Washington State University
School of Electrical Engineering and Computer Science
Fall 2018

CptS 440/540 Artificial Intelligence
Homework 7 Solution
Due: October 18, 2018 (11:59pm)

General Instructions: Put your answers to the following problems into a PDF document and submit as an attachment under Content → Homework 7 for the course CptS 440 Pullman (all sections of CptS 440 and 540 are merged under the CptS 440 Pullman section) on the Blackboard Learn system by the above deadline. Note that you may submit multiple times, but we will only grade the most recent entry submitted before the above deadline.

1. The table below shows the joint probability distribution over three random variables: Sea ∈ {blue, gray}, Sun ∈ {true, false}, and Sky ∈ {clear, cloudy, overcast}. Based on this distribution, answer the following questions. Show your work.

Sea		blue		gray	
Sun		true	false	true	false
Sky	clear	0.12	0.07	0.06	0.08
	cloudy	0.07	0.10	0.07	0.09
	overcast	0.05	0.09	0.08	0.12

   a. P(Sea=blue, Sun=true, Sky=clear)?
   b. P(Sky=clear)?
   c. P(Sun=true)?
   d. P(Sun=true | Sky=clear)?
   e. P(Sea=blue | Sky=clear ∨ Sky=cloudy)?

Solution:

   a. P(Sea=blue, Sun=true, Sky=clear) = 0.12
   b. P(Sky=clear) = 0.12 + 0.07 + 0.06 + 0.08 = 0.33
   c. P(Sun=true) = 0.12 + 0.07 + 0.05 + 0.06 + 0.07 + 0.08 = 0.45
   d. P(Sun=true | Sky=clear) = P(Sun=true, Sky=clear) / P(Sky=clear) = (0.12 + 0.06) / 0.33 = 0.55
   e. P(Sea=blue | Sky=clear ∨ Sky=cloudy)
   = P(Sea=blue ∧ (Sky=clear ∨ Sky=cloudy)) / P(Sky=clear ∨ Sky=cloudy)
   = (0.12 + 0.07 + 0.07 + 0.10) / (0.12 + 0.07 + 0.06 + 0.08 + 0.07 + 0.10 + 0.07 + 0.09)
   = 0.36 / 0.66 = 0.55
   Or, to completely remove the ∨'s…
   = P((Sea=blue ∧ Sky=clear) ∨ (Sea=blue ∧ Sky=cloudy)) / P(Sky=clear ∨ Sky=cloudy)

$$= \frac{P(Sea{=}blue \wedge Sky{=}clear) + P(Sea{=}blue \wedge Sky{=}cloudy) - P(Sea{=}blue \wedge Sky{=}clear \wedge Sky{=}cloudy)}{P(Sky{=}clear) + P(Sky{=}cloudy) - P(Sky{=}clear \wedge Sky{=}cloudy)}$$

   = [(0.12 + 0.07) + (0.07 + 0.10) – 0] / [0.33 + (0.07 + 0.10 + 0.07 + 0.09) – 0]
   = 0.36 / 0.66 = 0.55

2. Consider the problem with three Boolean random variables: EatRight, Exercise, Healthy. Assume you know only the following information:
    - P(Healthy=true) = 0.8
    - P(EatRight=true $\wedge$ Exercise=true | Healthy=true) = 0.6
    - P(EatRight=true $\wedge$ Exercise=true | Healthy=false) = 0.3

    Using Bayes rule and normalization, compute **P**(Healthy | EatRight=true $\wedge$ Exercise=true). Note the "**P**" is boldfaced, so we want a distribution.

    Solution:

    P(Healthy | EatRight=true $\wedge$ Exercise=true)
    = **P**(EatRight=true $\wedge$ Exercise=true | Healthy) * **P**(Healthy) / P(EatRight=true $\wedge$ Exercise=true)
    = α **P**(EatRight=true $\wedge$ Exercise=true | Healthy) * **P**(Healthy)
    = α $<$ P(EatRight=true $\wedge$ Exercise=true | Healthy=true) * P(Healthy=true),
    P(EatRight=true $\wedge$ Exercise=true | Healthy=false) * P(Healthy=false) $>$
    = α $<$ (0.6)(0.8), (0.3)(0.2) $>$
    = α $<0.48, 0.06>$
    = $<0.89, 0.11>$

3. Consider the problem with three Boolean random variables: EatRight, Exercise, Healthy. Assume you know only the following information:
    - P(Healthy=true) = 0.8
    - P(EatRight=true | Healthy=true) = 0.6
    - P(EatRight=true | Healthy=false) = 0.2
    - P(Exercise=true | Healthy=true) = 0.8
    - P(Exercise=true | Healthy=false) = 0.3
    - EatRight and Exercise are conditionally independent given Healthy.

