

1、给出命题的谓词表示：不是每个计算机专业的学生都喜欢编程,凡是喜欢编程的人都喜欢计算机专业。有人喜欢梅花，有人喜欢菊花，有人既喜欢梅花又喜欢菊花。

定义谓词 $S(x)$ ：x 是计算机系学生 $L(x, programming)$ ：x 喜欢编程 $U(x, computer)$ ：x 使用计算机

将知识用谓词表示为： $\neg (\forall x) (S(x) \rightarrow L(x, programming) \wedge U(x, computer))$

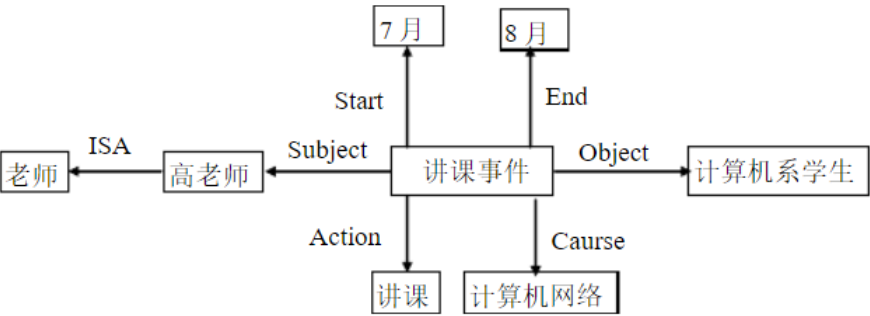
定义谓词 $P(x)$ ：x 是人 $L(x, y)$ ：x 喜欢 y

将知识用谓词表示为： $(\forall x) (P(x) \wedge L(x, programming) \rightarrow L(x, computer))$

定义谓词 $P(x)$ ：x 是人 $L(x, y)$ ：x 喜欢 y 其中，y 的个体域是{梅花，菊花}。

将知识用谓词表示为： $(\exists x)(P(x) \rightarrow L(x, 梅花) \vee L(x, 菊花) \vee L(x, 梅花) \wedge L(x, 菊花))$

2、给出命题的语义网络：高老师从 3 月到 7 月给计算机系学生讲《计算机网络》课

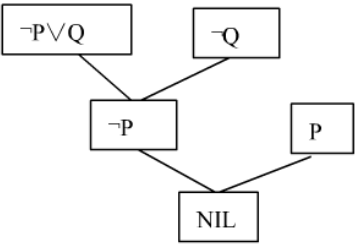


3、假设有以下一段天气预报：“北京地区今天白天晴，偏北风 3 级，最高气温 12°，最低气温 -2°，降水概率 15%。”请用框架表示这一知识。

Frame<天气预报>

地域：北京
时段：今天白天
天气：晴
风向：偏北
风力：3 级
气温：最高：12 度
最低：-2 度
降水概率：15%

解：(1) 不可满足，其归结过程为：



4、判断下列子句集中哪些是不可满足的，写出归结过程：

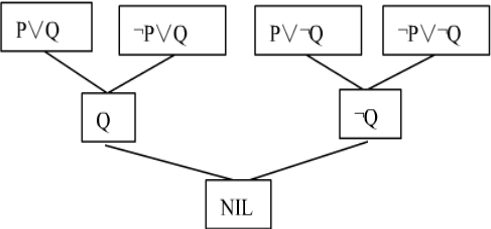
(1){ $\neg P \vee Q, \neg Q, P, \neg P$ }

(2){ $P \vee Q, \neg P \vee Q, P \vee \neg Q, \neg P \vee \neg Q$ }

(3){ $P(y) \vee Q(y), \neg P(f(x)) \vee R(a)$ }

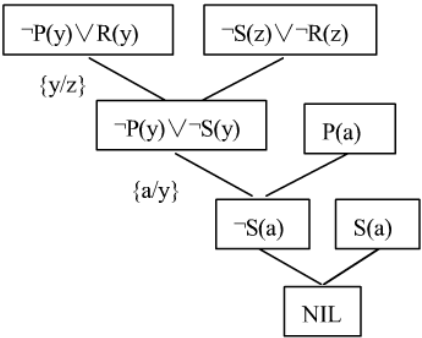
(4){ $\neg P(x) \vee Q(x), \neg P(y) \vee R(y), P(a), S(a), \neg S(z) \vee \neg R(z)$ }

