Com S 472/572 Principles of Artificial Intelligence

Final Exam

Fall 2020

Honor statement: *I affirm that I am the assigned student taking the test, and this is entirely my own work. I affirm my acceptance of these rules: 1) closed-book and closed-notes during the exam; 2) no online search for information during the exam; and 3) no discussion or sharing in any form with others during or after the exam.*

Name: __Sean Gordon__________________________________

ISU NetID (username): ___Sgordon4__________________________________@iastate.edu

1	2	3	4	5	6	Total
47	8	11	10	12	12	100

1. **[47 pts]** *Short Questions*

(a) [**12 pts**] Determine if the following statements are true or false. For each statement, mark only the answer you think is correct.

(i) The propositional logic (PL) sentence $\neg P \vee Q \vee R$, where P, Q, and R are atomic sentences, is a Horn clause.

true ___________ false ___X___

(ii) A knowledge base *KB* entails a PL sentence α if and only if *KB* $\wedge \neg\alpha$ is unsatisfiable.

true ___X____ false __________

(iii) The two first order logic (FOL) sentences $\forall x \exists y$ *Love*(x, y) and $\exists x \forall y$ *Love*(x, y) are logically equivalent.

true ___________ false ___X___

(iv) To find one root of an equation $g(x) = 0$, where g is differentiable, the Newton-Raphson method conducts local search by changing the current state (i.e., the root estimate) x by an amount which depends on the value of $g(x)$ only.

true ___X____ false __________

(v) Any constraint satisfaction problem (CSP) can be transformed into a binary CSP.

true ___X____ false _________

(vi) In a Bayesian network, the Markov blanket of a node includes its children, their parents, and its own parents.

true ___X____ false _________

(b) [**3 pts**] What are the input and output of an agent program?

Input is provided through percepts, and output is given in the form of actions.

(c) [**4 pts**] Two admissible heuristic functions h_1 and h_2 are separately used for solving the same search problem. What does it mean to say that h_1 dominates h_2? Which of the two may end up expanding more nodes?

h1 dominates h2 if at every node h1 >= h2.
h2 will not expand more nodes than h1, so h1 may end up expanding more nodes.

(d) [**3 pts**] In the FOL sentence $(\forall x\ P(x) \Rightarrow \exists y\ Q(y, z)) \wedge R(z)$, which variables are bound and which are free?

x and y are bound
z is free

(e) [**3 pts**] Apply the technique of Skolemization to eliminate all the quantifiers from the FOL sentence given below.
(You need only write down the final quantifier-free form. In case you introduce any new terms, please briefly explain them.)

$$\exists s \forall t \exists u \forall v \exists x \forall y \exists z\ \ P(s, t, u, v, x, y, z)$$

P(S1, t, f1(t), v, f2(t, v), y, f3(t, v, y))

(f) [**4 pts**] Assuming the predicate *Parents*(u, v, w), functions *mom*(u) and *dad*(u), constant *John*, and variables x, y, z, give the most general unifier of the following two atomic sentences:

$$Parents(x, dad(y), mom(John)) \qquad Parents(y, z, mom(x))$$

$Parents(x, dad(y), mom(John)),\ Parents(y, z, mom(x))\ :\ \{x/John\}$
$Parents(John, dad(y), mom(John)),\ Parents(y, z, mom(John))\ :\ \{x/John$, y/John}
$Parents(John, dad(John), mom(John)),\ Parents(John, z, mom(John))\ :\ \{x/John$, y/John, z/$dad(John)\}$
$Parents(John, dad(John), mom(John)),\ Parents(John, dad(John), mom(John))\ :\ \{x/John$, y/John, z/$dad(John)\}$

(g) [**4 pts**] Two random variables X and Y have sample spaces $\{x_1, x_2, \ldots, x_m\}$ and $\{y_1, y_2, \ldots, y_n\}$, respectively. What do we mean by saying that X and Y are independent of each other?

