

# CPSC 322, Spring 2006

## Homework #5

---

### 1. [25 points] Variable Elimination

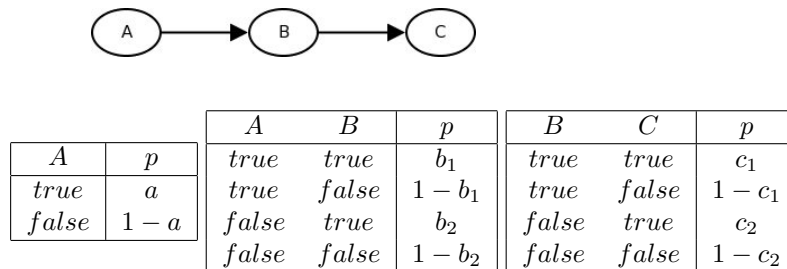


Figure 1: Simple belief network

Given the belief network in Figure 1, use variable elimination to find expressions for the following probabilities. Show any factors created in the process. You can re-use factors where appropriate. Because the factors contain variables instead of values, your answers will have to contain variables as well. This also means you'll need to solve these by hand, though you can check your answers by picking values for  $a, b_1, b_2, c_1, c_2$  and comparing the numbers you get with answers from CIspace.

- (a)  $p(B = \text{true} \wedge C = \text{true})$
- (b)  $p(C = \text{true} | A = \text{false})$
- (c)  $p(A = \text{true} | C = \text{true})$

## 2. [25 points] Decision Networks

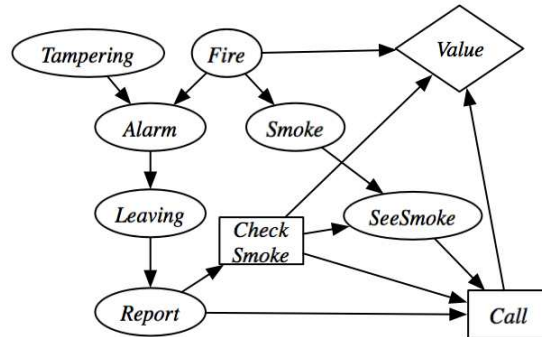


Figure 2: Decision network for the alarm problem (Figure 12.8 in the textbook)

Consider the decision network in Figure 2, which is one of the sample graphs included in CISpace under the name “Fire Alarm Decision Problem.” Using the CISpace Decision Network applet, you can find optimal policies and expected utilities for this problem.

- What is the optimal policy and its expected utility?
- Describe how changing  $p(\text{fire})$  affects the optimal policy and its expected utility. Give examples of values that produce different policies, explain why the policy changes.
- What is the value of control for *tampering*? How do you calculate this?
- What is the value of information for *alarm* (knowing it for both decisions)? How do you calculate this?

### 3. [25 points] Markov Decision Processes

Consider the following indefinite-horizon MDP: a robot is exploring a 4x3 maze, starting from the bottom-left corner. There are two terminal states on the right side, one with a reward of +100 and one with -100. This reward is constant regardless of the previous state and action. ( $\forall s \forall a, A4 = 100$ ,  $\forall s \forall a, B4 = -100$ ) All the other states have a reward of 0. The robot can try to move up, left, right or down. There is a 70% chance that the robot will actually go in its intended direction, and a 10% chance for each of the other directions. When the robot collides with a wall, it stays where it is and suffers no penalty. (See Figure 3.)

A	0	0	0	+100 (End)
B	0		0	-100 (End)
C	0 (Start)	0	0	0
	1	2	3	4

Figure 3:  $V_0^*(s)$

In Figures 4 and 5, fill in the value function for each state after the first and second cycles of value iteration. We initialize the value function  $V_0^*(s)$  to equal the reward for each state. For a discount factor, use  $\gamma = 0.75$ .

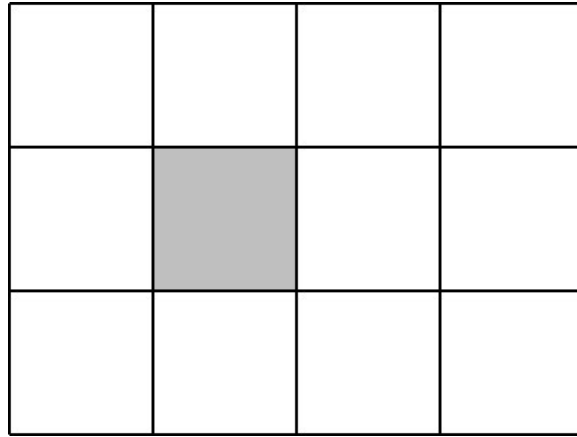


Figure 4:  $V_1^*(s)$

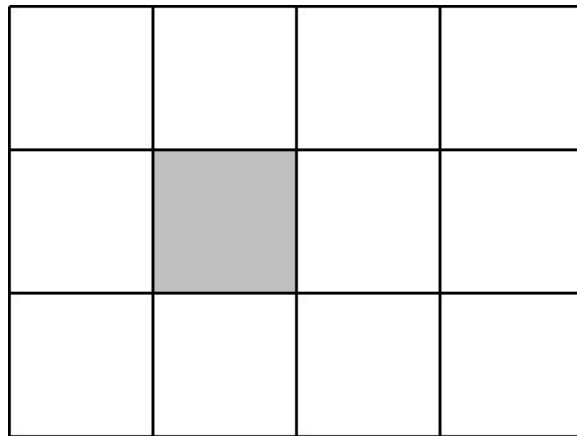


Figure 5:  $V_2^*(s)$

## 4. [25 points] Game Theory

Consider the following game:

		Player 2	
		L	R
Player 1	T	8,8	-12,2
	B	2,-12	4,4

where the first number in each square is the payoff of player 1 and the second number is player 2's payoff.

- Find all Pareto optimal pure strategy profiles.
- Find all pure strategy Nash equilibria.
- Suppose that player 2 plays L with probability  $p$ , and R with probability  $1 - p$ , and player 1 plays T and B with probabilities  $q$  and  $1 - q$ . Write expressions for the expected utility of player 1 for playing T and for playing B.
- Use these equations to derive a condition, in terms of  $p$ , which states when player 1 would play a mixed strategy that assigns positive probability to both T and B.
- Reasoning similarly, find a condition in terms of  $q$  under which player 2 would assign positive probability to playing both L and R.
- Find a mixed strategy Nash equilibrium of this game in which both players randomize, or prove that no such mixed strategy equilibrium exists.
- Which of the game's equilibria do you prefer? What would you play as player 1?