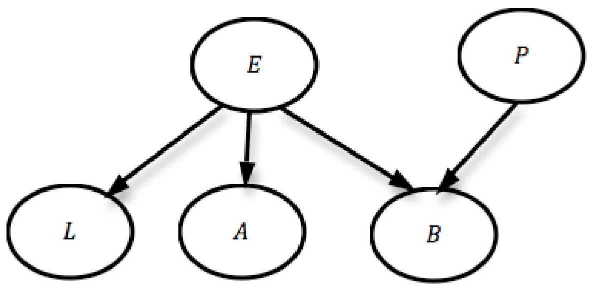
COMP 131 Midterm 2 – November 17, 2008 – Solutions

Question 1 – Bayes Nets – 30 points

In a mountainous town, a building collapsed. The town geography makes earthquakes more likely. We estimate that a mild earthquake can occur with probability .1. When an earthquake occurs, there is a 90% chance of an avalanche, and a 20% chance of a landslide. Avalanches also happen without earthquakes to cause them, about 30% of the time. Landslides don't happen without earthquakes. A building can collapse as a result of an earthquake, but also because of poor construction. In fact, we estimate that a poorly constructed building will collapse 80% of the time when even a mild earthquake occurs, and 10% of the time even if no earthquake is registered. A well-constructed building, however, has only a mall chance (probability .15) of collapsing during an earthquake, and virtually no chance (probability .001) otherwise. In general, we believe that the buildings in this town are well-constructed with probability .9.

1. (5 points) Draw a Bayes net that represents this problem, including all the relevant probability tables. Use the boolean random variables $E = Earthquake$, $L = Landslide$, $A = Avalanche$, $B = BuildingCollapse$ and $P = PoorConstruction$.



$P(E = e) = .1$
$P(P = p) = .1$

E	$P(L = l \mid E)$
e	.2
$\neg e$	0

E	$P(A = a \mid E)$
e	.9
$\neg e$	.3

P	E	$P(B = b \mid P, E)$
p	e	.8
p	$\neg e$	.1
$\neg p$	e	.15
$\neg p$	$\neg e$	.001

2. (10 points) A building collapsed, and we know that there were no landslides. What is the probability that the building was poorly constructed?

Answer: We need to calculate $P(P = p|b, \neg l, E, A)$. To do that we remember that $P(X|e) = \alpha \sum_y P(X, e, y)$, so we will sum over all values of E and A. We also remember that the full joint probability of a Bayes net is just the product of all CPTs and priors. This gives us:

$$\begin{aligned} P(P = p|b, \neg l, E, A) &= \alpha \sum_{e,a} P(P = p, b, \neg l, e, a) = \alpha P(p) \sum_e P(e)P(\neg l|e)P(b|e, p) \sum_a P(a|e) \\ &= \alpha \times .1 \times \sum_e \begin{matrix} e = true \\ e = false \end{matrix} \begin{pmatrix} .1 \times .8 \times .8 \\ .9 \times 1.0 \times .1 \end{pmatrix} \times \langle 1,\ 1 \rangle \\ &= \alpha \times .1 \times (.064 + .09) = \alpha \times .0154 \end{aligned}$$

To get the numeric answer, find also $P(P = \neg p|b, \neg l, E, A) = \alpha \times .01161$, so $P(P|b, \neg l, E, A) = \langle .57,\ .43 \rangle$.

3. (5 points) Consider a simpler problem, with only the variables E, L, A and B. Suppose we don't know anything about the probabilities in this network, but we do have a data set with examples of co-occurring events. In the past 100 years, there have been 10 earthquakes, 6 of which co-occurred with avalanches, and one with a landslide. Once a building had collapsed when an earthquake occurred. Avalanches were registered during 42 out of the 100 years of data. (Assume one data point per year.)

   Use this data set to learn a naive Bayes model. How many parameters do you need to estimate with maximum-likelihood learning? What assumption are you making?

   Answer: A naive Bayes model assumes conditionally independent effects given a cause. Here, the cause variable is E (earthquake) and the effects are L, A, and B. In general, you need $2N + 1$ parameters for N effects: one estimating the prior $P(Cause)$, and then two per effect: $P(Effect_i|cause)$ and $P(Effect_i|\neg cause)$. Here, we need to estimate 7 parameters.

   The assumption we make by using maximum-likelihood learning is that of a uniform prior over the hypothesis space (that is, no sets of parameters are a priori any likelier than others).

