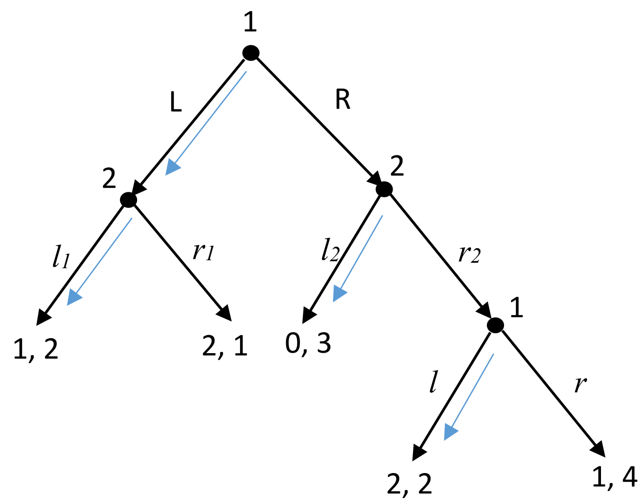
社会博弈论 2022 春季
期末考试参考答案

回答下面各题。总分 100 分。时间 110 分钟。

1. （16 分）考察以下博弈



a. 画出该博弈对应的策略型(即收益矩阵)。(8 分)

	l_1l_2	l_1r_2	r_1l_2	r_1r_2
Ll	1, 2	1, 2	2, 1	2, 1
Lr	1, 2	1, 2	2, 1	2, 1
Rl	0, 3	2, 2	0, 3	2, 2
Rr	0, 3	1, 4	0, 3	1, 4

b. 找出该博弈所有的单纯策略的纳什均衡。（4 分）

两个纳什均衡： ***（Ll, l_1l_2）,(Lr, l_1l_2)***

c. 使用逆推归纳法，找出本博弈所有的子博弈完善均衡。（4 分）

（Ll, l_1l_2）

2. （12 分）考察以下博弈：

博弈者 2

博弈者 1		L	C	R
	T	3， 2	4， 0	0， 0
	M	2， 0	3， 3	-1， 0
	B	0， 0	0， 0	3， 3

a. 找出本博弈单纯策略的纳什均衡。（4 分）

（T,L）和（B，R）

b. 本博弈中是否存在劣势策略，是谁的什么策略；是严格劣势策略还是弱劣势策略？（2 分）

有，是 1 的 M 策略，是严格劣势策略。

c. 找出本博弈混合策略的纳什均衡(提示：可利用 b 部分的答案，有帮助)。（6 分）

由于 1 不会使用 M 策略，可删去 M。（1 分）之后可以看出，在混合策略均衡中，2 不会使用 C 策略（因为只要 1 使用 T 的概率为正，L dominates C）。（1 分）那么令 1 使用 T、B 的概率是 p，1-p，2 使用 L、R 的概率是 q，1-q。

1 混合的条件是：U（T）= 3q = U（B）=3（1-q），可解出 q=1/2 （1.5 分）
2 混合的条件是：U（L）= 2p = U（R）=3（1-p），可解出 p=3/5 （1.5 分）

即：均衡为（3/5，0，2/5），(1/2，0，1/2) （1 分）

另一解题思路：如果只看出 1 不会用 M，但是没有马上看出 2 不会用 C，那么令 1 使用 T、B 的概率是 p，1-p，令 2 使用 L、C、R 的概率是 q_L，q_C，$1-q_L-q_C$

1 混合的条件是使用 T 和 B 的收益相同：
$U(T) = 3q_L+4q_C$
$U(B) = 3(1-q_L-q_C)$
两者相等，可解出：$6q_L+7q_C=3$

考虑 2 混合的条件（2 使用 L、C、R 的收益其中有某些相同）
U（L）= 2p
U（C）= 0
U（R）= 3（1-p）
显然，无论 p 的取值（1>p>0），2 使用 C 的收益肯定比 L、R 小，所以他不会用 C。剩下两个策略的收益相等意味着 U（L）= U（R），可解出 p=3/5。而且，既然 2 不会用 C，就是说 $q_C=0$，那么根据 1 的混合条件也可以解出 $q_L=1/2$