(2) 不可满足，其归结过程为：



(3) 不是不可满足的，原因是不能由它导出空子句。

(4) 不可满足，其归结过程为：

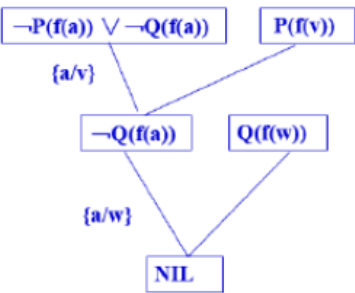
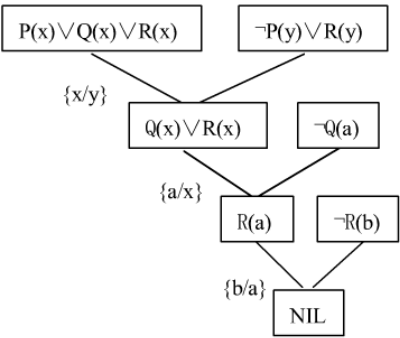


(5) { $\neg P(x) \vee Q(f(x), a), \neg P(h(y)) \vee Q(f(h(y)), a) \vee \neg P(z)$ }

(6) { $P(x) \vee Q(x) \vee R(x), \neg P(y) \vee (y), \neg Q(a), \neg R(b)$ } R

(5) 不是不可满足的，原因是不能由它导出空子句。

(6) 不可满足，其归结过程为：



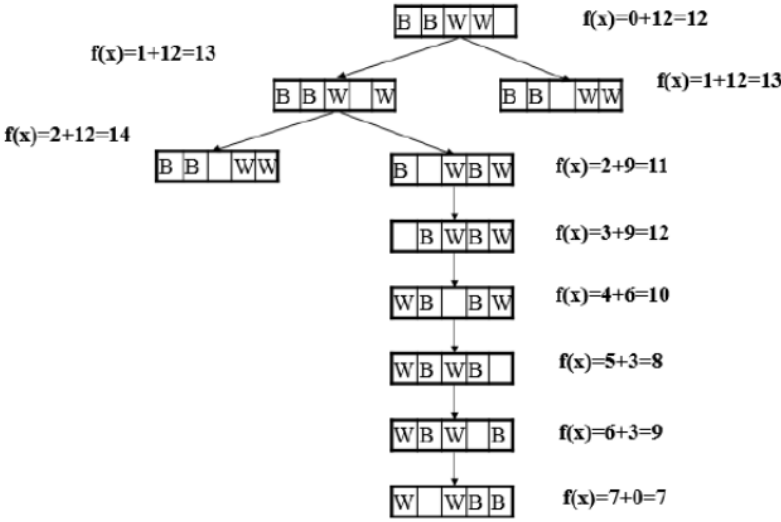
4、证明 G 是 F 的逻辑结论 F: $(\exists x)(\exists y)(P(f(x)) \wedge (Q(f(y))))$ G: $P(f(a)) \wedge P(y) \wedge Q(y)$

证：先转化成子句集。对 F，进行存在固化，有 $P(f(v)) \wedge (Q(f(w)))$
得以下两个子句 $P(f(v)), Q(f(w))$ 对 $\neg G$ ，有 $\neg P(f(a)) \vee \neg P(y) \vee \neg Q(y)$
先进行内部合一，设合一 $\{f(a)/y\}$ ，则有因子 $\neg P(f(a)) \vee \neg Q(f(a))$
再对上述子句集进行归结演绎推理。其归结树如下图所示，即存在一个到空子句的归结过程。 因此 G 为真。

