# Artificial Intelligence Qualifying Exam

Fall 2020

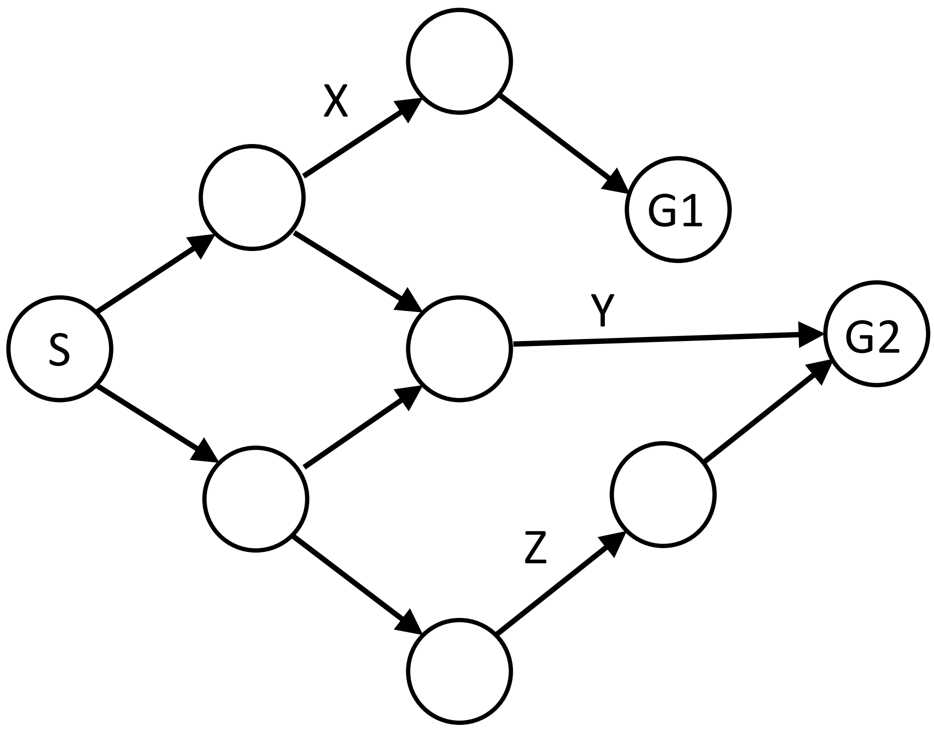
**ID:**

This exam is closed book. You may use a calculator for question 8, but it should not be needed for the other questions. Try to answer as many questions correctly as you can. It is not necessary to answer all of the questions correctly to pass, so if you are stumped by a question, move on to the next question.

# 1 Search (20 points)

The questions in this section all involve variations of search and all use the figure below. The node labeled $S$ is the start state. The nodes labeled $G1$ and $G2$ both satisfy the goal test. Assume all unlabeled edges have a cost of $1$. You will be asked to provide costs for the edges labeled $X$, $Y$, and $Z$ for various parts of this question. **The costs of these nodes must be greater than 0.** For this question, you may assume that all algorithms apply a goal test *after* popping a node off the queue. As in all search algorithms, our notion of optimality refers to whether the path returned by the algorithm is the lowest cost path.

If any of your answers requires an assumption about how the algorithm handles the fact that the state space is a DAG rather than a tree, you should state those assumptions explicitly.




## 1.1 DFS (5 points)

Assuming that Depth First Search (DFS) pushes nodes that point **UP** onto the queue before nodes that point down, provide values of $X$, $Y$ and $Z$ for which DFS will return an optimal solution.

### 1.2 BFS (5 points)

Assuming that Breadth First Search (BFS) pushes nodes that point **UP** onto the queue before nodes that point down, provide values of $X$, $Y$ and $Z$ for which BFS will return a suboptimal solution.

### 1.3 A* 1 (5 points)

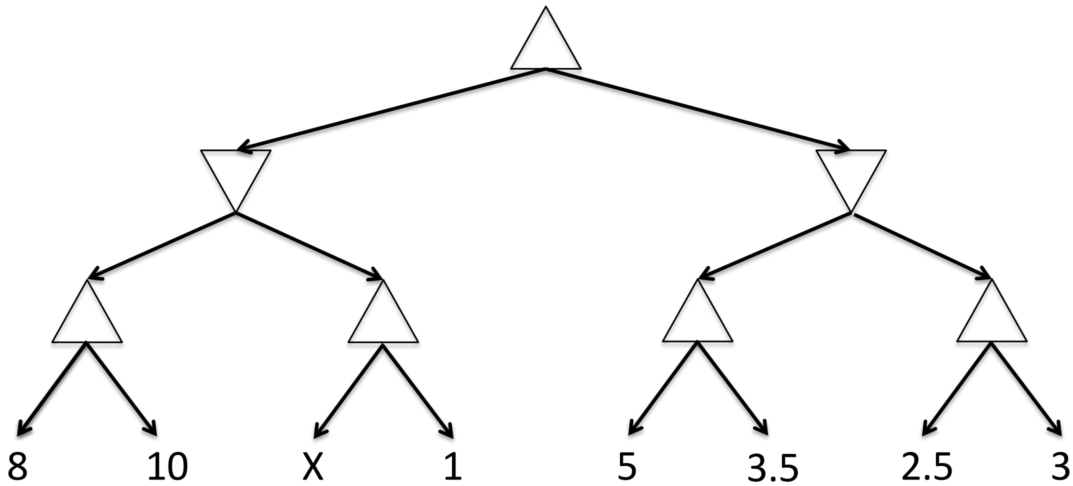
Assume the heuristict function $h(\cdot)$ is uniformly $1$, except at goal states, where it is $0$. Provide values of $X$, $Y$ and $Z$ for which A* will return a path to $G1$. Your answer should *not* require any assumptions about how ties are broken.

