

伦敦国王学院

这篇论文是学院授予学位考试的一部分。所有考试均以《书院考试规则》为准
学术委员会的权威。

学位课程	MSc,摩根 士丹利资 本国际 (MSci)
模块代码模块职称 考试周期	7 ccsmain Arti ?2018 年 1 月(第一阶段)
时间允许标题	两个小时 回答所有的问题。在答题本的新一页上回答每一个问题， 并在空白处写下答案号。 可以使用计算器。卡西欧 fx83 /卡西欧 fx85。 书本、笔记或其他书面材料不得带入本次考试
计算器的 笔记	

请 做 不 删除 这 纸 从 的
考场

1. 英国的火车旅行可以用三个二元随机变量来描述: 晚点, 火车是否会晚点; 人多拥挤, 无论火车是否拥挤; “热”, 不管火车会不会太热。下面的联合概率分布描述了这三个变量之间的关系:

	晚些时候		¬晚	
	拥挤, 拥挤, 拥挤			
热	0.168	0.009	0.336	0.028
¬热	0.042	0.081	0.084	0.252

利用这些信息, 计算以下概率, 确保你的工作:

$\hat{P}(\text{晚})$ (2 分)

$P_{\perp}(\text{热}|\text{拥挤})$ (三是)

2. 想象一下，你是政府道德委员会的一员，该委员会的职责是制定法规，规范自动驾驶汽车的开发和使用。找出你认为道德委员会应该考虑的两个问题，解释为什么这些问题与自动驾驶汽车的开发和使用有关。

(5 分)

3. 举一个智能代理的例子(可以来自现实世界，也可以来自现实世界)。

a. 用几行文字描述这个例子:i. 它使用的是哪种传感器;

2 它能采取什么行动;

3 它在什么环境下工作。

(1 分)

b. 确保你提供了你的答案的原因，分类你的代理操作的环境是根据它是:i. 情景或顺序;

(2 分)

2 静态或动态。

(2 分)

4.您正在构建一个贝叶斯网络模型，编码三个变量的医学信息:中风、高血压和吸烟。

一个。你知道高血压和吸烟都是导致中风的原因。画一个与中风、高血压和吸烟有关的贝叶斯网络。

(2 分)

b. 已知条件概率:

P(高血压)P(卒中=真 高血压)	
真正的	0
假的	.6
	0
	.1

和

P(吸烟)P(中风=真 吸烟)	
真正的	0
假的	.7
	0.05

展示如何使用 noise - or 模型来构建一个与中风、高血压和吸烟相关的条件概率表，并解释为什么要使用这样的模型。

(8 标志)

5. 一个。设 $hS Ri$ 是一个抽象的论证框架。

i. 如果一组论点 $T \subseteq S$ 是可接受的，这意味着什么?(2 分)

b. 考虑抽象的论证框架 hS, Ri ，其中：

• $S = \{a1, a2, a3, a4, a5\}$ 和

• $R = \{(a1, a2), (a3 a2), (a3、a4), (a3, a5), (a4 a2), (a5, a3)\}$ 。把所有的缺点都写下来?S 的无 ict 子集。

(2 分)

2 写下 S 的所有容许子集。

(三是)

3 写下所有 S 的首选扩展。

(三是)

6. 您正在进行一个涉及到为机器人编程的项目，突然，在运行代码的过程中，机器人停止了工作。你不久前给电池充电，所以电池运行的先验概率是 0.01(电池是新的，所以它是相当可靠的)。你也知道，如果电池是?at，机器人完全按照刚才的方式失败的概率是 0.99。你也知道还有一些其他的東西会导致机器人失败，所以如果电池没有?，它失败的概率是 0.1。

a. 电池，at 的概率是多少?

(6 分)

b. 请将你对(a)的回答与电池处于?at 的先验概率进行比较。

(4 分)