5、设已知：(1)能阅读的生物是识字的；(2)海豚不识字；(3)有些海豚是很聪明的。

请用归结演绎推理证明：有些很聪明的人并不识字。课本和 PPT 有

解：设 $h(x)$ =每个 W 左边的 B 的个数， $f(x)=d(x)+3*h(x)$ ，其搜索树如下：

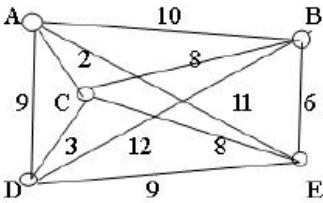


5 设有如下结构的移动将牌游戏：B B W W E 其中，B 表示黑色将牌，W 表是白色将牌，E 表示空格。游戏的规定走法是：

- (1) 任意一个将牌可移入相邻的空格，规定其代价为 1；
- (2) 任何一个将牌可相隔 1 个其它的将牌跳入空格，其代价为跳过将牌的数目加 1。

游戏要达到的目标是把所有 W 都移到 B 的左边。对这个问题，请定义一个启发函数 $h(n)$ ，并给出用这个启发函数产生的搜索树。你能否判别这个启发函数是否满足下界要求？在求出的搜索树中，对所有节点是否满足单调限制？

6、如右边 5 个城市的交通图，城市之间的连线旁边的数字是城市之间路程的费用。要求从 A 城出发，经过其它各城市一次且仅一次，最后回到 A 城，请画出搜索树，找出一条最优线路



解：这个问题又称为旅行商问题（travelling salesman problem, TSP）或货郎担问题，是一个较有普遍性的实际应用问题。根据数学理论，对 n 个城市的旅行商问题，其封闭路径的排列总数为：

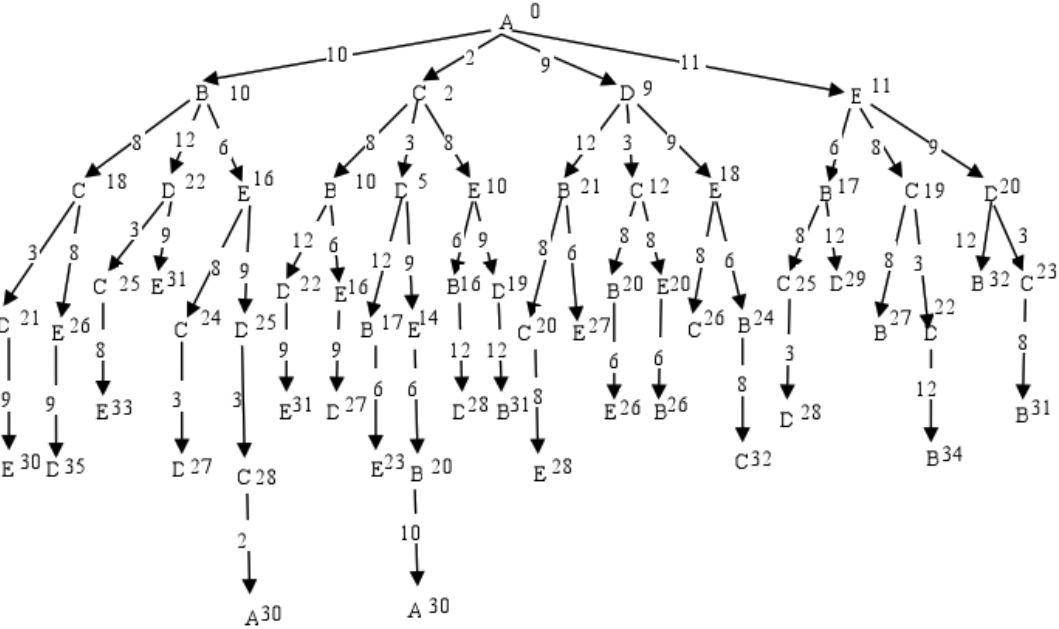
$$(n!)/n=(n-1)!$$

其计算量相当大。例如，当 $n=20$ 时，要穷举其所有路径，即使用一个每秒一亿次的计算机来算也需要 350 年的时间。因此，对这类问题只能用搜索的方法来解决。

下图是对图 4-32 按最小代价搜索所得到的搜索树，树中的节点为城市名称，节点边上的数字为该节点的代价 g。其计算公式为

$$g(n_{i+1})=g(n_i)+c(n_i, n_{i+1})$$

其中， $c(n_i, n_{i+1})$ 为节点 n_i 到 n_{i+1} 节点的边代价。



可以看出，其最短路径是
A-C-D-E-B-A
或 A-B-E-D-C-A
其实，它们是同一条路径。

图 4.32 的最小代价搜索树

7、设有如下推理规则

r1: IF E1 THEN (2, 0.00001) H1 r2: IF E2 THEN (100, 0.0001) H1
r3: IF E3 THEN (200, 0.001) H2 r4: IF H1 THEN (50, 0.1) H2
且已知 P(E1)= P(E2)= P(H3)=0.6, P(H1)=0.091, P(H2)=0.01 P(E1| S1)=0.84, P(E2|S2)=0.68, P(E3|S3)=0.36
请用主观 Bayes 方法求 P(H2|S1, S2, S3)=?

