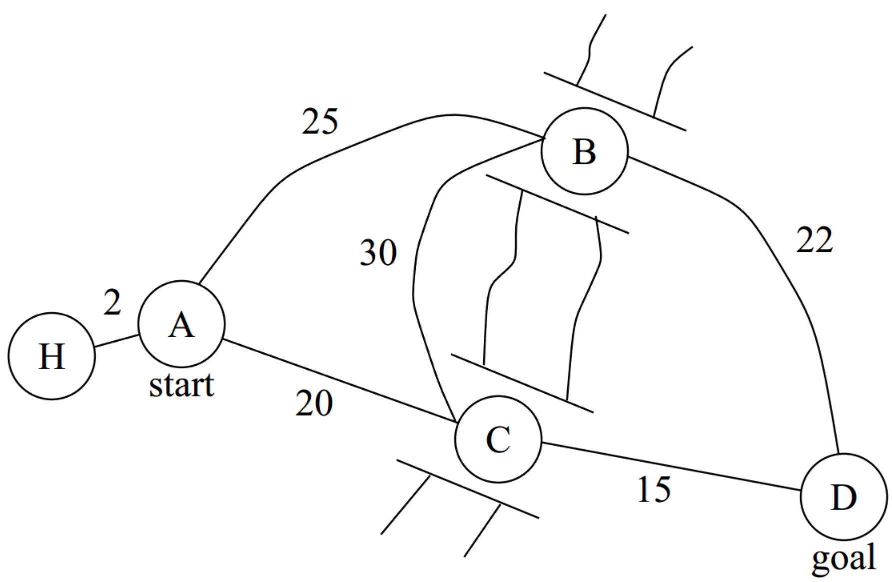
CS360 – Homework #12

Value of Information

1) A mobile robot is trying to get from his current position A to a destination D as quickly as possible. There is a river separating A from D and there are two bridges, B and C, spanning the river. The robot must design a strategy to move from A to D via one of the bridges.



Though the robot knows that one (and only one) of the bridges is inoperable, it is uncertain regarding which one of the two bridges is out. From its start position, position A, the robot can climb a hill to position H and use sensors to obtain information regarding which bridge is out. Collecting this information will take time, since it must go to H and return back to A, and the information gained is uncertain because the sensors will not be able to tell precisely which bridge is out.

The numbers in the above graph give the distance between locations in miles. In similar situations in the past, the robot experienced that 4 out of 5 times bridge C was out and only 1 out of 5 times bridge B was out. The robot has a short-range sensor that tells it with 100 percent reliability whether a bridge is out. The sensor can only be used when the robot is directly in front of the bridge. The long-range sensor of the robot is unreliable. It errs with a probability of 10 percent, that is, suggests that the broken bridge is operable and the other bridge is broken.

Design a strategy for the robot that minimizes the expected execution time.

Expectation Maximization

2) Similar to the example in our favorite textbook, assume that your favorite Surprise candy comes in two flavors: cherry and lime. The manufacturer has a peculiar sense of humor and wraps each piece of candy in the same opaque wrapper, regardless of flavor. Surprise candies are sold in very large bags, of which there are known to be two kinds - again, indistinguishable from the outside. A bag of Type 1 contains $p_1 \times 100$ percent cherry candies and $(1 - p_1) \times 100$ percent lime candies, while a bag of Type 2 contains $p_2 \times 100$ percent cherry candies and $(1 - p_2) \times 100$ percent lime candies. When buying a bag of Surprise candy, one gets a bag of Type 1 with probability p and a bag of Type 2 with probability $1 - p$.

You buy four bags of Surprise candy at different stores and sample two candies from each bag. The following table shows the results:

Bag	Candy
Bag 1	One Cherry candy and one Lime candy
Bag 2	One Lime candy and one Cherry candy
Bag 3	Two Lime candies
Bag 4	Two Lime candies

