Jacob Kruse

# Assignment 3

1a) Player 1 Actions: {A or B or C, X or Y, a or b}
Player 2 Actions: {L or R, ℓ or r, λ or ϒ}

Player 2

| Player 1 | L,ℓ,λ | L,ℓ,ϒ | L,r,λ | L,r,ϒ | R,ℓ,λ | R,ℓ,ϒ | R,r,λ | R,r,ϒ |
|---|---|---|---|---|---|---|---|---|
| A,X,a | 3,0 | 3,0 | 3,0 | 3,0 | 0,2 | 0,2 | 0,2 | 0,2 |
| A,X,b | 3,0 | 3,0 | 3,0 | 3,0 | 0,2 | 0,2 | 0,2 | 0,2 |
| A,Y,a | 1,5 | 1,5 | 1,5 | 1,5 | 0,2 | 0,2 | 0,2 | 0,2 |
| A,Y,b | 1,5 | 1,5 | 1,5 | 1,5 | 0,2 | 0,2 | 0,2 | 0,2 |
| B,X,a | 1,2 | 1,2 | 2,1 | 2,1 | 1,2 | 1,2 | 2,1 | 2,1 |
| B,X,b | 1,2 | 1,2 | 2,1 | 2,1 | 1,2 | 1,2 | 2,1 | 2,1 |
| B,Y,a | 1,2 | 1,2 | 2,1 | 2,1 | 1,2 | 1,2 | 2,1 | 2,1 |
| B,Y,b | 1,2 | 1,2 | 2,1 | 2,1 | 1,2 | 1,2 | 2,1 | 2,1 |
| C,X,a | 1,0 | 2,1 | 1,0 | 2,1 | 1,0 | 2,1 | 1,0 | 2,1 |
| C,X,b | 0,1 | 0,0 | 0,1 | 0,0 | 0,1 | 0,0 | 0,1 | 0,0 |
| C,Y,a | 1,0 | 2,1 | 1,0 | 2,1 | 1,0 | 2,1 | 1,0 | 2,1 |
| C,Y,b | 0,1 | 0,0 | 0,1 | 0,0 | 0,1 | 0,0 | 0,1 | 0,0 |

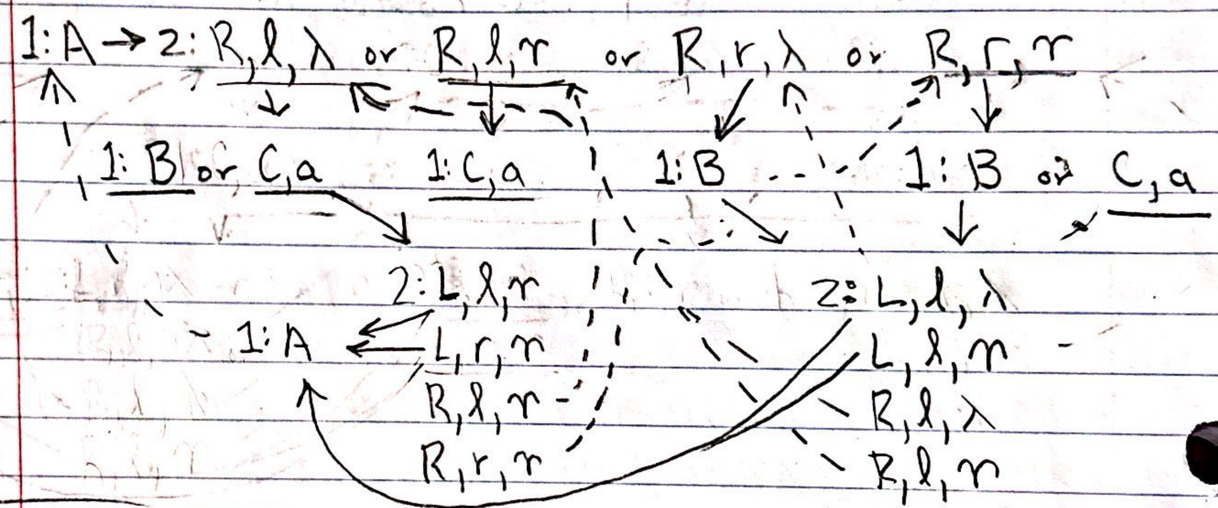
b) C,X,b and C,Y,b are strictly dominated for Player 1
Then A,Y,a and A,Y,b are strictly dominated for Player 1
Finally L,r,λ is strictly dominated for Player 2

| | L,ℓ,λ | L,ℓ,ϒ | L,r,ϒ | R,ℓ,λ | R,ℓ,ϒ | R,r,λ | R,r,ϒ |
|---|---|---|---|---|---|---|---|
| A,X,a | 3,0 | 3,0 | 3,0 | 0,2 | 0,2 | 0,2 | 0,2 |
| A,X,b | 3,0 | 3,0 | 3,0 | 0,2 | 0,2 | 0,2 | 0,2 |
| B,X,a | 1,2 | 1,2 | 2,1 | 1,2 | 1,2 | 2,1 | 2,1 |
| B,X,b | 1,2 | 1,2 | 2,1 | 1,2 | 1,2 | 2,1 | 2,1 |
| B,Y,a | 1,2 | 1,2 | 2,1 | 1,2 | 1,2 | 2,1 | 2,1 |
| B,Y,b | 1,2 | 1,2 | 2,1 | 1,2 | 1,2 | 2,1 | 2,1 |
| C,X,a | 1,0 | 2,1 | 2,1 | 1,0 | 2,1 | 1,0 | 2,1 |
| C,Y,a | 1,0 | 2,1 | 2,1 | 1,0 | 2,1 | 1,0 | 2,1 |

The strategies above are what remain

1c)

| | L,l,λ | L,l,r | L,r,r | R,l,λ | R,l,r | R,r,λ | R,r,r |
|---|---|---|---|---|---|---|---|
| A,x,_ | 3,0 | | | 0,2 | | | |
| B,_,_ | 1,2 | | 2,1 | 1,2 | | 2,1 | |
| C,_,a | 1,0 | 2,1 | | 1,0 | 2,1 | 1,0 | 2,1 |




Nash Equilibria ⟨B,x,a ; R,l,λ⟩ ⟨B,y,a ; R,l,λ⟩ ⟨B,x,b ; R,l,λ⟩ ⟨C,x,a ; R,l,r⟩ ⟨C,x,a ; R,r,r⟩ ⟨B,y,b ; R,l,λ⟩ ⟨C,y,a ; R,l,r⟩ ⟨C,y,a ; R,r,r⟩

d) No, the game is <u>not</u> exhibiting perfect recall. The choice node for Player 1 after Player 2 chooses λ or r is an imperfect recall node (denoted by oval in figure). This means Player 1 does not know which action

2)

Husband

| Wife | Opera | Football |
|---|---|---|
| Opera, Opera | $2p, 1$ | $2(1-p), 0$ |
| Opera, Football | $2p+(1-p), p$ | $0, 2(1-p)$ |
| Football, Opera | $0, 1-p$ | $p+2(1-p), 2p$ |
| Football, Football | $1-p, 0$ | $p, 2$ |

↓

| | O | F |
|---|---|---|
| OO | $2p, 1$ | $2-p, 0$ |
| O,F | $p+1, p$ | $0, 2-p$ |
| F,O | $0, 1-p$ | $2-p, 2p$ |
| F,F | $1-p, 0$ | $p, 2$ |

When $p=0$

$0,0 = F,0$ and $0,F = F,F$

$EU(0,0) = 2x$ $\qquad 2x = 1-x$

$EU(F,F) = 1-x$ $\qquad 3x = 1 \rightarrow x = 1/3$

$EU(O) = y$ $\qquad y = 2-2y$

$EU(F) = 2(1-y) = 2-2y$ $\qquad 3y = 2 \rightarrow y = 2/3$

$$\left\{ \begin{array}{ccc} 0,0 \mid F,0 & 0,F \mid F,F & O \quad F \\ \langle \ 2/3, & 1/3 \ \rangle, & \langle 1/3, 2/3 \rangle \end{array} \right\}$$

When $p=1$

$0,0 = F,0$ and $0,F = F,F$

$EU(0,0) = 2(1-x) = 2-2x$ $\qquad 2-2x = x$

$EU(F,F) = x$ $\qquad 3x = 2 \rightarrow x = 2/3$

$EU(O) = y$ $\qquad 2-2y = y$

$EU(F) = 2(1-y) = 2-2y$ $\qquad 3y = 2 \rightarrow y = 2/3$

$$\{(0,0 \mid F,0), O\}, \ \{(0,F \mid F,F), F\}, \ \left\{ \begin{array}{ccc} 0,0 \mid F,0 & 0,F \mid F,F & O \quad F \\ \langle \ 2/3, & 1/3 \ \rangle & \langle 2/3, 1/3 \rangle \end{array} \right\}$$

These are special cases for when game collapses into "one-type" for wife

2 F,F is strictly dominated except when $p=0$ or $p=1$. This is because the "two-type" game collapses into one type.

| | O | F |
|---|---|---|
| O,O | $2p, 1$ | $2-p, 0$ |
| O,F | $p+1, p$ | $0, 2-p$ |
| F,O | $0, 1-p$ | $2-p, 2p$ |

If Husband chooses O, the Wife will always choose O,F. If the Wife chooses O,F, the Husband will always choose F. If the Husband chooses F, then the Wife will always choose F,O or O,O. If the Wife chooses O,O, the Husband will choose O, and both will continue to deviate. However, if the Wife chooses F,O, there is a range of p values that will result in an equilibrium at (F,O), F.

$$1-p = 2p \rightarrow 3p = 1 \rightarrow p = 1/3$$

When $p \geq 1/3$
$\langle (F,O), F \rangle$

$EU(O,O) = x(2p) + (1-x)(2-p) = 2xp + 2 - 2x - p + xp = 3xp - 2x - p + 2$
$EU(O,F) = x(p+1) = xp + x$
$EU(F,O) = (1-x)(2-p) = 2 - 2x - p + xp$

| OO vs OF | OF vs FO | OO vs FO |
|---|---|---|
| $3xp-2x-p+2 = xp+x$ | $xp+x = 2-2x-p+xp$ | $3xp-2x-p+2 = 2-2x-p+xp$ |
| $2xp - 3x = p-2$ | $3x = 2-p$ | $2xp = 0$ |
| $x(2p-3) = p-2$ | $x = \frac{2-p}{3}$ | $x = 0$ |
| $x = \frac{p-2}{2p-3}$ | | |

OO vs OF

$EU(O) = y(1) + (1-y)p = y + p - yp$

$EU(F) = (1-y)(2-p) = 2 - 2y - p + yp$

$y + p - yp = 2 - 2y - p + yp$

$-2yp + 3y = 2 - 2p$

$y(3-2p) = 2 - 2p$

$$y = \frac{2-2p}{3-2p}$$

$$\frac{3-2p}{3-2p} - \frac{2-2p}{3-2p} = \frac{1}{3-2p}$$

OF vs FO

$EU(O) = z(p) + (1-z)(1-p) = zp + 1 - z - p + zp$

$EU(F) = z(2-p) + (1-z)(2p) = 2z - zp + 2p - 2zp$

$zp + 1 - z - p + zp = 2z - zp - 2p - 2zp$

$2zp - z + 1 - p = 2z - 2p - 3zp$

$5zp - 3z = -p - 1$

$z(5p-3) = -(p+1)$

$$z = \frac{-(p+1)}{5p-3}$$

$$\frac{5p-3}{5p-3} - \frac{-p-1}{5p-3} = \frac{6p-2}{5p-3}$$

For all p

$\left\{ \begin{matrix} (O,O) & , & (O,F) & & O & , & F \\ \langle {}^{2-2p}/_{3-2p}, & & {}^{1}/_{3-2p} \rangle & & \langle {}^{p-2}/_{2p-3}, & & {}^{p-1}/_{2p-3} \rangle \end{matrix} \right\}$

For $p \leq 1/3$

$\left\{ \begin{matrix} (O,F) & , & (F,O) & & O \\ \langle {}^{-p-1}/_{5p-3}, & & {}^{6p-2}/_{5p-3} \rangle & & \langle {}^{2-p}/_{3}, & {}^{1+p}/_{3} \rangle \end{matrix} \right\}$

So for the special case of $p = 0.5$

$\{(F,O), F\}$

$\left\{ \begin{matrix} (O,O), (O,F) & & O, & F \\ \langle 1/2, 1/2 \rangle & & \langle 3/4, & 1/4 \rangle \end{matrix} \right\}$

Scanned with CamScanner

3 Each stage game in an infinitely repeated game will have a reward for each player $i \in I$ and every strategy profile $s$

↳ $r_i(s)$

If $s^*$ is a strategy profile which is a Nash equilibrium of the stage game, then all $I$ players will have no incentive to deviate because of the following...

