

中国科学技术大学计算机学院  
2020~2021 学年第二学期考试试卷

课程名称：人工智能基础

课程代码：\_\_\_\_\_

开课院系：计算机学院

考试形式：开卷

姓名：\_\_\_\_\_ 学号：\_\_\_\_\_

专业：\_\_\_\_\_

题号										总分
得分										

(本试卷必须与答题纸一并上交)

一、判断题 (10分):

- 如果  $h_1$  和  $h_2$  是两个可采纳 (admissible) 启发式函数, 那在 A\* 算法中, 使用启发式函数  $h_3(n) = \max(h_1(n), h_2(n))$  一定比使用启发式函数  $h_4(n) = \min(h_1(n), h_2(n))$  更好。 ( )
- 对于任何决策树上的根节点, 其信息增益 (Information Gain) 一定不比树中其他节点的信息增益少。 ( )
- 随着 k-NN 分类器中使用的 k 值从 1 逐渐增加到 n (即训练样本的总数), 训练集上的分类精度将始终增加。 ( )
- K-Means 聚类保证收敛 (即算法会终止)。 ( )
- 一个使用 quadratic kernel  $K(\mathbf{x}, \mathbf{y}) = (1 + \mathbf{x}^T \cdot \mathbf{y})^2$  的 SVM, 可以正确分类这个数据集  $\{(1,1), +\}, \{(1,-1), -\}, \{(-1,1), -\}, \{(-1,-1), +\}$ 。 ( )

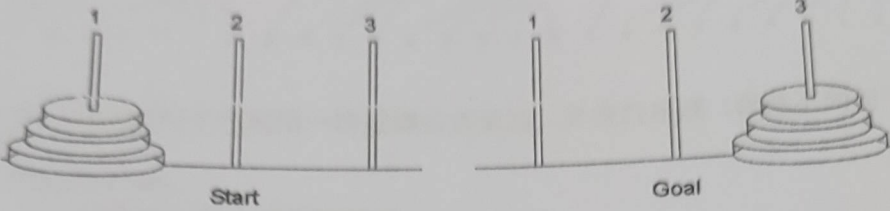
二、单项选择题 (10分):

- 假设有一个状态空间, 其中所有边 (状态转移) 的代价都是 1, 同时我们没有启发式函数。我们想要使用节省空间的搜索算法 (仅考虑边界 (Fringe) 在搜索期间存储的最大节点数), 同时保证我们能找到最优解。在这种情况下, 最好使用以下哪一种搜索方法? ( )
  - 广度优先搜索 (Breadth-First Search)
  - 深度优先搜索 (Depth-First Search)
  - 迭代加深的深度优先搜索 (Iterative-Deepening Search)
  - 一致代价搜索 (Uniform-Cost Search)
- 给定布尔随机变量 A、B 和 C, 并且它们之间没有独立性或条件独立性假设, 以下哪一项等于  $P(A, B, C)$ ? ( )
  - $P(A|B)P(B|C)P(C|A)$
  - $P(C|A, B)P(A)P(B)$
  - $P(A, B|C)P(C)$
  - $P(A|B)P(B|C)P(C)$
- Naive Bayes 分类器做了以下哪项假设? ( )
  - 所有属性都是独立的。
  - 给定输出标签, 所有属性都是条件独立的。

(C) 所有的属性都相互依赖。

(D) 部分属性相互依赖。

4. 给定如下图汉诺塔问题(从初始状态到目标状态, 每次只能移动一个顶端的圆盘, 并且任何时候大盘不能叠在小盘上面), 有三个盘子重量从小到大分别是 1, 2, 3, 每步行动的代价(cost)为盘子的重量乘以盘子移动的距离(第一个杆子与第三个杆子距离为 2, 相邻杆子距离为 1)。下列定义中不是可采纳(admissible)启发式函数的是? ( )

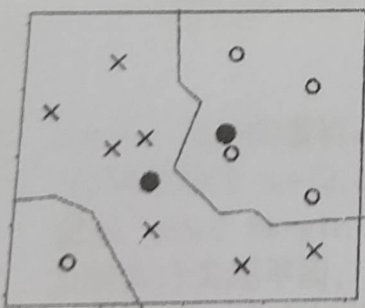


- (A) 在第三个杆上的盘子数目。  
(B) 不在第三个杆上的盘子数目。  
(C) 不在第三个杆上的盘子重量。  
(D) 不在第三个杆上的盘子重量乘以它们与第三个杆子的距离。
5. 使用 K-Means 聚类算法将 7 个点聚类为 3 类。假设第一次迭代后, 簇 C1、C2 和 C3 包含以下二维点: C1 包含两个点  $\{(0, 6), (6, 0)\}$ , C2 包含三个点  $\{(2, 2), (4, 4), (6, 6)\}$ , C3 包含两个点  $\{(5, 5), (7, 7)\}$ 。哪些中心点可以生成这三簇? ( )
- (A) C1: (3, 3), C2: (4, 4), C3: (6, 6)  
(B) C1: (3, 3), C2: (6, 6), C3: (12, 12)  
(C) C1: (0, 0), C2: (48, 48), C3: (35, 35)  
(D) 以上都不对。

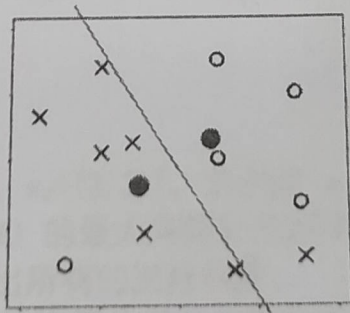
### 三、填空题 (10 分):

1. 将下面图中的决策边界与最有可能产生此类决策边界的分类器算法相匹配。每个类的均值由实心圆表示(红点是圆圈的均值、绿点是叉号的均值)。

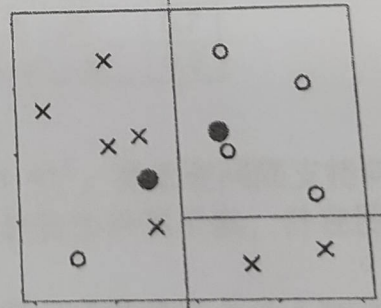
- (1) Decision Tree Classifier (选择 A, B, 或 C) \_\_\_\_\_  
(2) Linear Classifier (选择 A, B, 或 C) \_\_\_\_\_  
(3) Nearest Neighbor Classifier (选择 A, B, 或 C) \_\_\_\_\_



(A)



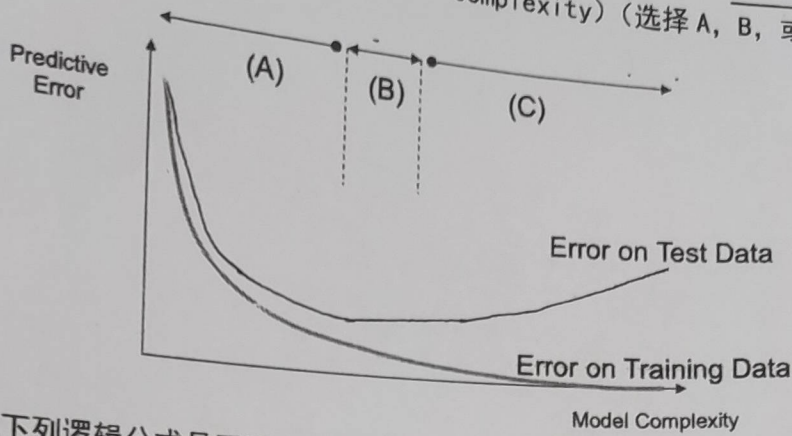
(B)



(C)

2. 考虑下图, y 轴为预测误差 (y 轴), x 轴为模型复杂度 (x 轴)。图中显示测试/训练数据上的误差如何随着模型复杂度的增加而变化。图中标记了 (A)、(B)、(C) 三个区域, 请将其分别对应为过拟合、欠拟合或理想模型复杂性:

- (1) 过拟合 (Overfitting) (选择 A, B, 或 C) \_\_\_\_\_  
 (2) 欠拟合 (Underfitting) (选择 A, B, 或 C) \_\_\_\_\_  
 (3) 理想模型复杂性 (Ideal Model Complexity) (选择 A, B, 或 C) \_\_\_\_\_



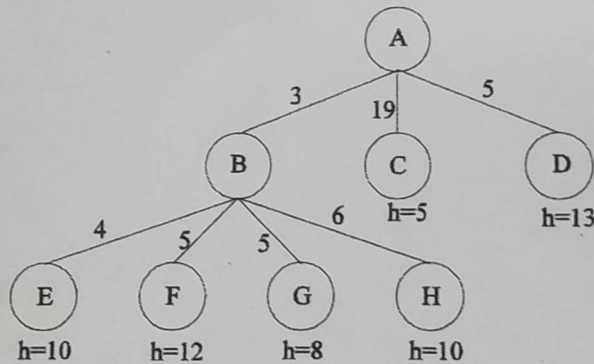
3. 考虑下列逻辑公式是否可以归结 (resolve)。如果可以, 请给出归结结果和所需的合一置换; 否则, 回答不可归结。

(1)  $\forall x \forall y (P(x) \wedge Q(y)) \rightarrow (R(x) \vee S(g(x), y))$  与  $\forall u \forall v S(g(u), h(u, v)) \rightarrow R(h(v, u))$

