Solutions to Take-Home Final Exam
PHIL 371
Winter 2023

Possible points: 120

1 Honor Statement (5 points)

1. (5 points) Copy the following statement: "I, [insert name here], promise that all the work on this midterm is my own. I may consult my notes from class, past problem sets I've submitted, and lecture slides, but I will not collaborate with others or copy work that is not my own."

2 Problems (80 points)

Complete the following four problems. Each is worth 20 points.

1. Richard Jeffrey was an evidentialist who nevertheless thought that 2-boxing in Newcomb's problem was rational. To make 2-boxing compatible with evidential decision theory, he proposed that rational acts must be *ratifiable*. An act A is ratfiable iff $V(A|dA) > V(B|dB)$ for all alternatives B. $V(A|dA)$ is the "news value" of A given that one performed A; it measures how happy one is to have performed A given that one performed it. (We can read $V(A|dA)$ as "the news value of A given that I actually **d**id A). $V(B|dA)$ is the news value of B given that one performed A; it measures how happy one would have been if one had performed B, given that one actually **d**id A.

   Here's now to calculate news value:

   $$V(A|dA) = p(S_1|A) \times u(S_1\&A) + p(S_2|A) \times u(S_2\&A)...$$
   $$V(B|dA) = p(S_1|A) \times u(S_1\&B) + p(S_2|A) \times u(S_2\&B)...$$

   Suppose you are in a prisoner's dilemma with a twin, where $p(\text{twin defects} \mid \text{you Defect}) = \frac{3}{4}$ and $p(\text{twin cooperates} \mid \text{you Cooperate}) = \frac{4}{5}$. Show that Cooperating is not ratifiable for you (i.e., show that $V(C|dC) \not> V(D|dC)$).

   Since $p(c|C) = \frac{4}{5}$, we can infer that $p(d|C) = \frac{1}{5}$.

   Suppose we're working with the following payoffs:

	Cooperate	Defect
cooperate	$-9, 9$	$0, -10$
defect	$-10, 0$	$-1, -1$

   Now we can calculate $V(C|dC)$ and $V(D|dC)$):

   $$V(C|dC) = p(d|C)u(d\&C) + p(c|C)u(c\&C) = (\frac{1}{5})(-10) + (\frac{4}{5})(-1) = -2.8$$
   $$V(D|dC) = p(d|C)u(d\&D) + p(c|C)u(c\&D) = (\frac{1}{5})(-9) + (\frac{4}{5})(0) = -1.8$$

   Since $-2.8 \not> -1.8$, Cooperating is not ratifiable.

2. Suppose a bag contains 90 balls. 30 of them are yellow and 60 of them are either red or blue. You are offered two choices between wagers based on the color of the ball that's drawn from the bag. The first choice is between L_1 and L_2, while the second choice is between L_3 and L_4.

	Yellow ($\frac{1}{3}$)	Red (x)	Blue ($\frac{2}{3} - x$)
L_1	\$500	\$0	\$0
L_2	\$0	\$500	\$0

	Yellow ($\frac{1}{3}$)	Red (x)	Blue ($\frac{2}{3} - x$)
L_3	\$500	\$0	\$500
L_4	\$0	\$500	\$500

EMPIRICAL FACT: most people prefer L_1 to L_2 and L_4 to L_3.

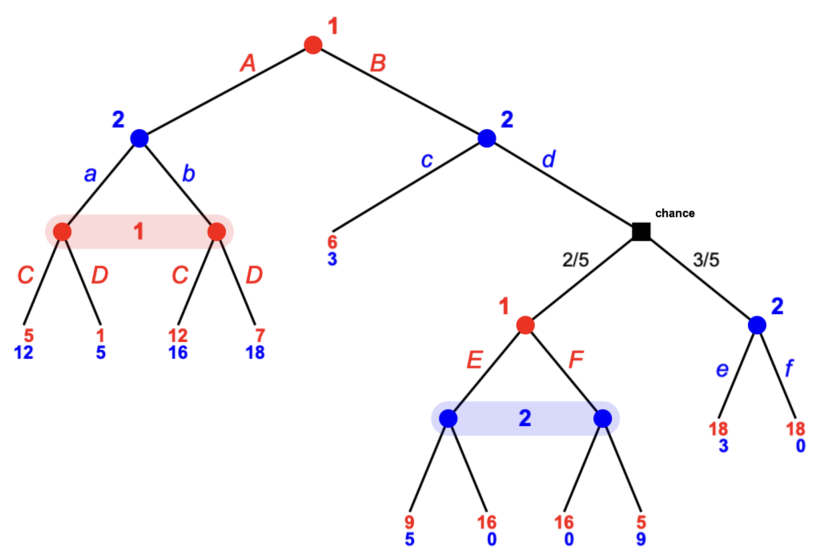
(a) (10 points) Explain why the preferences described in EMPIRICAL FACT violate the Sure Thing principle.

The Sure Thing Principle says that when two acts are sure to produce the same outcome if E, then one's preference between the two outcomes should depend solely on what happens if $\sim E$. The preferences described in EMPIRICAL FACT violate the Sure Thing Principle because L_1 and L_2 are sure to produce the same thing if a blue ball is drawn. This means that replacing the \$0 outcomes with \$500 outcomes should not alter one's preferences. However, since $L_4 \succ L_3$, the change did alter one's preferences.

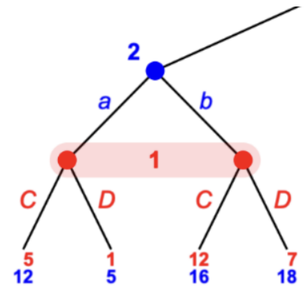
(b) (10 points) In what way do the preferences described in EMPIRICAL FACT differ from Allais preferences?

Allais preferences suggest that agents prefer acts that produce an outcome *for certain*, even if those acts do not maximize expected utility. On the other hand, the preferences described above suggest that agents prefer acts where the likelihood of risk is clear, rather than instances in which the likelihood of risk is unknown, even if those acts do not maximize expected utility.

3. (30 points) Find the subgame-perfect Nash equilibrium of the following game:



Start with the following minimal proper subgame:

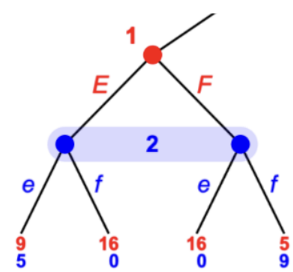


To determine each player's rational strategy, we find the Nash equilibrium:

	a	b
C	$5, 12$	$12, 16$
D	$1, 5$	$7, 18$

In the Nash equilibrium, Player 1 will play C and Player 2 will play b, giving a result of (12,16).

Now look to the next minimal proper subgame:



There is no pure-strategy Nash equilibrium in this subgame. So we find the mixed-strategy Nash equilibrium:

	e	f
E	$9, 5$	$16, 0$
F	$16, 0$	$5, 9$

Let $p(E) = x$ and $p(F) = 1 - x$. To calculate x, we set Player 2's expected utilities equal to one another:

$$EU(e) = EU(f)$$
$$5x + 0(1-x) = 0x + 9(1-x)$$
$$5x = 9 - 9x$$
$$x = \frac{9}{14}$$
$$1 - x = \frac{5}{14}$$

Let $p(e) = y$ and $p(f) = 1 - y$. To calculate y, we set Player 1's expected utilities equal to one another:

$$\begin{aligned} EU(E) &= EU(F) \\ 9y + 16(1-y) &= 16y + 5(1-y) \\ 9y + 16 - 16y &= 16y + 5 - 5y \\ -7y + 16 &= 11y + 5 \\ y &= \frac{11}{18} \\ 1 - y &= \frac{7}{18} \end{aligned}$$

In this mixed strategy Nash equilibrium, Player 1 will play $< \frac{9}{14}E, \frac{5}{14}F >$. To find the Player 1's expected utility in this subgame, we need to calculate the expected utility of each strategy and multiply it by the probability that strategy will be played.

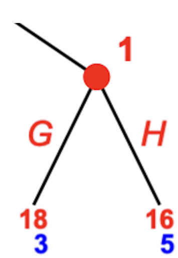
$$\begin{aligned} EU(< \frac{9}{14}E, \frac{5}{14}F >) &= EU(E)p(E) + EU(F)p(F) \\ &= [9(\frac{11}{18}) + 16(\frac{7}{18})](\frac{9}{14}) + [16(\frac{11}{18}) + 5(\frac{7}{18})](\frac{5}{14}) \\ &= 11.7222 \end{aligned}$$

Player 2 will play $< \frac{11}{18}e, \frac{7}{18}f >$. To find his expected utility in the subgame, we need to calculate the expected utility of each strategy and multiply it by the probability that the strategy will be played.

$$\begin{aligned} EU(< \frac{11}{18}e, \frac{7}{18}f >) &= EU(e)p(e) + EU(f)p(f) \\ &= [5(\frac{9}{14}) + 0(\frac{5}{14})](\frac{11}{18}) + [0(\frac{9}{14}) + 9(\frac{5}{14})](\frac{7}{18}) \\ &= 3.21428 \end{aligned}$$

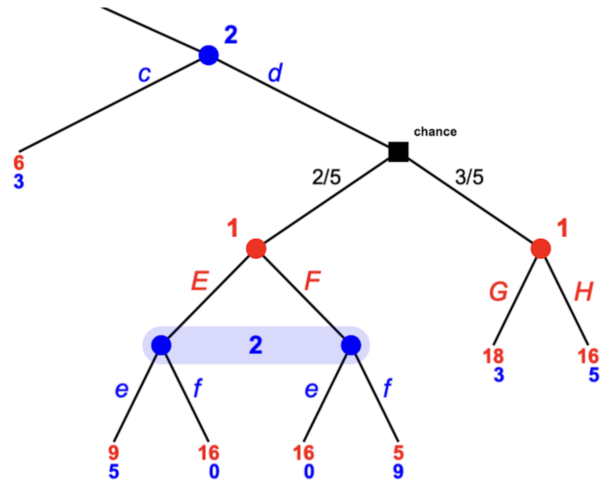
So, the solution to the second subgame is $(11.7222, 3.21428)$, with Player 1 playing $< \frac{9}{14}E, \frac{5}{14}F >$ and Player 2 playing $< \frac{11}{18}e, \frac{7}{18}f >$.

Now we look at the third proper subgame:



Player 1 will play G with a result of $(18, 3)$.

At this point, we look at the final subgame:



Here, Player 2 must choose between playing c and getting a payoff of 3 for sure, or playing d and getting a gamble between two different games. In the first game, he has a $\frac{2}{5}$ chance of getting a payoff of 3.21428 (his payoff in the MSNE solution). In the second game, he has a $\frac{3}{5}$ chance of getting a payoff of 3 (which is what he would get when player one played G). So we calculate $EU(d)$ as follows:

$$\begin{aligned}
EU_{P2}(d) &= EU(\text{Branch 1})p(\text{Branch 1}) + EU(\text{Branch 2})p(\text{Branch 2}) \\
&= (3.214)(\frac{2}{5}) + (3)(\frac{3}{5}) \\
&= 3.0857
\end{aligned}$$

In this subgame, Player 2 is facing a choice between playing c, where $EU(c) = 3$, and playing d where $EU(d) = 3.0856$. Since $EU(d) > EU(c)$, Player 2 will play d, giving him a payoff of 3.0857. However, we also need to find Player 1's payoff in order to finish the backward induction.

If Player 2 plays d, Player 1 gets a $\frac{2}{5}$ chance at a payoff of 11.478 and a $\frac{3}{5}$ chance of a payoff of 18:So her expected payoff is $(11.478)(\frac{2}{5}) + (18)(\frac{3}{5}) = 15.4712$.

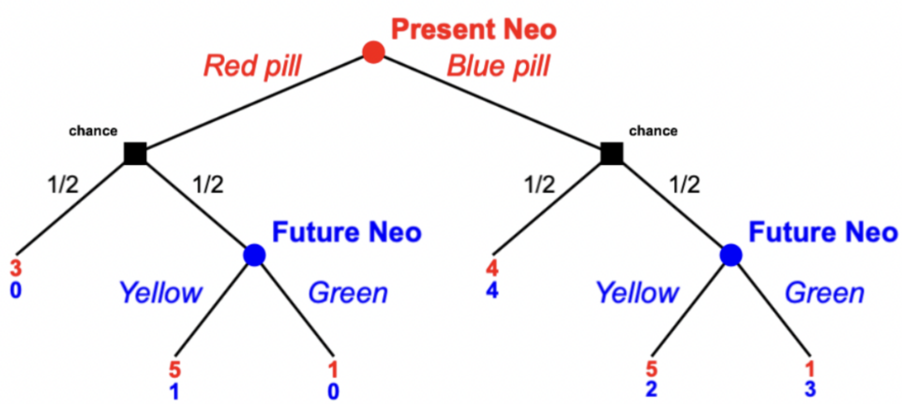
$$\begin{aligned}
EU_{P1}(d) &= EU(\text{Branch 1})p(\text{Branch 1}) + EU(\text{Branch 2})p(\text{Branch 2}) \\
&= (11.7222)(\frac{2}{5}) + (18)(\frac{3}{5}) \\
&= 15.4888
\end{aligned}$$

Now let's zoom back out to the full game. At the first choice node, Player 1 must make a decision between playing A and getting an expected payoff of 12 (while Player 2 gets a payoff of 16), and playing B and getting an expected payoff of 15.4888 (while Player 2 gets a payoff of 3.0857). Player 1 will choose A.

The full backward induction to the game is as follows. Players 1 and 2 will end up with payoffs of (15.4888, 3.0857), and will play the following strategies:

$$S_{P1} = \{B, C, < \frac{9}{14}E, \frac{5}{14}F >, G\}$$
$$S_{P2} = \{b, d, < \frac{11}{18}e, \frac{7}{18}f >\}$$

4. Neo faces a choice between taking a red pill or a blue pill. Each pill will alter his preferences in different ways. After he takes a pill, a fair coin will be tossed. If it lands heads, Neo will receive a new shirt that is the color of the pill he took. If it lands tails, he will be offered the choice between a yellow shirt and a green shirt. Before taking the pill, Neo's utilities for sweaters are $u(\text{Blue}) = 4, u(\text{Red}) = 3, u(\text{Yellow}) = 5, u(\text{Green}) = 1$. If he takes the red pill, his utilities become $u_R(\text{Blue}) = 4, u_R(\text{Red}) = 0, u_R(\text{Yellow}) = 1, u_R(\text{Green}) = 0$. If he takes the blue pill then his utilities become $u_B(\text{Blue}) = 4, u_B(\text{Red}) = 0, u_B(\text{Yellow}) = 2, u_R(\text{Green}) = 3$. Draw a decision tree that represents Neo's situation and use it to explain what he should to do maximize his expected utility.



Let PN refer to Present Neo and FN refer to Future Neo. Using backward induction, we can show that PN maximizes expected utility by choosing the red pill, but FN maximizes expected utility by choosing the blue pill.

If PN takes the red pill and the coin lands heads, PN and FN will end up with payoffs of $(3, 0)$. If PN takes the red pill and the coin lands tails, FN would maximize his utility by choosing the yellow shirt, yielding payoffs of $(5, 1)$.

$$\begin{aligned} EU_{PN}(\text{Red pill}) &= p(\text{Heads})u_{PN}(\text{Heads \& Red}) + p(\text{Tails})u_{PN}(\text{Tails \& Red}) \\ &= (0.5)(3) + (0.5)(5) \\ &= 1.5 + 2.5 \\ &= 4 \end{aligned}$$

If PN takes the blue pill and the coin lands heads, PN and FN will end up with payoffs of $(4, 4)$. If PN takes the blue pill and the coin lands tails, FN would maximize his expected utility by choosing the green shirt, yielding payoffs of $(1, 3)$.

$$\begin{aligned} EU_{PN}(\text{Blue pill}) &= p(\text{Heads})u_{PN}(\text{Heads \& Blue}) + p(\text{Tails})u_{PN}(\text{Tails \& Blue}) \\ &= (0.5)(4) + (0.5)(1) \\ &= 2 + 0.5 \\ &= 2.5 \end{aligned}$$

Since $EU_{PN}(\text{Red pill}) > EU_{PN}(\text{Blue pill})$, PN should take the red pill to maximize his expected utility.

However, since $EU_{FN}(\text{Blue pill}) > EU_{FN}(\text{Red pill})$, FN must take the blue pill to maximize his expected utility:

$$\begin{aligned} EU_{FN}(\text{Red pill}) &= p(\text{Heads})u_{FN}(\text{Heads \& Red}) + p(\text{Tails})u_{FN}(\text{Tails \& Red}) \\ &= (0.5)(0) + (0.5)(1) \\ &= 0 + 0.5 \\ &= 0.5 \end{aligned}$$

$$\begin{aligned} EU_{FN}(\text{Blue pill}) &= p(\text{Heads})u_{FN}(\text{Heads \& Blue}) + p(\text{Tails})u_{FN}(\text{Tails \& Blue}) \\ &= (0.5)(4) + (0.5)(3) \\ &= 2 + 1.5 \\ &= 3.5 \end{aligned}$$

At this point, there are multiple acceptable answers to the question. You might say that PN should prioritize his future wellbeing and choose the blue pill. Or you might say that we should think about PN and FN as different "players" who should each act in their own self-interest. In that case, PN should take the red pill.

3 Essay (35 points)

Write a 3 paragraph essay on the following topic. The essay will be graded on the basis of clarity, comprehension, and persuasiveness. (Do not spend too much time on this!)

> Are there situations where a group knows that *p* but few people in the group know that *p*? Are there situations where a group knows that *p* but *no one* in the group knows that *p*? Defend your answers with reference to the papers we've read on group knowledge.