

## 第一章 习题 1

### 1-1 什么是人工智能？试从学科和能力两方面加以说明。

答：人工智能（学科）是计算机学科中设计研究、设计和应用智能机器的一个分支，它的近期主要目标在于研究用机器来模仿和执行人脑的某些智力功能，并开发相关理论和技术。

（第 2 页，定义 1.3）

人工智能（能力）是智能机器所执行的通常与人类智能有关的智能行为，这些智能行为涉及学习，感知，思考，理解，识别判断，推理，证明，通信，设计，规划，行动和问题求解等活动。（第 2 页，定义 1.4）

### 1-6 现在人工智能有哪些学派？它们的认知观是什么？现在这些学派的关系如何？

答：1、符号主义：符号主义认为人工智能源于数理逻辑。（第 12 页，最后 1 段，第 1 行）

2、连接主义：连接主义认为人工智能源于仿生学，特别是人脑模型的研究。（第 13 页，第 2 段，第 1 行）

3、行为主义：行为主义认为人工智能源于控制论。（第 13 页，第 3 段，第 1 行）

以上三个人工智能学派将长期共存于合作，取长补短，并走向融合和集成，为人工智能的发展做出贡献。（Page 13，第四段）

### 1-10 你是如何理解人工智能的研究目标的？

答：

人工智能的一般目标为：（第 24 页，第 2 段，第 2,3 行）

（1）更好地理解人类智能 通过编写程序来模仿和检验有关人脸智能的理论。

（2）创造有用的灵巧程序 该程序能够执行一般需要人类专家才能实现的任务。

人工智能的近期目标是建造智能计算机以代替人类的某些智力活动。（第 24 页，第 5 段，第 1 行）

人工智能的远期目标是用自动机模仿人类的思维活动和智力功能。（第 24 页，第 6 段，第 1 行）

### 1-11 人工智能研究包括哪些内容？这些内容的重要性如何？

答：人工智能研究的基本内容：（第 24-27 页）

1、认知建模；2、知识表示；3、知识推理；4、计算智能；5、知识应用；6、机器感知；7、机器思维；8、机器学习；9、机器行为；10、智能系统构建

### 1-12 人工智能的基本研究方法有几类？它们与人工智能学派的关系如何？

答：人工智能研究的主要方法：（第 27-28 页）

1、功能模拟法：符号主义学派也可称为功能模拟学派。

2、结构模拟法：连接主义学派也可称为结构模拟学派

3、行为模拟法：行为主义学派也可称为行为模拟学派

4、集成模拟法：是上述 3 种人工智能方法取长补短，实现优势互补。

**1-13 人工智能的主要研究和应用领域是什么？其中，哪些是新的研究热点？**

答：人工智能研究与应用领域：**（第 30-37 页）**

1、问题求解与博弈；2、逻辑推理与定理证明；3、计算智能；4、分布式人工智能与 Agent；  
5、自动程序设计；6、专家系统；7、机器学习；8、自然语言理解；9、机器人学；10 模式识别；11、机器视觉；12、神经网络；13、智能控制；14、智能调度与指挥；15 智能检索；  
16、系统与语言工具

人工智能新产业的一些领域：（1）智能制造；（2）智慧医疗；（3）智慧农业；（4）智慧金融；  
（5）智能交通与智能驾驶；（6）智慧城市；（7）智能家居；（8）智能管理；（9）智能经济

## 第二章 习题 2

### 2-1 状态空间法，问题归约法，谓词逻辑法个语义网络法的要点是什么？它们有何本质上的联系及异同点？

答：状态空间法是一种基于解答空间的问题表示和求解的方法，它是以状态和操作符号为基础的。由于状态空间法需要扩展过多的节点，容易出现“组合爆炸”，因而只适合于表示比较简单的问题。(第 77 页，第 1 段)

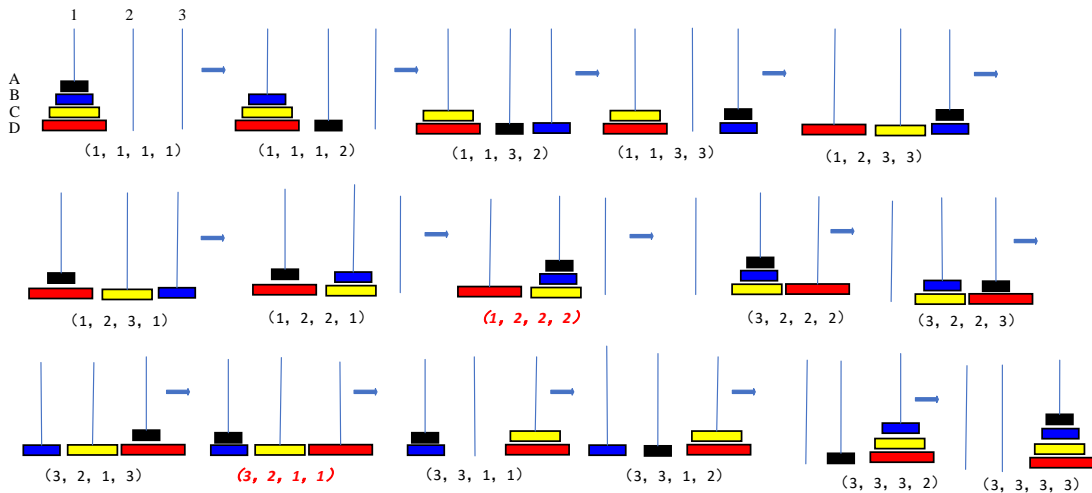
问题归约法从目标（要解决的问题）出发，逆向推理，通过一系列变换把初始问题变换为子问题集合和子子问题集合，直至最后归约为一个平凡的本原问题集合。问题归约法能够比状态空间法更有效地表示问题。状态空间法是问题归约法的一种特例。在问题归约法的与或图中，包含有与节点和或节点，而在状态空间法中只含有或节点。(第 77 页，第 2 段)

谓词逻辑法采用谓词合式公式和一阶谓词演算把要解决问题变为一个有待证明的问题，然后采用消解定理和消解反演来证明一个新语句是从已知的正确语句导出的，从而证明这个新语句也是正确的。谓词逻辑法常与其他表示方法混合使用，灵活方便，可以表示比较复杂的问题。(第 77 页，第 3 段)

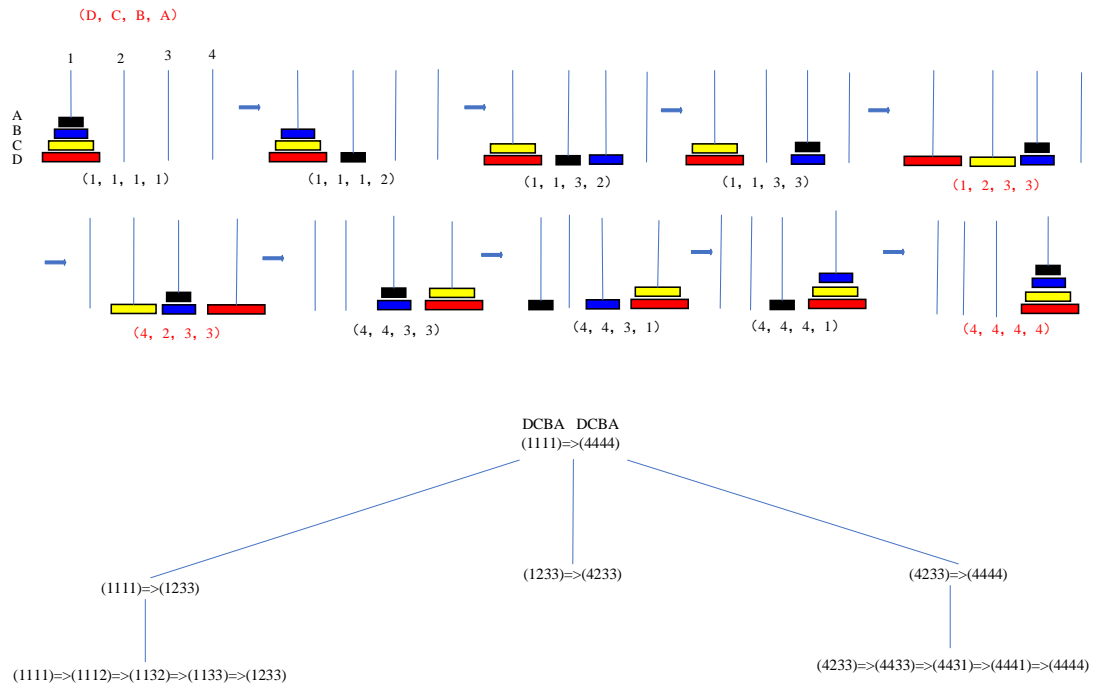
语义网络是一种化表示方法，它由节点和弧线或链线组成。语义网络的解答是一个经过推理和匹配而得到具体明确结果的新的语义网络。语义网络可用于表示多元关系，扩展后可以表示更复杂的问题。(第 77 页，第 4 段)