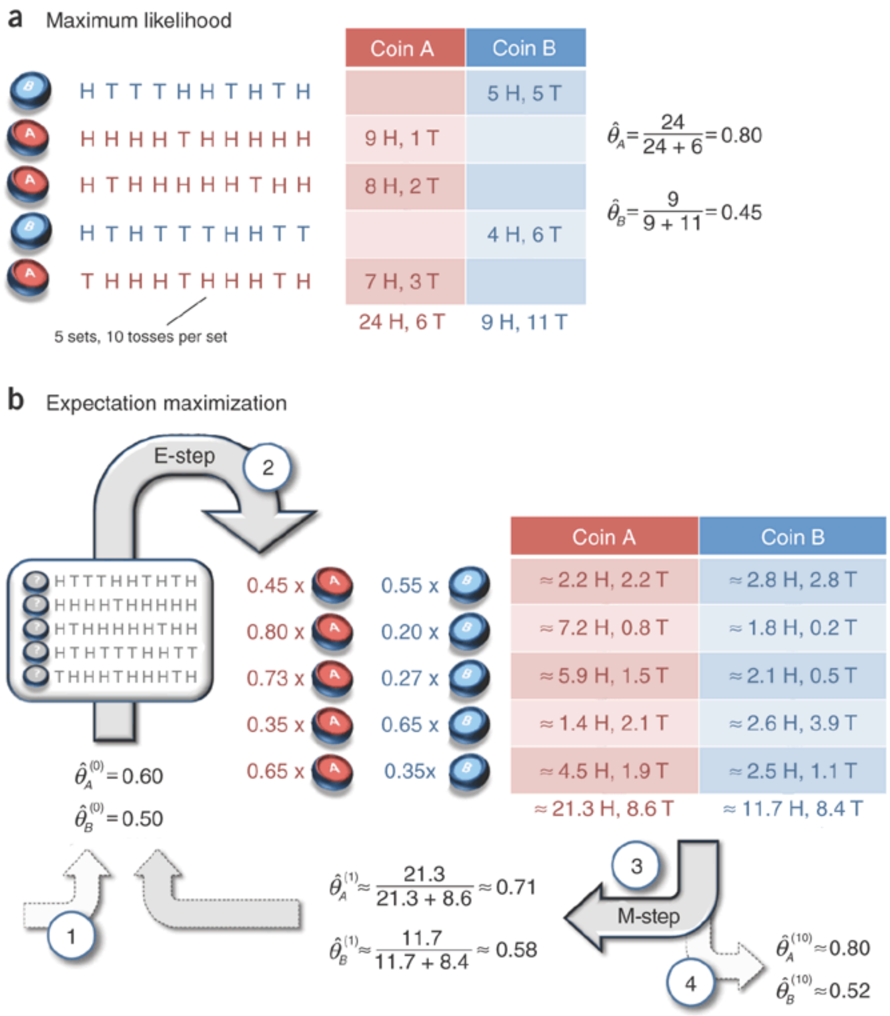
a) You learned that bags 1 and 2 are of Type 1 and bags 3 and 4 are of Type 2. You don't know the values of p, p_1 and p_2. Develop a maximum-likelihood algorithm (that uses frequencies to estimate the corresponding probabilities). What estimates of p, p_1 and p_2 does it compute?

b) You don't know the values of p, p_1 and p_2. Develop an expectation-maximization algorithm and run it until it converges, using $p = 0.1$, $p_1 = 0.2$ and $p_2 = 0.3$ as initial guesses. What estimates of p, p_1 and p_2 does it compute?

c) You know that $p_1 = 0.4$ but don't know the values of p and p_2. Develop an expectation-maximization algorithm and run it until it converges, using $p = 0.1$ and $p_2 = 0.3$ as initial guesses. What estimates of p, p_1 and p_2 does it compute?

3) Consider Figure 1 from "What is the expectation maximization algorithm" by Do and Batzoglou, which appeared in Nature Biotechnology, Volume 26, Number 8, August 2008, pages 897ff:

http://www.nature.com/nbt/journal/v26/n8/full/nbt1406.html



You can read only the following part of the text: "As an example, consider a simple coin-flipping experiment in which we are given a pair of coins A and B of unknown biases, Θ_A and Θ_B, respectively (that is, on any given flip, coin A will land on heads with probability Θ_A and tails with probability $1 - \Theta_A$ and similarly for coin B). Our goal is to estimate $\Theta = (\Theta_A, \Theta_B)$ by repeating the following procedure five times: randomly choose one of the two coins (with equal probability) and perform ten independent coin tosses with the selected coin. Thus, the entire procedure involves a total of 50 coin tosses (Fig. 1a)."

a) Explain the calculation that yields $\hat{\Theta}_A = 24/(24{+}6) = 0.80$ in Figure 1a. Is it important for this calculation that the coins were chosen with equal probability? Is this important anywhere in the maximum likelihood calculation?

b) Explain how you can verify in Figure 1b that the two coins were chosen with equal probability. In this context, also explain how to calculate the value 0.45 shown in Figure 1b.

c) Explain why Figure 1b states that "$0.45 \times A$ $0.55 \times B = 2.2H, 2.2T$ for Coin A and $= 2.8H, 2.8T$ for Coin B." What does this mean? In this context, also explain why $\hat{\Theta}_A^{(1)} = 21.3/(21.3 + 8.6)$.