### 1.4 A* 2 (5 points)

Assume the heuristict function $h(\cdot)$ is uniformly $1$, except at goal states, where it is $0$. Provide values of $X$, $Y$ and $Z$ for which A* will return a path to $G2$ that traverses the edge with cost $Z$. Your answer should *not* require any assumptions about how ties are broken.

## 2 Game Search (15 points)

The questions in this section will refer to the following game tree. Upward-pointing triangles refer to *max* nodes. Downward-pointing triangles refer to min nodes. Assume that nodes are expanded in depth-first, left-to-right order. The leaf values are shown at the bottom of the tree. One of these is left as a variable, the domain of which is the natural numbers between 1 and 12, inclusive.




### 2.1 Root Value (5 points)

What is the minimax value of the root node as a function of $X$ and the other leaf values?

### 2.2 Alpha-Beta Pruning 1 (5 points)

As the value of $X$ increases, the number of nodes that Alpha-Beta pruning can prune should be monotonically non-decreasing. What is the *largest* value of $X$ for which no pruning is possible? Recall that $X \in \{1...12\}$.

### 2.3 Alpha-Beta Pruning 2 (5 points)

As the value of $X$ increases, the number of nodes that Alpha-Beta pruning can prune should be monotonically non-decreasing. What is the *smallest* value of $X$ at which pruning is maximized? Provide the value for $X$, and also indicate which leaves will not be considered by Alpha-Beta by listing the corresponding leaf values. Recall that $X \in \{1...12\}$.

## 3 Bayesian Networks (15 points)

Consider a Bayesian network defined over 3 binary random variables, $A$, $B$, and $C$. Following convention, use lower case for the domains of these variables, e.g., $A \in \{a, \overline{a}\}$ with $A = a$ corresponding to the true case and $A = \overline{a}$ corresponding to the false case. Assume that $A$ and $B$ are unconditionally independent, and that $C$ is a deterministic function of $A$ and $B$.

### 3.1 Bayes Net Structure (5 points)

Write down the graphical structure of the network that is implied by the description above.

### 3.2 Marginal Probabilities (5 points)

Write out $P(C)$ as a function of the conditional probability tables (CPTs) in your network. This answer does not require any numbers, but should demonstrate your understanding of how a marginal probability can be computed from the conditional probabilities in a Bayesian network.

### 3.3 Conditional Probability Table Values (5 points)

Assume $P(A = a) = \frac{1}{2}$ and $P(B = b) = \frac{3}{4}$, and that $C$ is the logical **OR** of $A$ and $B$, do the following: 1) Compute $P(C = c)$ by any means you want, and 2) Show that your equation form the previous part gives the correct answer when the corresponding conditional probabilities are plugged in. (If you are confident in your use of the marginal probability calculation, then you don't need to answer this question two parts. You will get partial credit if you provide the correct probability, computed by some other method, and your marginal probability calculation isn't fully correct.)

## 4 HMMs (15 points)

Consider a Hidden Markov Model (HMM) with the following property: After every j steps, the model is guaranteed to reveal what the state is. (This can be done within the standard HMM framework by having states with distinctive outputs and structuring the transition probabilities so that one of these states is visited on regular intervals. Alternatively, you could assume that there is an oracle that simply tells you the state after every j transitions.) For example, for j=0, there is never any uncertainty about the state. For j=1, you alternate between known states and states with some uncertainty. For j=infinity, we have a standard HMM.

For this question, we will consider whether this property can be exploited. Assume that there are n states, and that there are t transitions and observations. Suppose our task is to compute the smoothed probabilities, $P(S_k|e_0...e_t)$, where $S_k$ is the state at time $k$, and $e_0 \ldots e_t$ correspond to the outputs (also called evidence) for time steps 0 through $t$, with $0 < k < t$. Note that we are interested in doing this for a single, given, $k$ where the state is not directly observed.