### 2-5 试用四元数列结构表示四圆盘梵塔问题，并画出求解该问题的与或图。

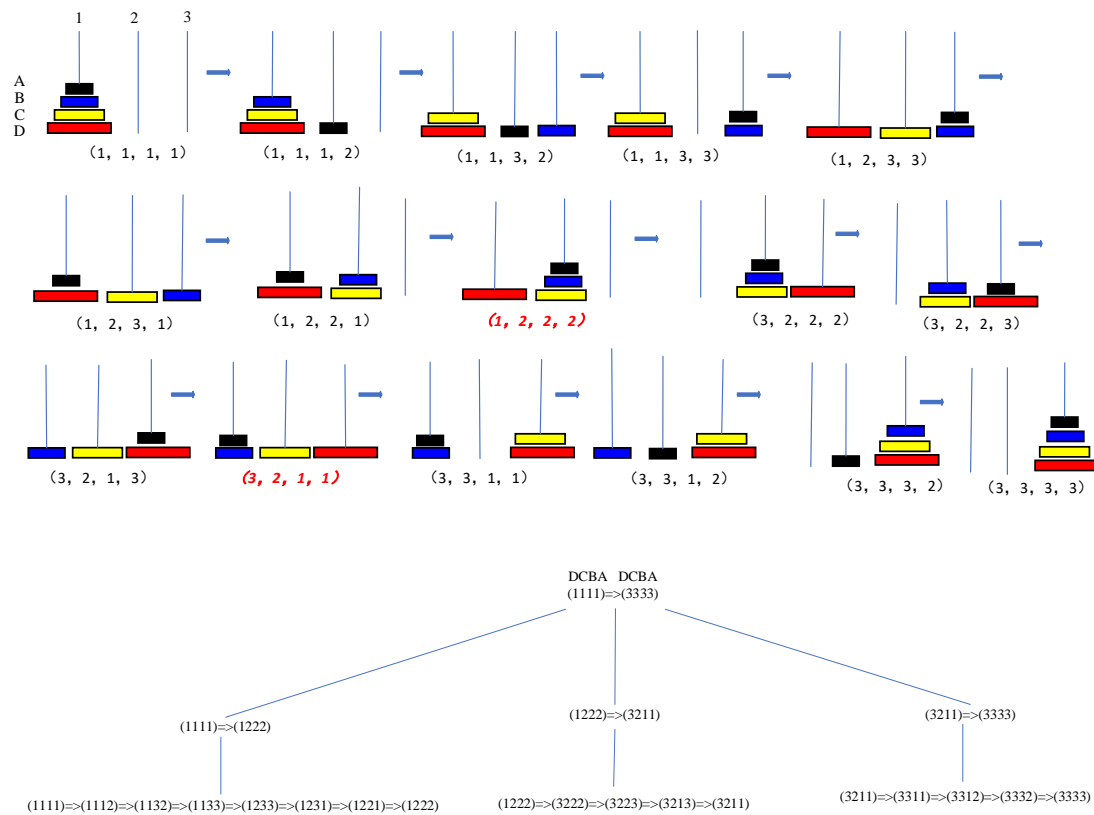
答：四个盘由大到小一次为 DCBA



答：方案 1 四根柱子和四个圆盘：



方案 2 三根柱子和四个圆盘：



2-6 用谓词演算公式表示下列英文句子（多用而不是省用不同谓词和项。例如不要用单一的

谓词字母来表示每个句子)。

A computer system is intelligent if it can perform a task which, if performed by a human, requires intelligence.

答:  $P(x,y)$ :  $x$  performs  $y$  task ( $x$  完成有任务)

$Q(y)$ :  $y$  requires intelligence ( $y$  需要智能)

$C(x)$ :  $x$  is a computer system ( $x$  是计算机系统)

$I(x)$ :  $x$  is intelligent ( $x$  是智能系统)

$$(\forall x)(\exists y)(C(x) \wedge P(x, y) \wedge P(\text{human}, y) \wedge Q(y) \rightarrow I(x))$$

2-7 把下列语句表示成语义网络描述:

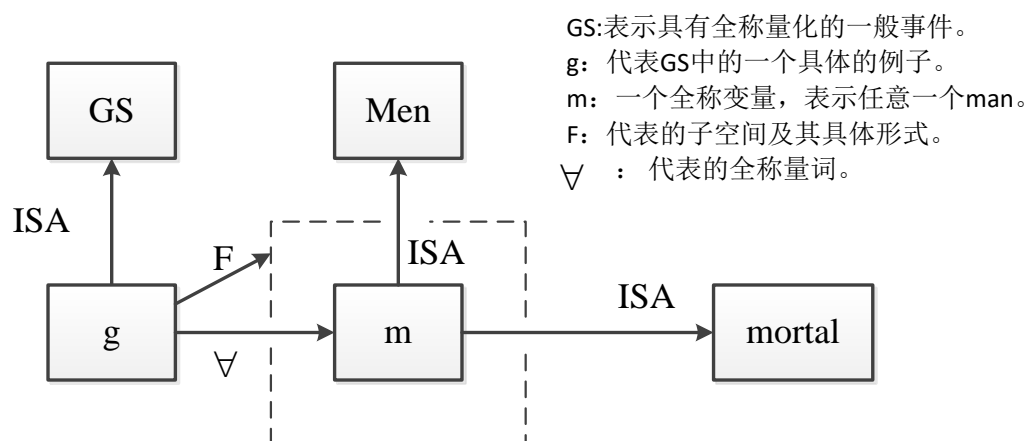
(1) All men are mortal.

(2) Every cloud has silver lining.

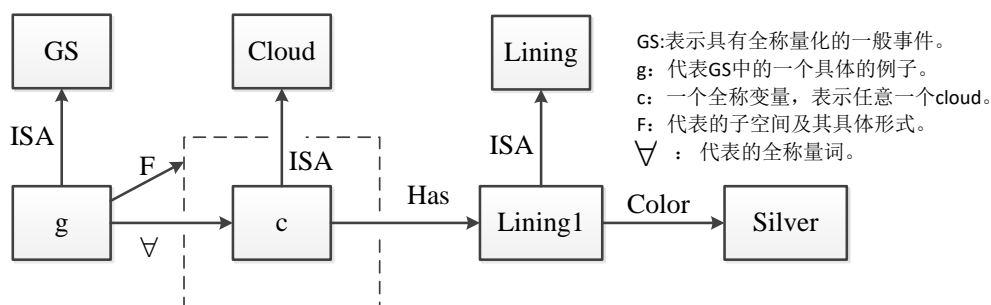
(3) All branch managers of DEC participate in a profit-sharing plan.

