需应用简单的与外界交互的规则时时刻刻指导它们的行为。反应型机器人技术的支持者认为：当计划一个长途汽车旅行时，人类不会预先制定全面而详细的计划；相反，他们仅是选择主要路线，而对于像到哪儿吃饭，走哪些出口，以及如何绕道行驶等细节到时候考虑。同样，一个需要通过一条拥挤的走廊或从一栋大楼走到另一栋大楼的反应式机器人不会预先制定非常详细的计划，但是当碰到障碍时，它会应用简单的规则避开每一个障碍。这是历史上最畅销的机器人——iRobot Roomba真空吸尘器所采用的方法，真空吸尘器以反应模式在地面上来回移动，而不会为记住家具的详细信息和其他障碍而费心。毕竟，家庭宠物下次不可能在同一个地方。

当然，单一的方法并不是对于所有情况都是最好的。真正的自主机器人最有可能是使用多层标准的推理和计划，应用高层技术设定和达到主要目标，低层反应系统完成次要目标。这种多层次推理的例子可在Robocup比赛（一个机器人足球队的国际性比赛）中发现，该比赛为到2050年开发能够对抗世界级人类足球队的机器人足球队的研究提供了一个场合。这里，重点不

是仅仅建造能够“踢”球的移动计算机，而是设计一个能够相互协作达到共同目标的机器人足球队。这些机器人不仅要移动和对自己的行为作出推断，而且它们还要对队友和对手的行为作出推断。

机器人学研究领域的另一个例子是称为进化机器人学的领域，把进化理论应用于开发低级反应规则和高级推理。这里我们发现，适者生存理论用到了设备的开发上，经过若干代的学习，这些设备能够自己获得平衡或移动的方法。关于这个领域的许多研究不同之处在于机器人的内部控制系统（很大程度上是软件）及其形体的物理结构。例如，一个能游泳的蝌蚪机器人的控制系统换成一个有腿的类似机器人。然后在控制系统中应用进化技术，得到一个能爬行的机器人。在其他的例子中，进化技术已经被应用在机器人的物理形体上，让传感器发现执行特定任务的最佳位置。更具有挑战性的研究正在寻求软件控制系统与形态结构同时进化的途径。

要列出机器人学研究带来的所有令人难忘的成果是一项太大的任务，当前的机器人与科幻电影和小说中的超能机器人相差甚远，但是在执行特定任务上已经取得了重大的成功。我们使机器人能够驾驶交通工具，像宠物狗一样表现，为武器导航。然而，享受这些成功的同时，我们应该注意，对人造宠物狗的钟情以及智能武器的可怕威力带来了社会问题和伦理问题，这些都向社会发出了挑战。我们的未来是我们自己造就的。

- 问题与练习**
1. 对于机器人行为的反应方法在何种方式上有别于更传统的“基于计划”的行为？
 2. 当前机器人学领域研究的几个主题是什么？
 3. 进化理论用在机器人开发有哪两个层次？

10.7 后果的思考

毫无疑问，人工智能的进展有造福人类的潜能，人们很容易热衷于这些潜在的好处。然而，将来也隐藏着潜在的危險，它的破坏性后果与其有利的那一面同样巨大。这种差异常常仅在于一个人的观点或一个人的社会地位的不同——彼之所得，此之所失。所以花一点时间从另外一个角度观察正在进步的技术对于我们来说是比较恰当的。

有些人把技术的进步看成是给与人类的一份厚礼——将人类从枯燥的、普通的任务中解放出来，为更愉悦的生活方式打开大门的一种方式。但对于同一个现象，另一些人则把它看作是剥夺公民就业机会、把财富引向权势人物的祸根。其实，这正是印度忠诚的人道主义者圣雄甘地所预言的。甘地再三地称辩，如果用农夫家庭手纺车来代替大型纺织工厂，那么印度人的生活将会变得更好。他断言，通过这个途径，可以用一个分散的大宗生产系统取代只能雇用少数人的集中式大宗生产，这将有利于平民大众。

历史上有很多因财富和权力分配不均而引起的革命。如果今天正在进步的技术使得这种悬殊更为巩固，那将产生灾难性的后果。

但是，建造越来越智能的机器的后果，比对付不同社会族群间的权力斗争的后果更加微妙——更加根本。这些问题震撼了人类自身形象的核心。19世纪，达尔文的进化论及人类可能由更低等的生命形式进化而来的想法震惊了整个社会。那么，面对机器的心智能力向人类挑战的冲击，社会将如何反应呢？