3.（20 分）两名博弈者一起生产一种公共品。博弈者 i (i = 1 or 2) 可以选择自己的努力程度x_i，$x_i \geq 0$。公共品的产量Q由两人努力的总和决定：$Q = x_1 + x_2$。对个体而言，付出努力意味着成本：$c_i = x_i^2$。博弈者的个体收益取决于公共品的产量和自己的成本：$u_i = 10Q - c_i$ 。

a. 找出博弈者 1 和 2 的最优反应函数（4 分）。

1 的收益：$10(x_1+x_2)-x_1^2$（1 分）

最大化的一阶条件（对 x_1 取导数）是：$10-2x_1=0$，（2 分）

最优反应函数为：$x_1=5$ （1 分）

同理，2 的最优反应函数为 $x_2=5$

b. 博弈的纳什均衡中，每位博弈者付出的努力是多少，收益是多少？（4 分）

显然，纳什均衡是（5，5），各自的收益是 10(10)-25=75

评分：纳什均衡 2 分，收益 2 分

c. 如果每一方投入的努力是 10，双方的收益是多少？（2 分）

每方的收益是 10*（20）-100 =100 （2 分）

现在考虑将以上博弈进行**无限重复**形成的一个新博弈，其中博弈者的时间折扣(time discounting)系数为δ，$0 < \delta < 1$。

d. 构建一个子博弈完善均衡（你可以按照需要指定一个δ值），其中双方的平均收益都为 100。请完整写出该策略组合，并清晰的证明它是一个子博弈完善均衡。（4 分）

构建策略组合：合作期双方各投入 10。若任意一方偏离，进入惩罚期，双方各投入 5，惩罚永远持续。（惩罚期也可以只持续有限个阶段，但需要的$\boldsymbol{\delta}$值就不同）

证明构成均衡：

先考虑合作期是否有动力偏离。

不偏离的收益流是：100，100，100，…，平均收益是 100。一次性偏离能获得的最高收益是（自己投入 5，给定对方投入 10）125，然后进入永远的惩罚期，每个阶段的收益是 75，收益流是 125，75，75，…，平均收益是 50(1-$\boldsymbol{\delta}$)+75。比较这两者可得，当$\boldsymbol{\delta} \geq \mathbf{1/2}$，就没有动力偏离。（这两个收益流也可以都折现成一次性收益，前者为 100/(1-$\boldsymbol{\delta}$)，后者为 50 + 75/(1-$\boldsymbol{\delta}$)，得出的结果也一样。）

再考虑惩罚期。因为惩罚期是永远重复阶段博弈的纳什均衡，所以显然是自我稳定的（每一方都没有动力偏离）。

注意：构建的均衡也可以采取其他惩罚策略。例如以上的惩罚持续有限个阶段而不是永远惩罚。只要惩罚期内使用的是纳什均衡策略（双方投入 5），那么惩罚期是自我稳定的。根据具体惩罚几个阶段，需要计算出相应的$\boldsymbol{\delta}$，以保证在合作期博弈者没有动力偏离。（惩罚期为 T 阶段，$\boldsymbol{\delta}$需满足$\boldsymbol{\delta} + \boldsymbol{\delta}^2 + \cdots + \boldsymbol{\delta}^{\boldsymbol{T}} \geq \mathbf{1}$，例如当 T=2，需要$\boldsymbol{\delta} \geq \frac{\sqrt{5}-1}{2} \approx \mathbf{0.62}$；当 T=3，需要$\boldsymbol{\delta} \geq \mathbf{0.55}$，等等）

如果惩罚期采用的不是纳什策略（双方投入 5），而是其他策略（例如双方都投入 0），那么惩罚期不能够永远持续，而是要经过几个阶段后回到合作期，而且需规定惩罚期内若有任何人偏离，惩罚期重新开始。如果惩罚期双方投入 0，那么惩罚一个阶段就足够了（当然也可以惩罚更长时间），只需要$\boldsymbol{\delta} > \mathbf{0}$即可。

e. 请问双方收益为 0 是否是该重复博弈的一个可行的(feasible)平均收益？如果是，请写出一个策略组合（不要求是均衡），其结果是双方平均收益为 0。（2 分）

是的。（1 分）

双方在每一阶段都投入 0。（1 分）

f. 是否存在这样的子博弈完善均衡，其中双方的平均收益为 0？如果存在，请构建这样一个均衡，并证明它是一个子博弈完善均衡。如果不存在，请说明理由。(4 分)