6 设有如下一组推理规则:

r1: IF E1 THEN E2 (0.6) r2: IF E2 AND E3 THEN E4 (0.7)
r3: IF E4 THEN H (0.8) r4: IF E5 THEN H (0.9)
且已知 CF(E1)=0.5, CF(E3)=0.6, CF(E5)=0.7。求 CF(H)=?
解: (1) 先由 r1 求 CF(E2) CF(E2)=0.6 × max{0,CF(E1)} =0.6 × max{0,0.5}=0.3
(2) 再由 r2 求 CF(E4)=0.7×max{0, min{CF(E2) , CF(E3) }} =0.7×max{0,min{0.3, 0.6}}=0.21
(3) 再由 r3 求 CF1(H) CF1(H)= 0.8 × max{0,CF(E4)} =0.8 × max{0, 0.21}=0.168
(4) 再由 r4 求 CF2(H) CF2(H)= 0.9 ×max{0,CF(E5)} =0.9 ×max{0, 0.7}=0.63
(5) 最后对 CF1(H)和 CF2(H)进行合成, 求出 CF(H)=CF1(H)+CF2(H)-CF1(H)×CF2(H)=0.692

10 设有如下一组知识:

r1: IF E1 THEN H (0.9) r2: IF E2 THEN H (0.6)
r3: IF E3 THEN H (-0.5) r4: IF E4 AND (E5 OR E6) THEN E1 (0.8)
已知: CF(E2)=0.8, CF(E3)=0.6, CF(E4)=0.5, CF(E5)=0.6, CF(E6)=0.8 求: CF(H)=?

解: 由 r4 得到:
CF(E1)=0.8×max{0, CF(E4 AND (E5 OR E6))} = 0.8×max{0, min{CF(E4), CF(E5 OR E6))} =0.8×max{0, min{CF(E4), m
ax{CF(E5), CF(E6))}} =0.8×max{0, min{CF(E4), max{0.6,0.8}} } =0.8×max{0, min{0.5, 0.8}} =0.8×max{0, 0.5} = 0.4
由 r1 得到: CF1(H)=CF(H, E1)×max{0, CF(E1)} =0.9×max{0, 0.4} = 0.36
由 r2 得到: CF2(H)=CF(H, E2)×max{0, CF(E2)} =0.6×max{0, 0.8} = 0.48
由 r3 得到: CF3(H)=CF(H, E3)×max{0, CF(E3)} =-0.5×max{0, 0.6} = -0.3
根据结论不精确性的合成算法, CF1(H)和 CF2(H)同号, 有:

$$\begin{aligned} CF_{1,2}(H) &= CF_1(H) + CF_2(H) - CF_1(H) \times CF_2(H) \\ &= 0.36 + 0.48 - 0.36 \times 0.48 \\ &= 0.84 - 0.17 = 0.67 \\ CF_{12}(H) \text{和 } CF_3(H) \text{异号, 有:} \\ CF_{1,2,3}(H) &= \frac{CF_{1,2}(H) + CF_3(H)}{1 - \min\{|CF_{1,2}(H)|, |CF_3(H)|\}} \\ &= \frac{0.67 - 0.3}{1 - \min\{0.67, 0.3\}} = \frac{0.37}{0.7} \\ &= 0.53 \end{aligned}$$

即综合可信度为 CF(H)=0.53

7 设训练例子集如下表所示:

序 号	属 性		分 类
	x ₁	x ₂	
1	T	T	+
2	T	T	+
3	T	F	-
4	F	F	+
5	F	T	-
6	F	T	-

请用 ID3 算法完成其学习过程。

在本题中, 当 x1=T 时, 有: ST={1, 2, 3} 当 x1=F 时, 有: SF={4, 5, 6}
其中, ST 和 SF 中的数字均为例子集 S 中例子的序号, 且有 |S|=6, |ST|=|SF|=3。
由 ST 可知: P(+)=2/3, P(-)=1/3 则有: H(ST)= - (P(+))log₂ P(+)- P(-)log₂ P(-))
= - ((2/3)log₂(2/3)- (1/3)log₂(1/3)) =0.9183 再由 SF 可知: PSF(+)=1/3, PSF(-)=2/3
则有: H(S_F)=-(P_{SF}(+)log₂ P_{SF}(+)-PSF(-)log₂ PSF(-)) = - ((2/3)log₂(2/3)- (1/3)log₂(1/3)) = 0.9183
将 H(S_T)和 H(S_F)代入条件熵公式, 有:

11 设有如下知识:

r1: IF E1 (0.6) AND E2 (0.4) THEN E5 (0.8)
r2: IF E3 (0.5) AND E4 (0.3) AND E5 (0.2) THEN H (0.9)
已知: CF(E1) =0.9, CF(E2) =0.8, CF(E3) =0.7, CF(E4) =0.6 求: CF(H) =? 解: CF(E1 AND E2)=0.9*0.6+0.8*0.4=0.86
CF(E5)=0.86*0.8=0.69 CF(E3 AND E4 AND E5)
=0.7*0.5+0.6*0.3+0.69*0.2=0.67 CF(H)=0.67*0.9=0.60

设根节点为 S, 尽管它包含了所有的训练例子, 但却没有包含任何分类信息, 因此具有最大的信息熵。即: H(S)= - (P(+))log₂ P(+)- P(-)log₂ P(-) 式中 P(+)=3/6, P(-)=3/6
即有 H(S)= - ((3/6)*log₂ (3/6) - (3/6)*log₂ (3/6)) = -0.5*(-1) - 0.5*(-1) = 1
按照 ID3 算法, 需要选择一个能使 S 的期望熵为最小的一个属性对根节点进行扩展, 因此我们需要先计算 S 关于每个属性的条件熵: H(S|x_i)= (|ST| / |S|)* H(ST) + (|SF| / |S|)* H(SF)
其中, T 和 F 为属性 x_i 的属性值, ST 和 SF 分别为 x_i=T 或 x_i=F 时的例子集, |S|、|ST|和|SF|分别为例子集 S、ST 和 SF 的大小。下面先计算 S 关于属性 x₁ 的条件熵:

$$H(S|x_1)=(|ST|/|S|)H(ST)+(|SF|/|S|)H(SF)=(3/6)*0.9183+(3/6)*0.9183=0.9183$$

下面再计算 S 关于属性 x2 的条件熵：

在本题中，当 x2=T 时，有：ST={1, 2, 5, 6} 当 x2=F 时，有：SF={3, 4}

其中，ST 和 SF 中的数字均为例子集 S 中的各个例子的序号，且有|S|=6，|ST|=4，|SF|=2。由 ST 可知：
PST (+) = 2/4 PST (-) = 2/4 则有：

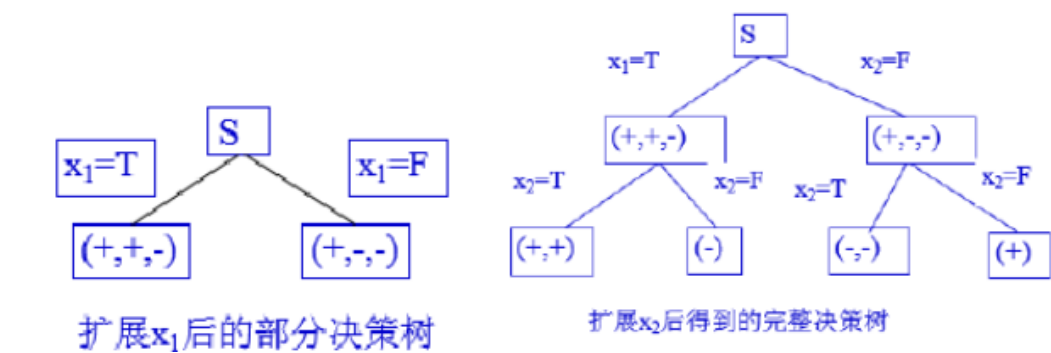
$$H(ST)=-(P_{ST}(+)\log_2 P_{ST}(+)-P_{ST}(-)\log_2 P_{ST}(-))=-((2/4)\log_2(2/4)-(2/4)\log_2(2/4))=1$$