答:

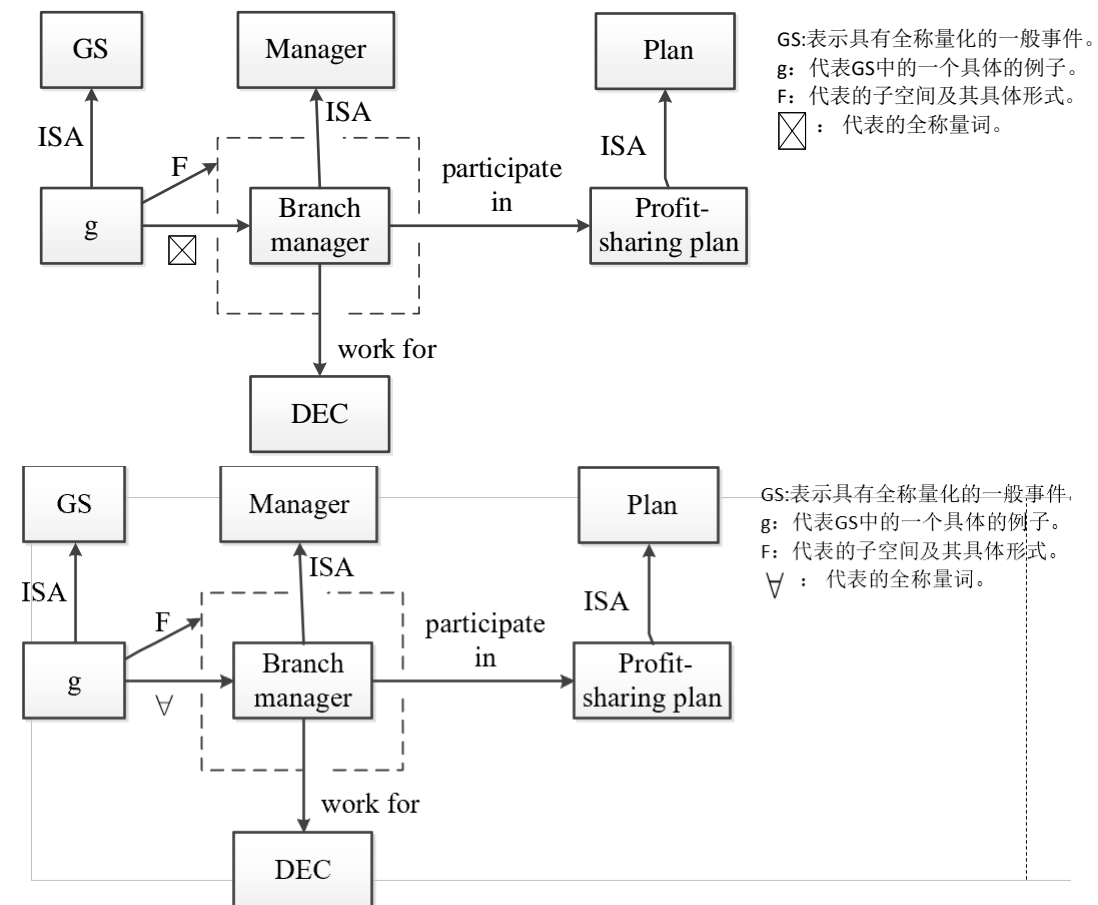
(1) All men are mortal; 所有人都是凡人



(2) Every cloud has a silver lining. 黑暗中总有一线光明。



(2) All branch managers of DEC participate in a profit-sharing plan. DEC 的所有分支机构经理都参与利润共享计划。



## 第三章 习题 3

### 3-1 什么是图搜索过程？其中，重排 OPEN 表意味着什么，重排的原则是什么？

答：图搜索（GRAPHSEARCH）的一般过程如下：。（第 83 页）

- (1) 建立一个只含有起始节点 S 的搜索图 G，把 S 放到一个 OPEN 表中。
- (2) 初始化 CLOSED 表为空表。
- (3) LOOP：若 OPEN 表是空表，则失败退出；
- (4) 选择 OPEN 表上的第一个节点，把它从 OPEN 表移出并放进 CLOSED 表中。称此节点为节点 n。
- (5) 若 n 为一目标节点，则有解并成功退出，此解是追踪图 G 中沿着指针从 n 到 S 这条路径而得到的（指针将在第（7）步中设置）
- (6) 扩展节点 n，生成一组子节点。把这些子节点中不是节点 n 先辈的那部分子节点计入后继节点集合 M，并把这些子节点作为节点 n 的子节点加入 G 中；
- (7) 针对 M 中子节点的不同情况，分别作如下处理：
  - 1) 对那些未曾在 G 中出现过的（既未曾在 OPEN 表上，也未在 CLOSED 表上出现过）M 成员设置其父节点指针指向 n 并加入 OPEN 表。（新生成的）
  - 2) 对那些原来已在 G 中出现过，但还没有被扩展的 M 成员，确定是否需要修改他指向父节点的指针。（原生成但未扩展的）
  - 3) 对于那些先前已在 G 中出现过，并已经扩展的 M 成员，确定是否需要修改其后继节点指向父节点的指针。（原生成也扩展过的）
- (8) 按某一任意方式或按某个试探值，重排 OPEN 表。
- (9) GO LOOP

**重排 OPEN 表意味着**，在第（8）步中，将优先扩展哪些节点，不同的排序标准对应着不同的搜索策略。（从图搜索的过程可以看出，是否重新安排 OPEN 表，即是否按照某个试探值（或准则、启发信息等）重新对未扩展节点进行排序，将决定该图搜索过程是无信息搜索或启发式搜索。（第 84 页，第 4 段））

**重排原则**视具体需求而定，不同的原则对应着不同的搜索策略，如果想尽快地找到一个解，则应当将最有可能达到目标节点的哪些节点排在 OPEN 表的前面部分，如果想找到代价最小的解，则应当按代价从小到达的顺序重排 OPEN 表。（过程第（8）步对 OPEN 表上的节点进行排序，以便能够从中选出一个“最好”的节点作为第（4）步扩展使用。这种排序可以是任意的（即盲目的，属于盲目搜索），也可以用以后要讨论的各种思想或其他准则为依据（属于启发式搜索）（第 84 页，第 3 段））

### 3-3 化为子句形有哪些步骤？请结合例子说明。

答：（第 97-98 页）

- (1) 消去蕴涵符号；(2) 减少否定符号的辖域；(3) 对变量标准化；(4) 消去存在量词；
- (5) 化为前束形；(6) 把母式化为合取范式；(7) 消去全称量词；(8) 消去连词符号；(9) 更换变量名称；

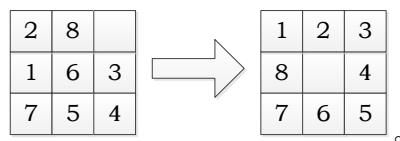
### 3-4 如何通过消解反演求取问题答案？

答：（第 101 页）

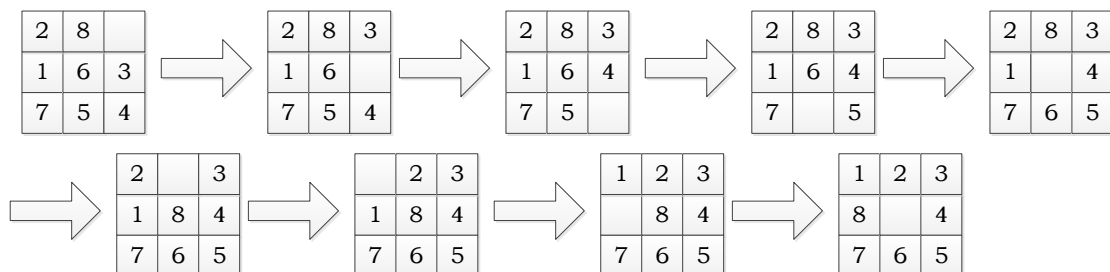
给出一个公式集  $S$  和目标公式  $L$ ，通过反证或反演来求证目标公式  $L$ ，其证明步骤如下：

