

学 号

班 级

姓 名

装  
订  
线  
内  
不  
要  
答  
题

- 答案必须写在答题纸上，否则无效！试题页（共 2 页）可以做演算纸。
- 一、选择题（每空 2 分，共 32 分）。将选项按照空格中的编号写在答题纸上。
- 1、 迄今为止，图灵测试仍然是最重要的机器智能判定标准，但也备受争议。例如，它有两个缺点\_(1)\_、\_(2)\_。  
A) 必须模拟人类的缺点，例如算术运算慢且易错，反应缓慢等  
B) 回避无法说清的“Can machine thinking?”这一问题  
C) 不能测试知觉等属于智能的其它属性
- 2、 若表达式 G 是不可满足的，当且仅当对所有的解释\_(3)\_。  
A) 必真      B) 必假      C) 真假不能断言
- 3、 若状态空间中的任意状态只有有限个后继状态，则下列搜索算法中，具备完备性的是\_(4)\_。      A) A\*算法    B) 一致代价搜索    C) 以上皆是
- 4、 用 A\*搜索算法求解某问题，已构造出 3 个不同的可纳启发函数  $h_1$ 、 $h_2$ 、 $h_3$ 。令  $h_4=\min\{h_1, h_2, h_3\}$ 、 $h_5=\max\{h_1, h_2, h_3\}$ ，下列说法正确的是\_(5)\_。  
A)  $h_4$  可纳      B)  $h_5$  可纳  
C) 选  $h_4$  作为启发函数，不可能比选  $h_5$  少扩展节点      D) 以上皆对
- 5、 在一般的树搜索算法的简略描述中，  
(6)、\_(7)\_、\_(8)\_分别是什么？  
A) 根据估值函数，从 OPEN 表选择一个结点  $n$   
B) 判断结点  $n$  是不是一个目标结点  
C) 扩展节点  $n$
- 初始化：  
循环，直到 OPEN 表为空：

\_(6)\_  
\_(7)\_  
\_(8)\_
- 6、 遗传算法充分体现了以下哪些优化措施\_(9)\_。（2 分）  
A) 多个搜索线程并行搜索      B) 不同搜索线程之间进行有效的信息交换  
C) 注意探索和利用上的均衡    D) 以上皆是
- 7、 对于用于分类的决策树学习算法，下列说法正确的是\_(10)\_。（2 分）  
A) 基于贪婪的思想      B) 基于分治的思想      C) 结点分裂标准以测试属性所减少的不纯度/不确定性/惊奇度为基本标准    D) 以上皆对
- 8、 在遗传算法的简略描述中，  
(11)、\_(12)\_、\_(13)\_分别是什么？  
A) 选择  
B) 交叉  
C) 变异
- 初始化：  
循环，直到满足停止条件：

\_(11)\_，模拟适者生存  
\_(12)\_，从而产生子代  
\_(13)\_

- 9、 假设全程跟踪记录了 200 名学生的某门课程的学习情况：20 次课出席情况、20 次课笔记是否完整、历次课后作业分数、编程作业完成情况、各门先修课成绩、期末成绩等信息。现随机抽取 80%同学的数据用来训练，训练后的模型用来预测其余 20% 同学的期末成绩（百分制）。关于上述机器学习问题的描述，可以判断这\_(14)\_ 监督学习问题；同时，它\_(15)\_ 回归问题。      A) 是      B) 不是
- 10、      对于前馈神经网络的一个单元，输入向量  $\mathbf{x} = [x_0=1, x_1, x_2, \dots, x_n]^T$ ，权值向量  $\mathbf{w} = [w_0, w_1, w_2, \dots, w_n]^T$ ，激励函数为  $f(x)=1/e^{-x}$ ，则其输出的计算公式为\_(16)\_。  
A)  $f(\mathbf{x}) = \mathbf{w}^T \cdot \mathbf{x}$       B)  $f(\mathbf{x}) = 1/(1 + e^{-\mathbf{w}^T \cdot \mathbf{x}})$       C) 以上皆错
- 二、简答题（共 18 分）。将答案写在答题纸上。
- 1、 （本题 6 分）列举人工智能成功应用的三个领域及其典型成就。
- 2、 （本题 6 分）启发式搜索中  $f(x)=g(x)+h(x)$  中，解释  $f(x)$ 、 $g(x)$ 、 $h(x)$  的含义。
- 3、 （本题 6 分）朴素爬山法（最陡上升）求解八皇后问题时，将会以 0.86 的概率陷入局部最优而无法找到解。请你给出一种方案改进算法，使之能够满足下列条件之一：  
（1）提高一次爬山就能成功求解的概率；或者（2）跳出局部最优，从而找到一个完整解。给出你的改进方法，并简述求解能力提升的原因。
- 三、综合应用题（共 50 分）。答案写在答题纸上。
- 1、 （本题 16 分）回答本题时，限定使用以下谓词和函数：  
谓词  $N(x)$  表示  $x$  是自然数；      谓词  $I(x)$  表示  $x$  是整数；  
谓词  $E(x)$  表示  $x$  是偶数；      谓词  $GZ(x)$  表示  $x$  是大于等于零的数；  
谓词  $O(x)$  表示  $x$  是奇数；      函数  $S(x)$  表示将  $x$  除以 2 得到  $x/2$ 。
- （1） 将事实  $F_1$ 、 $F_2$ 、 $F_3$  分别用谓词公式表示出来。（6 分）  
事实  $F_1$ ：自然数是大于等于零的整数；    事实  $F_2$ ：所有整数不是偶数就是奇数；  
事实  $F_3$ ：偶数除以 2 是整数。
- （2） 仿照事实  $F_1$ 、 $F_2$ 、 $F_3$  的自然语言陈述，将结论  $G$  用自然语言描述出来。（2 分）  
结论  $G$ ： $(\forall x) (N(x) \rightarrow (O(x) \vee I(S(x))))$ 。
- （3） 将  $F_1$ 、 $F_2$ 、 $F_3$ 、 $\neg G$  化成子句集。（4 分）
- （4） 用归结反演的方法证明  $G$  是  $F_1$ 、 $F_2$ 、 $F_3$  的逻辑结论。（4 分）
- 2、 （本题 10 分）考虑以熵的增益为结点分裂标准的决策树。训练集如表 1 所示，包含了 7 个训练样例，分别属于 no、yes 两类，每个训练样例都由 A、B、C 三个属性描述，目标属性为 F。