再由 SF 可知：PSF (+)=1/2 PSF (-)=1/2 则有：

$$H(SF)=-(P(+)\log_2 P(+)-P(-)\log_2 P(-))=-((1/2)\log_2(1/2)-(1/2)\log_2(1/2))=1$$

将 H(ST)和 H(SF)代入条件熵公式有 H(S|x2)=(|ST|/|S|)H(ST)+(|SF|/|S|)H(SF)=(4/6)*1+(2/6)*1=1

可见，应该选择属性 x1 对根节点进行扩展。用 x1 对 S 扩展后所得到的部分决策树如下图。



8、给出训练集表格，采用 ID3 算法完成其学习过程。P24 补充

在该树中，节点“x3=1, y3 ”为决策方案 y3。由于 y3 已是具体的决策方案，故该节点的信息熵为 0，已经为叶节点。

节点“x3=2, x1, x2 ?”的含义是需要进一步考虑学历和专业这两个属性，它是一个中间节点，还需要继续扩展。至于其扩展方法与上面的过程类似。

通过计算可知，该节点对属性 x1 和 x2，其条件熵均为 1。由于它对属性 x1 和 x2 的条件熵相同，因此可以先选择 x1，也可以先选择 x2。

依此进行下去，若先选择 x1 可得到如图 6.6 所示的最终决策树；若先选择 x2 可得到如图 7.7 所示的最终决策树。

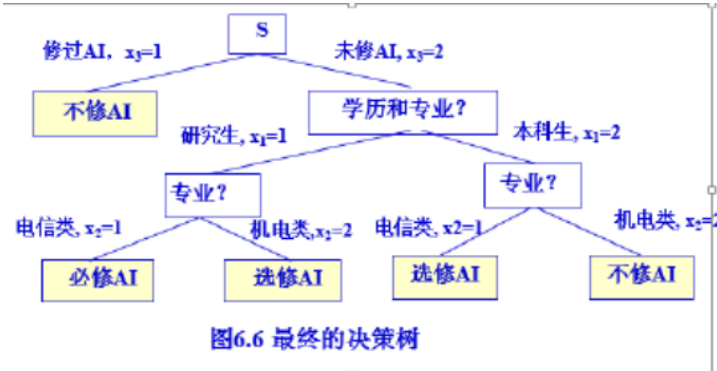


表1 天气预报数据集例子

Outlook	Temperature	Humidity	Windy	Play?
sunny	hot	high	false	no
sunny	hot	high	true	no
overcast	hot	high	false	yes
rain	mild	high	false	yes
rain	cool	normal	false	yes
rain	cool	normal	true	no
overcast	cool	normal	true	yes
sunny	mild	high	false	no
sunny	cool	normal	false	yes
rain	mild	normal	false	yes
sunny	mild	normal	true	yes
overcast	mild	high	true	yes
overcast	hot	normal	false	yes
rain	mild	high	true	no

可以看出，一共 14 个样例，包括 9 个正例和 5 个负例。那么当前信息的熵计算如下

$$Entropy(S)=-\frac{9}{14}\log_2\frac{9}{14}-\frac{5}{14}\log_2\frac{5}{14}=0.940286$$

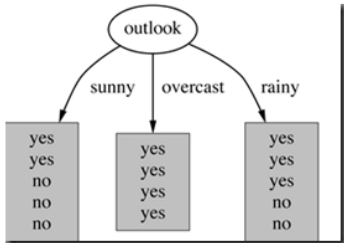
在决策树分类问题中，信息增益就是决策树在进行属性选择划分前和划分后信息的差值。假设利用属性 Outlook 来分类，那么如下图

划分后，数据被分为三部分了，那么各个分支的信息熵计算如下

$$Entropy(sunny)=-\frac{2}{5}\log_2\frac{2}{5}-\frac{3}{5}\log_2\frac{3}{5}=0.970951$$

$$Entropy(overcast)=-\frac{4}{4}\log_2\frac{4}{4}-0\cdot\log_20=0$$

$$Entropy(rainy)=-\frac{3}{5}\log_2\frac{3}{5}-\frac{2}{5}\log_2\frac{2}{5}=0.970951$$



