

学号

班级

姓名

装订线内不要答题

东北大学秦皇岛分校

课程名称：人工智能 试卷：(A) 考试形式：闭卷  
授课专业：计算机科学与技术 考试日期：2017 年 5 月 26 日 试卷：共 5 页

题号	一	二	三	四	五
得分					
阅卷人					
题号	六	七	八	九	总分
得分					
阅卷人					

提醒：在灰色底纹留空处写下（选择、填空、简答等）问题答案。

一、（启发式搜索，共 20 分）K×K 方格里的 Flood-it 游戏如图 1。

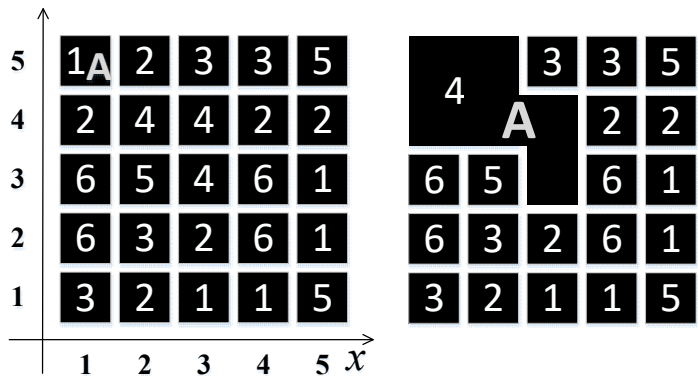


图 1 Flood-it 状态

- 区域。将所有的、在水平或垂直方向有公共邻边的、且标注了数字  $d$  的方块所构成的方块集合 A，定义为标注了数字  $d$  的一块区域 A。用坐标  $[x, y]$  表示方块位置，如图 1(a)中方块集合  $\{[3, 5], [4, 5]\}$  构成了标注“3”的一块区域。
- 吞并。区域 A 吞并与之有公共邻边的，标注指定数字  $d$  的所有区域：区域 A 与每一个与之有公共邻边，且标注了数字  $d$  的区域合并，且将合并

后的新区域 A'标注为  $d$ 。例如，图 1(a)中，区域 A 可一次吞并标注“2”的两个相邻区域  $\{[2, 1]\}$  和  $\{[1, 2]\}$ 。

- 任务。最左上角（含方块  $[1, 1]$ ）区域 A 用最少的步数合并所有区域。
- 1. （选择）在 A\*搜索中，对于任意结点  $node$ ,  $f(node) = g(node) + h(node)$ 。  
 $f(node) = (2 \text{ 分})$ ,  $g(node) = (2 \text{ 分})$ ,  $h(node) = (2 \text{ 分})$ 。  
A) 从初始节点  $s_0$  到最近的目标结点  $g$  的最短路径的一个乐观估计  
B) 从初始节点  $s_0$  到当前节点  $node$  的各个已搜索路径的最小代价  
C) 从当前节点  $node$  到最近的目标结点  $g$  的最短路径的乐观估计
- 2. （选择）就图 1(b)的具体状态  $n$  而言， $g(n) = (2 \text{ 分})$ 。  
A) 2      B) 3      C) 4      D) 以上都不对

问题 3~5 会用到下表中的启发函数。

$h(node)$	$h(node)$ 函数的说明	举例
$h_1$	状态 $node$ 中待合并区域的个数	$h_1(s_0) = 16$ $h_1(n) = 13$
$h_2$	状态 $node$ 中各待合并区域所标注的不同数字的数目	$h_2(s_0) = 6$ $h_2(n) = 5$
$h_3$	合并最右上角（含方块 $[5, 5]$ 的）区域所需的最少步数	$h_3(n) = 2$
$h_4$	合并最右下角（含方块 $[5, 1]$ 的）区域所需的最少步数	$h_4(n) = 5$
$h_5$	合并最左下角（含方块 $[1, 1]$ 的）区域所需的最少步数	$h_5(n) = 3$

3. （选择）上表中不可纳的启发函数为 (2 分)。  
A)  $h_1(node)$       B)  $h_2(node)$       C)  $h_3(node)$       D) 以上都不对
4. （选择）以下复合的启发函数中，不可纳的是 (2 分)。  
A)  $\frac{1}{2}(h_2(node)+h_3(node))$       B)  $\min\{h_2(node), h_3(node)\}$   
C)  $3h_3(node)$       D)  $\frac{1}{3}h_3(node)$
5. （填空）设计一个启发函数，使其剪枝效率都不比  $h_1 \sim h_5$  中任何一个