- (1) 否定  $L$ ，得  $\sim L$
- (2) 把  $\sim L$  添加到  $S$  中去；
- (3) 把新产生的集合  $\{\sim L, S\}$  化成子句集；
- (4) 应用消解原理，力图推导出一个表示矛盾的空子句。

### 3-7 有界深度优先搜索方法求解图 3.30 所示八数码难题。



答：



### 3-9 试比较宽度优先搜索、有界深度优先搜索及有序搜索的搜索效率，并以实例数据加以说明。

答（第 128 页，第 4-5 段）

盲目搜索包括宽度优先搜索，深度优先搜索和等代价搜索，其中有界深度优先搜索在某种意义上讲，具有一定的启发性。从搜索效率看，一般来说，有界深度优先搜索较好，宽度优先搜索次之，深度优先搜索较差。有序搜索（或启发式搜索）要比盲目搜索（有界深度优先搜索，宽度优先搜索）有效得多。

### 3-12 把下列句子变换成子句形式：

- (1)  $(\forall x)(P(x) \rightarrow P(x))$   
 $\Rightarrow (\forall x)(\sim P(x) \vee P(x))$   
 $\Rightarrow \sim P(x) \vee P(x)$
- (2)  $\forall x \forall y (On(x, y) \rightarrow Above(x, y))$   
 $\Rightarrow \forall x \forall y (\sim On(x, y) \vee Above(x, y))$   
 $\Rightarrow \sim On(x, y) \vee Above(x, y)$



$$\begin{aligned}
(3) \quad & \forall x \forall y \forall z (Above(x, y) \wedge Above(x, y) \rightarrow Above(x, y)) \\
& \Rightarrow \forall x \forall y \forall z (\sim (Above(x, y) \wedge Above(x, y)) \vee Above(x, y)) \\
& \Rightarrow \forall x \forall y \forall z (\sim Above(x, y) \vee \sim Above(x, y) \vee Above(x, y)) \\
& \Rightarrow \sim Above(x, y) \vee \sim Above(x, y) \vee Above(x, y)
\end{aligned}$$

蕴涵 $\rightarrow$ 的优先级低于与 $\wedge$

$$(4) \sim \left( (\forall x) \left( P(y) \rightarrow \left( (\forall y) (P(y) \rightarrow P(f(x, y))) \wedge (\forall y) (Q(x, y) \rightarrow P(y)) \right) \right) \right)$$

$$\sim \left\{ (\forall x) \left\{ P(x) \rightarrow \left\{ (\forall y) [P(y) \rightarrow P(f(x, y))] \wedge (\forall y) [Q(x, y) \rightarrow P(y)] \right\} \right\} \right\}$$

$$\sim \left\{ (\forall x) \left\{ \sim P(x) \vee \left\{ (\forall y) [\sim P(y) \vee P(f(x, y))] \wedge (\forall y) [\sim Q(x, y) \vee P(y)] \right\} \right\} \right\}$$

$$(\exists x) \left\{ P(x) \wedge \left\{ (\exists y) [P(y) \wedge \sim P(f(x, y))] \vee (\exists y) [Q(x, y) \wedge \sim P(y)] \right\} \right\}$$

消除存在量词，按存在量词范围消除

$$(\exists x) \left\{ P(x) \wedge \left\{ [P(B1) \wedge \sim P(f(x, B1))] \vee [Q(x, B2) \wedge \sim P(B2)] \right\} \right\}$$

$$P(A) \wedge \left\{ [P(B1) \wedge \sim P(f(A, B1))] \vee [Q(A, B2) \wedge \sim P(B2)] \right\}$$

$$P(A) \wedge \left\{ [P(B1) \vee [Q(A, B2) \wedge \sim P(B2)]] \wedge [\sim P(f(A, B1)) \vee [Q(A, B2) \wedge \sim P(B2)]] \right\}$$

$$P(A) \wedge \left\{ [[P(B1) \vee Q(A, B2)] \wedge [P(B1) \vee \sim P(B2)]] \wedge [\sim P(f(A, B1)) \vee Q(A, B2)] \wedge [\sim P(f(A, B1)) \vee \sim P(B2)] \right\}$$

$$P(A) \wedge [P(B1) \vee Q(A, B2)] \wedge [P(B1) \vee \sim P(B2)] \wedge [\sim P(f(A, B1)) \vee Q(A, B2)] \wedge [\sim P(f(A, B1)) \vee \sim P(B2)]$$

### 3-13 非经典逻辑、非经典推理与经典逻辑、经典推理有何不同？

答：

- (1) 在推理方法上，经典逻辑采用**演绎**逻辑推理，而非经典逻辑采用**归纳**逻辑推理。
- (2) 在辖域取值上，经典逻辑都是**二值**逻辑，而非经典逻辑都是**多值**逻辑。
- (3) 在运算法则上，属于经典逻辑的形式逻辑和数理逻辑，它们的许多运算法则在非经典逻辑中就不能成立。
- (4) 在逻辑算符上，非经典逻辑具有更多的逻辑算符。
- (5) 在是否单调上，经典逻辑是单调的，非经典逻辑是非单调的。

### 3-14 什么是不确定性推理？为什么需要采用不确定性推理？

答：**什么是不确定性推理？（第 116 页，第 5 段）**

**不确定性推理**（reasoning with uncertainty）也称不精确推理，是一种建立在非经典逻辑基础上的基于不确定性知识的推理，它从不确定性的初始证据出发，通过运用不确定性知识，推出具有一定程度的不确定性的和合理的或近乎合理的结论。

**为什么需要采用不确定性推理？**

现实世界中打的问题和事物间的关系，往往比较复杂，客观事物存在的随机性、模糊性、不完全性和不精确性，往往导致人们认识上一定程度的不确定性。这是，若**仍然采用经典的精确推理方法进行推理，必然无法反映事物的真实性**。为此，需要在不完全和不确定的情况下运用不确定性知识进行推理，即进行不确定性推理。

### 3-15 不确定性推理可分为哪几种类型？

答：（第 116 页，第 7 段）

不确定性推理中存在三种不确定性，即关于知识的不确定性、关于证据的不确定性和关于结论的不确定性。

3-16 设有三个独立的结论 H1, H2, H3 及两个独立的证据 E1, E2，它们的先验概率和条件概率分别为：

$$P(H1) = 0.4, P(H2) = 0.3, P(H3) = 0.3;$$

$$P(E1|H1) = 0.5, P(E1|H2) = 0.3, P(E1|H3) = 0.5;$$

$$P(E2|H1) = 0.7, P(E2|H2) = 0.9, P(E2|H3) = 0.1。$$

利用概率方法分别求出：

（1）当只有证据 E1 出现时的  $P(H1|E1)$ ,  $P(H2|E1)$ ,  $P(H3|E1)$  值；说明 E1 的出现对证据 H1, H2 和 H3 的影响。

（2）当 E1 和 E2 同时出现时的  $P(H1|E1E2)$ ,  $P(H2|E1E2)$ ,  $P(H3|E1E2)$  值；说明 E1 和 E2 同时出现对证据 H1, H2 和 H3 的影响。