7. 你正在旅行，在机场有一些时间可以打发。你可以花时间做三件事：

你可以叫个服务生。

这个概率是 0.8，让你有时间喝一杯可口的饮料放松一下，效用是 10。它也有 0.2 的概率为你提供了一杯过度烘烤的咖啡豆，但却有 -5 的效用。

你可以买衣服。

这有一个概率是 0.1，你会找到一个很棒的？价格便宜，实用。然而，你最终把钱浪费在价格过高的垃圾产品上的概率是 0.9，即效用 -10。

你可以吃点东西。

有 0.8 的概率，你找到了一些很普通的东西，让你不会饿，效用是 2，有 0.2 的概率，你找到了一些好吃的东西，效用是 5。

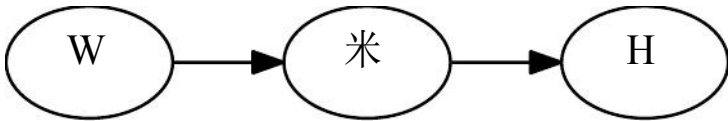
一个。如果你把最大期望效用原则作为你的决策标准，你应该怎么做？

(6 分)

b. 写下另一个决策标准的名称，并解释该标准将选择什么行动。

(4 分)

8. 下面的贝叶斯网络描述了我的工作习惯。努力工作(W)意味着我往往会错过午餐(M)，而错过午餐意味着我在下午 4 点就会饿到不能再工作(H):



$P(W=\text{true}) = 0.8$ ，该网络的条件概率表如下:

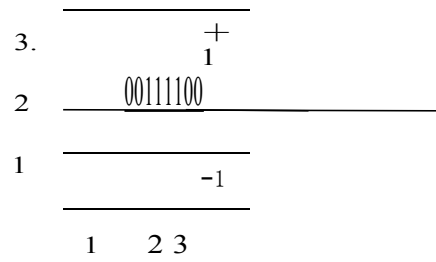
W	P (M = true W)	米	P (H = true 米)
真	0	真	0
正	.9	正	.7
的	0	的	0
假	.5	假	.3

使用网络上的重要性抽样计算 $P(M=\text{真}|H=\text{假})$ 的近似。在你的答案中，对重要性抽样方法进行两次迭代，并使用下列从 0 到 1 之间均匀分布的随机数序列来创建必要的样本:

0.42 0.95 0.75 0.31 0.16 0.23

(10 分)

9.agent 在以下环境中进行强化学习:



agent 可以进行不确定性动作:上(U)、下(D)、左(L)、右(R)，每次动作的代价为 0.05。在学习过程中，agent 从(1,1)开始在环境中运行三次:

- (1, 1) → (1, 2) → (1, 3) → (2, 3) → (1, 3) → (2, 3) → (3, 3)
- (1, 1) → (2, 1) → (3, 1)
- (1, 1) → (1, 2) → (1, 1) → (1, 2) → (1, 3) → (2, 3) → (3, 3)

一个。使用直接效用估计，说明如何效用的 $U_{(1,1)}$ 每次运行后的变化。
(6 分)

b.解释一个关于直接效用估计的问题，以及如何实现它。(4 分)

10. 考虑以下范式游戏:

	l	R
U	4 2	3. 1
D	0 5	2 2

一个。确定任何纯策略纳什均衡，并解释为什么它们是均衡解。
(三是)

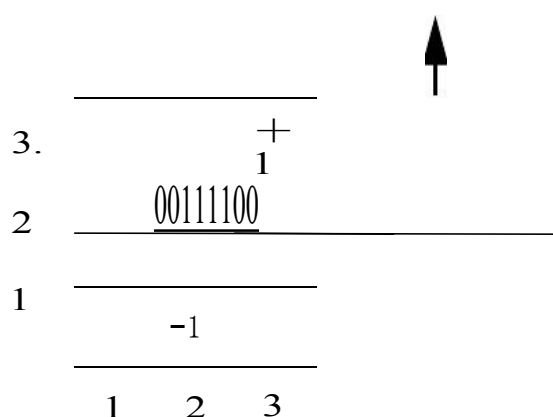
b. 解释游戏的哪些结果能使社会福利最大化。(三是)

c. 一般来说，博弈的纳什均衡和社会福利最大化的结果之间的关系是什么?
(4 分)

11. a. 写下贝尔曼方程。

(2 分)

b. 下面是一个非常简单的世界，在这个世界中，一个代理操作：北方



阴影正方形代表障碍。在每个状态下的奖励都是-0.05，除了那些以? 代理人可以东、南、西、北移动。世界是不确定性的:当 agent 采取行动时，成功的概率是 0.9，而 agent 不行动的概率是 0.1。所有试图进入障碍物或世界边界的尝试都失败了，所以代理将不会移动。

解释如何使用值迭代来确定(1,1)的值(取折扣率为 1)。

(6 分)

c. 解释在这种情况下如何应用策略迭代，以及为什么要这样做。

(7 标志)