

Assignment Eight: Sequential Decisions Solutions

Question One

- (a) In the following the actual numbers are not significant. However, they have the following characteristics that is important: A is preferred to B which is preferred to C which is much preferred to F for all efforts. Working little is much better than working a lot. The best outcome is to get an A with little effort. The worst outcome is to work hard and get an F.

Grade	Total Effort	Utility
A	Lot	50
A	Little	100
B	Lot	40
B	Little	90
C	Lot	30
C	Little	80
F	Lot	0
F	Little	20

- (b) For this student, getting an A is much better than getting a B, which is a little bit better than a C which is a little bit better than F. Arguably they will feel better getting a low grade if they worked hard than if they didn't (although we won't take marks off for the other way around). What is important is the big gap between the A's and the rest.

Grade	Total Effort	Utility
A	Lot	95
A	Little	100
B	Lot	40
B	Little	30
C	Lot	20
C	Little	10
F	Lot	10
F	Little	0

Question Two

- (a) At first, cheat. At the second time, cheat if you didn't get caught the first time. The expected utility is 81.69.
- (b) When the probability of Watched is increased a small amount (e.g., from 0.3 to 0.5), the policy changes to cheat at the first opportunity and don't cheat at the second opportunity.

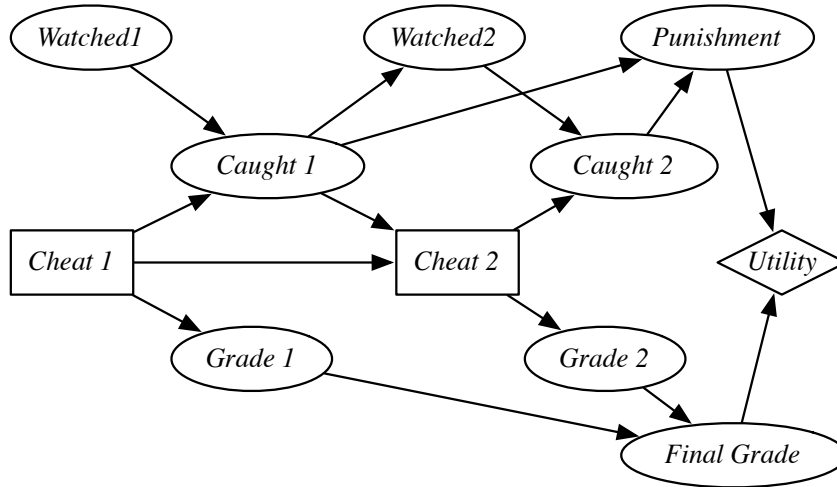
When the probability of increased more (e.g., to 0.8) the policy becomes to not cheat at either time.

- (c) We can change $P(\text{Grade1} \mid \text{Cheat1})$ and $P(\text{Grade2} \mid \text{Cheat2})$.

If the probability are changed to $P(\text{Grade2} = F \mid \text{Cheat2} = T) = 0.1$ and $P(\text{Grade2} = A \mid \text{Cheat2} = T) = 0.4$ (with the other probabilities for $\text{Cheat2} = T$ remaining at 0.25 each), the policy changes to not cheat at the second opportunity.

If the probabilities of $P(\text{Grade1} \mid \text{Cheat1})$ are changed to be $P(\text{Grade2} = F \mid \text{Cheat2} = T) = 0.1$, $P(\text{Grade2} = C \mid \text{Cheat2} = T) = 0.25$, $P(\text{Grade2} = B \mid \text{Cheat2} = T) = 0.25$ and $P(\text{Grade2} = A \mid \text{Cheat2} = T) = 0.4$ the policy changes to not cheat at the first opportunity.

- (d) See:

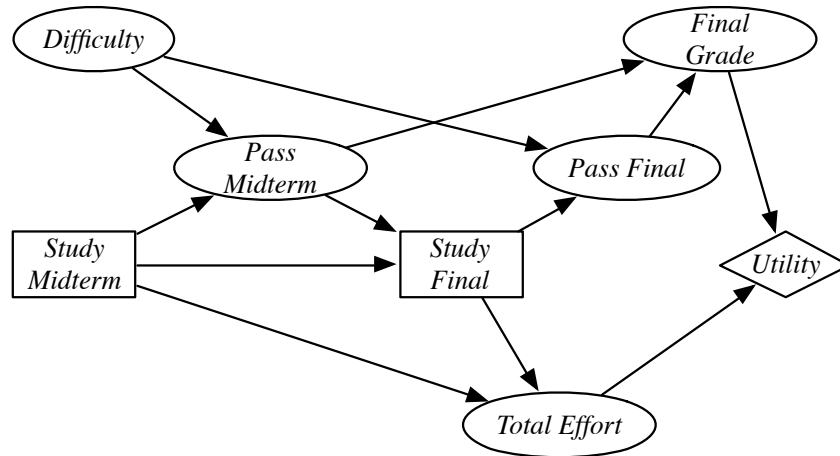


If the probability of watched1 goes up, we can even get to the situation where the policy is to not cheat at the first opportunity and to cheat at the second opportunity. See: https://www.cs.ubc.ca/~poole/cs322/2020/as8/cheat_decision_watched2.xml for such an example (which didn't happen by changing the probabilities in the first model).

- (e) Making punishment more severe (suspension or recorded on transcript for first offences) makes cheating less advantageous even if the probability of getting caught is 0.2. See e.g.,: https://www.cs.ubc.ca/~poole/cs322/2020/as8/cheat_decision_honour.xml

Question Three

The decision network is:



Question Four

It should not have taken more than a few hours. Most of this should have been in understanding the material and playing, not in doing busy work. I hope it was reasonable, and you learned something.