答：（第 120-121 页，公式 (3.17)，公式 (3.18)，例 3.10 和例 3.11）

(1)

$$\begin{aligned} P(H1|E1) &= \frac{P(H1) \times P(E1|H1)}{P(H1) \times P(E1|H1) + P(H2) \times P(E1|H2) + P(H3) \times P(E1|H3)} \\ &= \frac{0.4 \times 0.5}{0.4 \times 0.5 + 0.3 \times 0.3 + 0.3 \times 0.5} = \frac{0.2}{0.2 + 0.09 + 0.15} = \frac{0.2}{0.44} = 0.4545 \end{aligned}$$

$$\begin{aligned} P(H2|E1) &= \frac{P(H2) \times P(E1|H2)}{P(H1) \times P(E1|H1) + P(H2) \times P(E1|H2) + P(H3) \times P(E1|H3)} \\ &= \frac{0.3 \times 0.3}{0.4 \times 0.5 + 0.3 \times 0.3 + 0.3 \times 0.5} = \frac{0.09}{0.2 + 0.09 + 0.15} = \frac{0.09}{0.44} = 0.2045 \end{aligned}$$

$$\begin{aligned} P(H3|E1) &= \frac{P(H3) \times P(E1|H3)}{P(H1) \times P(E1|H1) + P(H2) \times P(E1|H2) + P(H3) \times P(E1|H3)} \\ &= \frac{0.3 \times 0.5}{0.4 \times 0.5 + 0.3 \times 0.3 + 0.3 \times 0.5} = \frac{0.15}{0.2 + 0.09 + 0.15} = \frac{0.15}{0.44} = 0.3409 \end{aligned}$$

计算结果表明： $P(H1|E1) > P(H1)$ ,  $P(H2|E1) < P(H2)$ ,  $P(H3|E1) > P(H3)$ ，由于证据 E1 的出现，H1, H3 成立的可能性略有增加，而 H2 成立的可能性下降。

(2)

$$\begin{aligned} P(H1|E1E2) &= \frac{P(H1) \times P(E1|H1) \times P(E2|H1)}{P(H1) \times P(E1|H1) \times P(E2|H1) + P(H2) \times P(E1|H2) \times P(E2|H2) + P(H3) \times P(E1|H3) \times P(E2|H3)} \\ &= \frac{0.4 \times 0.5 \times 0.7}{0.4 \times 0.5 \times 0.7 + 0.3 \times 0.3 \times 0.9 + 0.3 \times 0.5 \times 0.1} = \frac{0.14}{0.14 + 0.081 + 0.015} = \frac{0.14}{0.236} = 0.5932 \end{aligned}$$

$$P(H2|E1E2) = \frac{P(H2) \times P(E1|H2) \times P(E2|H2)}{P(H1) \times P(E1|H1) \times P(E2|H1) + P(H2) \times P(E1|H2) \times P(E2|H2) + P(H3) \times P(E1|H3) \times P(E2|H3)}$$

$$= \frac{0.3 \times 0.3 \times 0.9}{0.4 \times 0.5 \times 0.7 + 0.3 \times 0.3 \times 0.9 + 0.3 \times 0.5 \times 0.1} = \frac{0.081}{0.14 + 0.081 + 0.015} = \frac{0.081}{0.236} = 0.3432$$

$$P(H3|E1E2) = \frac{P(H3) \times P(E1|H3) \times P(E2|H3)}{P(H1) \times P(E1|H1) \times P(E2|H1) + P(H2) \times P(E1|H2) \times P(E2|H2) + P(H3) \times P(E1|H3) \times P(E2|H3)}$$

$$= \frac{0.3 \times 0.5 \times 0.1}{0.4 \times 0.5 \times 0.7 + 0.3 \times 0.3 \times 0.9 + 0.3 \times 0.5 \times 0.1} = \frac{0.015}{0.14 + 0.081 + 0.015} = \frac{0.015}{0.236} = 0.0636$$

计算结果表明： $P(H1|E1E2) > P(H1)$ ， $P(H2|E1E2) > P(H2)$ ， $P(H3|E1E2) < P(H3)$ ，由于证据 E1 和 E2 同时出现，H1，H2 成立的可能性增加，而 H3 成立的可能性下降。

### 3-17 在主观贝叶斯推理中，LS 和 LN 的意义是什么？

答：（第 122 页，第 3 段）

在主观贝叶斯方法中，用 IF E THEN (LS, LN) H 表示知识，(LS, LN) 表示知识的静态强度，称 LS 为该式成立的充分性因子，LN 为该式成立的必要性因子，它们分别衡量证据（前提）E 对结论 H 的支持程度和~E 对 H 的支撑程度。

### 3-18 设有如下推理规则：

E1 : LS1 =500, LN1= 0.01 P(H1) = 0.1

E2 : LS2 =500, LN2= 0.01 P(H1) = 0.1

E3 : LS3 =500, LN3= 0.01 P(H2) = 0.1

H1 : LS4 =500, LN4= 0.01 P(H2) = 0.1

$P(E1|S1)=0.5$ ， $P(E2|S2)=0$ ， $P(E3|S3)=0.8$ ，用主观贝叶斯方法求 H2 的后验概率  $P(H2|S1, S2, S3)$ 。

$$P(H2|S1, S2, S3) = \frac{O(H2|S1, S2, S3)}{1 + O(H2|S1, S2, S3)}$$

$$O(H2|S1, S2, S3) = \frac{O(H2|S1, S2)}{O(H2)} \times \frac{O(H2|S3)}{O(H2)} \times O(H2)$$

(1)

$$O(H2) = \frac{P(H2)}{1 - P(H2)} = \frac{0.1}{1 - 0.1} = \frac{1}{9}$$

(2)

$$O(H2|S3) = \frac{P(H2|S3)}{1 - P(H2|S3)}$$

采用 EH 公式计算  $P(H2|S3)$ ， $P(E3|S3)=0.8$ ， $P(E3)$ ，

$$P(H2|S3) =$$

(3)

$$O(H2|S1, S2) = \frac{P(H2|S1, S2)}{1 - P(H2|S1, S2)}$$

$$P(H2|S1, S2) = P(H2) + \frac{P(H2|H1) - P(H2)}{1 - P(H1)} \times [P(H1|S1, S2) - P(H1)]$$

需要求： $P(H1)$ ， $P(H2)$ ， $P(H2|H1)$ ， $P(H1|S1, S2)$

$$P(H2|H1) = \frac{O(H2|H1)}{1 + O(H2|H1)} = \frac{LS3O(H2)}{1 + LS3O(H2)} = \frac{500 \times (1/9)}{1 + 500 \times (1/9)} = \frac{500}{509}$$

$$P(H1|S1, S2) = \frac{O(H1|S1, S2)}{1 + O(H1|S1, S2)}$$

$$O(H1|S1, S2) = \frac{O(H1|S1)}{O(H1)} \times \frac{O(H1|S2)}{O(H1)} \times O(H1)$$

需要求:  $O(H1|S1)$ ,  $O(H1|S2)$

$$O(H1|S1) = \frac{P(H1|S1)}{1 - P(H1|S1)}$$

$$O(H1|S2) = \frac{P(H1|S2)}{1 - P(H1|S2)}$$

需要求:  $P(H1|S1)$ ,  $P(H1|S2)$

$$P(E1|S1) = 0.5$$

**EH公式前半部分**  $P(H1|S1) = P(H1|\sim E1) + \frac{P(H1) - P(H1|\sim E1)}{1 - P(E1)} \times P(E1|S1)$

EH 公式后半部分(无效)  $P(H1|S1) = P(H1) + \frac{P(H1|E1) - P(H1)}{1 - P(E1)} \times [P(E1|S1) - P(E1)]$

$$P(E2|S2) = 0$$

**EH公式后半部分**  $P(H1|S2) = P(H1|\sim E2) + \frac{P(H1) - P(H1|\sim E2)}{1 - P(E2)} \times P(E2|S2)$

**要求:  $P(H1|E1)$ ,  $P(E1)$ ;  $P(H1|\sim E2)$ ,  $P(E2)$**

(无效概率值为负)  $P(H1|E1) = \frac{O(H1|E1)}{1 - O(H1|E1)} = \frac{500}{9 - 500}$

$$O(H1|E1) = LS1 \times O(H1) = 500 \times \frac{1}{9}$$

(有效概率) **$P(H1|\sim E1)$**   $= \frac{O(H1|\sim E1)}{1 - O(H1|\sim E1)} = \frac{0.01}{9 - 0.01} = \frac{1}{899}$

$$O(H1|\sim E1) = LN1 \times O(H1) = 0.01 \times \frac{1}{9} = \frac{1}{900}$$

**下面计算 $P(E1)$**

$$LS1 = \frac{P(E1|H1)}{P(E1|\sim H1)} \Rightarrow P(E1|\sim H1) = \frac{P(E1|H1)}{LS1}; \quad P(E1|\sim H1) = \frac{1 - 0.01}{500 - 0.01} = \frac{99}{49999}$$

$$LN1 = \frac{1 - P(E1|H1)}{1 - P(E1|\sim H1)} = \frac{1 - P(E1|H1)}{1 - \frac{P(E1|H1)}{LS1}} \Rightarrow P(E1|H1) = \frac{LN1 - 1}{\frac{LN1}{LS1} - 1} = \frac{0.01 - 1}{\frac{0.01}{500} - 1} = \frac{500 - 5}{500 - 0.01} = \frac{49500}{49999}$$