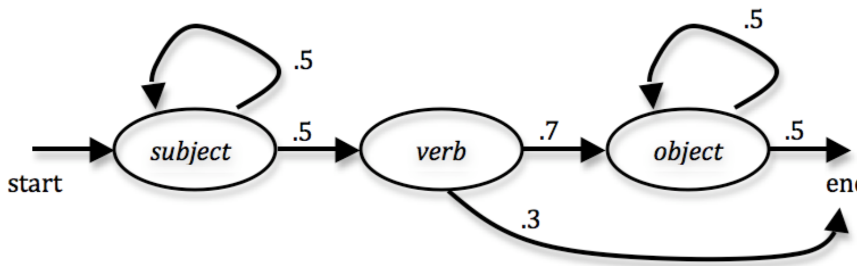
4. (10 points) Give the maximum likelihood estimates for all of this model's parameters.

   Answer: Maximum likelihood estimates are just frequencies of occurrence:

$$
\begin{aligned}
\theta = \hat{P}(E = e) &= \frac{\#\ earthquakes}{total\ \#\ observations} = \frac{10}{100} = .1 \\
\theta_{11} = \hat{P}(L = l|e) &= \frac{\#\ landslides\ during\ earthquakes}{\#\ earthquakes} = \frac{1}{10} = .1 \\
\theta_{12} = \hat{P}(L = l|\neg e) &= \frac{\#\ landslides\ with\ no\ earthquake}{\#\ observations\ with\ no\ earthquake} = \frac{0}{90} = 0 \\
\theta_{21} = \hat{P}(A = a|e) &= \frac{\#\ avalanches\ during\ earthquakes}{\#\ earthquakes} = \frac{6}{10} = .6 \\
\theta_{22} = \hat{P}(A = a|\neg e) &= \frac{42 - 6}{90} = .4 \\
\theta_{31} = \hat{P}(B = b|e) &= \frac{1}{10} = .1 \\
\theta_{32} = \hat{P}(B = b|\neg e) &= \frac{0}{90} = 0
\end{aligned}
$$

Question 2 – Temporal Reasoning with Uncertainty – 25 points

Consider this very simple model of a grammatical structure of a sentence of English. A sentence always starts with a subject phrase, which can be made of one or several words, at some point transitions to a verb. The sentence can stop there (with probability .3) but is more likely to transition to an object phrase, which has the same possible composition as the subject phrase. The figure below gives a compact HMM representation of this model, with all transition and output (aka sensor) conditional probabilities specified. (This figure is reproduced on p. 10, the tear-off sheet for your convenience.)



output word probabilities:	*subject*	*verb*	*object*
boat	.2	0	.2
is	0	.5	0
man	.2	.1	.2
old	.2	0	.2
rows	.1	.4	.1
the	.3	0	.3

The output probabilities take into account that some words (like *man* and *rows*) can be both nouns (and so part of the subject and object phrases) and verbs. We observe the following beginning of a sentence "The old man...".

1. (5 points) What is the hidden variable and the observed variable (evidence) in this HMM? What are their respective domains?

   Answer: The hidden variable X loosely represents parts of speech, and takes values from the set $\{subject,\ verb,\ object\}$. The evidence variable E represents the word generated and observed and takes values from the set $\{boat,\ is,\ man,\ old,\ rows,\ the\}$.

2. (10 points) What word is most likely to appear next in this sentence, given that we've observed "The old man...". (Hint: filter to the most likely next state first)

   Answer: First, we compute the probability of X_4 (the state that comes after "The old man") using filtering and one-step prediction: $P(X_4|e_1...e_3) = ...$

$$\begin{aligned}\mathbf{f}_{1:1} &= \alpha P(the|X_1)P(X_1|start) = \alpha\langle .3\times 1,\ 0\times 0,\ .3\times 0\rangle = \alpha\langle .3,\ 0,\ 0\rangle = \langle 1,\ 0,\ 0\rangle \\ \mathbf{f}_{1:2} &= \alpha P(old|X_2)\sum_{x_1}(P(X_2|x_1)\mathbf{f}_{1:1})\end{aligned}$$

$$
\begin{aligned}
&= \alpha P(old|X_2) \times \langle .3 \times .5 + 0 \times 0 + 0 \times 0,\ .3 \times .5 + 0 \times 0 + 0 \times 0,\ .3 \times 0 + 0 \times .7 + 0 \times .5\rangle \\
&= \alpha\langle .2,\ 0,\ .2\rangle \times \langle .15,\ .15,\ 0\rangle = \langle 1,\ 0,\ 0\rangle \\
\mathbf{f}_{1:3} &= \alpha P(man|X_3) \sum_{x_2} (P(X_3|x_2)\mathbf{f}_{1:2}) \\
&= \alpha\langle .2,\ .1\ .2\rangle \times \langle .5,\ .5,\ 0\rangle = \alpha\langle .1,\ .05,\ 0\rangle = \langle .67,\ .33,\ 0\rangle
\end{aligned}
$$

Note: if you figure out the correct probabilities without writing out the algorithmic steps in equations, simply by describing them in English (correctly), you get full marks.