过去，技术发展缓慢，有时间让我们重新调整智能的概念，维护人类的自我形象。19世纪，我们的老祖宗会认为当时的机械装置具有超自然的能力，而今天我们决不会认为这些机械有什么智能。但是，如果机器真的挑战了人类的智能，或者更有可能的是机器能力的进步超过了我们的适应能力，那么人类将如何应对呢？

考虑一下20世纪中期社会对当时IQ测试的反响，也许可以从中得到一些线索，看看人类面对挑战我们智能的机器时的潜在反应。这些测试被认为可以用来确定孩童的智力水平。美国的孩童常常依据他们在测试中的表现来分类，并据此制定相应的教育计划。随之，受教育的机会向那些在测试中表现良好的孩子开放，而那些测试表现差的孩子只能安排去参加补习计划。简而言之，当给出一种尺度来衡量个体的智能时，社会会倾向于漠视被认为是低于这个尺度的那些人的能力。那么，如果机器的“智能”能力已经变得可与人类相匹敌，甚至只是看上去可以相匹敌，社会将怎样应对这种局面呢？抑或社会也漠视那些能力看上去不如机器的人？如果这样，对于社会的这些成员来说，后果是什么？难道一个人的尊严取决于他（她）与机器比较的结果吗？

我们已经开始看到，在一些特定场合，人类的智力正面临机器的挑战。机器现在有能力打败棋王；计算机化的专家系统能够给出治疗意见；管理证券投资的简单程序常常比投资专家做得更好。这样的一些系统怎样影响所涉及人员的自身形象？随着在越来越多的领域里个人被机器胜出，个人的自尊心会受到怎样的影响？

由于人是生物，而机器不是，所以许多人认为机器拥有的智能与人类的智能有着本质的区别。因此他们认为机器永远不会再生出人类的决策过程。机器也许会得到与人同样的结论，但是得到这些结论所依赖的基础与人类并不相同。那么在怎样的程度上存在不同类型的智能？对于社会来说，如果遵循非人类智能提出的道路运作，是否合乎道德？

Joseph Weizenbaum在他的*Computer Power and Human Reason*一书中坚决反对不加抑制地应用人工智能，他这样写到：

计算机能够作出司法判决，计算机能够作出精神病判定。它们能够以比最有耐心的人更加老练的方式投掷硬币。关键在于不应当给它们这些任务。在某些场合，它们甚至能够得到“正确的”决策——但是其依赖的基础总是而且一定不是人类所乐于接受的。

已经有很多有关“计算机与人脑”的争辩。我在这里的结论是，问题的实质不在于技术，甚至也不在于教学，而在于伦理道德。设定计算机去做什么，不能用“能够不能够”这样的问题，计算机适用性的限制最终只能根据“应当”来表达。最基本的认识应该是：因为我们现在还没有办法让计算机有智慧，那么我们现在就不应当让计算机去做有智慧的工作。

也许你会认为本节所述的许多内容近乎科幻小说，而不是计算机科学。就在不久前，许多人因抱有同样的“这永远不会发生”的态度，而拒绝考虑“如果计算机操纵了社会，会发生什么？”。但从许多方面来看，这一天现在已经来临。如果一个计算机化的数据库错报了你有不良的信用度、有犯罪记录或是银行账户透支，那么，是计算机的报告会奏效还是你自己的清白申诉会奏效？如果一个不正常的导航系统错误指示了大雾笼罩的跑道位置，那么飞机将降落在何处？如果一个机器用来预测公众对不同政治决策的反应，那么一个政治家应采取何种决策？你遇到过多少次因为“计算机坏了”所以服务员无法为你服务的情形？那么，究竟谁掌管着这个社会？我们还没有准备让社会屈从于机器吗？

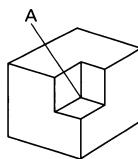
问题 与练习

1. 如果把过去100年来发明的所有机器都去掉，那么今天的人还有多少能幸存？如果是过去50年呢？20年呢？幸存者会在何处？
2. 你的生活在多大程度上被机器所控制？谁又控制着这些影响你生活的机器？
3. 你从哪里获得那些你的日常决策赖以为基础的信息？对于你的重大决策呢？对这些信息的准确度你有多少把握？为什么？