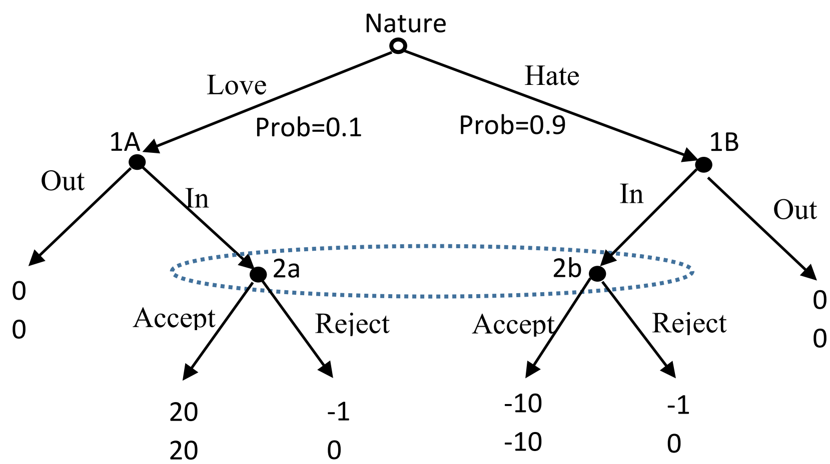
不存在。（2 分）

因为每一方的 minmax 值是 25（即当自己使用最优反应 5，对方使用的最严厉的 minmax 策略 0，自己能获得 25）。博弈者只要每个阶段使用 5，就可以保证自己至少获得这个值。（2 分）

4.（18 分）北大社会学系决定改革招收研究生的制度，废除考试，希望录取到真正热爱社会学的同学。假定所有的潜在申请者中，有 10%的同学是真正热爱社会学的（类型 A），90%不热爱（类型 B）；这是公共信息。每一位同学知道自己是否热爱社会学。但是社会学系不知道申请者的真实类型。考察以下博弈：

自然先行动，按以上前提概率决定同学的类型。同学（博弈者 1）选择是否提出申请(In or Out)。如果不提出申请(Out)则博弈结束，双方收益为 0。若提出申请(In)，社会学系（博弈者 2）选择是接受还是拒绝(accept or reject)。如果拒绝，同学收益为-1，社会学系收益为 0。如果接受，那么双方收益取决于申请者的真实类型：当申请者是类型 A，双方收益各为 20；当申请者是类型 B，双方收益各为-10。

a. 按照以上描述，画出博弈的延展型。（6 分）



b. 构建这样一个的均衡，其中 A 型的同学提出申请，B 型的同学不申请。请写出各方的策略，以及社会学系在其信息集上的信念。并清晰的证明它构成一个弱序贯均衡。（6 分）

同学的策略：A 型申请，B 型不申请。社会学系的信念是申请者是 A 型的，策略是录取。（2 分）

证明：

首先，社会学系的信念是合理的（符合同学的策略）。（1 分）

给定信念（相信申请者是 A 型），社会学系选录取是最优的。（1 分）

给定社会学系的策略，A 型同学选申请是最优的（申请得 20，否则得到 0），B 型同学选不申请也是最优的（不申请得 0，否则得到-10）。（每个类型 1 分）

c. 构建这样一个的均衡，其中两种类型的同学都不提出申请。请写出各方的策略，以及社会学系在其信息集上的信念。并清晰的证明它构成一个弱序贯均衡。（6 分）

同学的策略：A 和 B 类型都不提出申请。社会学系在信息集上的信念是 A 型是 10%，B 型是 90%,策略是拒绝。（2 分）

证明：

信念合理性：给定同学策略，社会学系在信息集上可以有任何信念。（1 分）

给定信念，社会学系的行动的合理性：接受的收益是 20*0.1+(-10)*0.9 = -7;拒绝的收益为 0。显然选拒绝是最优的。（2 分）

给定社会学系的行动，显然两种类型的同学都选择不申请是合理的（对每一种类型而言，申请的收益都是-1，不申请收益为 0）（1 分）

5.（20 分）设想有一天你作了老师，教社会博弈论这门课。期末时，班上的同学可以分为三种类型：t=0（没有学懂），t=6（学的不错），t=12（融会贯通），这三种类型的比例各是 1/3。这是你和所有人都知道的公共信息。每一个同学知道自己确切的类型；但你不知道同学的类型。由于你对于参加考试的痛苦记忆犹新，所以决定废除考试，让大家自报类型，你根据报告给出分数。具体的说，你和同学进行以下博弈（为简单起见，可想象你在和一位同学进行两人博弈）：“自然”先走，决定同学的类型 t。同学给出一个报告,记为 m，m 可以是 0、6、或者 12（m 不一定等于该同学的真实类型 t）。你观察到所报的 m，然后决定给该同学的分数g，g可以是[0,16]区间上的任意整数。双方收益如下：

- 你希望所给的分数越接近同学的真实类型越好，你的收益为：$u_1 = -(g-t)^2$。
- 同学希望拿高分，但是作为有自尊心的北大学生，也不希望分数离自己的真实水平差的太远，最理想的分数是自己的真实类型加上 b 分，同学的收益表达为：$u_2 = -[g-(t+b)]^2$。这里先设定 b=4。

a. 考察如下“分离型”策略组合：同学报出他的真实类型（即 m=t）；你收到报告后，相信他的报告（即 t=m），并给他分数$g = m$。这是否构成一个弱序贯均衡？解释为什么。（4 分）

