# Final Exam

Perfect score: 100.

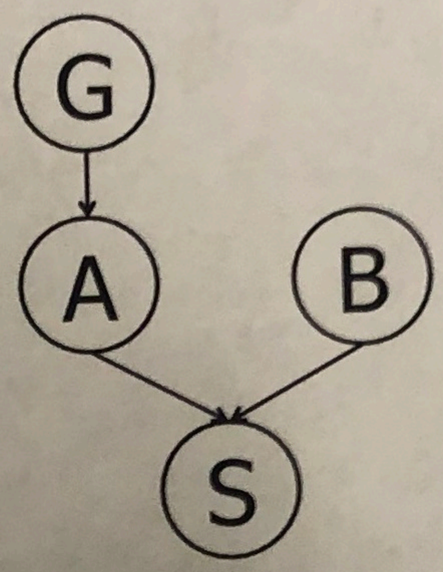
Duration: 120 minutes.

**Problem 1** (*30 points*):

Suppose that a patient can have a symptom ($S$) that can be caused by two different diseases ($A$ and $B$). It is known that the variation of gene $G$ plays a big role in the manifestation of disease $A$. The Bayes' Net and corresponding conditional probability tables for this situation are shown below. ("+" means true and "−" means false).

| $\mathbb{P}(G)$ | |
|---|---|
| $+g$ | 0.1 |
| $-g$ | 0.9 |

| $\mathbb{P}(A\|G)$ | | |
|---|---|---|
| $+g$ | $+a$ | 1.0 |
| $+g$ | $-a$ | 0.0 |
| $-g$ | $+a$ | 0.1 |
| $-g$ | $-a$ | 0.9 |




| $\mathbb{P}(B)$ | |
|---|---|
| $+b$ | 0.4 |
| $-b$ | 0.6 |

| $\mathbb{P}(S\|A,B)$ | | | |
|---|---|---|---|
| $+a$ | $+b$ | $+s$ | 1.0 |
| $+a$ | $+b$ | $-s$ | 0.0 |
| $+a$ | $-b$ | $+s$ | 0.9 |
| $+a$ | $-b$ | $-s$ | 0.1 |
| $-a$ | $+b$ | $+s$ | 0.8 |
| $-a$ | $+b$ | $-s$ | 0.2 |
| $-a$ | $-b$ | $+s$ | 0.1 |
| $-a$ | $-b$ | $-s$ | 0.9 |

1. What is the probability that a patient has disease $A$? (Remember: Only the parents matter when no evidence about the descendents is given, this should be very short, two or three lines at most).
2. What is the probability that a patient has disease $A$ given that they have symptom $S$ and disease $B$?

**Problem 2** (*20 points*):

Consider the Markov Decision Process (MDP) with transition probabilities and reward function as given in the tables below. Assume the discount factor $\gamma = 1$ (i.e., there is no actual discounting).

| $s$ | $a$ | $s'$ | $T(s,a,s')$ |
|---|---|---|---|
| $A$ | 1 | $A$ | 1 |
| $A$ | 1 | $B$ | 0 |
| $A$ | 2 | $A$ | 0.5 |
| $A$ | 2 | $B$ | 0.5 |

| $s$ | $a$ | $R(s,a)$ |
|---|---|---|
| $A$ | 1 | 0 |
| $A$ | 2 | $-1$ |

| $s$ | $a$ | $s'$ | $T(s,a,s')$ |
|---|---|---|---|
| $B$ | 1 | $A$ | 0 |
| $B$ | 1 | $B$ | 1 |
| $B$ | 2 | $A$ | 0 |
| $B$ | 2 | $B$ | 1 |

| $s$ | $a$ | $R(s,a)$ |
|---|---|---|
| $B$ | 1 | 5 |
| $B$ | 2 | 0 |

We follow the steps of the Policy Iteration algorithm as explained in the class.

1. Write down the Bellman equation.
2. The initial policy is $\pi(A) = 1$ and $\pi(B) = 1$. That means that action 1 is taken when in state $A$, and the same action is taken when in state $B$ as well. Calculate the values $V_2^\pi(A)$ and $V_2^\pi(B)$ from two iterations of policy evaluation (Bellman equation) after initializing both $V_0^\pi(A)$ and $V_0^\pi(B)$ to 0.
3. Find an improved policy $\pi_{new}$ based on the calculated values $V_2^\pi(A)$ and $V_2^\pi(B)$.
4. This question is unrelated to the previous ones. Assume we have a stochastic policy $\pi$ where $\pi(s,a) = P(a|s)$ is equal to the probability of taking action $a$ when in state $s$. Write the equivalent of the Bellman equation for the value of this stochastic policy.

**Problem 3** (*40 points*):

Consider a two-bit register. The register has four possible states: 00, 01, 10 and 11. Initially, at time $t = 0$, the contents of the register is chosen at random to be one of these four states, each with equal probability. At each time step, the register is randomly manipulated as follows: with probability $1/2$, the register is left unchanged; with probability $1/4$, the two bits of the register are exchanged (01 becomes 10, 10 becomes 01, 00 remains 00, and 11 remains 11 ); and with probability $1/4$, the right bit is flipped (01 becomes 00, 00 becomes 01, 11 becomes 10, and 10 becomes 11). After the register has been manipulated in this fashion, the left bit is observed.
Suppose that on the first two time steps ($t = 1$ and $t = 2$), we observe the sequence $0, 0$. In other terms, the observed bit at time step $t = 1$ is 0, and the observed bit at time step $t = 2$ is also 0.

1. Show how the register can be formulated with a temporal model (a Hidden Markov Model): What is the probability of transitioning from every state to every other state? What is the probability of observing each output (0 or 1) in each state?

2. Use the filtering equation (forward algorithm) to determine the probability of being in each state at time $t$ after observing only the first $t$ bits, for $t = 1, 2$.

3. Use the smoothing equation to determine the probability of being in each state at time $t = 1$ given both observed bits.

4. Use the prediction equation to determine the probability of observing 1 in the next time step $t = 3$.

**Problem 4** (*10 points*):

Which ones of the statements below are **true** and which ones are **false**? You do NOT need to provide any explanation.

(a) If the only difference between two MDPs is the discount factor $\gamma$ then they must have the same optimal policy.

(b) For an MDP with a finite number of states and actions and with a discount factor $\gamma$ with $0 < \gamma < 1$, policy evaluation (the recursive updates of the value vector $V^\pi$ using the Bellman equation) is guaranteed to converge.

(c) The optimal action in a given state is the same regardless of the planning horizon $H$ (number of time-steps ahead).

(d) In the Q-learning algorithm, the transition function (transition probabilities between states) is not required to be known.

(e) In Bayesian networks, a variable is conditionally independent of all other variables in the network, given its parents and children.

(f) In Bayesian networks, a variable is conditionally independent of its non-descendants, given its parents.

(g) In Bayesian networks, a variable that has no parents is independent of its non-descendants.

(h) Rejection sampling wastes a lot of samples (all the samples that do not agree with the provided evidence). It can be improved by simply forcing the evidence variables to agree with the provided values and sampling only the non-evidence variables.

(i) The time and space complexity of exact inference in a poly-tree Bayesian network is linear in the number of conditional probability table entries.

(j) In a temporal model (Hidden Markov Model), the observed evidence is the cause of the hidden variable's value.