学号

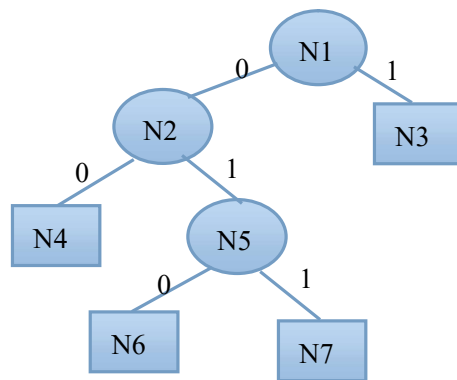
班级

姓名

装订线内不要答题

表 1

实例序号 L	A	B	C	F
1	0	0	0	no
2	0	0	1	yes
3	0	1	0	no
4	0	1	1	yes
5	1	0	0	yes
6	1	0	1	yes
7	1	1	0	no



- (1) 哪个属性做为根结点 N1 的测试属性？简要给出你的推理根据。(5 分)
- (2) 假设 N2 用属性 A 作为测试属性，请给出中间结点 N5 上的测试属性、叶结点 N4 上的类别标签、叶结点 N6 上的类别标签。(3 分)
- (3) 用上述决策树预测新实例 (1, 1, 1) 所属的类别。(2 分)