$f_{1:3}$ gives the probabilities of X_3 given the evidence "the old man". To get the prediction $P(X_4|the, old, man)$ we sum over the transition probabilities from X_3 to X_4:

$$
\begin{aligned}
P(X_4|the, old, man) &= \alpha \sum_{x_3} P(X_4|x_3)P(x_3|the, old, man) = \sum_{x_3} P(X_4|x_3)\mathbf{f}_{1:3} \\
&= \alpha\langle .67 \times .5 + .33 \times 0 + 0 \times 0,\ .67 \times .5 + .33 \times 0 + 0 \times 0,\ .67 \times 0 + .33 \times .7 + 0 \times .5\rangle \\
&= \alpha\langle .33,\ .33,\ .23\rangle = \langle .37,\ .37,\ .26\rangle
\end{aligned}
$$

Finally, to find the most likely word, we sum over X_4 the observation model:

$$
\begin{aligned}
P(E_4|X_4, the, old, man) &= \alpha \quad \sum_{x_4} P(E_4|x_4)P(x_4|the, old, man) \\
&= \alpha\langle \quad .2 \times .37 + 0 \times .37 + .2 \times .26, \\
&\qquad 0 \times .37 + .5 \times .37 + 0 \times .26, \\
&\qquad .2 \times .37 + .1 \times .37 + .2 \times .26, \\
&\qquad .2 \times .37 + 0 \times .37 + .2 \times .26, \\
&\qquad .1 \times .37 + .4 \times .37 + .1 \times .26, \\
&\qquad .3 \times .37 + 0 \times .37 + .3 \times .26\rangle \\
&= \alpha \quad \langle .126,\ .185,\ .163,\ .126,\ .237,\ .189\rangle
\end{aligned}
$$

The most likely next word is "rows".

3. (10 points) The next two words we observe are actually "the boat" (the whole word sequence is: "The old man the boat"). What is now the most likely value of the hidden variable that generated the word "man"? Assume that the sentence ends after "boat".

Answer: The question is one of smoothing or updating our belief in past states given more recent evidence. In order to compute that we need to find:

$$P(X_3|the, old, man, the, boat) \quad = \quad \alpha\, \mathbf{f}_{1:3}\mathbf{b}_{4:5}$$

We already know the solution to the forward part of this calculation, so we only need the backward part. In general:

$$\mathbf{b}_{k:t} = P(e_{k:t}|X_{k-1}) = \sum_{x_k} P(e_k|x_k)P(x_k|X_{k_1})\mathbf{b}_{k+1:t}$$

In our case, notice that the could be a transition to the sentence end either from $X_5 = verb$ or from $X_5 = object$, so we need to incorporate that into our backward message.

$$\begin{aligned}
\mathbf{b}_{5:5} &= \sum_{x_5} P(boat|x_5)P(x_5|X_4)P(end|x_5) \\
&= \sum_{x_5} \begin{pmatrix} .2 \times \langle .5,\ 0,\ 0\rangle \times 0 \\ 0 \times \langle .5,\ 0,\ 0\rangle \times .3 \\ .2 \times \langle 0,\ .7,\ .5\rangle \times .5 \end{pmatrix} = 0 + 0 + \langle 0,\ .07,\ .05\rangle = \langle 0,\ .07,\ .05\rangle \\
\mathbf{b}_{4:5} &= \sum_{x_4} P(the|x_4)P(x_4|X_3)\mathbf{b}_{5:5} \\
&= \sum_{x_4} \begin{pmatrix} .3 \times \langle .5,\ 0,\ 0\rangle \times 0 \\ 0 \times \langle .5,\ 0,\ 0\rangle \times .07 \\ .3 \times \langle 0,\ .7,\ .5\rangle \times .05 \end{pmatrix} = 0 + 0 + \langle 0,\ .0105,\ .0075\rangle = \langle 0,\ .07,\ .05\rangle
\end{aligned}$$

