CS 188   Introduction to
Spring 2014   Artificial Intelligence
Written HW6

INSTRUCTIONS

- **Due:** Monday, March 17th 2014 11:59 PM

- **Policy:** Can be solved in groups (acknowledge collaborators) but must be written up individually. However, we strongly encourage you to first work alone for about 30 minutes total in order to simulate an exam environment. Late homework will not be accepted.

- **Format:** You must solve the questions on this handout (either through a pdf annotator, or by printing, then scanning; we recommend the latter to match exam setting). Alternatively, you can typeset a pdf on your own that has answers appearing in the same space (check edx/piazza for latex templating files and instructions). **Make sure that your answers (typed or handwritten) are within the dedicated regions for each question/part. If you do not follow this format, we may deduct points.**

- **How to submit:** Go to www.pandagrader.com. Log in and click on the class CS188 Spring 2014. Click on the submission titled Written HW 6 and upload your pdf containing your answers. If this is your first time using pandagrader, you will have to set your password before logging in the first time. To do so, click on "Forgot your password" on the login page, and enter your email address on file with the registrar's office (usually your @berkeley.edu email address). You will then receive an email with a link to reset your password.

Last Name	Chen
First Name	Jianzhong
SID	23478230
Email	chenjianzhong@berkeley.edu
Collaborators	None

For staff use only

Q. 1	Q. 2	Total
/12	/18	/30

Q1. [12 pts] Probability: Chain Rule Madness

Note that the set of choices is the same for all of the following questions. You may find it beneficial to read all the questions first, and to make a table on scratch paper.

(a) [3 pts] Fill in the circles of **all** expressions that are equivalent to $\mathbf{P(A, B \mid C)}$, **given no independence assumptions**:

○ $\frac{P(C|A)\ P(A|B)\ P(B)}{P(C)}$

○ $\frac{P(A|C)\ P(C|B)\ P(B)}{P(B,C)}$

○ $\frac{P(C|A,B)\ P(B|A)\ P(A)}{P(B|C)\ P(C)}$

○ $\frac{P(B,C|A)\ P(A)}{P(B,C)}$

○ $\frac{P(A|C)\ P(B,C)}{P(C)}$

● $P(A \mid B, C)\ P(B \mid C)$

○ None of the above.

(b) [3 pts] Fill in the circles of **all** expressions that are equivalent to $\mathbf{P(A, B \mid C)}$, **given that $\mathbf{A \perp\!\!\!\perp B \mid C}$:**

○ $\frac{P(C|A)\ P(A|B)\ P(B)}{P(C)}$

○ $\frac{P(A|C)\ P(C|B)\ P(B)}{P(B,C)}$

○ $\frac{P(C|A,B)\ P(B|A)\ P(A)}{P(B|C)\ P(C)}$

○ $\frac{P(B,C|A)\ P(A)}{P(B,C)}$

● $\frac{P(A|C)\ P(B,C)}{P(C)}$

● $P(A \mid B, C)\ P(B \mid C)$

○ None of the above.

(c) [3 pts] Fill in the circles of **all** expressions that are equivalent to $\mathbf{P(A \mid B, C)}$, **given no independence assumptions**:

○ $\frac{P(C|A)\ P(A|B)\ P(B)}{P(C)}$

○ $\frac{P(A|C)\ P(C|B)\ P(B)}{P(B,C)}$

● $\frac{P(C|A,B)\ P(B|A)\ P(A)}{P(B|C)\ P(C)}$

● $\frac{P(B,C|A)\ P(A)}{P(B,C)}$

○ $\frac{P(A|C)\ P(B,C)}{P(C)}$

○ $P(A \mid B, C)\ P(B \mid C)$

○ None of the above.

(d) [3 pts] Fill in the circles of **all** expressions that are equivalent to $\mathbf{P(A \mid B, C)}$, **given that $\mathbf{A \perp\!\!\!\perp B \mid C}$:**

○ $\frac{P(C|A)\ P(A|B)\ P(B)}{P(C)}$

● $\frac{P(A|C)\ P(C|B)\ P(B)}{P(B,C)}$

● $\frac{P(C|A,B)\ P(B|A)\ P(A)}{P(B|C)\ P(C)}$

● $\frac{P(B,C|A)\ P(A)}{P(B,C)}$

○ $\frac{P(A|C)\ P(B,C)}{P(C)}$

○ $P(A \mid B, C)\ P(B \mid C)$

○ None of the above.