学 号

班 级

姓 名

装  
订  
线

装  
订  
线  
内  
不  
要  
答  
题

差  $h_6(node) =$  (2 分) 。

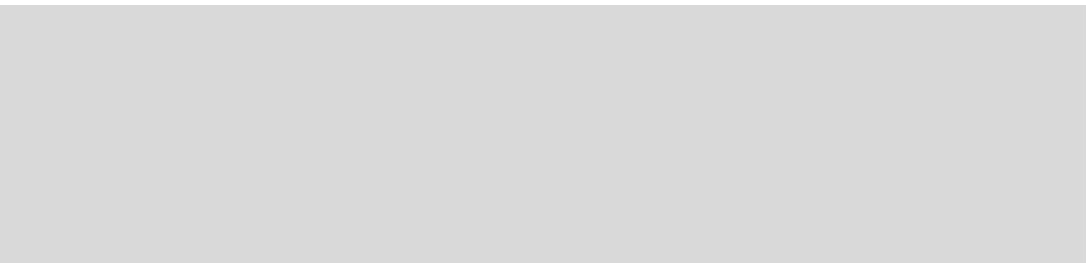
6. (填空) 考虑当  $K=100$  时, 每个方块随机标注数字 1~6, 生成一个  $K \times K$  Flood-it 初始状态。此初始状态下, 在普通 PC 机上, 不限时间, A\*搜索程序失败的最大风险是 (1 分) ; 给出解决该问题的一个方法 (1 分) 。

7. (简答) 在图 1(b)的状态  $n$  中, 下一步可合并的候选区域标记数字分别是: 6、5、3、2。此时, 最佳选择是合并标记“6”的区域, 进而得到新状态  $n'$ 。试说明: 假设  $n$  在最佳路径上, 则新状态  $n'$ 必然也在最佳路径上。

答: (2 分)

8. (简答) 随着搜索获得的经验的增加, 启发函数  $h(node)$ 总给出相同的启发值恐非最佳方案。在 Flood-it 问题中, 请设计一个可利用历史经验改进启发函数  $h(node)$ 的设计和实现方案。

答: (2 分)



二、(局部搜索, 10 分)

1. (选择) 若给定的八数码问题有解, 则最陡爬山法 (2 分) 找到解; 基于随机重爬的最陡爬山策略 (2 分) 找到解。  
A) 能                      B) 不能                      C) 未必能
2. (选择) 以下两种方法相比, (2 分) 更容易陷入局部最优。  
A) 束搜索                      B) 遗传算法
3. (选择) 遗传算法中, (2 分) 操作更倾向于“探索”; 交叉操作更倾向于“利用”; 基于适应度的 (2 分) 操作模拟了自然选择。  
A) 交叉                      B) 变异                      C) 选择

三、(约束满足问题, 10 分)

1. (选择) 给定三个变量 A、B、C, 每个变量的初始值域都是  $\{1, 2, 3\}$ , 约束为  $\{A \neq B, A \neq C, B > C\}$ 。为了保证弧相容, A 的值域变为 (2 分) ; B 的值域变为 (2 分) ; C 的值域变为 (2 分) 。
2. (选择) 在约束满足问题中用回溯法时, 变量选择的顺序为 (2 分) ; 变量的值的选择顺序为 (2 分) 。
- A) 剩余值最少的变量优先                      B) 最多剩余值的变量优先  
C) 变量的最少约束值优先                      D) 变量的最多约束值优先