10.8 章复习题

(带*的题目涉及选读小节的内容)

1. 正如10.2节说明的那样,人类会用一个问题来表达某个目的,而不是提问。另一个例子,“你知道你的轮胎漏气了吗?”,这也是用来提醒而不是问。给出一些问题的例子,目的是用来表达安慰、警告或责备。
2. 把一个苏打水剂量器当作一个智能体来进行如下分析:它的传感器是什么?它的效应器是什么?它可以展现什么级别的反应(本能反应、基于知识或基于目标)?
3. 确定下列每一个反应,是本能反应基于知识的反应还是基于目标的反应。论证你的回答。
 - a. 门一开,电冰箱里的灯就亮了。
 - b. 一个计算机程序把文本从德文翻译成英文。
 - c. 一个登山者计划登山路径。
4. 如果一个研究人员使用计算机模型来研究人脑的记忆能力,那么为机器所开发的程序必定要达到机器的最佳存储能力?请解释。
5. 举出几个陈述性知识的例子。举出一个过程性知识的例子。
- *6. 在面向对象程序设计的环境中,一个对象的哪些部分是用来存储陈述性知识?哪些部分用来存放过程性知识?
7. 下列活动中,你认为哪些是面向性能的?哪些是面向模拟的?
 - a. 一个飞行模拟器的设计。
 - b. 一个自动导航系统的设计。
 - c. 一个处理图书馆藏书的数据库的设计。
 - d. 一个用于理论测试的国家经济模型的设计。
 - e. 一个用于监视病人生命体征的程序的的设计。
8. 当今,许多商用电话的呼叫都采用了自动应答系统,系统根据打电话人的选项直接呼叫。这些系统通过了图灵测试吗?请解释你的答案。
9. 确定能用来区分符号O、G、C、Q的一组几何特征。
- *10. 请描述通过与模板比较鉴别特性的技术与第1章介绍的利用纠错码鉴定特性的技术之间的相似之处。
11. 根据下面线绘图中标记A的那个角是凸起还是凹下,说明这个图两种解读。



12. 比较下列两个句子中介词短语的作用(仅有一个词的不同)。如何对一台机器编程让它做这样的区分?

The pigpen was built by the barn.

The pigpen was built by the farmer

13. 下面两个句子的语法分析的结果有什么不同?语义分析的结果有什么不同?

Theodore rode the zebra.

The zebra was ridden by Theodore

14. 下面两个句子的语法分析的结果有什么?区语义法分析的结果有什么不同?

If X=5 then add 1 to X else subtract 1 from X.

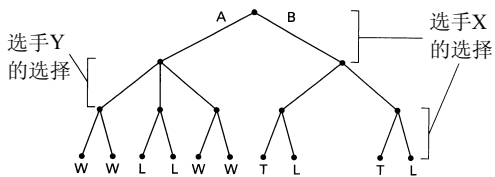
If X<>5 then subtract 1 from X else add 1 to X.

15. 正文中,我们与形式程序设计语言相比,简要讨论了理解自然语言的问题,作为讨论自然语言案例所涉及的复杂性,给出问题“Do you know what time it is?”有不同含义的情形。
16. 一个句子上下文的改变能够改变这个句子的含义以及意思。图10-4的上下文中,如果两人都出生于20世纪60年代,那么句子“Mary hit John.”的含义怎样改变?如果一个出生在20世纪60年代,而另一个出生在20世纪90年代,那么含义如何改变?
17. 画一个语义网,把下列段落的意思表示出来。

Donna threw the ball to Jack, who hit it into center field. The center fielder tried to catch it, but it bounced off the wall instead.

18. 有时候回答一个问题的能力,更多依赖于对该知识的限度的了解,而不是对事实本身的知晓。例如,假定数据库A和B都包含一个完整的雇员名单,该名单与公司健康保险程序相关联。但是只有数据库A知道名单是完整的。那么关于一个不在名单里的员工,数据库A能够推断出什么信息是数据库B做不到的?