    Using Bayes rule and normalization, compute **P**(Healthy | EatRight=true $\wedge$ Exercise=true). Note the "**P**" is boldfaced, so we want a distribution.

    Solution:

    P(Healthy | EatRight=true $\wedge$ Exercise=true)
    = **P**(EatRight=true $\wedge$ Exercise=true | Healthy) * **P**(Healthy) / P(EatRight=true $\wedge$ Exercise=true)
    = α **P**(EatRight=true $\wedge$ Exercise=true | Healthy) * **P**(Healthy)
    By conditional independence information…
    = α **P**(EatRight=true | Healthy) * P(Exercise=true | Healthy) * **P**(Healthy)
    = α $<$ P(EatRight=true | Healthy=true) * P(Exercise=true | Healthy=true) * **P**(Healthy=true),
    P(EatRight=true | Healthy=false) * P(Exercise=true | Healthy=false) * **P**(Healthy=false) $>$
    = α $<$ (0.6)(0.8)(0.8), (0.2)(0.3)(0.2) $>$
    = α $<$ 0.384, 0.012 $>$
    = $<0.97, 0.03>$

4. Suppose your agent has just started playing a 4x4 Wumpus world game. The agent does not observe a breeze in location (1,1). The agent moves right to location (2,1) and observes a breeze. Given this information, we want to compute the probability there is a pit in location (3,1). You may use $p_{x,y}$ and $\neg p_{x,y}$ as shorthand notation for $Pit_{x,y}$=true and $Pit_{x,y}$=false, respectively. Similarly, you may use $breeze_{x,y}$ and $\neg breeze_{x,y}$ as shorthand notation for $Breeze_{x,y}$=true and $Breeze_{x,y}$=false, respectively. Specifically:
   a. Define the sets: *breeze*, *known*, *frontier* and *other*.
   b. Following the method in the textbook and lectures, compute the probability distribution $\mathbf{P}(Pit_{3,1} \mid breeze, known)$. Show your work.

Solution:

a. Sets
   - *breeze* = breeze information we know for certain
     $= \{ \neg\, breeze_{1,1}, breeze_{2,1} \}$
   - *known* = pit information we know for certain
     $= \{\neg\, pit_{1,1}, \neg\, pit_{1,2}, \neg\, pit_{2,1}\}$
   - *frontier* = unknown pit random variables for locations adjacent to known, minus *query*
     $= \{ P_{1,3}, P_{2,2}, P_{3,1} \} - query$
   - *other* = unknown pit random variables for locations other than *known*, *frontier*, and *query*
     $= \{ P_{1,4}, P_{2,4}, P_{3,4}, P_{4,4}, P_{2,3}, P_{3,3}, P_{4,3}, P_{3,2}, P_{4,2}, P_{4,1} \}$

b. $\mathbf{P}(Pit_{3,1} \mid breeze, known)$

$= \mathbf{P}(P_{3,1}, breeze, known) / \mathbf{P}(breeze, known)$

$= \alpha\, \mathbf{P}(P_{3,1}, breeze, known)$

$= \alpha \sum_{unknown} \mathbf{P}(P_{3,1}, breeze, known, unknown)$

$= \alpha \sum_{frontier} \sum_{other} \mathbf{P}(P_{3,1}, breeze, known, frontier, other)$

$= \alpha \sum_{frontier} \sum_{other} \mathbf{P}(breeze \mid P_{3,1}, known, frontier, other)\, \mathbf{P}(P_{3,1}, known, frontier, other)$

Since *breeze* is independent of *other* given $P_{3,1}$, *known* and *frontier*:

$= \alpha \sum_{frontier} \sum_{other} \mathbf{P}(breeze \mid P_{3,1}, known, frontier)\, \mathbf{P}(P_{3,1}, known, frontier, other)$

$= \alpha \sum_{frontier} \mathbf{P}(breeze \mid P_{3,1}, known, frontier) \sum_{other} \mathbf{P}(P_{3,1}, known, frontier, other)$

Since $P_{3,1}$, *known, frontier, other* are independent of each other:

$= \alpha \sum_{frontier} \mathbf{P}(breeze \mid P_{3,1}, known, frontier) \sum_{other} \mathbf{P}(P_{3,1})\, P(known)\, P(frontier)\, P(other)$

$= \alpha\, \mathbf{P}(P_{3,1})\, P(known) \sum_{frontier} \mathbf{P}(breeze \mid P_{3,1}, known, frontier)\, P(frontier) \sum_{other} P(other)$

Letting $\alpha' = \alpha * P(known)$, and since $\sum_{other} P(other) = 1$:

$= \alpha' \,\mathbf{P}(P_{3,1}) \sum_{frontier} \mathbf{P}(breeze \mid P_{3,1}, known, frontier)\, P(frontier)$

$= \alpha' \,\mathbf{P}(P_{3,1}) \sum_{frontier} \mathbf{P}(breeze \mid P_{3,1}, known, frontier)\, P(frontier)$

$= \alpha' < P(pit_{3,1})\, [\, \sum_{frontier} P(breeze \mid pit_{3,1}, known, frontier)\, P(frontier)\,]\,,$
$\quad P(\neg\, pit_{3,1})\, [\, \sum_{frontier} P(breeze \mid \neg\, pit_{3,1}, known, frontier)\, P(frontier)\,] >$

$= \alpha' < P(pit_{3,1})\, [\, P(breeze \mid pit_{3,1}, known, pit_{2,2}, pit_{1,3})\, P(pit_{2,2}, pit_{1,3}) +$
$\quad P(breeze \mid pit_{3,1}, known, pit_{2,2}, \neg\, pit_{1,3})\, P(pit_{2,2}, \neg\, pit_{1,3}) +$
$\quad P(breeze \mid pit_{3,1}, known, \neg\, pit_{2,2}, pit_{1,3})\, P(\neg\, pit_{2,2}, pit_{1,3}) +$
$\quad P(breeze \mid pit_{3,1}, known, \neg\, pit_{2,2}, \neg\, pit_{1,3})\, P(\neg\, pit_{2,2}, \neg\, pit_{1,3})\,]\,,$
$\quad P(\neg\, pit_{3,1})\, [\, P(breeze \mid \neg\, pit_{3,1}, known, pit_{2,2}, pit_{1,3})\, P(pit_{2,2}, pit_{1,3}) +$

$P(breeze \mid \neg pit_{3,1}, known, pit_{2,2}, \neg pit_{1,3})\, P(pit_{2,2}, \neg pit_{1,3})\, +$

$P(breeze \mid \neg pit_{3,1}, known, \neg pit_{2,2}, pit_{1,3})\, P(\neg pit_{2,2}, pit_{1,3})\, +$ *not possible*

$P(breeze \mid \neg pit_{3,1}, known, \neg pit_{2,2}, \neg pit_{1,3})\, P(\neg pit_{2,2}, \neg pit_{1,3})\ ]\ >$ *not possible*

Given that $\mathbf{P}(Pit_{x,y}) = \langle 0.2, 0.8 \rangle$:

$= \alpha' \langle (0.2)\ [\ (1)(0.2)(0.2) + (1)(0.2)(0.8) + (1)(0.8)(0.2) + (1)(0.8)(0.8)\ ],$

$(0.8)\ [\ (1)(0.2)(0.2) + (1)(0.2)(0.8)\ ] \rangle$

$= \alpha' \langle 0.2, 0.16 \rangle$

$= \langle 0.56, 0.44 \rangle$

5. *CptS 540 Students Only*. Consider a variant of the Wumpus world game in which the agent may reveal whether or not any one location has a breeze. Given the scenario in Problem #4, what one location should the agent reveal, and what would the result need to be, in order to know for sure that there is a pit in location (3,1).

   Solution: If there is not a breeze in location (1,2), then there must be a pit in location (3,1).