

座位号

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： (B) 考试形式： 闭卷

考试对象： 计算机科学与技术 考试日期： 2018 年 05 月 31 日 试卷：共 4 页

题号	一	二	三	四	五	六	七	八	总分
得分									

得分

一、基础题（每题 3 分，共 27 分，单选，答案写在圆括号内）

1. () 关于人工智能，较科学的说法是_____。
- A) 和机器学习一回事，文字游戏而已。
- B) 只会给人类社会带来积极影响，如减轻人类负担，增加就业等。
- C) 关于什么是，以及怎样能，不同观点互相争鸣。
2. () 考虑迭代次数上限为 N 的最陡爬山法，其空间复杂度为_____。
- A) $O(1)$
- B) $O(N)$
- C) $O(N^2)$
3. () 在模拟退火算法中，正确的说法是_____。
- A) 算法的目标函数值，既考虑“上山”趋势，又考虑“下山”趋势。
- B) 评价函数必须可导。
- C) 在退火时间 T （最外层循环变量）增大时，接受劣解的概率始终维持不变。
4. () 约束满足搜索中，状态内部_____；A*搜索中，状态内部_____。
- A) 是无结构的；是有结构的
- B) 是有结构的；无结构的
- C) 是无结构的；无结构的
- D) 是有结构的；有结构的
5. () 给定概率公式：① $P(X,Y|Z) = P(X|Z)P(Y|Z)$ ；② $P(X|Y,Z) = P(X|Z)$ ；③ $P(Y|X,Z) = P(Y|Z)$ 。关于三者所断言的条件独立性，有_____。
- A) 只有②和③相同
- B) 都不相同
- C) 都相同
6. () 关于决策树，不正确的说法是_____。

A) 采用分治法 B) 采用贪婪策略 C) 选项 A 和 B 都对 D) 不会过拟合

7. () 关于马尔科夫决策过程的策略 π ，正确的说法是_____。

A) $P(s'|s,a)$ B) 一个行动序列

C) 从状态集到行动集的映射 D) 每个状态 s 的动作集

8. () 既属于带隐变量的时序模型，又在课堂上讲授过的是_____。

A) 隐马尔科夫过程 B) 卡尔曼滤波 C) 动态贝叶斯网络 D) 选项 A 和 B

9. () 教师上课时多次强烈推荐过一门英文人工智能公开课，还采用了其强化学习课件中的部分内容。这门课是_____。

A) MIT 的 6.034 B) Berkeley 的 CS188 C) Udacity 的《人工智能入门》

得分

二、灵活运用题（每题 3 分，共 63 分，单选，答案写在圆括号内）

关于对抗搜索和盲目搜索算法，回答问题 10~11。

10. () 在不采用其它优化技术的情况下，初始状态相同，最大深度相同，分别采用 $Alpha$ - $Beta$ 搜索（初始窗口用 $(-\infty, +\infty)$ ）和 $Mini$ - Max 搜索，其返回值_____。

A) 相等 B) 不等 C) 相等或不相等，不能一概而论

11. () 按自左向右的子结点访问顺序在图 1 中的带权树，权值越小越好。在找到解时，关于结点计数（包括 $Start$ ）的正确说法是_____。

- A) 深度优先搜索访问结点的数量最多
- B) 迭代加深搜索访问结点的数量最多
- C) 宽度优先搜索访问结点的数量比一致代价搜索多
- D) 选项 A、B、C 提到的各搜索算法访问结点的数量一样多

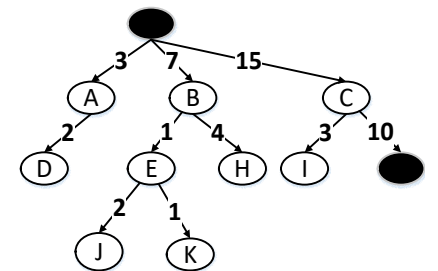


图 1 搜索树

根据 A*搜索算法的一致性启发函数的相关定理及其证明，回答问题 12~17。

任务描述。考虑在带权图中执行 A*图搜索。在带权图中，任意边上的代价（或权值） $\geq \epsilon$ (ϵ 满足 (1))，任意结点的直接后继个数为有限个，代价越小越好。初始状态为 s ，最优解为路径“ $s \rightarrow \dots \rightarrow x \rightarrow \dots \rightarrow s_g$ ”；最优解所对应的目标状态为 s_g ，对应的路径代价为有限的正值 C^* ， C^* 未