19. 举出一个封闭世界假设导致矛盾的例子。
20. 举出两个例子，共用一个封闭世界假设。
21. 在产生式系统中，状态图和搜索树有什么区别？
22. 依照一个产生式系统分析解决魔方问题的任务（什么是状态？什么是产生式，等等？）。
23. a. 假定搜索树是一个二叉树，达到目标需要10个产生式。如果该树是以广度优先的方式构建的，那么当达到目标状态时，树中最大的结点数是多少？
b. 解释通过同时构建两个搜索如何能够减少搜索过程中考虑的全部结点数——一个搜索从初始状态开始，同时另一个搜索从目标状态逆向进行直到这两个搜索会合。（假设记录在逆向搜索过程中发现的状态的搜索树也是一个二叉树，并且两个搜索以相同速度进展。）
24. 正文中我们提到，产生式系统通常被用来作为从已知事实中得出结论的一种技术。系统的状态是推理过程的每一个阶段认为是真的事实，产生式对于操纵已知事实来说是逻辑规则。标识几个逻辑规则，使从事实“John is a basketball player”，“Basketball player are not short”，以及“John is either short or tall”中能够得出结论“John is tall”。
25. 下面的树表示一个竞赛游戏中可能的移动，选手X当前可在移动A和移动B中选择其一。选手X移动后，选手Y跟着选一移动，然后选手X在跟着选择最后一步。树的叶子结点标记为W、L、T，分别代表选手X最后是赢、输还是平局。选手X应选择移动A还是移动B？为什么？在一个竞赛性的游戏中选取一个“产生式”和一个如8数码游戏这种单人游戏中的选取有什么不同？



26. 按照产生式系统分析跳棋游戏，并描述一个用来在两个状态中确定一个更接近目标的启发。这种情况中的控制系统与一个如8数码游戏这种单人游戏中的控制系统有什么区别？
27. 把代数定律看作产生式，代数式简化的问题就能在产生式系统的上下文中解决。确定一组代

数产生式，使等式 $3/(2x+1)=2/(2x-2)$ 简化为 $x=4$ 。当进行这种代数简化时，一些经验法则（即启发的法则）是什么？

28. 不用任何启发信息的帮助，画出利用广度优先搜索方法解决如下初始状态的8数码游戏生成的搜索树。

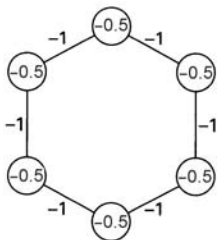
	1	3
4	2	5
7	8	6

29. 利用图10-10的算法解决第28题的8数码游戏，用未到位的方块的数目作为启发信息，画出搜索树。
30. 利用图10-10的算法解决如下初始状态的8数码游戏，假设使用与10.3节中一样的启发信息，画出搜索树。

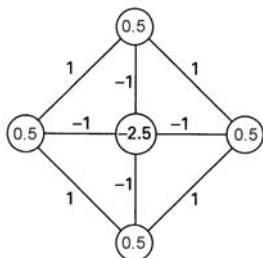
1	2	3
5	7	6
4		8

31. 当解决8数码游戏时，为什么用未到位的方块的数目作为启发信息不如10.3节用的那种好？
32. 执行二叉树搜索（5.5节）时决定考虑哪一半列表的技术，和执行一个启发时决定要执行哪个分支的技术，二者有什么不同？
33. 注意，如果一个产生式系统的状态图中有一个状态的启发值与其他状态相比极其低，并且如果从这个状态到自己有一个产生式，那么图10-10的算法会陷入一个循环，一遍又一遍地考虑这个状态。说明如果执行该系统中任何产生式的代价至少为1，那么把启发值加上沿正遍历的路径到达该状态的代价，通过这样计算预测的代价，就可以避免这种无限循环。
34. 在一幅大的交通图上寻找两个城市间的道路，你会用怎样的启发。
35. 列出可用于产生式系统的启发所具有的两个特性。
36. 假定有两个桶，一个容量是3升，一个容量是5升。任何时候你都可以把水从一个桶倒入另一个桶，把一个桶倒空，或把一个桶倒满。问题是要将正好4升的水注入5升的那个桶。说明这个问题如何可以设计成一个产生式系统。