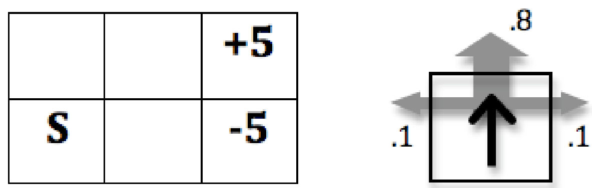
Which gives our required probabilities:

$$\begin{aligned}
P(X_3|the, old, man, the, boat) &= \alpha\, \mathbf{f}_{1:3}\mathbf{b}_{4:5} \\
&= \alpha\langle .1,\ .05,\ 0\rangle\langle 0,\ .07,\ .05\rangle = \langle 0,\ 1,\ 0\rangle
\end{aligned}$$

The word “man” has to have been a verb.

Question 3 – MDPs and Reinforcement Learning – 30 points

This gridworld MDP operates similarly to the one we saw in class. The states are grid squares, identified by their row and column number (row first). The agent always starts in state (1,1), marked with the letter S. There are two terminal goal states, (2,3) with reward +5 and (1,3) with reward -5. Rewards are 0 in non-terminal states. The transition function is such that the intended agent movement happens with probability .8. With probability .1 each, the agent ends up in one of the states at right angles to the intended direction. If a collision with a wall happens, no motion is executed (the agent stays in the same state).



(a) (b)

Figure 1: (a) Gridworld MDP. (b) Transition function.

1. (5 points) Intuitively, what should the optimal policy be in this world?

   Answer: The agent should go up when in state S (1,1), should go right when in one of the upper-row states (2,1) or (2,2), and should go left when in state (1,2).

2. (10 points) Suppose the agent knows the transition probabilities. Give the first three updates of value iteration for each state.

   Answer: Assume $U^0(s)$ for all s is initialized to 0 and $\gamma = 1$ (as it was not specified). The general value iteration update rule is:

   $$U^{i+1}(s) = R(s) + \gamma max_a \sum s' T(s,a,s') U^i(s)$$

   During the first iterations, since initial values are zero, and so is reward in non-terminal states, only the values for states (1,3) and (2,3) are updated:

   $$\begin{aligned} U^1(1,3) &= +5 \\ U^1(2,3) &= -5 \end{aligned}$$

   On the second iteration, these values will propagate to adjacent states (actions are considered clockwise from "up"):

   $$\begin{aligned} U^2(1,2) &= 0 + max_a(.8 \times 0 + .1 \times (-5) + .1 \times 0,\ .8 \times (-5) + .2 \times 0,\ .8 \times 0 + .1 \times (-5) + .1 \times 0,\ 0) = 0 \\ U^2(2,2) &= 0 + max_a(.8 \times 0 + .1 \times 5 + .1 \times 0,\ .8 \times 5 + .2 \times 0,\ .8 \times 0 + .1 \times 5 + .1 \times 0,\ 0) = 4 \end{aligned}$$

   The other values remain the same. On the third iteration, these values again will propagate to adjacent states ($U^3(1,1)$ remains zero):

   $$U^3(1,2) = 0 + max_a(.8 \times 4 + .1 \times (-5),\ .8 \times (-5) + .1 \times 4,\ .1 \times (-5),\ .1 \times 4)$$

$$
\begin{aligned}
&= max(2.7, -3.6, .5, .4) = 2.7 \\
U^3(2,1) &= = 0 + max_a(.1 \times 4,\ .8 \times 4, .1 \times 4, 0) = 3.2 \\
U^3(2,2) &= 0 + max_a(.8 \times 4 + .1 \times 5,\ .8 \times 5 + .1 \times 4 + .1 \times 0,\ .1 \times 5,\ .1 \times 4) = 4.4
\end{aligned}
$$

The utility values of (1,3) and (2,3) remain the same.

3. (5 points) Suppose the agent does not know the transition probabilities. What does it need to be able do (or have available) in order to learn the optimal policy?

   Answer: Without knowing the transition probabilities, the agent needs to be able to sample the MDP instead. This is done by acting in the environment (or having rial trajectories available). So the agent needs to be able to execute some policy in the environment and update estimates of state utilities, using adaptive dynamic programming or temporal-difference learning.

4. (5 points) The agent starts with the policy that always chooses to go right, and executes the following three trials: 1) (1,1)–(1,2)–(1,3), 2) (1,1)–(1,2)–(2,2)–(2,3), and 3) (1,1)–(2,1)–(2,2)–(2,3). What are the direct utility estimates of states at the end of the first trial?