3、（本题 9 分）贝叶斯网络。已知由 5 个随机变量(取图中单词的首字母)构成的贝叶斯网络，各个条件概率表如图 1 所示。

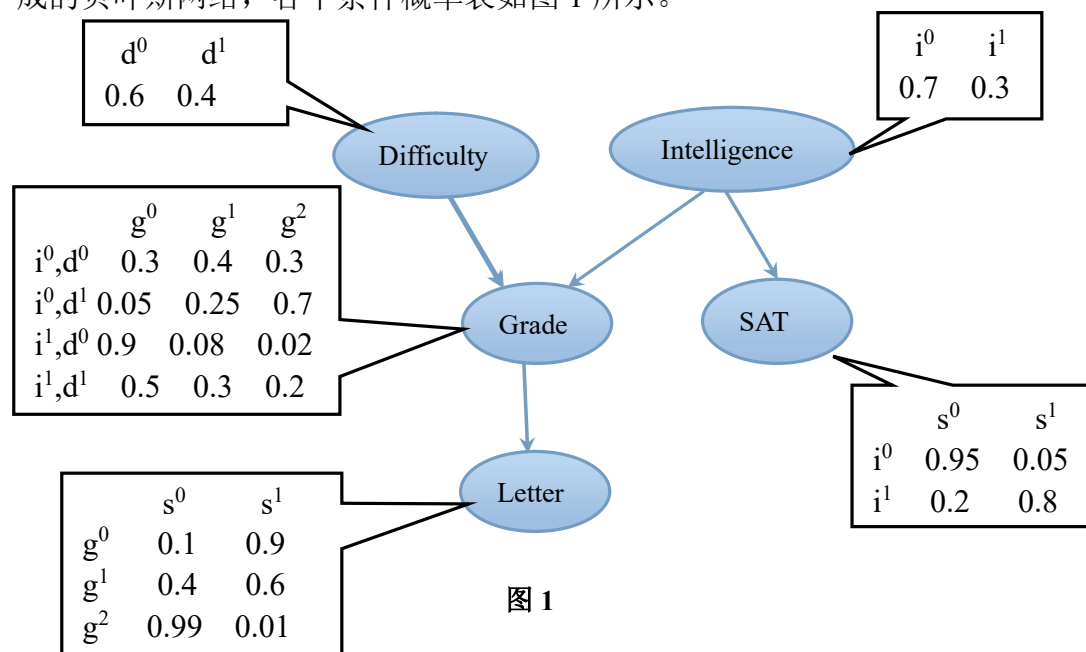


图 1

- (1) 请写出联合概率  $P(D,I,G,S,L)$  分解为条件概率的分解式。(3 分)
- (2) 求  $P(i^1, s^1, g^1) = ?$  (2 分)
- (3) 全联合概率分布至少要存储多少个概率值？贝叶斯网络用若干个条件概率表表示全联合概率表以后，一共需要存储多少个概率值？(4 分)

4、（本题 15 分）已有三个用于求解八数码问题的正确程序：程序  $P_1$  采用深度优先的迭代加深搜索；程序  $P_2$  采用启发函数为  $h_1$  的 A\* 搜索， $h_1$  表示不在位数字的总数目（错位数之个数和）；程序  $P_3$  采用启发函数为  $h_2$  的 A\* 搜索， $h_2$  表示所有数字到目标位置的曼哈顿距离之和（错位数之曼哈顿距离和）。

有人拿到了分别实现上述算法的三个程序，但不知道到底哪个程序实现了哪种算法。于是，他首先将三个程序分别标记成 X、Y、Z；然后，又随机地生成了 1000 个八数码问题（对应于解路径长度  $d = 2, 4, 6, \dots, 20$ ，各有 100 个八数码问题）作为测试集；最后，在测试集上分别执行了三个程序，最终得到的实验数据如表 2 所示。请简短回答下列问题：

- (1) 请你帮他推断出：X、Y、Z 与  $P_1$ 、 $P_2$ 、 $P_3$  的对应关系。说明你依据什么将 Y、Z 区分开的。(6 分)
- (2) 若每个结点  $n$  的后继结点的集合都避免包括结点  $n$  的直接祖先，且每个数码在棋盘上任意位置等概率出现。在采用无信息的宽度优先搜索算法时，给出八数码问题的平均分支因子 R 的一个估计值，并给出估算过程。(3 分)

表 2

解的深度 d	求解每个问题，平均扩展的结点数			有效的平均分支因子		
	程序 X	程序 Y	程序 Z	程序 X	程序 Y	程序 Z
2	10	6	6	2.45	1.79	1.79
4	112	13	12	2.87	1.48	1.45
6	680	20	18	2.73	1.34	1.30
8	6384	39	25	2.80	1.33	1.24
10	47127	93	39	2.79	1.38	1.22
12	3644035	227	73	2.78	1.42	1.24
14	-----	539	113	-----	1.44	1.23
16	-----	1301	211	-----	1.45	1.25
18	-----	3056	363	-----	1.46	1.26
20	-----	7276	676	-----	1.47	1.27

- (3) 从表中观察到，Y、Z 程序的平均分支因子比 R 的估计值还小（当  $d > 2$  时），请解释原因。(2 分)
- (4) 假设程序 IZ 在 Z 的基础上采用了迭代加深技术。对于较复杂问题，通常会采用迭代加深搜索，请给出 IZ 相比较 Z 更实用的一个重要原因。(2 分)
- (5) 若不能恰当处理重复结点，则迭代加深程序将反复展开相同结点，浪费大量的时间和空间资源。假设是你正在写一个 24 数码问题（提示：状态空间异常大），你用何种技术避免重复结点的重复展开问题？(2 分)

学 号

班 级

姓 名

装  
订  
线

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： 计算机科学技术 考试日期： 2015 年 7 月 3 日 试卷： 共 2+2 页

题号	一	二	三	总分
得分				
阅卷人				

答题纸

一、 选择题（每空 2 分，共 32 分）。将选择题选项按空格编号填入下表。

(1)	(2)	(3)	(4)	(5)
(6)	(7)	(8)	(9)	(10)
(11)	(12)	(13)	(14)	(15)
(16)				

二、 简答题（共 18 分）。

1、（本题 6 分）

2、（本题 6 分）

3、（本题 6 分）

三、 综合应用题（共 50 分）

1、 （本题 16 分）

2、（本题 10 分）

3、（本题 9 分）

4、（本题 15 分）