由公式 (4.13)

$$P(H1|\sim E1) = \frac{P(\sim E1|H1)P(H1)}{P(\sim E1)} \Rightarrow P(\sim E1) = \frac{P(\sim E1|H1)P(H1)}{P(H1|\sim E1)} = \frac{\left(1 - \frac{49500}{49999}\right) \times 0.1}{1/899}$$

**$P(E1) = 1 - P(\sim E1) = 1 - 89.9 \times (1 - 49500/49999) = 0.10278$ , 不满足  $P(E1|S1) < P(E1)$**

$$(\text{有效概率})P(H1|\sim E1) = \frac{O(H1|\sim E1)}{1 - O(H1|\sim E1)} = \frac{0.01}{9 - 0.01} = \frac{1}{899}$$

$$O(H1|\sim E1) = LN1 \times O(H1) = 0.01 \times \frac{1}{9} = \frac{1}{900}$$

下面计算P(E1)

$$LS1 = \frac{P(E1|H1)}{P(E1|\sim H1)} \Rightarrow P(E1|\sim H1) = \frac{P(E1|H1)}{LS1}; \quad P(E1|\sim H1) = \frac{1 - 0.01}{500 - 0.01} = \frac{99}{49999}$$

$$LN1 = \frac{1 - P(E1|H1)}{1 - P(E1|\sim H1)} = \frac{1 - P(E1|H1)}{1 - \frac{P(E1|H1)}{LS1}} \Rightarrow P(E1|H1) = \frac{LN1 - 1}{\frac{LN1}{LS1} - 1} = \frac{0.01 - 1}{\frac{0.01}{500} - 1} = \frac{500 - 5}{500 - 0.01}$$

$$= \frac{49500}{49999}$$

由公式 (4.13)

$$P(H1|\sim E1) = \frac{P(\sim E1|H1)P(H1)}{P(\sim E1)} \Rightarrow P(\sim E1) = \frac{P(\sim E1|H1)P(H1)}{P(H1|\sim E1)} = \frac{\left(1 - \frac{49500}{49999}\right) \times 0.1}{1/899}$$

$$P(E1) = 1 - P(\sim E1) = 0.10278.$$

## 第四章 习题 4

### 4-1 计算智能的含义是什么？它涉及哪些分支？

答：（第 137 页，第 2 段）

**计算智能**是一种智力方式的低层认知，它与人工智能的区别只是认知层次从中层下降至低层而已。中层系统含有知识（精品），低层系统没有。

若一个系统只涉及数值（低层）数据，含有模式识别部分，不应用人工智能意义上的知识，而且能够呈现出：（1）计算适应性；（2）计算容错性；（3）接近人的速度；（4）误差率与人相近，则该系统就是计算智能系统。

（第 135 页，第 1 段）

**计算智能**涉及神经计算、模糊计算、进化计算、粒群计算、蚁群算法、自然计算、免疫计算和人工生命等领域，它的研究和发展反映了当代科学技术多学科交叉与集成的重要发展趋势。

### 4-2 试述计算智能（CI）、人工智能（AI）和生物智能（BI）的关系？

答：（第 135-136 页）

把计算智能理解为智力的低层认知，它主要取决于数值数据而不依赖于知识。人工智能是在计算智能的基础上引入知识而产生的智力中层认知。生物智能，尤其是人类智能，则是最高层的智能。也就是， $CI < AI < BI$ 。

### 4-3 人工神经网络为什么具有诱人的发展前景和潜在的广泛应用领域？

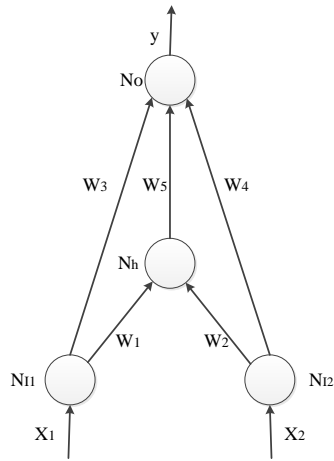
答：（第 138 页，第 1-6 段）

人工神经网络有如下至关重要的特性：（1）并行分布处理；（2）非线性映射；（3）通过训练进行学习；（4）适应与集成；（5）硬件实现。显然，神经网络由于其学习和适应、自组织、函数逼近和大规模并行处理能力，因而具有用于智能系统的潜力。

### 4-6 构作一个神经网络，用于计算含有两个输入的 XOR 函数。指定所用神经网络单元的种类。

答：（第 145 页）

方案一



神经网络实现异或逻辑

$y = f(x_1 \cdot w_3 + x_2 \cdot w_4 + z \cdot w_5)$ ，按照 XOR 函数的增值表有下列不等式：  
 $z = f(x_1 \cdot w_1 + x_2 \cdot w_2)$

$$1 \cdot w_3 + 1 \cdot w_4 + f(1 \cdot w_1 + 1 \cdot w_2) \cdot w_5 < t$$

$$0 \cdot w_3 + 0 \cdot w_4 + f(0 \cdot w_1 + 0 \cdot w_2) \cdot w_5 < t$$

$$0 \cdot w_3 + 1 \cdot w_4 + f(0 \cdot w_1 + 1 \cdot w_2) \cdot w_5 \geq t$$

$$1 \cdot w_3 + 0 \cdot w_4 + f(1 \cdot w_1 + 0 \cdot w_2) \cdot w_5 \geq t$$

$\Rightarrow$

$$t = 0.5, f(< 0.5) = 0, f(> 0.5) = 1$$

$$w_3 + w_4 + f(w_1 + w_2) \cdot w_5 < 0.5$$

$$f(0) \cdot w_5 < 0.5$$

$$w_4 + f(w_2) \cdot w_5 \geq 0.5$$

$$w_3 + f(w_1) \cdot w_5 \geq 0.5$$

$\Rightarrow$

$$w_3 + w_4 + f(w_1 + w_2) \cdot w_5 < 0.5$$

$$w_3 + w_4 + f(w_2) \cdot w_5 + f(w_1) \cdot w_5 \geq 1$$

$\Rightarrow$

$$w_3 = w_4 = 1$$

$$w_1 + w_2 > 0.5$$

$$w_5 = -2$$

$\Rightarrow$

$$w_3 = w_4 = 1$$

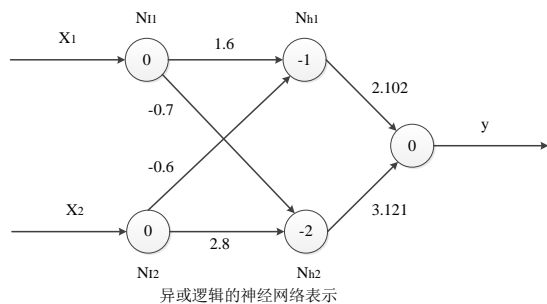
$$w_1 = w_2 = 0.3$$

$$w_5 = -2$$

$(w_1, w_2, w_3, w_4, w_5)$  为  $(0.3, 0.3, 1, 1, -2)$ 。

$z$  为隐含节点 Nh 的输出。  $f(a) = \begin{cases} 0, & a < 0.5 \\ 1, & a \geq 0.5 \end{cases}$  为输入输出关系函数。