$\forall i \in I, s \neq s^* \quad r_i(s) \leq r_i(s^*)$

Because an infinitely repeated game means that the stage game will be played over and over again, the strategy profile $s^*$ is a subgame perfect equilibrium due to each player receiving the maximum payoff at each stage or subgame as defined by the Nash equilibrium

4

| | Row Belief | Column Belief | Row Action | Column Action |
|---|---|---|---|---|
| Round 0 | (0.5, 0.25, 0.25) | (0.5, 0.25, 0.25) | | |
| 1 | (0.5, 1.25, 0.25) | (0.5, 1.25, 0.25) | Paper | Paper |
| 2 | (0.5, 1.25, 1.25) | (0.5, 1.25, 1.25) | Scissors | Scissors |
| 3 | (0.5, 1.25, 2.25) | (0.5, 1.25, 2.25) | Scissors | Scissors |
| 4 | (1.5, 1.25, 2.25) | (1.5, 1.25, 2.25) | Rock | Rock |
| 5 | (2.5, 1.25, 2.25) | (2.5, 1.25, 2.25) | Rock | Rock |
| 6 | (3.5, 1.25, 2.25) | (3.5, 1.25, 2.25) | Rock | Rock |
| 7 | (3.5, 2.25, 2.25) | (3.5, 2.25, 2.25) | Paper | Paper |
| 8 | (3.5, 3.25, 2.25) | (3.5, 3.25, 2.25) | Paper | Paper |
| 9 | (3.5, 4.25, 2.25) | (3.5, 4.25, 2.25) | Paper | Paper |
| 10 | (3.5, 5.25, 2.25) | (3.5, 5.25, 2.25) | Paper | Paper |
| 11 | (3.5, 5.25, 3.25) | (3.5, 5.25, 3.25) | Scissors | Scissors |
| 12 | (3.5, 5.25, 4.25) | (3.5, 5.25, 4.25) | Scissors | Scissors |
| 13 | (3.5, 5.25, 5.25) | (3.5, 5.25, 5.25) | Scissors | Scissors |
| 14 | (3.5, 5.25, 6.25) | (3.5, 5.25, 6.25) | Scissors | Scissors |
| 15 | (3.5, 5.25, 7.25) | (3.5, 5.25, 7.25) | Scissors | Scissors |
| 16 | (4.5, 5.25, 7.25) | (4.5, 5.25, 7.25) | Rock | Rock |
| 17 | (5.5, 5.25, 7.25 | (5.5, 5.25, 7.25) | Rock | Rock |
| 18 | (6.5, 5.25, 7.25) | (6.5, 5.25, 7.25) | Rock | Rock |
| 19 | (7.5, 5.25, 7.25) | (7.5, 5.25, 7.25) | Rock | Rock |
| 20 | (8.5, 5.25, 7.25) | (8.5, 5.25, 7.25) | Rock | Rock |
| 21 | (9.5, 5.25, 7.25) | (9.5, 5.25, 7.25) | Rock | Rock |
| 22 | (9.5, 6.25, 7.25) | (9.5, 5.25, 7.25) | Paper | Paper |

⋮

4 From the simulated table for fictitious play, it can be observed that the actions will cycle in blocks of Rock -> Paper -> Scissors of increasing length. Although in the short term example the empirical distribution doesn't converge, if the game were to be simulated infinitely, it would converge to (1/3, 1/3, 1/3) for both agents. This is a property of zero-sum games where the agents use fictitious play.

## Question 5 (Bonus)

Similarities

- Both RMMs and I-POMDPs take the view of a single agent and model the type/strategy of other agents.
- Both methods ascribe the agent to one belief, which gives the agent the ability to make rational decisions.
- Both use recursion/nested beliefs to effectively model the multi-agent environment.

Differences

- RMMs use the reward function to encode the type/strategy of the other agents, while I-POMDPs encode this in the interactive state.
- RMMs treat sub-intentional models as a termination to the recursion process. On the other hand, I-POMDPs still continue when these models are encountered.
- RMMs are used for one-step decision making, while I-POMDPs are specifically designed to handle sequential decision making.