

## HW3: Game Playing

## 1 Non Zero Sum Games

Alice is taking a class taught by Bob called “Artificial Intelligence.” Bob has three ways he can teach the class: “Hard,” “Medium” or “Easy.” Alice has three ways she can take the class: “Hard Working,” “Working” and “Hardly Working.” For each of them, there are pros and cons. For instance, it’s easy for Bob to teach an Easy class or a Hard class, but hard to balance the two. Obviously Alice doesn’t like to work hard, but she realizes that she might have to in order to learn something. Bob is happy to teach a hard class to students who are willing to work hard, but if the students don’t work hard, they punish Bob by giving him bad teaching evals!<sup>1</sup> All of these things considered gives rise to the following table of rewards. These are written as  $(A, B)$  where  $A$  is Alice’s reward and  $B$  is Bob’s reward:

	Hard	Medium	Easy
Hard	(9, 9)	(6, 6)	(2, 1)
Working	(5, 8)	(8, 7)	(4, 2)
Hardly	(4, 1)	(5, 2)	(4, 3)

1. If Bob assumes that Alice will optimize her own reward (i.e., Bob assume’s Alice is an optimal agent), how should he teach the class, supposing that Bob plays first? If Alice assumes Bob is an optimal agent, how hard should she work?
2. Draw a game tree for this problem supposing that Bob goes first. Propagate values up through the tree using (the non-zero-sum variant of) minimax search.
3. Alice is clearly a good student (see question one), but once in a while we get students who aren’t quite as dilligent :( . It makes sense for Bob to model his class as a distribution over types of students. Suppose Bob believes that 30% of his class will work hard, 50% will work, and 20% will hardly work. Draw the expectimax tree for this setting, concentrating only on Bob’s reward, and compute expected node values. What is Bob’s expected reward for this setting and which type of class should he teach?
4. Some faculty (Bob not included!) follow the model of trying to scare away all students who won’t work hard on the first day of class. The idea is that if they drop the class, then the professor won’t get a bad teaching review! Suppose Bob decided to follow this policy and all 20% of the hardly working students dropped. What is Bob’s new expected reward and what type of class should he teach?
5. **(6300 only)** Suppose students were somewhat adversarial and were only happy if Bob weren’t happy. Operationally, if  $(A, B)$  was the old reward pair, suppose the new reward pair is  $(A - \beta B, B)$ , for some  $\beta > 0$ . Is there a value of  $\beta$  that will make this a zero sum game? If so, what is it; if not, why not? What is the *minimal* value of  $\beta$  that would cause the answer to problem 1 to change (if there is no such  $\beta$ , why not?)?

---

<sup>1</sup>Yes, this implies that any bad teaching reviews *must* be due to shortcomings of students, not of professors!

## 2 Probability and Utility

I don't have a cute story to wrap around the following questions, so just answer them :).

1. I flip a fair coin but don't let you see how it came up. I tell you that if you guess right, I'll give you \$10. What is your expected reward (write out the computation!)?
2. Coins are boring. Now I roll a fair six-sided die but don't let you see how it came up. I tell you that if you guess right, I'll give you \$10. What is your expected reward?
3. Let's say that now I tell you that the die isn't fair, but that the probabilities are as follows:  $p(1) = 0.3$ ,  $p(2) = 0.1$ ,  $p(3) = 0.1$ ,  $p(4) = 0.2$ ,  $p(5) = 0.2$ ,  $p(6) = 0.1$ . Again, I'll give you \$10 if you guess right. For each of the six possible guesses you could make, compute your expected reward. Which would you guess?
4. Now, I make you the following offer. Keep the same die as before. But now, I tell you that if you guess right, I'll give you \$10 *times* the number you guess. I.e., if you guess 2 and you're right, I'll give you \$20. Now what is your best option to guess? Is this the same or different from the previous problem? Explain why or why not.
5. **(6300 only)** Suppose now I roll two dice, each weighted as in the previous questions. You are allowed to make two guesses and if *either* one matches one of the die I rolled, you win \$10 times that guess. (So, if I roll a 2 and 6 and you guess 2 and 3, you get \$20; but if you have guessed 3 and 6 you would have gotten \$60.) If you guess both right, you get the larger amount. Rigorously compute your optimal action (i.e., guess). (Hint: if you get something that says you should guess the same number twice, you've done something wrong!) Next, suppose that we play the game again but under slightly different rules. You still win if you're able to guess one of the die values, but now I give you \$10 times your first guess times your second guess. So if I roll a 2 and 6 and you guess 2 and 4, then you win \$80. Now what is your best action choice?