知。搜索算法从初始状态 s 出发, 推进到 \forall 结点 n 时, 设 $f^*(n) = g^*(s \rightarrow n) + h^*(n \rightarrow s_g)$ 为必须经过结点 n 的约束最优解的真实代价。其中, s_g 是离 n 最近的目标状态; $g^*(s \rightarrow n)$ 为从 s 到 n 的所有路径中最短路径的代价, 记为 $g^*(n)$; $h^*(n \rightarrow s_g)$ 为从 n 到 s_g 的最短路径 $n \rightarrow s_g$ 的代价, 记为 $h^*(n)$ 。 $g^*(s \rightarrow n)$ 和 $h^*(n \rightarrow s_g)$ 均难以计算, 视为未知, 因而 $f^*(n)$ 也未知。在给定的问题有解的情况下, A* 的最终任务可转化为找出 $f^*(s) = C^*$ 所对应的一条路径。

A* 用估计的方法简化问题, 即, 用 $f(n)$ 来逼近 $f^*(n)$, $f(n) = g(s \rightarrow n) + h(n \rightarrow s_g)$ 。其中, $g(s \rightarrow n)$ 为当前路径 $s \rightarrow n$ 实际代价, 记为 $g(n)$, 极易计算, 可视为已知, 因此用 $g(n)$ 估计 $g^*(n)$ 是自然而然的想法。显然, $g(n) \geq g^*(n)$ 。A* 的巧妙处在于: 找一个 $h^*(n)$ 的乐观估计 $h(n \rightarrow s_g)$, 记为 $h(n)$; 在数学上, 必须满足 $h(n) \leq h^*(n)$, $h(s_g) = 0$ 和 $h(s_g) = 0$, 而在计算上, $h(n)$ 务必远比 $h^*(n)$ 容易, 方有实用价值。 $h(n)$ 结合树搜索, 可高效剪枝。若能找到更好性质的 $h()$, 如满足单调性 (见引理 1), 结合图搜索, A* 剪枝的效率将更上一层楼。

一般的图搜索过程概述。先初始化。在搜索结束前, 反复执行以下三个步骤: 选择 OPEN 表中 $f()$ 值最小的结点 n 、(2)、扩展结点 n 。

引理 1: $h()$ 满足一致性 (单调性) 时, 对于任意经过图搜索 “选择步骤” 从 OPEN 表选出的结点 n , 及其任意直接后继结点 n' , 总有 $f(n) \leq f(n')$ 成立。

证明: 根据 (3), $f(n) = g(s \rightarrow n) + h(n) \leq g(s \rightarrow n) + g(n \rightarrow n') + h(n') = g(n \rightarrow n') + h(n') = f(n')$ 。因此, 新扩展出来的结点 n' , 其 $f(n')$ 值不小于扩展路径上的任意祖先 $f(n)$, 即单调不减。 □

推论 1: $h()$ 满足一致性时, 在扩展完任意结点 n 之后, 图搜索在 “选择步骤” 选出的新结点 x , 必有 (4)。即, 从 OPEN 表选择出来的结点, 其 $f()$ 值按时间顺序不减。

引理 2: 在任何时候, OPEN 表中总有最优路径上的结点。

证明: 根据一般的图搜索过程、OPEN 表定义、(5), 引理 2 必然成立。 □

推论 2: $h()$ 满足一致性时, 在最优解路径上的所有结点当中, $f()$ 值最小的是 (6); $f()$ 值最大的是 s_g , 其值为 C^* 。

定理: 启发函数满足一致性 (或单调性) 的 A* 图搜索是最优的。

证明: 先用反证法证明 A* 图搜索必然终止于最优状态。假设 A* 终止于非最优目标状态 s_g , 即 s_g 满足 $f(s_g) > C^*$ 。根据图搜索过程, 算法终止于 s_g 必须经过以下两个步骤: 第一, 先从 OPEN 表中选择出 s_g , 这要求 s_g 的 $f()$ 值任何比 OPEN 表中的任意其它结点的小; 第二, 再经过目标检测确认 s_g 为目标状态。然而, 第一步不成立。根据推论 1 和

引理 2 和推论 2: 任何时刻 OPEN 表中总存在最优路径上的结点, 其 $f()$ 值不大于 C^* , 所以图搜索的 “选择步骤” 无法选择出 s_g 。这与 A* 终止于 s_g 相矛盾! 假设不成立, 故 A* 只能终止于最优解。