Q2. [18 pts] Extending the Forward Algorithm

Consider the HMM graph structure shown below. The joint distribution is listed to the right of the graph.
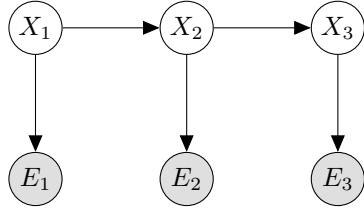
$$P(X_1, E_1, X_2, E_2, ..., X_T, E_T) = P(X_1)P(E_1|X_1) \prod_{t=2}^{T} P(X_t|X_{t-1})P(E_t|X_t)$$

Recall the Forward algorithm is a two step iterative algorithm used to approximate the probability distribution $P(X_t|e_1, \ldots, e_t)$. The two steps of the algorithm are as follows:
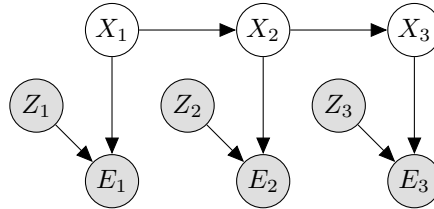
Elapse Time $P(X_t|e_1, \ldots, e_{t-1}) = \sum_{x_{t-1}} P(X_t|x_{t-1})P(x_{t-1}|e_1, \ldots, e_{t-1})$

Observe $P(X_t|e_1, \ldots, e_t) = \frac{P(e_t|X_t)P(X_t|e_1, \ldots, e_{t-1})}{\sum_{x_t} P(e_t|x_t)P(x_t|e_1, \ldots, e_{t-1})}$

For this problem we will consider modifying the forward algorithm as the HMM graph structure changes. Our goal will continue to be to create an iterative algorithm which is able to compute the distribution of states, X_t, given all available evidence from time 0 to time t.

Note: If the probabilities required can be computed without *any* change to original update equations, write **NO CHANGE**.

Consider the graph below where new observed variables, Z_i, are introduced and influence the evidence. The joint distribution is listed to the below the graph.

$$P(X_1, Z_1, E_1, ..., X_T, Z_T, E_T) = P(X_1)P(Z_1)P(E_1|X_1, Z_1) \prod_{t=2}^{T} P(X_t|X_{t-1})P(Z_t)P(E_t|X_t, Z_t)$$
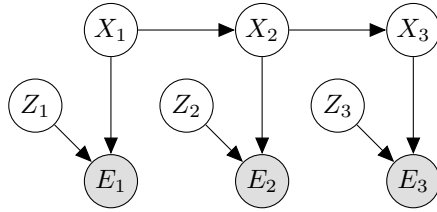
(a) [3 pts] State the modified Elapse Time update.

$P(X_t|e_1, \ldots, e_{t-1}, z_1, \ldots, z_{t-1}) = \sum_{x_{t-1}} P(X_t|x_{t-1})P(x_{t-1}|e_1, ..., e_{t-1}, z_1, ..., z_{t-1})$

(b) [3 pts] State the modified Observe update.

$$P(X_t|e_1,\ldots,e_t,z_1,\ldots,z_t) = \frac{P(e_t|z_t,X_t)P(X_t|e_1,\ldots,e_{t-1},z_1,\ldots,z_{t-1})}{\sum_{x_t}P(e_t|x_t,z_t)P(x_t|e_1,\ldots,e_{t-1},z_1,\ldots,z_{t-1})}$$

Next, consider the graph below where the Z_i variables are unobserved. The joint distribution is listed to the below the graph.

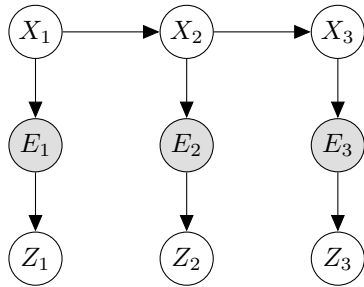$$P(X_1,Z_1,E_1,\ldots,X_T,Z_T,E_T) = P(X_1)P(Z_1)P(E_1|X_1,Z_1)\prod_{t=2}^{T}P(X_t|X_{t-1})P(Z_t)P(E_t|X_t,Z_t)$$

(c) [3 pts] State the modified Elapse Time update.

$$P(X_t|e_1,\ldots,e_{t-1}) = \sum_{x_{t-1}}P(X_t|x_{t-1})P(x_{t-1}|e_1,\ldots,e_{t-1})$$

(d) [3 pts] State the modified Observe update.

$$P(X_t|e_1,\ldots,e_t) = \frac{\sum_{x_t}P(z_t)P(e_t|z_t,X_t)P(X_t|e_1,\ldots,e_{t-1})}{\sum_{x_t}P(x_t|e_1,\ldots,e_{t-1})\sum_{x_t}P(z_t)P(e_t|z_t,x_t)}$$

Finally, consider a graph where the newly introduced variables are unobserved and influenced by the evidence nodes. The joint distribution is listed to the right of the graph.

$$P(X_1, Z_1, E_1, ..., X_T, Z_T, E_T) = P(X_1)P(E_1|X_1)P(Z_1|E_1)\prod_{t=2}^{T} P(X_t|X_{t-1})P(E_t|X_t)P(Z_t|E_t)$$

(e) [3 pts] State the modified Elapse Time update.

$$P(X_t|e_1, \ldots, e_{t-1}) = \sum_{x_{t-1}} P(X_t|x_{t-1})P(x_{t-1}|e_1, ..., e_{t-1})$$

(f) [3 pts] State the modified Observe update.

$$P(X_t|e_1, \ldots, e_t) = \frac{P(e_t|X_t)P(X_t|e_1, ..., e_{t-1})}{\sum_{x_t} P(e_t|x_t)P(x_t|e_1, ..., e_{t-1})}$$