37. 假设你的任务是监督两辆卡车装货,每辆车最多可载14吨货。货物装在不同的筐里,总重28吨,但每一筐的重量不一样,都标在各自的筐边上。为了在两辆车上分装这些货物,你会采用什么样的启发?
38. 下列哪些是元推理的例子?
- 他还没走多长时间所以没走远。
 - 因为我经常做出错误的决定,而所作的最后两个决定是正确的,那么我将逆转下一个决定。
 - 我有些疲倦了,所以我想我将要打个盹。
 - 我有些疲倦了,所以我可能不会清晰地思考。
39. 描述人类解决框架问题的能力如何帮助人类找到丢失的项目。
40. a. 通过模仿学习与通过监督学习在何种意义上相似?
b. 通过模仿学习与通过监督学习在何种意义上不同?
41. 下图表示一个用于10.5节讨论的联想记忆的一个神经网络。如果模式中只有2个单元兴奋,而这2个单元被一个单元分开,那么它与什么模式相关联?如果网络初始时所有单元都兴奋,会发生什么情况?



42. 下图表示一个用于10.5节讨论的联想记忆的一个神经网络。如果初始模式中至少有3个单元兴奋,而中央单元抑制,那么它与怎样的稳定配置相关联?如果初始模式中只有2个相对的周边单元兴奋,那么将会发生什么情况?



43. 设计一个用于联想记忆(10.5节所讨论的)的人工神经网络,它由一个处理单元矩形队列组成,要移动到这样的稳定模式,其中一个纵列的单元都兴奋。
44. 调整图10-19所示的人工神经网络中的权值和阈值,使其在2个输入相同(全为0或全为1)时输出为1,2个输入不同(一个为0另一个为1)时的输入为0。
45. 画一个与图10-6类似的图,表示把代数式 $7x+3=3x-5$ 简化为 $x=-2$ 的过程。
46. 详述上题的答案,说明其他路径,解题时可遵循一个控制系统。
47. 画一个与图10-6类似的图,表示从初始事实“Polly is a parrot”,“A parrot is a bird”以及“All birds can fly”中得到结论“Polly can fly”的推理过程。
48. 与上题中的句子不同,有些鸟不会飞,如鸵鸟或是折了翅膀的鸟。但是,要建立一个演绎推理系统,其中把对陈述“All birds can fly”的所有例外都明确列出,看来并不合理。那么,我们作为人类如何确定一只鸟是能飞还是不能飞?
49. 详述句子“I read the new tax law”在不同上下文中的不同含义。
50. 说明怎样能够把从一个城市旅行到另一个城市的问题设计成一个产生式系统。什么是状态?什么是产生式?
51. 假定你要执行A、B和C3个任务,它们可以以任何次序执行(但不能同时)。说明这个问题如何设计成一个产生式系统,并画出其状态图。
52. 对于上一题中状态图,如果任务C一定要在任务A之前执行,那么怎样改变状态图?
53. a. 如果记号 (i, j) 用来表示“若一个列表中第 i 位置的项大于第 j 位置的项,则把两项交换”,其中 i 和 j 为正整数,那么下面两个序列中哪一个更好地完成一个长度为3的列表的排序?
(1, 3)(3, 2)
(1, 2)(2, 3)(1, 2)
- b. 注意,通过这种方式表示交换序列,序列能够侵入后继序列,然后重新结合形成新的序列。使用该方法,描述一种遗传算法,用于开发一个为长度为10的列表排序的程序。
54. 假定一组机器人的每个成员都配备有一对传

感器，每个传感器都能探测到正前方2m范围的物体。每个机器人的形状都像一个废物筒，能在任何方向移动。试设计一系列实验，用来确定传感器装在哪里，使得造就的机器人能成功地将一个篮球直线抛出。你的一系列实验如