再用构造性方法证明 A* 图搜索在有限步之内终止。 C^* 为有限正数, 每个循环从 OPEN 表选出的结点, 至少比上一次选出的结点的 $f()$ 值大 ϵ ; 再根据推论 1 和推论 2, 从 OPEN 表中经过 “步骤” 选出来的所有结点中, 结点 s 的 $f()$ 值最小为 $f(s)$, 结点 s_g 的 $f()$ 值最大为 C^* 。故结点经过有限步即可终止。 □

12. () 前文中, 空 (1) 应选择 ____。 A) $\epsilon > 0$ B) $\epsilon \geq 0$ C) $\epsilon \leq 0$ D) 无限制
13. () 前文中, 空 (2) 应选择 ____。
A) 检测 n 的某个儿子 n' 是否是目标结点 B) 检测 n 是否是目标结点
14. () 前文中, 空 (3) 应选择 ____。
A) 一致/单调性 B) 可纳性 C) 非负性
15. () 前文中, 空 (4) 应选择 ____。 A) $f(n) \geq f(x)$ B) $f(n) \leq f(x)$
16. () 前文中, 空 (5) 应选择 ____。 A) 图分隔性质 B) CLOSED 表定义
17. () 前文中, 空 (6) 应选择 ____。 A) 初始结点 s B) 路径中的任意结点 n

关于约束满足搜索, 回答问题 18 ~ 21。

18. () 构成约束满足问题的要素是:
A) 变量集, 值域集 B) 变量集, 约束集 C) 变量集, 值域集, 约束集 D) 以上全错
19. () 在约束满足问题中, 进行回溯搜索的同时执行弧相容。当约束图是 n 个结点的树结构时, 每个变量的值域最多有 k 个值, 断言 ____ 是正确的。
A) 没有关于变量个数的多项式时间复杂度的约束满足算法 B) 无需回溯
C) 该问题中的约束数量可能大于 n 个 D) 以上全错
20. () 用基本回溯搜索和强制弧相容策略, 解决图 2 的约束满足问题。假设每个变量的值域至少有两个以上的值, 按变量顺序 A, B, C, D, E, F 搜索直到算法结束时, 在变量 ____ 处, 可能产生回溯。

- A) A, B B) A, D C) B, E D) 以上全错

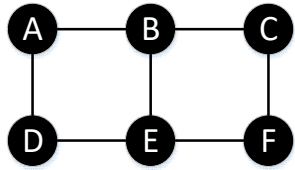


图 2 约束图

21. () 用基本回溯搜索和强制弧相容策略, 解决图 2 的约束满足问题。假设每个变量的值域至少有两个以上的值, 按变量顺序_____搜索直到算法结束时, 可能的回溯次数最少。

- A) A, D, B, E, C, F B) A, B, C, F, E, D C) B, A, D, E, F, C

关于图 3 的贝叶斯网络及其精确推理, 回答问题 22~24。

22. () 假设图 3 中的 A~F 都是布尔型随机变量, 该贝叶斯网络所需存储的概率值有_____个。

- A) 6 个 B) 15 个 C) 11 个 D) 64 个

23. () 最全面的选项是_____。

- A) $P(E|B) = P(E|B, A)$ B) $P(E|B) = P(E|B, F)$
C) $P(A, B|C) = P(C|B) / [\sum_c P(c|B)]$
D) 只有选项 A 和 B 对 E) 选项 A, B 和 C 都对

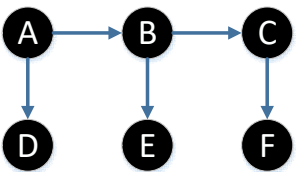


图 3 贝叶斯网络

24. () 用精确推理的算法, 相对而言, 选项_____最难计算。

- A) $P(E|A = -a)$ B) $P(F|A = -a)$ C) $P(D|F = +f)$ D) 以上全错

关于图 4 的贝叶斯网络及其非确定性推理, 回答问题 25~27。

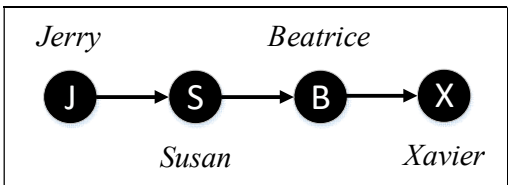


图 4 贝叶斯网络

Jerry, Susan, Beatrice 和 Xavier 意外地得到了一笔钱, 四人分钱的方式很别致。首先, 从 1 到 100 之间 (包括 1 和 100) 等概率地取某整数 J , Jerry 会最先分到 J 元钱。接下来, 从 1 到 J 等概率地取某个整数 S (包括 1 和 J), Jerry 再把刚分到的 J 元中的 S 元钱给 Susan。依此类推, Beatrice 从 Susan 得到 B 元钱; Xavier 从 Beatrice 得到 X 元钱。图 4 的贝叶斯网络描述了这一过程。