那么划分后的信息熵为

$$Entropy(S|T) = \frac{5}{14} \cdot 0.970951 + \frac{4}{14} \cdot 0 + \frac{5}{14} \cdot 0.970951 = 0.693536$$

$Entropy(S|T)$ 代表在特征属性 T 的条件下样本的

条件熵。那么最终得到特征属性 T 带来的信息增益为 $IG(T) = Entropy(S) - Entropy(S|T) = 0.24675$ 信息增益的

计算公式如下 $IG(S|T) = Entropy(S) - \sum_{value(T)} \frac{|S_v|}{S} Entropy(S_v)$ 其中 S 为全部样本集合, $value(T)$ 是属性 T 所有取值

的集合, v 是 T 的其中一个属性值, S_v 是 S 中属性 T 的值为 v 的样例集合, $|S_v|$ 为 S_v 中所含样例数。

在决策树的每一个非叶子结点划分之前, 先计算每一个属性所带来的信息增益, 选择最大信息增益的属性来划分, 因为信息增益越大, 区分样本的能力就越强, 越具有代表性, 很显然这是一种自顶向下的贪心策略。以上就是 **ID3 算法** 的核心思想。

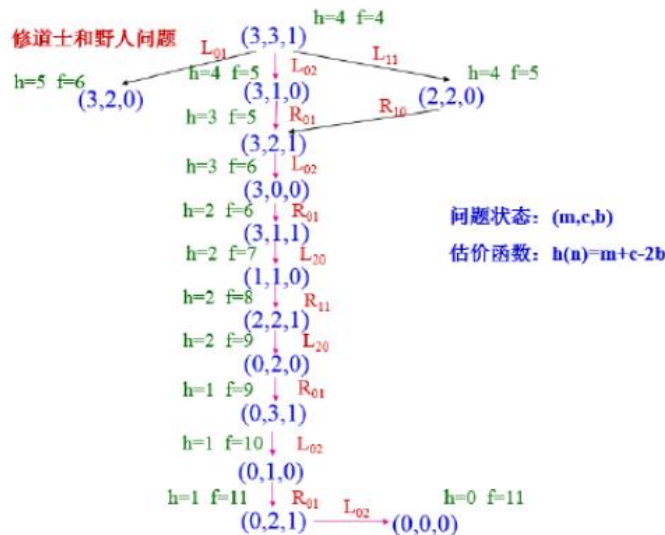
9、A*算法求解八数码问题

9、修道士和野人问题

解: 用 m 表示左岸的修道士人数, c 表示左岸的野人数, b 表示左岸的船数, 用三元组 (m, c, b) 表示问题的状态。

对 A* 算法, 首先需要确定估价函数。设 $g(n)=d(n)$, $h(n)=m+c-2b$, 则有 $f(n)=g(n)+h(n)=d(n)+m+c-2b$

其中, $d(n)$ 为节点的深度。通过分析可知 $h(n) \leq h^*(n)$, 满足 A* 算法的限制条件。 M-C 问题的搜索过程如下图所示。



估计、生物、医学、经济等领域已成功地解决了许多现代计算机难以解决的实际问题，表现出了良好的智能特性。神经网络的研究可以分为理论研究和应用研究两大方面。理论研究可分为以下两类：1、利用神经生理与认知科学研究人类思维以及智能机理。2、利用神经基础理论的研究成果，用数理方法探索功能更加完善、性能更加优越的神经网络模型，深入研究网络算法和性能，如：稳定性、收敛性、容错性、鲁棒性等；开发新的网络数理理论，如：神经网络动力学、非线性神经场等。应用研究可分为以下两类：1、神经网络的软件模拟和硬件实现的研究。2、神经网络在各个领域中应用的研究。这些领域主要包括：模式识别、信号处理、知识工程、专家系统、优化组合、机器人控制等。随着神经网络理论本身以及相关理论、相关技术的不断发展，神经网络的应用定将更加深入。

(1) 对智能和机器关系问题的认识将进一步增长。(2) 神经计算和进化计算将有重大的发展。(3) 神经网络结构和神经元芯片的作用将不断扩大。

12、什么是遗传算法，核心流程及步骤？和粒子群优化的异同？进化算法(EA)

通过实验和文献分析，我们对遗传算法 GA、粒子群算法 PSO、差分进化算法 DE 的一些指标分别进行分析现归纳如下：(1) 编码标准：GA 采用二进制编码，PSO、DE 都采用实数编码，近年来许多学者通过整数编码将 GA 算法、PSO 算法应用与求解离散型问题，特别是 0-1 非线性优化为题，整数规划问题、混合整数规划问题，而离散的 DE 算法则研究的比较少，而采用混合编码技术的 DE 算法则研究更少。

(2) 参数设置问题：DE 算法主要有两个参数要调整，而且参数设置对结果影响不太明显，因此更容易使用.相对于 GA 和 PSO 算法的参数过多，不同的参数设置对最终结果影响也比较大，因此在实际使用中，要不断调整，加大了算法的使用难度.高维问题在实际问题中，由于转化为个体的向量维数非常高，因此算法对高维问题的处理，将是很重要的.只有很好的处理高维问题，算法才能很好的应用于实际问题。

(3) 高维问题：GA 对高维问题收敛速度很慢甚至很难收敛，但是 PSO 和 DE 则能很好解决.尤其是 DE 算法，收敛速度很快而且结果很精确。

(4) 收敛性能：对于优化问题，相对 GA, DE 和 PSO 算法收敛速度比较快，但 PSO 容易陷入局部最优解，且算法不稳定。

(5) 应用广泛性：由于 GA 算法发明比较早，因此应用领域比较广泛，PSO 算法自从发明以来，已成为研究热点问题，这方面应用也比较多，而 DE 算法近几年才引起人们的关注而且算法性能好，因此应用领域将会增多。

13、简述机器学习的过拟合问题，怎么解决？

过拟合 (overfitting) 是指在模型参数拟合过程中的问题，由于训练数据包含抽样误差，训练时，复杂的模型将抽样误差也考虑在内，将抽样误差也进行了很好的拟合。具体表现就是最终模型在训练集上效果好；在测试集上效果差。模型泛化能力弱。它通常发生在模型过于复杂的情况下，如参数过多等。overfitting 会使得模型的预测性能变弱，并且增加数据的波动性。过拟合主要是有两个原因造成的：数据太少+模型太复杂。

1. 获取更多数据。从数据源头获取更多数据、根据当前数据集估计数据分布参数，使用该分布产生更多数据、数据增强。2. 使用合适的模型。通过使用合适复杂度的模型来防止过拟合问题，让其足够拟合真正的规则，同时又不至于拟合太多抽样误差。对于神经网络从以下四个方面来限制网络能力：网络结构、训练时间、限制权值，也叫正则化、增加噪声（在输入中、在权值上、对网络的响应加噪声）3. 结合多种模型：Bagging、Boosting、Dropout4. 贝叶斯方法。

14、专家系统基本结构，架构？各核心模块功能。15、简述上下文无关文法原理，给出实例。文法分析树？16、机器学习的基本概念，思想，步骤，原理，过程，测试划分？

1、公司雇佣问题归结反演

已知：“张和李是同班同学，如果 x 和 y 是同班同学，则 x 的教室也是 y 的教室，现在张在 302 教室上课。”问：“现在李在哪个教室上课？”解：首先定义谓词：

$C(x, y)$ x 和 y 是同班同学； $At(x, u)$ x 在 u 教室上课。

把已知前提用谓词公式表示如下： $C(zhang, li)$ $(\forall x)(\forall y)(\forall u)(C(x, y) \wedge At(x, u) \rightarrow At(y, u))$ $At(zhang, 302)$

把目标的否定用谓词公式表示如下： $\neg(\exists v)At(li, v)$

把上述公式化为子句集： $C(zhang, li)$ $\neg C(x, y) \vee \neg At(x, u) \vee At(y, u)$ $At(zhang, 302)$

把目标的否定化成子句式，并用重言式 $\neg At(li, v) \vee At(li, v)$ 代替之。

把此重言式加入前提子句中，得到一个新的子句集，对这个新的子句集，应用归结原理求出其证明树。其求解过程如下图所示。