   Answer: At the end of the first trial, the agent has seen one transition from (1,1) to (1,2) and one transition from (1,2) to (1,3). The terminal state (1,3) has the reward -5, so $U^\pi(1,3) = -5$ and since the reward are 0 in non-terminal states, so estimating these utilities directly gives $U^\pi(1,1) = U^\pi(1,2) = -5$

5. (5 points) Using a learning rate of .1, what updates does the TD-learning agent make after trials 2 and 3?

   Answer: The general TD-learning update is:

$$U^\pi(s) \leftarrow U^\pi(s) + \alpha(R(s) + \gamma U^\pi(s') - U^\pi(s))$$

After trial 2, the TD agent will update the visited states, that is

$$
\begin{aligned}
U^\pi(2,3) &= 5 \\
U^\pi(1,1) &= -5 + .1\,(0 + 1 \times (-5) + 5) = -5 \\
U^\pi(1,2) &= -5 + .1\,(0 + 1 \times 0 - (-5)) = -4.5 \\
U^\pi(2,2) &= 0 + .1\,(0 + 1 \times 5 - 0) = .5
\end{aligned}
$$

After the third trial, the TD-learning agent makes non-zero updates to all the visited states:

$$
\begin{aligned}
U^\pi(2,3) &= 5 \\
U^\pi(1,1) &= -5 + .1\,(0 + 1 \times 0 + 5) = -0.45 \\
U^\pi(2,1) &= 0 + .1\,(0 + 1 \times .5 - 0) = .05 \\
U^\pi(2,2) &= .5 + .1\,(0 + 1 \times 5 - .5) = .95
\end{aligned}
$$

Question 4 – Statistical Machine Learning – 15 points

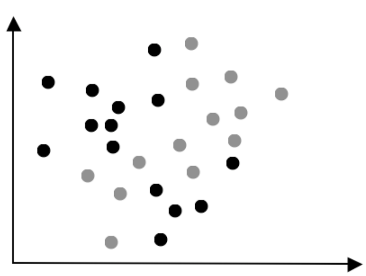
1. (5 points) In your own words: what is a hypothesis space? What are the important considerations in choosing a hypothesis space? What does this choice affect?

   Answer: A hypothesis space is the class of hypothesis functions from which the learning agent will choose one. For example, "polynomials of degree at most 3" is a hypothesis space. The important considerations when choosing a hypothesis space are its expressiveness and potential for overfitting. For example, a hypothesis space consisting only of linear separators will not contain any hypothesis that could classify XOR data correctly (it will not be expressive enough). On the other hand, a hypothesis space that is too expressive can result in overfitting to the training data set, where the learner fits its hypothesis function to the noise in the data – a situation that will result in poor generalization to new data. The choice of the hypothesis space will affect the ability of the learning agent to find a consistent hypothesis and to generalize.

2. (5 points) A Bayesian learner makes classification decisions based on all possible hypotheses, which is optimal. Why would you ever want to use MAP instead? Explain specifically.

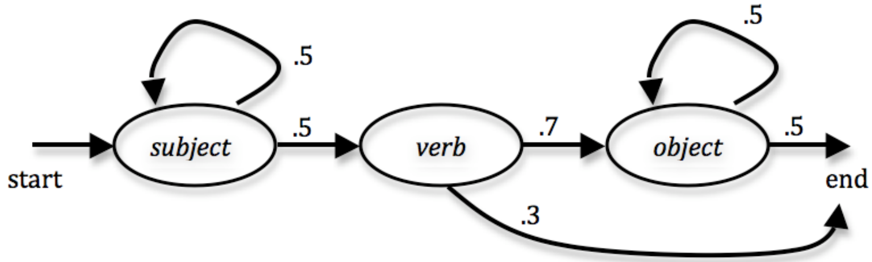
   Answer: You use MAP learning when the hypothesis space is very large or infinite, and it is not practical to sum over (or integrate over) the entire hypothesis space.

3. (5 points) Can a perceptron be trained to perform this classification problem? Why or why not? Be specific.



Answer: No, it cannot, because the data is not linearly separable.

Figure 2: Sentence HMM with conditional probabilities for question 2 (tear-off).



output word probabilities:	*subject*	*verb*	*object*
boat	.2	0	.2
is	0	.5	0
man	.2	.1	.2
old	.2	0	.2
rows	.1	.4	.1
the	.3	0	.3