四、(对抗搜索, 6 分)  $\alpha$ - $\beta$ 搜索自左向右, 按深度优先访问博弈树。

学号

班级

姓名

装订线内不要答题

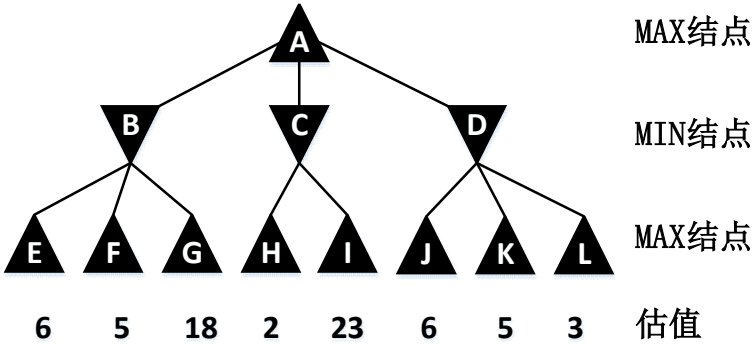


图2 博弈树

1. （选择）结点 B 的倒推值是（2 分）。 A) 6 B) 5 C) 18
2. （选择）被剪枝的结点为（2 分）。  
A) I, L B) I, K, L C) G, I D) J, K, L
3. （选择）为剪去更多结点，应调整（2 分）。  
A) A 的子结点顺序为 C、D、B B) B 的子结点顺序为 G、E、F  
C) C 的子结点顺序为 I、H D) D 的子结点顺序为 L、J、K

五、（贝叶斯网络，20 分）考虑 3 个布尔变量构成的贝叶斯网络。  
[符号约定：正体大写 A 表示随机变量；正体小写 +a, -a 表示 A 值域中的常量；斜体小写 *a* 表示用于遍历 A 的整个值域的迭代变量]

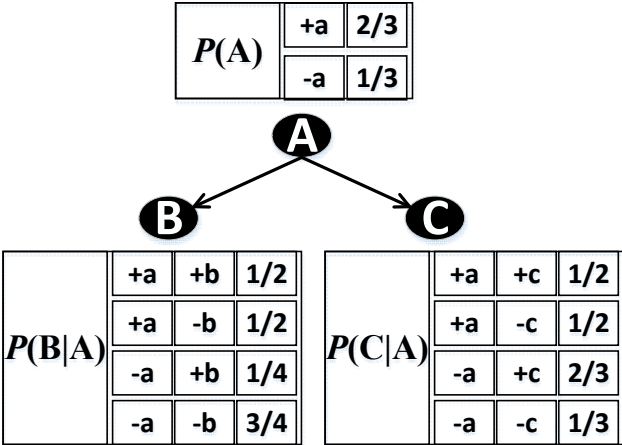


图3 贝叶斯网络

1. （选择题）依照上述拓扑结构，错误的是（2 分）。  
A)  $P(A,B,C) = P(A)P(B)P(C)$  B)  $P(A,B,C) = P(A)P(B|A)P(C|A)$

C)  $P(A, B, C)=P(A)P(B,C|A)$  D)  $P(A,B,C) = P(A,B) P(C|A)$

2. （每空 2 分，共 4 分）根据贝叶斯网络补全联合概率分布。

A	B	C	P(A, B, C)	A	B	C	P(A, B, C)
+a	+b	+c	1/6	-a	+b	+c	1/18
+a	+b	-c	1/6	-a	+b	-c	(2 分)
+a	-b	+c	(2 分)	-a	-b	+c	1/6
+a	-b	-c	1/6	-a	-b	-c	1/6

3. （选择题，每空 2 分，共 6 分）依照网络拓扑，精确推理概率  $P(+c)$  的过程见下表，请从 A) ~ C) 选出依据，直接填入下表对应位置。

贝叶斯网络的精确推理	将 A) ~ C) 写入正确的位置
$P(+c)$	-----
$= \sum_a \sum_b P(+c, a, b)$	依据：(2 分)
$= \sum_{a \in A} \sum_{b \in B} P(a) P(+c   a) P(b   a)$	依据：(2 分)
$= \sum_{a \in A} P(a) P(+c   a) \sum_{b \in B} P(b   a)$	依据：提取公因子
$= \sum_{a \in A} P(a) P(+c   a)$	依据：(2 分)

A) 加法公式 B) 乘法公式 C) 概率公理(归一化公理)

4~6 题，采用基于采样的近似推理计算  $P(+c | +b)$ 。给出以下样本：

Sample1	Sample2	Sample3	Sample4
(+a, +b, +c)	(+a, -b, -c)	(+a, +b, -c)	(-a, +b, +c)

4. （填空）直接采样的方法， $P(+c | +b) =$ （2 分）。
5. （选择）要估计  $P(+c | +b)$ ，下列子集中，（2 分）可能是由拒绝采样，或似然加权采样得到的样本集合。  
A) Sample1, Sample2, Sample3 B) Sample1, Sample3, Sample4  
C) Sample1, Sample2 D) Sample2, Sample3, Sample4