Two values are independent if the value of one does not affect the value of the other.
P(X|Y) = P(Y) and P(X|Y) = P(X)

(h) [**4 pts**] In a discrete-time model, we let $\boldsymbol{X}_t$ be the set of state variables, $\boldsymbol{E}_t$ be the set of evidence variables, and $\boldsymbol{e}_t$ be the observed values of the evidence variables, all at time t. Explain in a succinct way the two tasks of filtering and smoothing. (You may denote by $\boldsymbol{X}_{1:t}$ the state sequence $\boldsymbol{X}_1, \boldsymbol{X}_2, \dots, \boldsymbol{X}_t$ and by $\boldsymbol{e}_{1:t}$ the state sequence $\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_t$.)

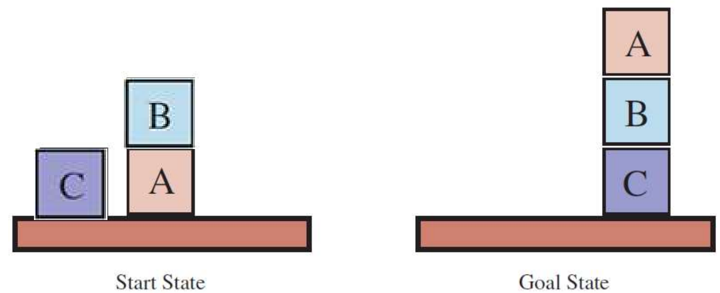
Filtering is computing the new state Xt given all observed evidence variables e1:t
Smoothing is computing a previous state Xk given all observed evidence variables e1:t, where 0 <= k < t

(i) [**4 pts**] A random variable X has three possible values with probabilities $\frac{1}{2}, \frac{1}{4}, \frac{1}{4}$, respectively. What is the entropy of X?

-((1/2)log2(1/2) + (1/4)log2(1/4) + (1/4)log2(1/4)) = 1.5

(j) [**6 pts**] Consider the planning domain of the blocks world that is specified in the planning domain definition language (PDDL) below. The condition *Clear*(x) means "there is a clear space on x to hold a block".

$$
\begin{aligned}
&Init(On(A, Table) \wedge On(B, Table) \wedge On(C, A) \\
&\quad \wedge Block(A) \wedge Block(B) \wedge Block(C) \wedge Clear(B) \wedge Clear(C) \wedge Clear(Table)) \\
&Goal(On(A, B) \wedge On(B, C)) \\
&Action(Move(b, x, y), \\
&\quad \text{PRECOND: } On(b, x) \wedge Clear(b) \wedge Clear(y) \wedge Block(b) \wedge Block(y) \wedge \\
&\qquad\qquad (b \neq x) \wedge (b \neq y) \wedge (x \neq y), \\
&\quad \text{EFFECT: } On(b, y) \wedge Clear(x) \wedge \neg On(b, x) \wedge \neg Clear(y)) \\
&Action(MoveToTable(b, x), \\
&\quad \text{PRECOND: } On(b, x) \wedge Clear(b) \wedge Block(b) \wedge Block(x), \\
&\quad \text{EFFECT: } On(b, Table) \wedge Clear(x) \wedge \neg On(b, x))
\end{aligned}
$$



Construct a sequence of actions to transform the start state into the goal state, both shown above on the right. (An action sequence is in the form of, say, *Move*(…), *MoveToTable*(…), …. Do not write down the precondition and effect of any action.)