(2)  $\forall x \forall y (P(x) \wedge Q(y)) \rightarrow (R(x) \vee S(g(x), y))$  与  $\forall w \forall z (P(h(w, z)) \wedge Q(h(z, w))) \rightarrow R(w)$

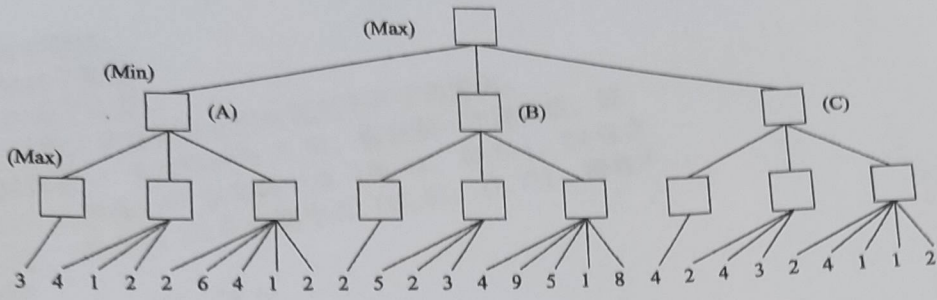
四、(12分) 考虑以下在扩展节点 A 和 B 后生成的搜索树, 其中每个边上都标注了相应的代价 (cost), 每个节点标有相应的启发式函数 h 的值。对于无信息搜索 (uninformed search), 假设子项从左到右展开。当有多个节点并列选择时, 按字母顺序选择。针对下列搜索算法给出下一个应该扩展的节点。

1. 深度优先搜索 (Depth-First search)
2. 一致代价搜索 (Uniform-Cost search)
3. 贪心最优搜索 (Greedy Best-First search)
4. A\* search



五、(12分) 在下面的博弈树中, 轮到 Max 决策。每个叶子结点的数值代表对相应状态的评估结果。

1. 执行 Mini-Max 搜索并用其值标记每个分支节点 (在分支结点对应的方框中填相应的数字)。
2. 采用 alpha-beta 剪枝, 列出所有被剪枝的叶子结点。
3. Max 的最佳走法是什么 (A, B, 或 C)? \_\_\_\_\_

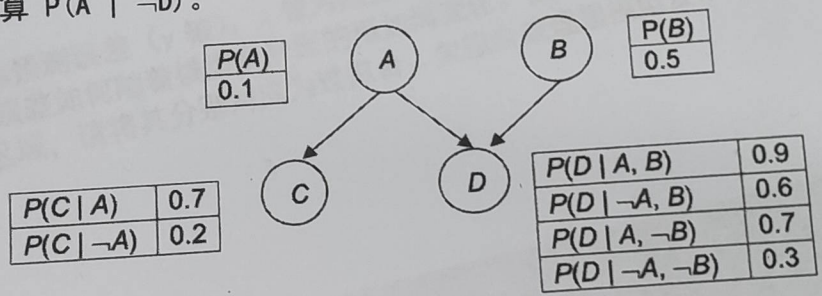


六、(14分) 将下列句子分别用一阶逻辑公式表达, 并进行推理 (保持不同句子中谓词一致)。

- 自然语言语句:
  - 所有的机器人都是聪明的。
  - 有些机器人是聪明的。
  - Robbie 是一个机器人。
  - 所有的机器人都对它的主人友好。
  - Bob 是 Robbie 的主人。
  - 有些拥有机器人的人是聪明的。
- 采用归结方法 (Resolution) 证明 “Robbie 对 Bob 友好”。
- 基于以上知识库, 能否证明 “Bob 是聪明的”? 如果成立, 请给出证明过程; 如果不成立, 请解释为什么。

七、(16分) 考虑如下具有 4 个布尔型随机变量的贝叶斯网络 (Bayesian network)。

- 计算  $P(\neg A, B, \neg C, D)$ 。
- 判断 A 与 B 在给定 D 时, 是否条件独立?
- 计算  $P(A | B, C, D)$ 。
- 计算  $P(A | \neg D)$ 。



八、(16分) 支持向量机:

- 已知正例点  $x_1=(2, 3)^T$ ,  $x_2=(3, 3)^T$ , 负例点  $x_3=(1, 1)^T$ , 试求硬间隔支持向量机 (Hard Margin SVM) 的最大间隔分离超平面和分类决策函数, 并在图上画出分离超平面, 指出所有的支持向量。
- 现额外有一个点能被 SVM 正确分类且远离决策边界, 如果将该点加入到训练集, SVM 的决策边界会受到影响吗? 为什么?
- 现额外有一个点不能被当前的 SVM 正确分类且远离决策边界, 如果将该点加入到训练集并将硬间隔支持向量机扩展为软间隔支持向量机 (Soft Margin SVM), SVM 的决策边界会受到影响吗? 为什么?