学号

班级

姓名

装订线内不要答题

6. （选择）用拒绝采样估计  $P(+c \mid +b) =$  （2分）。

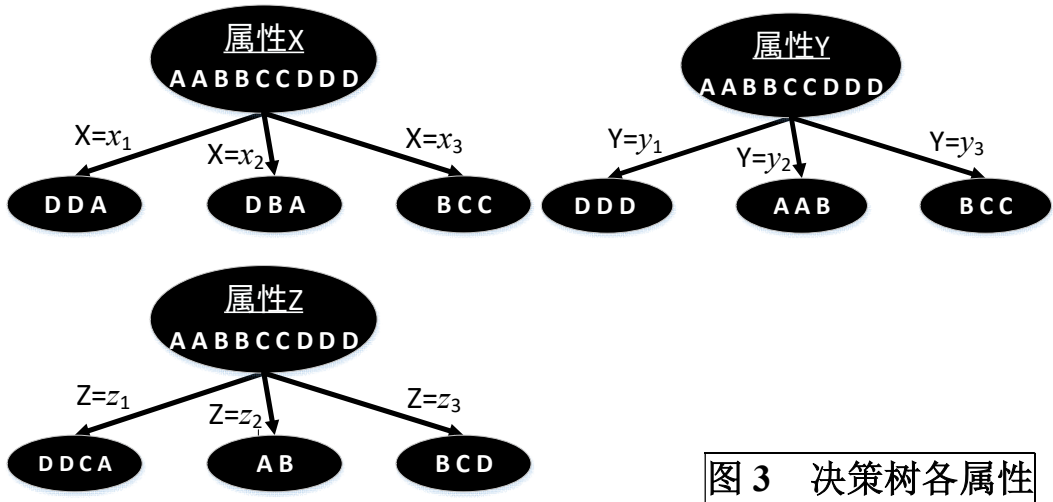
- A) 2/3      B) 3/4      C) 3/5      D) 以上都不对

7. （选择）用拒绝采样估计  $P(+c \mid +b) =$  （2分）。

- A) 2/3      B) 3/4      C) 3/5      D) 以上都不对

六、（决策树、提升， 14分）

测试集包含 9 个训练样例，分别属于 A、B、C、D 四类，每个训练样例都由 X、Y、Z 三个输入属性描述。考虑 ID3 决策树学习。



1. （填空）列出集合 {A, A, B, B, C, C, D, D, D} 的熵的算术表达式（不必计算数值） $H([2, 2, 2, 3]) =$  （2分）

2. （选择）给定下列熵， （2分） 的数值最大。

- A)  $H(2, 1)$       B)  $H(2, 2)$       C)  $H(1, 1, 1)$       D)  $H(1, 2)$

3. （选择）选项 A) ~ C) 为属性 X、Y、Z 的信息增益，其中属性 Y 的信息增益为选项 （2分）。

- A)  $H(2, 2, 2, 3) - \{(1/3) H(2, 1) + (1/3) H(1, 1, 1) + (1/3) H(1, 2)\}$   
B)  $H(2, 2, 2, 3) - \{(4/9) H(2, 1, 1) + (2/9) H(1, 1) + (1/3) H(1, 1, 1)\}$

C)  $H(2, 2, 2, 3) - \{(1/2) H(3, 0) + (1/3) H(2, 1) + (1/3) H(1, 2)\}$

4. （选择）最大化信息增益的属性应该作为属性选择的标准。因此，属性 （2分） 应作为 ID3 决策树的根节点。 A) X    B) Y    C) Z

5. （简答）给出一个属性 M，使得属性 M：

- M 的信息增益比 X、Y、Z 的信息增益都大
- M 基本不具备泛化能力，也就是几乎无法用于预测。

仿照 X、Y、Z 属性，画图表达属性 M 如何对集合 {A, A, B, B, C, C, D, D, D} 进行划分。

答：（2分）

（判断对错）关于提升方法 Adaboost 的说法错误的是 （4分）。

（1分）（    ）用弱学习器组合出强学习器。

（1分）（    ）在旧学习器所犯的误差上训练新学习器。

（1分）（    ）每学到一个新的基学习器以后，都根据误差修改训练样例抽取的概率。

（1分）（    ）每个基学习器都有相等权重。

七、（神经网络， 8 分）BP 神经网络拓扑结构如图 4 所示。

