Assignment Nine: Optimizing Sequential Decisions

Due: 11:59pm, Monday 30 November 2020.

Solving following problems requires using the `http://aispace.org` belief and decision networks applet (you may need to download the jar file) and/or the Python code at `http://aipython.org/aipython_322.zip` (which now includes reasoning with uncertainty and planning with uncertainty).

This can be done in groups of size 1, 2 or 3. Working alone is not recommended. All members of the group need to be able to explain the group's answer.

Submit your answers in pdf using Canvas. Use proper sentences in your answers.

Ask questions on Canvas discussion board. Feel free to answer them too.

Question One

Consider the decision network from the last assignment shown in Figure 1. This is available as `https://artint.info/code/aispace/cheat_decision.xml` for the AISpace Belief and Decision network tool. (Do a "Load from URL" in the "File" menu), and as *cheat_dn* in decnNetworks.py in the Python distribution.

(a) Suppose we want to compute the optimal policy. Which variables need to be summed out before any maximization?

(b) After summing out these variables, show the resulting factor. Just show what variables it is a factor of and give two rows of the table (where only the value for the variable that needs to be maximized differs).

(c) What is the resulting decision function after the first maximization? Give the variables it depends on and the row corresponding to the rows of part (b).

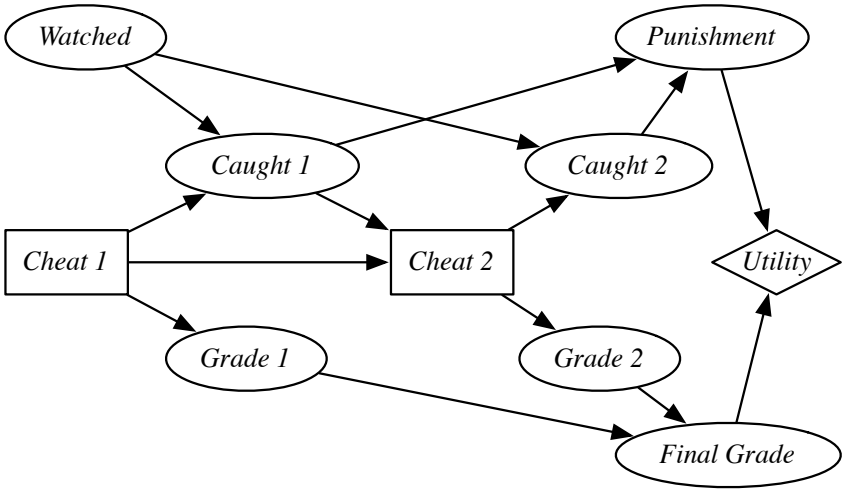

Figure 1: Cheat Decision Network

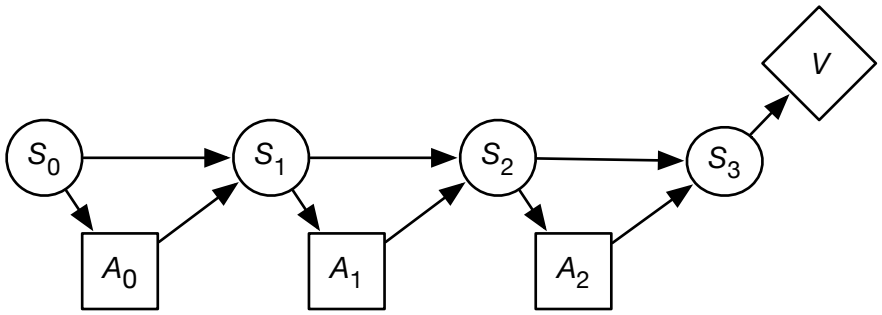

Figure 2: Decision Chain

(d) What is the resulting factor created by the first maximization? Give the variables it depends on and the row corresponding to the rows of part (b).

(e) What variable(s) need to be summed out before the next maximization? What is the resulting factor after summing out these variable(s).

(f) What is the resulting decision function after the second maximization?

(g) What is the resulting factor after the second maximization?

(h) What is the value of information of *Watched* for the second cheat decision?

(i) What is the value of information of *Watched* for the first cheat decision?

Question Two

Consider the decision network of Figure 2.

Suppose the domain of each S_i is $\{0, 1, 2, 3, 4\}$ and the domain of each A_i is $\{\mathit{flip}, \mathit{hold}, \mathit{reset}\}$. The conditional probability $P(S_{i+1} \mid A_i S_i)$ is described by:

- If $A_i = \mathit{hold}$ then S_{i+1} has the same value as S_i (with probability 1).
- If $A_i = \mathit{reset}$ then S_{i+1} has value 0 (with probability 1).
- If $A_i = \mathit{flip}$ then S_{i+1} is $\mathit{min}(S_i + 1, 4)$ with probability 0.3, $\mathit{min}(S_i + 2, 4)$ with probability 0.4, and $\mathit{min}(S_i + 3, 4)$ with probability 0.3.

The value function $V(S_3)$ is given by

S_3	*Value*
0	1
1	2
2	5
3	10
4	0

(a) What arcs need to be added to make this a no-forgetting decision network?

(b) For each variable eliminated, in order, fill in the following table, where "How eliminated" is either "sum" or "max". Use the original name for the input factors. Call each factor of one variable V, each factor of two variables Q, and each factor of more than two variables S. When applicable, give a decision function and call it d_i for action i. You only need to say what variables each factor or decision function depends on.

Variable Eliminated	How eliminated	Factors Removed	Factor added	Decision Function

(c) Explain the optimal policy in English. Warning: the AIspace applet gives the correct results in verbose mode for "Optimize decisions", but is broken when doing "View/Modify Decision".

(d) How does the probability distribution $P(S_0)$ affect the policy or the expected utility?

(e) Suppose instead of doing 3 actions, we were to continue and do n actions, and suppose there were s states and d possible decisions. What is the space and time complexity of finding the optimal policy?

Question Three

For each question, specify how long you spend on it, and what you learned. How was the work in the team allocated? Was the question reasonable?