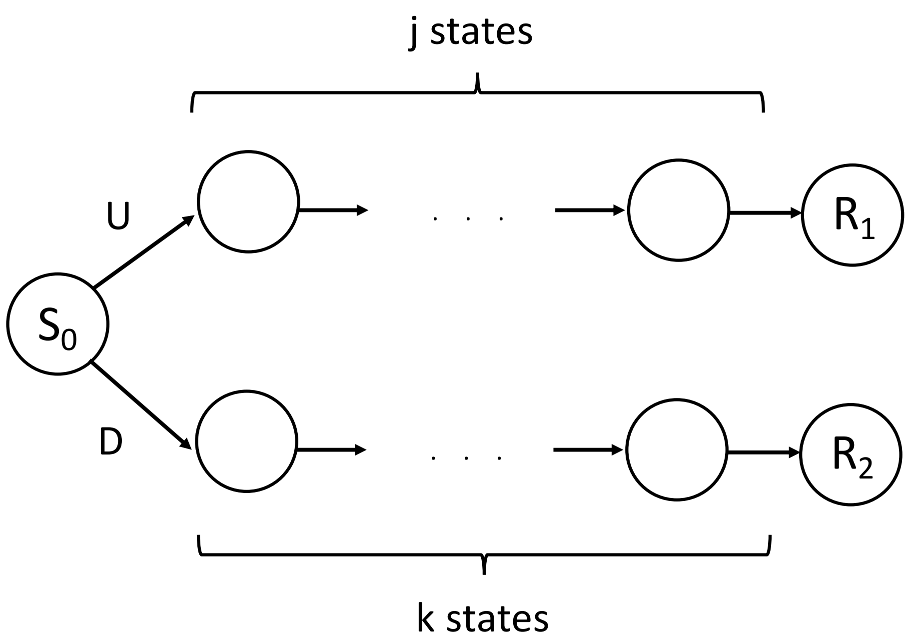
Using O() notation, what is the computational complexity of the most efficient algorithm you can think of for doing this in terms of n, j, and t, and 2) If your answer is different from the j=infinity case then you should explain any modifications to the standard algorithm. If your answer is the same as the j=infinity case, then you should justify why you can't take advantage of the extra information.

If you are really stumped, consider that you will get some partial credit for correctly stating and justifying the computational complexity of the original algorithm for the j=infinity case.

## 5 MDPs (15 points)

Consider the Markov Decision Process (MDP) shown in the figure below. There is only one state, $S_0$, which has a choice of action (U or D). All other state have no choice of action and make the transitions shown. All states have reward 0 except the two states marked $R_1$ and $R2$. These are terminal states with reward R1 and R2, respectively. For the purposes of this question, this means that these state values are locked at $V(R1) = R1$, and $V(R2) = R2$.

Between S0 and the terminal states there are two paths of states. The upper path is j states long, while the lower path is k states long.




For a discount of $\gamma = 0.5$, $j = 4$, $R_1 = 128$, $k = 6$, $R_2 = 256$, compute the optimal policy and V(S0). The first line of your answer should be: "$V(S0) = \ldots$". The second line should be "Policy = ...". You should use the space below that to show your work to arrive upon your answer. Note that we cannot give any partial credit if some work is not shown. The numbers were selected so that you should not need a calculator.

# 6 Value of Information (20 points)

The question is based upon a scenario some of you may have faced at the end of the last academic year. Suppose that you are trying to consider whether to sign a new lease through the summer $(N = n)$ or extend your existing lease $(N = \overline{n})$ for a minimum of 6 months. Signing a new lease is cheaper if you plan to stay in the same place for a while, but there is a cost for breaking your lease. Not signing a new lease and going on a month-to-month plan with your landlord is more expensive per month, but there is no cost associated with stopping.

The random variable that influences things is whether we are still on quarantine in Fall $(Q = q)$ or if we have returned to normality $(Q = \overline{q})$. The following are the utilities for various actions and outcomes. Please don't worry about whether these numbers make much intuitive sense.

- $U(n|q) = 100$ (You save money by signing the lease)
- $U(\overline{n}|q) = 0$ (You pay more by going month to month)
- $U(n|\overline{q}) = 100$ (You're happier because there's no quarantine, but it costs you to break the lease)
- $U(\overline{n}|\overline{q}) = 200$ (You're happier because there's no quarantine, and you don't have to break your lease)

Assume that there is a $\frac{1}{3}$ probability that there will be a quarantine in the Fall, i.e., $P(q) = \frac{1}{3}$.

You have the option of paying a consultant to do some fancy modeling and give a more accurate prediction of the chances that there will be a quarantine in the Fall. The information from the consultant will take the form $C = c$ if the consultant believes there will be a quarantine in the fall. $P(c|q) = \frac{3}{4}$, and $P(c|\overline{q}) = \frac{1}{3}$.

For this problem, we'll ignore the subtleties involved in converting utility to money and assume that you can pay in units of utility to keep things simple. You should compute the maximum you would be willing to pay the consultant, i.e., the value of information, expressed in units of your utility. To help us give you partial credit, we have broken this down into two subproblems. You should show as much work as you reasonably can.

You may want to use a calculator for this question since things not work out cleanly to be whole numbers.

## 6.1 Optimal Action Without Information (10 points)

Compute the optimal action and expected utility without any information from the consultant.

### 6.2 Value of information (10 points)

Compute the value of the information offered by the consultant.