现在已经知道 $X = 5$, 想知道: $J > 50$ 的概率有多大? 为此, 对图 4 的贝叶斯网络尝试采样的方法。其中, 条件概率表已经蕴含于分享金钱的过程中。用直接采样得到的样本序列如表 1 所示; 用似然加权采样得到的样本序列, 如表 2 所示。

25. () 基于表 1 直接采样的结果, 计算 $P(J > 50 | X = 5) =$ _____。

- A) 1 B) 2/3 C) 1/2 D) 以上全错

表 1 从贝叶斯网络用直接采样得到的样本序列。

序号	样本点
1	$J = 52, S = 21, B = 10, X = 5$
2	$J = 34, S = 21, B = 6, X = 3$
3	$J = 96, S = 48, B = 12, X = 2$
4	$J = 13, S = 12, B = 10, X = 1$
5	$J = 54, S = 12, B = 11, X = 6$
6	$J = 91, S = 32, B = 31, X = 29$

26. () 表 2 给出了似然加权采样, 表中空出的 (1) 和 (2) 应取_____。

- A) 0.0; 0.2 B) 0.1; 0.1 C) 0.2; 0.0 D) 以上都不对

表 2 从贝叶斯网络用似然加权采样得到的样本序列。

序号	样本点	似然
1	$J = 52, S = 21, B = 10, X = 5$	0.1
2	$J = 34, S = 21, B = 4, X = 5$	(1)
3	$J = 87, S = 12, B = 10, X = 5$	0.1
4	$J = 41, S = 12, B = 5, X = 5$	(2)
5	$J = 91, S = 32, B = 3, X = 5$	0.0

27. () 基于表 2 似然加权采样的结果, 计算 $P(J > 50 | X = 5) =$ _____。

- A) 1 B) 2/3 C) 1/2 D) 以上全错

关于分类问题中的决策树, 回答问题 28~30。

28. () 在决策树的训练过程中, 使用信息增益的目的在于_____。

- A) 从剩余属性中, 选择一个相当“好”的属性;
B) 被属性的不同值分割成不同子集, 各子集更“纯”的属性更好;
C) 选项 A 和 B 都对。

29. () 可以基于多种不确定性度量来描述“好”的属性, 课堂讲授了其中的_____。

- A) Gini 指数 B) 熵 C) 分类误差率

30. () 决策树的输入是 3 个二元属性 (A_1, A_2, A_3), 输出是 1 个二元属性 y 。训练集有 5 个样例, 如表 3 所示。其中, 样例 4 的属性 A_2 的值设为 x , 则说法不正确的是_____。
- A) 无论 x 为 0 或为 1, 根结点上的测试属性都应该是 A_2
- B) 若 x 为 0, 则根结点上的测试属性应该是 A_1
- C) 若 x 为 1, 则根结点上的测试属性应该是 A_1

表 3 决策树的训练集

样例序号	A_1	A_2	A_3	y
1	1	0	0	0
2	1	0	1	0
3	0	0	0	0
4	1	x	1	1
5	1	1	0	1

得分		三、简答题 (2 小题, 10 分)
----	--	--------------------

31. 学习一门课程, 应熟练掌握其知识体系。按教材, 本课程主要讲授了以下几大部分 (已乱序): 不确定知识与推理, 问题求解, 学习等。(本题 6 分)
- 1) 请按教材目录的组织顺序, 重新排列上述三大部分。(3 分)
- 答:
- 2) 请问“决策树”是在上述三大部分中的哪个部分讲授的? (3 分)
- 答:
32. 本课程的知识要点回顾。以下为贝叶斯网络的知识要点介绍, 参考该例,

回顾本课程讲授过的要点。(本题 4 分)

贝叶斯网络的知识要点
贝叶斯网络的定义。贝叶斯网络是一个有向无环图, 其中: 1、每个结点对应一个随机变量; 2、一组有向边链接结点对; 3、每个结点有一个条件概率分布 $P(X_i \mid Parents(X_i))$ 。 (2 分)
贝叶斯网络的语义。贝叶斯网络是对联合概率分布的一种表示。 (2 分)
贝叶斯网络的语义。贝叶斯网络可看作一组条件独立性的集合。 (2 分)
.....

请从局部搜索算法、马尔科夫决策过程中任意选定一个主题。介绍主题下任意两个重要的知识点 (每个知识点 2 分, 最多 4 分)。

答: