24.118 – Paradox and Infinity
Problem Set 9: Paradoxes of Common Knowledge

How this problem set will be graded:

You will be graded both on the basis of whether your answers are correct and on the basis of whether they are properly justified. There is a limit of 150 words per question. Each question is worth five points.

Problems:

1.) According to Carroll, the backward induction argument for finite prisoner's dilemma sequences relies on premises {(G), (B1'), (B2'), …}. What analogous premises would be needed to ground the backward induction in the case of the surprise exam (or surprise cookies)?

2.) Evaluate the following argument:

No matter what finite value of *n* is plugged into condition (a) of (G), a valid backward induction argument can be constructed using {(B1'), (B2')…} Therefore, if condition (a) were replaced with the following condition

(a') Players 1 and 2 are playing a finite sequence of prisoner's dilemmas.

then a valid backward induction argument could be constructed using (a'), (b), (c), and appropriately modified versions of {(B1'), (B2')…}.

3.) Suppose I make the following threat:
   a. One of these days, I'll hit you with a water balloon.
   b. I'll only do it once.
   c. I'll do it on the first day when you're not expecting it.

In what circumstances (if any) can the two of us have common knowledge of (a), (b), and (c)?

4.) There once was a small island whose inhabitants had a strange tradition: anyone who discovered his or her own eye color was obligated to commit ritual suicide by jumping off a cliff at midnight on the day the discovery was made. Due to this tradition, the topic of eye color on the island was taboo. Everyone on the island knew everyone else's eye color. (In fact, because it was such a small island, everyone knew just about *everything* about everyone else.) But no one knew his or her own eye color, and so the island's seven blue-eyed inhabitants lived in peace and happiness with their fifteen green-eyed compatriots and their three hundred brown-eyed compatriots.

All went smoothly until one visiting stranger, unfamiliar with the taboos of the island, made a shocking declaration: "Some of you have blue eyes." The stranger spoke loudly, in the middle of the town square, where it was clear everyone could hear. At midnight six days later, all the blue-eyed people killed themselves.

What the stranger said wasn't news: even before he spoke, everyone knew that some of the island's inhabitants had blue eyes. So how did his utterance enable the blue-eyed people to learn their own eye color? (You may assume that strangers are commonly known to be truthful.)

5.) The mayor of the island gathered everyone together and declared a new policy: strangers were banned from speaking in public. If they wished to communicate with the island's inhabitants, they could use the mail service, which (it was commonly known) always delivered letters within three days.

All went well until a second stranger began using the mail service to distribute information about eye color. Letters arrived at every household which said, "Some of you have green eyes. This letter has been sent to every household on the island."

What happened next? (Again, you may assume that strangers are commonly known to be truthful.)

Imagine that the surprise exam is converted into a game with two players called "Teacher" and "Student". The game lasts for *n* rounds, where *n* is some natural number known to both Teacher and Student. At each round, the players must bet on whether an exam will take place. Student chooses one side of the bet (either "yes" or "no"). Without knowing Student's choice, Teacher decides whether to hold the exam. Both Teacher and Student's choices are then revealed, and if the game has not ended, then players proceed to the next round. If Student has chosen the winning side of the bet, she pays Teacher $10; otherwise, Teacher pays Student $10. Play ends whenever the exam takes place, or at the *n*th round. If Teacher fails to set the exam by the *n*th round, she must pay Student $30.

6.) What is smallest *n* such that, no matter what Student does, it is always possible for Teacher to end up at least $10 richer?

7.) Let a *strategy* be an algorithm that determines which move a player will make on each round. (For example, Student might pick the strategy, "Guess 'yes' on all even-numbered days and 'no' on odd-numbered days", while Teacher might pick the strategy: "Set the exam on the first day after Student says 'no'.") Show that no matter which strategy Student picks, if Student were to find out that Teacher had learned his (Student's) strategy, it would be rational for him (Student) to switch strategies.