KOÇ UNIVERSITY

COLLEGE OF ENGINEERING

COMP 341: INTRODUCTION TO ARTIFICIAL INTELLIGENCE

Midterm Examination 2

FALL 2016, 05/12/2016

DURATION: 120 MINUTES

Name: Solutions

ID: 00110001 00110000 00110000

- This exam contains 9 pages including this cover page and 7 questions. Check to see if any pages are missing. Put your initials on the top of every page, in case the pages become separated.
- By submitting this exam, you **agree** to fully comply with Koç University Student Code of Conduct, and accept any punishment in case of failure to comply. If you do not submit this exam on time, you will receive 0 credits.
- The exam is **closed book** and **closed notes**. You are **not** allowed to use any additional material including any electronic equipment such as computers and mobile phones.
- You are expected to be able provide clear and concise answers. Gibberish will not receive any credit.
- Your answers need to be readable by a human, illegible writing is not gradable hence there is a strong chance that such answers will not get any credit.
- Read each question carefully and make sure to follow instructions. The written instructions take precedence over an answer that the instructor might give you during the exam, unless the instructor makes a class wide announcement.
- Do not write in the table below.

Question:	1	2	3	4	5	6	7	Total
Points:	16	16	10	13	10	15	25	105
Score:								

1. (16 points) True or False :
   KB: A Logical Knowledge Base formed with a set of logic sentences
   $M(L)$: Set of all models where the set of logic sentences, L, is true
   $\models$: Entails operator
   PL: Propositional Logic, FOL: First order Logic
   BN: Bayesian Network, MC: Markov Chain, HMM: Hidden Markov Model

   False $KB \models \alpha$ if and only if $M(\alpha) \subseteq M(KB)$

   True Only one pair at a time can be resolved with the Resolution algorithm for PL

   False Inference algorithms for FOL always terminate

   True There are ways to represent uncertainty other than probability

   False BNs always represent causality

   True BNs: Prior sampling and Gibbs sampling can be used calculate the same thing

   True All MCs with a fixed transition model have a stationary state distribution

   True HMMs can be used to estimate the state path of the underlying MC given evidence

2. (16 points) Propositional Logic
   (a) (6 points) Convert the following into CNF, show all your work and simplify the result.
      1. $(X \Rightarrow Z) \wedge (Y \Rightarrow Z)$
         $= (\neg X \vee Z) \wedge (\neg Y \vee Z)$

      2. $(X \vee Y) \Rightarrow Z$
         $= \neg(X \vee Y) \vee Z$
         $= (\neg X \wedge \neg Y) \vee Z$
         $= (\neg X \vee Z) \wedge (\neg Y \vee Z)$

      3. $(P \vee R) \Leftrightarrow ((R \wedge P) \Rightarrow \neg Q)$
         $= ((P \vee R) \Rightarrow ((R \wedge P) \Rightarrow \neg Q)) \wedge (((R \wedge P) \Rightarrow \neg Q) \Rightarrow (P \vee R))$
         $= (\neg(P \vee R) \vee (\neg(R \wedge P) \vee \neg Q)) \wedge (\neg(\neg(R \wedge P) \vee \neg Q) \vee (P \vee R))$
         $= ((\neg P \wedge \neg R) \vee (\neg R \vee \neg P \vee \neg Q)) \wedge ((R \wedge P \wedge Q) \vee (P \vee R))$
         $= (\neg P \vee \neg R \vee \neg Q) \wedge (\neg P \vee \neg R \vee \neg Q) \wedge (R \vee P) \wedge (P \vee R) \wedge (Q \vee P \vee R)$
         $= (\neg P \vee \neg R \vee \neg Q) \wedge (P \vee R)$

(b) (10 points) Show whether C is entailed by the following knowledge base using **resolution**. Generate new sentences by referring to the existing ones and writing them out, e.g., given, {S1, ..., Sn}, then "Sn+1: Resolve Si and Sj to get ...". Make sure you write out any additional information that is needed, e.g., to clarify any ambiguities. You are allowed to have *smart* picks. You do not have to follow any particular order and can pick or ignore any sentences.

S1: $A \vee \neg B \vee C \vee F$
S2: $D \vee B$
S3: $\neg D \vee A$
S4: $E \vee \neg F$
S5: $\neg A \vee D$
S6: B
S7: $\neg D$
S8: $\neg E$

One possible solution:
S9: Add the negation of the query variable $\neg C$
S10: S4 and S8 to get $\neg F$
S11: S1 and S10 to get $A \vee \neg B \vee C$
S12: S6 and S11 to get $A \vee C$
S13: S5 and S12 to get $C \vee D$
S14: S7 and S13 to get C
S15: S9 and S14 to get $\emptyset$

Thus we conclude that C is entailed by the knowledge base.

3. (10 points) Consider a vocabulary with the following symbols:
Angry(x): Person x is angry
Rich(x): Person x is rich
Occupation(x,y): Person x has occupation y
Boss(x,y): Person x is a boss of person y
Murder(x,y): Person x murdered y
InLove(x,y): Person x loves y
Married(x,y): Person x is married to y
Doctor, Butler: Constants denoting occupations
Jane, Alfred: Constants denoting people

Use these symbols to write the following assertions in first order logic:

- All doctors are rich or angry
  $\forall x Doctor(x) \Rightarrow (Rich(x) \lor Angry(x))$

- Alfred is a butler and his boss is a doctor
  $Occupation(Alfred, Butler) \land \exists x Boss(x, Alfred) \land Occupation(x, Doctor)$

- Jane loves a doctor who is not in love with Jane
  $\exists x InLove(Jane, x) \land Occupation(x, Doctor) \land \neg InLove(x, Jane)$

- Jane's boss is not a doctor
  $\neg(\exists x Occupation(x, Doctor) \land Boss(x, Jane))$

- Angry people who are in love with their boss' spouse, kill their boss
  $\forall x Angry(x) \land \exists y InLove(x, y) \land \exists z Married(y, z) \land Boss(z, x) \Rightarrow Murder(x, z)$

4. (13 points) Show whether $F7(C)$ is entailed by the following first order logic knowledge base using **resolution**. Generate new sentences by referring to the existing ones, writing them out and specifying the *substitution* e.g., given, {S1, ..., Sn}, then "Sn+1: Resolve Si and Sj to get ... with the substitution $\{x1 \setminus Alfred, \dots\}$". Make sure you write out any additional information that is needed, e.g., to clarify any ambiguities. You are allowed to have *smart* picks. You do not have to follow any particular order and can pick or ignore any sentences.

Symbols: A, B, C
Predicates: $F1, F2, F3, F4, F5, F6, F7$
Known Facts: $F1(A), F2(B), F5(C)$
Rules:
R1: $\forall x, y F1(x) \wedge F2(y) \Rightarrow F3(x, y)$
R2: $\forall x F1(x) \Rightarrow F4(x)$
R3: $\forall x F5(x) \Rightarrow F6(x)$
R4: $\forall x, y, z F3(x, y) \wedge F4(x) \wedge F6(z) \Rightarrow F7(z)$

We must put the KB in CNF, standardize apart and add $\neg F_7(C)$:
S1: $\neg F1(x1) \vee \neg F2(y1) \vee F3(x1, y1)$
S2: $\neg F1(x2) \vee F4(x2)$
S3: $\neg F5(x3) \vee F6(x3)$
S4: $\neg F3(x4, y4) \vee \neg F4(x4) \vee \neg F6(z4) \vee F7(z4)$
S5: $F1(A)$
S6: $F2(B)$
S7: $F5(C)$
S8: $\neg F7(C)$

One possible solution:
S9: S3 and S7 with $\{x3 \setminus C\}$ to get $F6(C)$
S10: S4 and S9 with $\{z4 \setminus C\}$ to get $\neg F3(x4, y4) \vee \neg F4(x4) \vee F7(C)$
S11: S2 and S5 with $\{x2 \setminus A\}$ to get F4(A)
S12: S10 and S11 with $\{x4 \setminus A\}$ to get $\neg F3(A, y4) \vee F7(C)$
S13: S1 and 12 with $\{x1 \setminus A, y2 \setminus y4\}$ to get $\neg F1(A) \vee \neg F2(y4) \vee F7(C)$
S14: S1 and S13 to get $\neg F2(y4) \vee F7(C)$
S15: S6 and S14 with $\{y4 \setminus B\}$ to get $F7(C)$
S16: S8 and S15 to get $\emptyset$

As a result, our KB entails $F7(C)$.

5. (10 points) Probability: Your parents want you to teach your 2 year old cousin about cats, dogs and robots, using a pedagogically appropriate data set. However, you need to study for the COMP341 final (this question is from the future!) and do not have the time to do so. With your knowledge in machine learning (told you it is from the future), you want to automate the process. You randomly take n images from the data set, annotate them as either cat, dog or robot and feed it to your machine learning algorithm. After learning, the program you wrote flashes the images to your baby cousin, classifies them and speaks its class.

   Your random selection picked n_c cat, n_d dog and n_r robot images. You go over these images, extract features from them and use these features to learn a model for their corresponding classes. Let's use x to denote an image's features. For learning, you estimate a likelihood model for each of the class, denoted by j, in the form of: $f_j(x) = exp(-(x-\mu_j)^T(x-\mu_j)/\sigma_j^2)$ from all the features of that class. (the learning part is not important in this problem). It is important to note that f_j's do not represent proper probabilities but can be considered as being proportional to them.

   After learning, when a new image comes in, you extract its features and calculate its likelihood value for each class. You then pick the class with the highest likelihood. Answer the following questions.

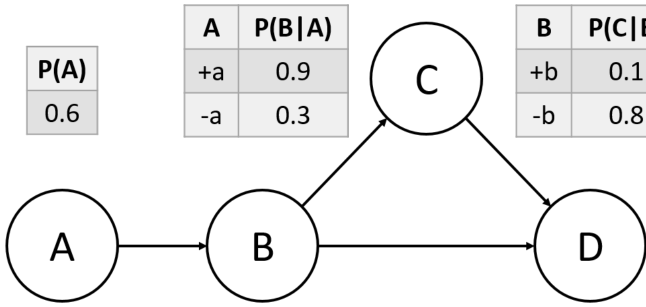
   1. (3) In terms of probability distributions, what does f_j represent out of these: $P(C=j,X)$, $P(C=j|X)$, or $P(X|C=j)$?
      $P(X|C=j)$

   2. (2) Formulate the existing decision procedure (picking the class with highest likelihood) in terms of the distribution you picked above. Hint: Use the *argmax* function.
      $d = arg_j max(f_j(x_i)) = arg_j max(P(X=x_i|C=j))$, where x_i is the input image features and d the decided class

   3. (3) Do you think your decision procedure is the best approach? If so, explain why in 20 words or less. If not, formulate a better one.
      $d = arg_j max(P(X=x_i|C=j)P(C=j)) = arg_j max(f_j(x_i)P(C=j))$

   4. (2) You had $n_c = 50$, $n_d = 40$ and $n_r = 10$ in your training set and learned f_c, f_d and f_r. You receive a new image. The likelihoods you calculate are: $f_c = 0.05$, $f_d = 0.02$ and $f_r = 0.1$. What is your classification result based on your answer you gave to the previous one?
      $d = arg_j max(0.05 \cdot 0.5, 0.02 \cdot 0.4, 0.1 \cdot 0.1) = arg_j max(0.025, 0.008, 0.01) = c$

6. (15 points) For this question, consider the Bayesian Network below. All variables are binary and *True* versions are denoted by a plus sign followed by a lower case letter. The CPTs are given for only the *True* values. You subtract this from 1 to get the probabilities for false. E.g. $P(-a) = 0.4, P(+c|-b) = 0.8, P(-d|+b,-c) = 0.4$ Answer the following questions and **show your work**.



P(A)
0.6

A	P(B|A)
+a	0.9
-a	0.3

B	P(C|B)
+b	0.1
-b	0.8

B	C	P(D|B,C)
+b	+c	0.9
+b	-c	0.6
-b	+c	0.3
-b	-c	0.8

(a) (2 points) Write the formula for the joint distribution, $P(A, B, C, D)$.
$P(A, B, C, D) = P(A)P(B|A)P(C|B)P(D|B, C)$

(b) (2 points) Calculate $P(+a, -b, +c, -d)$. You do not have to evaluate the final expression.
$P(+a, -b, +c, -d) = 0.6 \cdot 0.1 \cdot 0.8 \cdot 0.7 = 0.0336$

(c) (5 points) Calculate $P(+a| -b, +c)$. You do not have to evaluate the final expression.
Given B, A and C are conditionally independent! Thus $P(+a| - b, +c) = P(+a| - b)$
At this point join $P(A)$ and $P(B|A)$, then normalize:
$f1(+a, -b) = P(-b| + a)P(+a) = 0.1 \cdot 0.6 = 0.06$
$f1(-a, -b) = P(-b| - a)P(-a) = 0.7 \cdot 0.4 = 0.28$
$P(+a| - b) = 0.06/(0.06 + 0.28) = 0.06/0.34 = 3/17$

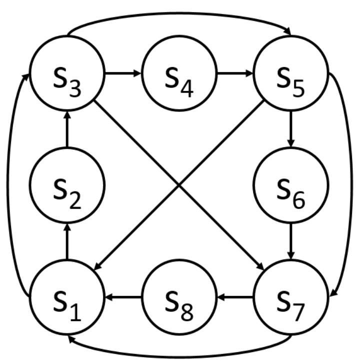
(d) (6 points) Calculate $P(+b| + d)$
The factors given the evidence are $P(A), P(+b|A), P(C| + b), P(+d| + b, C)$
Eliminate A by joining $P(A)$ and $P(+b|A)$ and summing out A. (Alternative is to start with C)
$f_1(+b) = \Sigma_{a\in A}P(a)P(+b|a) = P(+a)P(+b| + a) + P(-a)P(+b| - a) = 0.6 \cdot 0.9 + 0.4 \cdot 0.3 = 0.66$
Now eliminate C
$f_2(+b, +d) = f_1(+b)\Sigma_{c\in C}P(c| + b)P(+d| + b, c) = 0.66(0.1 \cdot 0.9 + 0.9 \cdot 0.6) = 0.66 \cdot 0.63$
We need to normalize this for B, to get our desired value. For this we need to calculate $f_2(-b, +d)$ which in turn makes us calculate $f_1(-b)$.
$f_1(-b) = P(+a)P(-b| + a) + P(-a)P(-b| - a) = 0.6 \cdot 0.1 + 0.4 \cdot 0.7 = 0.34$
$f_2(-b, +d) = 0.34(0.8 \cdot 0.3 + 0.2 \cdot 0.8) = 0.34 \cdot 0.40$
$\Rightarrow P(+b| + d) = f_2(+b, +d)/(f_2(+b, +d) + f_2(-b, +d)) = 0.66 \cdot 0.63/(0.66 \cdot 0.63 + 0.34 \cdot 0.40) = 0.7535$

7. (25 points) Reasoning over Time: Partially-Observable Mini-Monopoly (POMM)

You want to play POMM with a friend. There are 8 states, $S = \{s_1, \dots, s_8\}$. You own a house in each even numbered state. Your friend moves randomly between states. You want to charge your friend rent if he stops at even states. You can only observe him with a noisy sensor. At each turn you can ask for rent. If he is at an even state, you get the full rent. If not, you need to pay him a penalty of 25% of the rent. You each start with the same amount of money.

Below is the underlying Markov Chain for the game. The directed arcs show the possible transitions between states. Your friend can stay at his current state as well (i.e. self-transition). The transitions are equally likely. No arc implies no possible transition. m_i denotes the number of neighbors of state s_i, excluding itself.

Answer the following questions and **show your work**. Only write non-zero probabilities.



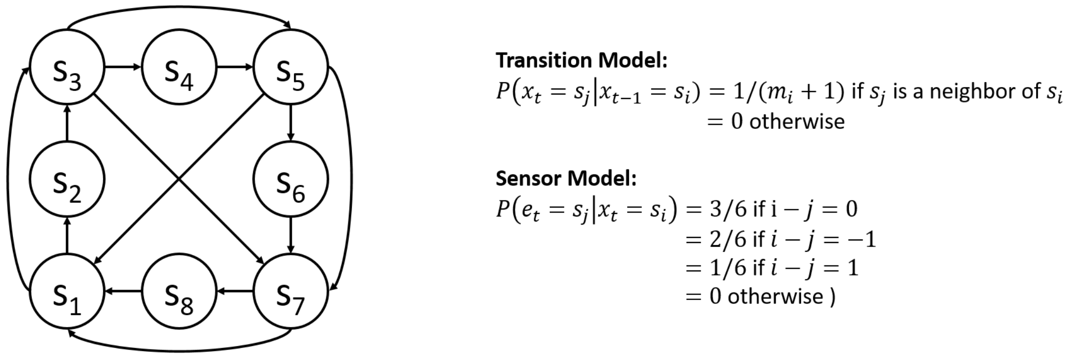
Transition Model:
$P(x_t = s_j | x_{t-1} = s_i) = 1/(m_i + 1)$ if s_j is a neighbor of s_i
$= 0$ otherwise

Sensor Model:
$P(e_t = s_j | x_t = s_i) = 3/6$ if $i - j = 0$
$= 2/6$ if $i - j = -1$
$= 1/6$ if $i - j = 1$
$= 0$ otherwise)

(a) (1 point) Your friend starts at s_1 and takes a turn without you observing him. What is his state distribution after his first turn? You can just write the numbers for this one. Hint: $1/(m_1 + 1)$
Let $P(x_t = s_j) = P_t(s_j), P(x_t = s_j | x_{t-1} = s_i) = P_t(s_j | s_i)$
$P_1(s_1) = P_1(= s_1 | s_1) P_0(s_1) = 1/3 \cdot 1.0 = 1/3$
$P_1(s_2) = P_1(= s_2 | s_1) P_0(s_1) = 1/3 \cdot 1.0 = 1/3$
$P_1(s_3) = P_1(= s_3 | s_1) P_0(s_1) = 1/3 \cdot 1.0 = 1/3$

(b) (9 points) He takes another turn (i.e. his 2^{nd} turn and you have not observed him yet). What is his state distribution?
Can be at states s_1, s_2, s_3, s_4, s_5 and s_7.
$P_2(s_1) = P_1(s_1|s_1)P_1(s_1) = 1/3 \cdot 1/3 = 1/9 = 4/36$
$P_2(s_2) = P_1(s_2|s_1)P_1(s_1) + P_1(s_2|s_2)P_1(s_2) = 1/9 + 1/2 \cdot 1/3 = 5/18 = 10/36$
$P_2(s_3) = P_1(s_3|s_1)P_1(s_1) + P_1(s_3|s_2)P_1(s_2) + P_1(s_3|s_3)P_1(s_3) = 1/9 + 1/6 + 1/4 \cdot 1/3 = 13/36$
$P_2(s_4) = P_1(s_4|s_3)P_1(s_3) = 1/12 = 3/36$
$P_2(s_5) = P_1(s_5|s_3)P_1(s_3) + P_1(s_5|s_4)P_1(s_4) = 1/12 + 0 = 3/36$
$P_2(s_7) = P_1(s_7|s_3)P_1(s_3) + P_1(s_7|s_5)P_1(s_5) + P_1(s_7|s_6)P_1(s_6) = 1/12 + 0 + 0 = 3/36$
1 point per correct entry, e.g., $P_2(s_3) : 3, P_2(s_7) : 1$

The figure is repeated here for your convenience.



(c) (8 points) After his 2^{nd} turn but before his 3^{rd} turn, you received a sensory reading at s_3. What is his state distribution now?

Let $B_t(s_j) = \alpha P(x_t = s_j | e_{1:t}), P(e_2 = s_i | x_2 = s_j) = P_2(e = s_i | s_j)$
$B_2(s_1) = P_2(e = s_3 | s_1) P_2(s_1) = 0 \cdot 1/9 = 0$
$B_2(s_2) = P_2(e = s_3 | s_2) P_2(s_2) = 2/6 \cdot 5/18 = 5/54 = 20/216$
$B_2(s_3) = P_2(e = s_3 | s_3) P_2(s_3) = 3/6 \cdot 13/36 = 13/72 = 39/216$
$B_2(s_4) = P_2(e = s_3 | s_4) P_2(s_4) = 1/6 \cdot 1/12 = 1/72 = 3/216$
$B_2(s_5) = P_2(e = s_3 | s_5) P_2(s_5) = 0$
$B_2(s_7) = P_2(e = s_3 | s_7) P_2(s_7) = 0$
Normalize
1 point per correct belief, 2 points for normalizing $P_2(s_2 | e = s_3) = 20/(20 + 39 + 3) = 20/62$
$P_2(s_3 | e = s_3) = 39/62$
$P_2(s_4 | e = s_3) = 3/62$

(d) (2 points) You asked for him to pay. He did not have to. What is the state distribution now?

This implies that he is not in an even state. The bad part is that you lost all your money. The good part is that, you know exactly where he is! In a real application, the information gained from an action might be more valuable than the action cost.
$P_2(s_3) = 1.0$

(e) (5 points) Bonus Question: How do you decide if this is a fair game or not?

An open ended question.