何与一个进化系统相比较？

55. 你做出某种决定是基于反应模式还是基于计划模式？你的回答是否依赖于你是决定中午吃什么还是作出求职决定？

10.9 社会问题

下面的问题有助于你分析一些与计算领域相关的伦理、社会和法律问题。回答这些问题不是唯一的目的，还应该考虑为什么这样回答，以及你的判断是否对每个问题都标准如一。

1. 核能、基因工程以及人工智能领域的研究者对于他们工作成果的利用方式应在多大程度上负有责任？科学家对其研究揭示的知识是否负有责任？若因此产生了意想不到的后果，怎么办？
2. 怎样区分智能和模拟的智能？你认为二者有区别吗？
3. 假定一个计算机化的专家系统因其给出好的建议而在医疗界享有盛誉。作为一个医生，在多大程度上可以让这个系统代替他（她）为病人作出治疗决定。如果医生的治疗方案与专家系统所提的治疗方案相对立，并且后来证实专家系统是正确的，那么那个医生是否应该对其不当治疗负有责任？一般说来，如果一个专家系统在某个领域内很有名，那么在多大程度上它会束缚而不是提高人类专家的判断力？
4. 许多人认为计算机的行为只不过是人对它怎样编程的结果，所以计算机不可能有自主意志。从而，计算机也不应对它的行为负责。人脑是计算机吗？人是否在出生的时候就事先被编程好了？人是否被他所处的环境编程？人是否要对自己的行为负责？
5. 是否有这样一些手段，科学即使能够去做，也不应当去做？例如，如果有朝一日可以造出具备能与人类相比拟的感知和推理能力的机器，那么建造这样的机器是否恰当？这样的机器的出现会带来什么样的问题？今天其他一些科学领域的进展正在引发哪些问题？
6. 历史上有许多例子表明，科学家、艺术家的创作活动受其所处时代的政治、宗教及其他社会势力的影响。这样的一些因素以何种方式影响着当今的科学成就？特别在计算机科学领域情况如何？
7. 当今，技术的进展造成了一些人的工作成为多余，许多文化至少应担负起一定的责任来帮助对一些人进行再教育。随着技术使我们越来越多的能力成为多余，社会应当或能够做些什么？
8. 假定你收到一张计算机处理的费用为\$0.00的账单。你该怎么办？假定你置之不理，30天后你又收到第二张\$0.00的催款通知单，你该怎么办？假定你依然不理睬，而30天后你又收到了一张\$0.00的催款通知单，而且还有提示，若不及时付款，将诉诸法律。谁将对此负责？
9. 是否有这样的情况，你会把个性与个人电脑联系在一起？你的计算机好像在施行报复或者固执难缠？你对你的计算机恼火生气过吗？对你的计算机恼火和对计算机所做的结果恼火有什么不同？你的计算机和你生过气吗？你与别的东西，如汽车、电视机、圆珠笔，有过类似的关系吗？
10. 根据你对上面问题的回答，人在多大程度上会把一个实体的行为与智能和意识的存在联系起来？在多大程度上，人应当做这样的关联？对于一个智能实体来说，是否可能

用有别于其他行为的方式来展现它的智能？

11. 许多人觉得，能通过图灵测试并不意味着机器有智能。一个论点是，智能的行为本身并不意味着智能。而进化论的基础是适者生存，这就是一种基于行为的测试。是否进化论意味着智能行为是智能的前身？机器能通过图灵测试，是否意味着它们正在变成有智能？

12. 医疗手段已经取得了很大的进步，人体的许多器官现在都能用人造器官或者捐赠人的器官来替代。可以设想，终究有一天连大脑也能换。如果这样的事能够做到，会产生什么样的道德问题？如果一个病人的神经细胞被人造神经细胞一点点换掉，那个病人还是同一个人吗？那个病人会觉察到有什么不同吗？那个病人还算人吗？

课外阅读

Banzhaf, W., P. Nordin, R.E.Deller, and E.D.Erancone. *Genetic Programming: An Introduction*. San Francisco, CA: Morgan Kaufmann, 1998.

Lu, J. and J. Wu. *Multi-agent Robotic Systems*. Boca Raton, FL: CRC Press, 2001.

Luger, G. *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*, 5th ed. Boston, MA: Addison-Wesley, 2005.

Mitchell, M. *An Introduction to Genetic Algorithms*. Cambridge, MA: MIT Press, 1998.

Nolfi, S. and D. Floreano, *Evolutionary Robotics*, Cambridge, MA: MIT Press, 2000.

Rumelhart, D.E. and J.L. McClelland. *Parallel Distributed Processing*. Cambridge, MA: MIT Press, 1986.

Russel, S. and P. Norvig. *Artificial Intelligence: A Modern Approach*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 2003.

Shapiro, L.G. and G. C. Stockman. *Computer Vision*. Englewood Cliffs, NJ: Prentice-Hall, 2001.

Weizenbaum, J. *Computer Power and Human Reason*. New York: W. H. Freeman, 1979.