不是均衡。例如 t=0 的同学，他有动力改变策略报告 m=6，这样（给定你的策略）他会获得 6，比起他现在获得 0，离他的最优点（t+b=4）更近，收益更高。

b. 考虑如下“混同型”策略组合：同学无论自己的类型是什么都报 m=12；你无论收到什么样的 m 都认为同学类型是 0、6、12 的概率各为 1/3，并且给他分数$g = 6$。这是否是一个弱序贯均衡？解释为什么。（6 分）

这是均衡。（2 分）

首先你的信念是合理的：由于同学的报告不提供任何额外信息，你根据前提概率来判断每个类型的概率，是合理的。（1 分）

给定你的信念，你的策略是最优的，因为你最希望的是 g=t，而现在在 t 的期望 E(t)=6。（2 分）

给定你的策略，任何类型的同学无论选择任何偏离（报告其他的 m），都仍然获得 g=6，和现在一样，所以没有动力偏离。（1 分）

c. 请你构建一个“部分分离型”均衡，其中存在一定的信息传递，即同学中的一种（或几种）类型采取的行动与其他类型的不同。请清晰的写出该均衡中每种类型的同学的策略，你的信念和策略。并证明它构成了一个弱序贯均衡。（6 分）

构建均衡

同学的策略：类型 0 的同学报 m=0，类型 6 和类型 12 的同学报告 m=12 （1 分）

你的信念和策略：你如果观察到 m=0，认为同学的类型是 0，并给他 0 分；如果观察到 m=12，认为同学的类型是 6 或 12 的概率各为 1/2，并给他 9 分。为完整起见，如果你观察到 m=6，可认为同学的类型是 0，并给他 0 分。（1 分；如果没有考虑到 m=6 的情况，可酌情扣 0.5 分）

证明

首先，给定同学的策略，你的信念是合理的。你观察到 m=0，认为同学的类型为 0，观察到 m=12 时，认为类型 6 和 12 的概率各为 1/2，这都符合同学的策略。观察到 m=6 时，因为不在博弈路径上，所以可以有任意信念。（1 分）

给定你的信念，你的策略是最优的。因为你最希望的是 g=t，当 m=0 时，你相信 t=0，所以选择 g=0 最优。而 m=12 时，你相信 E(t)=9，所以最优的选择是 g=9。（1 分）

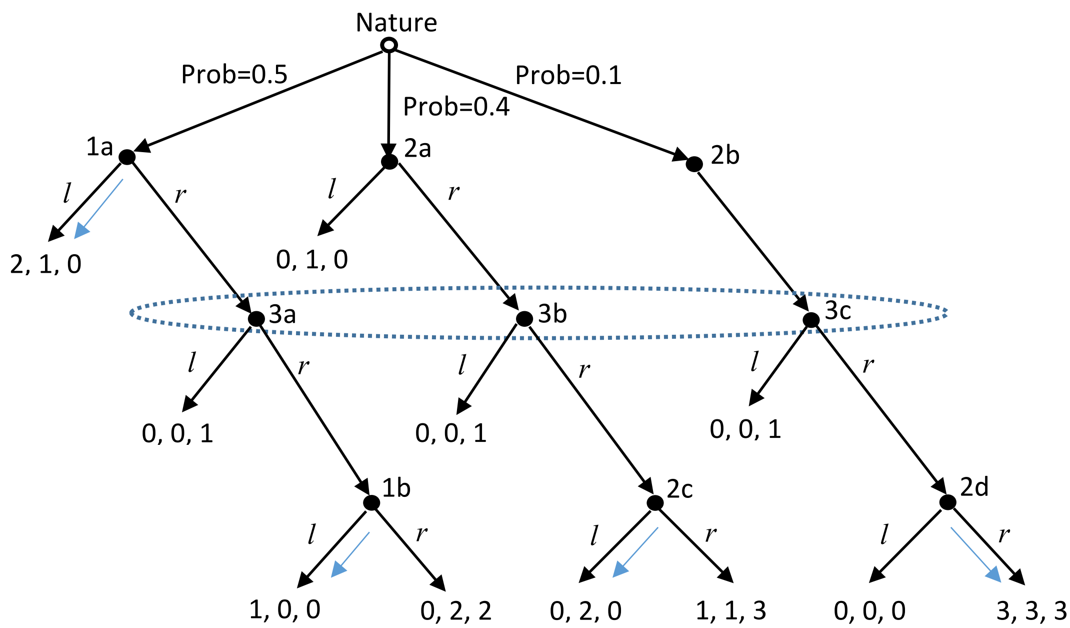
给定你的策略，类型为 0 的同学现在获得 0 分，他最希望获得的是 4 分(t+b=4),他如果改变策略报告 m=12,他将获得 9 分，离开 4 分更远（如果报 m=6，仍然获得 0 分，和现在一样），所以他没有动力偏离。类型为 6 的同学现在获得 9 分，而他最希望获得的是 10 分（t+b=10），如果报 m=6 或 m=0，他将获得 0 分，离开 10 分更远。类型为 12 的同学现在获得 9 分，他最希望获得的是 16 分（t+b=12），如果报 m=6 或者 m=0，他将获得 0 分，显然离 16 分更远。所以每一种类型都没有动力偏离。（2 分）