方案二：



$$N_{h1} = f(x_1 \times 1.6 + x_2 \times (-0.6) - 1)$$

$$N_{h2} = f(x_1 \times (-0.7) + x_2 \times 2.8 - 2.0) \quad (\text{由网络结构可得出}). \quad f(a) = \begin{cases} 0, & a < 0.5 \\ 1, & a \geq 0.5 \end{cases} \text{为输入输出关}$$

$$y = N_o = f(N_{h1} \times 2.102 + N_{h2} \times 3.121)$$

系函数。

#### 4-14 试述遗传算法的基本原理，并说明遗传算法的求解步骤？

答：遗传算法类似于自然进化，通过作用于染色体上的基因寻找好的染色体来求解问题。与自然界相似，遗传算法对求解问题的本身一无所知，它所需要的仅仅是对算法所产生的每个染色体进行评价，并基于适应值来选择染色体，使适应性好的染色体进行评价，并基于适应值来选择染色体，使适应性好的染色体有更多的繁殖机会。在遗传算法中，通过随机方式产生若干个所求问题的数字编码，即染色体，形成初始种群；通过适应度函数给每个个体一个数值评价，淘汰低适应度的个体，选择高适应度的个体参加遗传操作，经过遗传操作后的个体集合形成下一代新的种群。再对这个新的种群进行下一轮进化。**这就是遗传算法的基本原理。其求解步骤如下：**

- (1) 初始化种群
- (2) 计算种群上每个个体的适应度值；
- (3) 按抽样个体适应度值所决定的某个规则选择将进入下一代的个体；
- (4) 按概率  $P_c$  进行交叉操作；
- (5) 按概率  $P_c$  进行变异操作；
- (6) 若没有满足某种停止条件，则转步骤 (2)，否则进入下一步；
- (7) 输出种群中适应度值最优的染色体作为问题的满意解或最优解。



## 第五章 习题 5

### 5-1 什么叫专家系统？它具有哪些特点与优点？

答：（第 187-188 页）

定义 6.3 专家系统使用人类专家推理的计算机模型来处理现实世界中需要专家做出解释的复杂问题，并得出与专家相同的结论。

专家系统由一些共同的特点：（1）启发性；（2）透明性；（3）灵活性。

专家系统由一些共同的优点：（1）高效率、准确的工作；（2）不受周围环境影响；（3）不受时间和空间限制；（4）促进各领域发展；（5）汇集和集成多领域专家知识和经验，协作解决重大问题；（6）军事专家系统水平是重要标志；（7）具有巨大经济和社会效益；（8）促进整个科学技术的发展。

### 5-2 专家系统由哪些部分构成？各部分的作用为何？

答：（第 190 页）

专家系统的主要组成部分如下：

- （1）**知识库**：用于存储某领域专家系统的专门知识，包括事实、可行操作与规则等。
- （2）**综合数据库**：用于存储领域或问题的初始数据和推理过程中得到的中间数据（信息），即被处理对象的一些当前事实。
- （3）**推理机**：用于记忆所采用的规则和控制策略的程序，使整个专家系统能够以逻辑方式协调地工作。
- （4）**解释器**：向用户解释专家系统的行为。
- （5）**接口**：能够使系统与用户进行对话，使用户能够输入必要的数据、提出问题和了解推理过程及推理结果等。

### 5-3 建造专家系统的关键步骤是什么？

答：（第 190-191 页）

建立专家系统的一般步骤：（1）设计初始知识库：问题知识化，知识概念化，概念形式化，形式规则化，规则合法化；（2）原型机的开发与实验；（3）知识库的改进与归纳。

### 5-4 专家系统程序与一般问题求解软件程序有何不同？开发专家系统与开发其他软件的任务有何不同？

答：（第 189-190 页）

一般应用程序与专家系统的区别在于：前者把问题求解的知识隐含地编入程序，而后者则把其应用领域的问题求解知识单独组成一个实体，即为知识库。知识库的处理是通过与知识库分开的控制策略进行的。更确切地说，一般应用程序把知识组织为两级：数据级和程序级；大多数专家系统则将知识组织成三级：数据、知识库和控制。

在**数据级**上，是已解决了的特定问题的说明性知识以及需要求解问题的有关事件的当前状态。在**知识库级**是专家系统的专门知识与经验。是否拥有大量知识是专家系统成功与否的关键，因而知识表示就成为设计专家系统的关键。在**控制程序级**，根据既定的控制策略和所求解问题的性质来决定应用知识库中的哪些知识。这里的控制策略是指推理方式。

### 5-5 基于规则的专家系统是如何工作的？其结构为何？

答：（第 191-192 页）

基于规则的专家系统是一个计算机程序，该程序使用一套包含在知识库内的规则对工作存储器内的具体问题信息（事实）进行处理，通过推理机推断出新的信息。

基于规则的专家系统的结构包括：（1）知识库；（2）工作存储器；（3）推理机；（4）用户界面（接口）；（5）开发（者）界面；（6）解释器；（7）外部程序；图 5.5（）

#### 5-6 什么是基于框架的专家系统？它与面向目标编程有何关系？

答：（第 194-195 页）

**定义 5.7** 基于框架的专家系统是一个计算机程序，该程序使用一组包含在知识库内的框架对工作存储器内的具体问题信息进行处理，通过推理机推断出新的信息。（第 195 页，第 3 段）

基于框架的专家系统就是建立在框架的基础之上的。一般概念存放在框架内，而该概念的一些特例则被表示在其他框架内并含有实际的特征值。基于框架的专家系统采用了面向目标编程技术，已提高系统的能力和灵活性。现在，基于框架的设计和面向目标的编程共享许多特征，以至于应用“目标”和“框架”这两个术语时，往往引起某些混淆。（第 194 页，倒数第 2 段）

面向目标编程涉及其所有数据结构均以目标形式出现。每个目标含有两种基本信息，即描述目标的信息和说目标能够做些什么的信息。应用专家系统的术语来说，每个目标具有陈述知识和过程知识。（第 194 页，最后 1 段）

#### 5-7 基于框架的专家系统的结构有何特点？其设计任务是什么？

答：（第 196-197 页）

基于框架的设计和面向目标的编程共享许多特征，面向目标系统中的一个目标，而在基于框架系统中看作一个框架。基于框架的专家系统能够提供基于规则专家系统所没有的特征，如继承、侧面、信息通信和模式匹配规则等，因而也就提供了一种更加强大的开发复杂系统的工具。也就是说，基于框架的专家系统具有比基于规则的专家系统更强的功能，适用于解决更复杂的问题。（第 197 页）

**基于框架的专家系统的主要任务：**（1）定义问题，包括对问题和结论的考察与综述；（2）分析领域，包括定义事物，事物特征，事件和框架结构；（3）定义类及其特征；（4）定义例及其框架结构；（5）确定模式匹配规则；（6）规定事物通信方法；（7）设计系统界面；（8）对系统进行评价；（9）对系统进行扩展，深化和扩宽知识。（第 196 页）

#### 5-8 为什么要提出基于模型的专家系统？试述神经网络专家系统的一般结构。

答：（第 200-201 页）

采用各种定性模型来设计专家系统，有两个方面的优点，一方面。它增加了系统的功能，提高了性能指标；另一方面，可独立地深入研究各种模型及其相关问题，把获得结果用于改进系统设计。综合各种模型专家系统无论在知识表示、知识获取还是知识应用上都比那些基于逻辑心里模型的系统具有更强的功能，从而有可能显著改进专家系统的设计。（第 200 页，第 4 段，第 6 段）

神经网络专家系统的一般结构：图 5.9 神经网络专家系统的基本机构。自动化获取模块输入、组织并存储专家提供的学习实例、选定神经网络的结构、调用神经网络的学习算法，为知识

库实现知识获取。当输入新的学习实例后，知识获取模块通过对新实例的学习，自动获得新的网络权值分布，从而更新知识库。**(第 201 页，第 8 段)**

## 第六章 习题 6

### 6-1 什么是学习和机器学习？为什么要研究机器学习？

答：（第 246，248 页）

按照人工智能大事西蒙的观点，学习就是系统在不断重复的工作中对本省能力的增强或改进，使得系统在下次执行同样任务或类似任务时，会比现在做的更好或效率更高。（第 246，第 2 段）

**定义 6.11 机器学习**是研究机器模拟人类的学习活动、获取知识和技能的理论和方法，以改善系统性能的学科。（第 248 页）

现有的计算机系统和人工智能系统大多数没有什么学习能力，至多也只是非常有限的学习能力，因而不能满足科技和生产提出的新要求。对机器学习的讨论和机器学习研究的进展，必将促使人工智能和整个科学技术的进一步发展。

### 6-5 试说明归纳学习的模型和学习方法

答：（第 253，255-256 页）

**归纳学习的一般模式为**

给定：（1）观察陈述（事实）F；（2）假定的初始归纳断言；（3）背景知识。

求：归纳断言（假设）H。能重言观察陈述，并满足背景知识。

假设 H 永真蕴涵事实 F，说明 F 是 H 的逻辑推理，则有：

$H \models F$ （读作 H 特殊化为 F） -- 演绎推理

或  $F \models H$ （读作 F 一般化或消解为 H） -- 归纳推理

**归纳学习方法：**（1）示例学习；（2）观察发现学习

### 6-6 什么是类比学习？其推理和学习过程为何？

答：（第 261-262 页）

**类比学习**（learning by analogy）就是通过类比，即通过对相似事物加以比较所进行的一种学习。

**类比推理过程**

(1) 回忆与联想

通过回忆与联想，在 S 中找出与当前情况相似的情况，这些情况是过去已经处理过的，有现成的解决方法及相关的知识。

(2) 选择

从找出的相似情况中选出与当前情况最相似的情况及其有关知识，以便提高推理的可靠性。

(3) 建立对应关系

在 S 与 T 的相似情况之间建立相似元素的对应关系，并建立起相应的映射。

(4) 转换

根据映射，把 S 中的有关知识引到 T 中来，从而建立起求解当前问题的方法或者学习到关于 T 的新知识。

类比学习主要包括如下 4 个过程：

- (1) 输入一组已知条件（已解决问题）和一组未完全确定的条件（新问题）。
- (2) 对输入的两组条件，根据其描述，按某种相似性的定义寻找两者可类比的对应关系。
- (3) 根据相似转换的方法，将已有问题的概念、特性、方法、关系等映射到新问题上，以获得待求解新问题所需要的新知识。
- (4) 对类推得到的新问题的知识进行校验。验证正确的知识存入知识库中，而暂时还无法验证的知识只能作为参考性知识，置于数据库中。

#### 6-11 考虑一个具有阶梯形阈值函数的神经网络，假设

- (1) 用一常数乘所有的权值和阈值；
- (2) 用一个常数加于所有权值和阈值。

试说明网络性能是否会变化

答：所谓阶梯形阈值函数就是激活函数为二值函数。(1) 不会；(2) 会

#### 6-13 什么是知识发现？知识发现与数据挖掘有何关系？

答：(第 276, 283 页)

定义 5.12 数据库中的**知识发现**是从大量数据中辨识出有效的、新颖的、潜在有用的、并可被理解的 模式的高级处理过程。(第 276)

**数据挖掘**是只是知识发现中的一个步骤，它主要是利用某些特定的知识发现算法，在一定的运算效率内，从数据中发现出有关的知识。(第 283 页，第 4 段)

#### 6-14 试说明知识发现的处理过程。

答：(第 277-278 页)

(1) **数据选择**。根据用户的需求从数据库中提取与 KDD 相关的数据。KDD 主要从这些数据中提取知识。

(2) **数据预处理**。主要是对(1)产生的数据进行在加工，检查数据的完整性及数据的一致性，对其中的噪声数据进行处理，对丢失的数据利用统计方法进行填补，形成发掘数据库。

(3) **数据变换**。即从发掘数据里选择数据。变换的方法主要是利用聚类分析和判别分析。指导数据变换的方式是通过人机交互让专家来指导数据的挖掘方向。

(4) **数据挖掘**。根据用户要求，确定 KDD 的目标是发现何种类型的知识，因为对 KDD 的不同要求会在具体的知识发现过程中采用不同的知识发现算法。

(5) **知识评价**。这一过程主要是用于对于所获取的规则进行价值评定，以决定所得的规则是否存入基础知识库。主要通过人机交互界面由专家依靠经验来评价。

#### 6-15 有哪几种比较常用的知识发现方法？试略加介绍。

答：(第 278 页，最后 1 段)

知识发现的方法有统计方法、机器学习方法、神经计算方法和可视化方法等。

#### 6-19 什么是深度学习？它有何特点？

答：(第 286-287 页)

**定义 6.13** 深度学习算法是一类基于生物学对人脑进一步认识，将神经-中枢-大脑的工作原理设计成一个不断迭代、不断抽象的过程，以便得到最优数据特征表示的机器学习算法；该算法从原始信号开始，先做底层抽象，然后逐渐向高层抽象迭代，由此组成深度学习算法的基本架构。

### 深度学习的一般特点

一般来说，深度学习算法具有如下特点：

（1）**使用多重非线性变换对数据进行多层抽象**。该算法采用级联模式的多层非线性处理单元来组织特征提取以及特征转换。在这种级联模型中，后继层的数据输入由其前一层的输出数据充当。按学习类型，该类算法又可归为有监督学习，例如分类；无监督学习，例如模式分析。

（2）**以寻求更适合的概念表示方法为目标**。这类算法通过建立更好的模型来学习数据表示方法。对于学习所用的概念特征值或者说数据的表征，一般采用多层结构进行组织，这也是该类算法的一个特色。高层的特征值由底层特征值通过推演归纳得到，由此组成了一个层次分明的数据特征或者抽象概念的表示结构；在这种特征值的层次结构中，每一层的特征数据对应着相关整体知识或者概念在不同程度或层次上的抽象。

（3）**形成一类具有代表性的特征表示学习（learning representation）方法**。在大规模无标识的数据背景下，一个观测值可以使用多种方式来表示，例如一幅图像、人脸识别数据、面部表情数据等，而某些特定的表示方法可以让机器学习算法学习起来更加容易。所以，深度学习算法的研究也可以看做是在概念表示基础上，对更广泛的机器学习方法的研究。深度学习一个很突出的前景便是它使用无监督的或者半监督的特征学习方法，加上层次性的特征提取策略，来代替过去手工方式的特征提取。

深度学习具有如下优点：

（1）采用非线性处理单元组成的多层结构，使得概念提取可以由简单到复杂。

（2）每一层中非线性处理单元的构成方式取决于要解决的问题；同时，每一层学习模式可以按需求调整为有监督学习或无监督学习。

（3）学习无标签数据优势明显。不少深度学习算法通常采用无监督学习形式来处理其他算法很难处理的无标签数据。现实生活中无标签数据存在更加普遍。

### 6-20 深度学习存在哪几种常用模型？

答：（第 288-292 页）

常见的深度学习模型包含以下几类：（1）卷积神经网络（2）循环神经网络（3）受限玻尔兹曼机（4）自动编码器（5）深度信念网络