d. 你上面构建的这个均衡，如果 b 太大就无法成立了，能使该均衡成立的最大的 b 是多少？（我们允许 b 为正的实数）（4 分）

最大的 b 值是 4.5。（2 分）

因为 t=0 的类型最希望得到的分数是 t+b=b；要保证他愿意报 m=0 并获得 0 分，而不是改报 m=12 以获得 g=9，就需要使得 b 离 0 比离 9 更近，b 需要小于等于 4.5。（2 分）

6. （14 分）考查如下的一个 3 人博弈。其中每个信息集（决策节点）标注了该处轮到哪个博弈者（1、2、或 3）决策。由于每个博弈者拥有不止一个信息集（决策节点），所以每个节点处还用字母加以区别。（例如，1a，1b 是轮到博弈者 1 决策的两个不同节点。



a. 请使用逆推归纳法，确定以下决策节点上博弈者的最优选择：1a，1b，2c，2d（4 分）

显然，如上图所示，在 1b，1 选择 *l*；2c 选择 *l*，2d 选择 r；这样 1a 时，1 会选择 *l*。

基于上面的结果，进一步回答下面的问题

b. 如果博弈者在所有的信息集（决策节点）上都使用单纯策略，是否存在相应的弱序贯均衡？如果存在，请指出该均衡和相应的信念系统。如果不存在，请说明理由。（提示：思考博弈者 2 的可能的单纯策略，看能否构成均衡。4 分）

不存在。

如果 2 在 2a 处使用 *l*（收益为 1），这意味着 3 在他的信息集上只能是选择 *l*（2 的收益=0）,不能选择 r（2 的收益为 2）。但是 3 在自己信息集上的信念（根据 1 和 2 的策略）此时是 3c，在 3c 处，他的最优选择是 *r*。矛盾。

如果 2 使用 r, 那么 3 在他的信息集上的信念是在 3b 的概率 0.4/0.4+0.1 = 4/5, 在 3c 的概率是 1/5。此时他选 *l* 的期望收益是 1，而选 r 的期望收益是 0.8*0+0.2*3=0.6，他会选择 *l*。给定 3 选 *l*，2 不可能选 r。矛盾

c. 如果博弈者在（某些）信息集（决策节点）上使用混合策略，是否存在相应的弱序贯均衡。如果存在，请找出该均衡和相应的信念系统。如果不存在，请说明理由。（6 分）

既然 2 只能在 2a 处选择混合策略，那么这意味着对于他来说，在该处选择 1 和选 r 的收益相同。

$U2(l) = 1 = U2(r)$

U2(r) 的收益取决于 3 在自己信息集上的选择。如果 3 选 *l*，2 的收益为 0，如果 3 选 *r*，2 的收益为 2。由于这两者都不等于 1，那么 3 需要使用混合策略。令 3 使用 r 的概率为 p，

u2（r）=（1-p）*0 + p*2 = u2(*l*) =1，那么 p=1/2.

而 3 要使用混合策略，他需要持有合适的信念，使得他选 l 和 r 的收益相同。令他相信自己在 3b 的概率为 q，在 3c 的概率为 1-q（根据 1 的策略，3 不可能在 3a），

U3(*l*)=1

U3(*r*)=q*0 +(1-q)*3 =3-3q

两者相等，要求 q=2/3.

而 3 要持有 q=2/3 的信念，这需要 2 在 2b 处选择 r 的概率 s 满足以下贝叶斯公式

Prob(3b|3b or 3c) = 0.4*s/(0.4*s+0.1) = 2/3

可以解出：s=1/2

综上，我们有

2 在 2a 处以 1/2 的概率混合 l 和 r，3 在信息集上的信念是 Prob(3b)=2/3，3 的策略是以 1/2 的概率混合 l 和 r