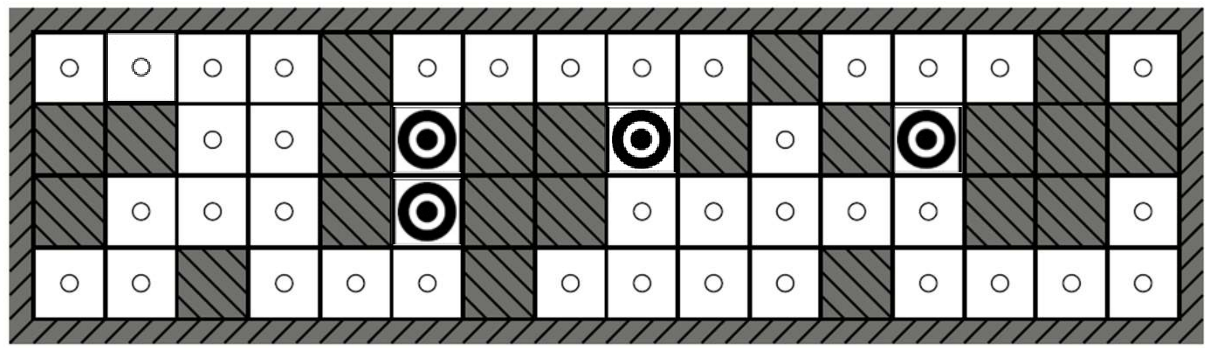
Move(B, A, C)
Move(A, Table, B)

2. **[8 pts]** *Robot Localization*

This is the same task we learned about during the first half of the course. A robot in a maze-like environment shown below wants to localize itself. The robot has an accurate map of the environment, as well as four perfect sonar sensors to tell whether there is an obstacle in the neighboring square in each of the four compass directions. The percept consists of four bits *NESW.* Values 1 of these bits indicate obstacles to the north, east, south, and west, respectively; and values 0 indicate no obstacles to the same four directions, respectively. So *NESW* = 1011 means there are obstacles to the north, south, and west, but not east.

Robot

(a) [**4 pts**] At the start, the robot does not know where it is. It receives the percept 0101. In the top map to the right, mark all possible locations of the robot at this moment. (If you are working on the PPTX version of the exam, simply cover those square locations with copies of the robot icon on the left. If you are working on the PDF version, you may mark or darken the squares.)

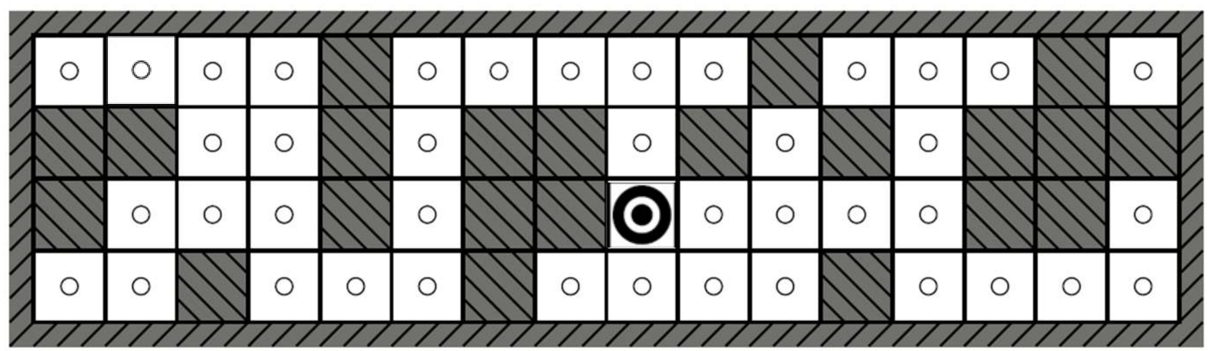


(a) Possible robot locations at the start.

(b) [**4 pts**] Based on the percept 0101, the robot executes a *Down* action to move to the south by one square. It receives a new percept 0001. In the bottom map to the right, mark the robot's possible locations after this movement.

Has the robot localized itself relative to the environment?

Yes



(b) Possible robot locations after one downward move.

3. **[11 pts]** *FOL Translation and CNF Conversion*

(a) [**4 pts**] Write down a first order logic sentence to represent the following English sentence:

A supplier is preferred if all the parts he supplies arrive on time.

You are required to use the vocabulary of symbols provided below.

$S(x)$: predicate. Person x is a supplier.

$Sps(x, y)$: predicate. Person x supplies part y.

$P(x)$: predicate. Person x is preferred.

$A(y)$: predicate. Part y arrives on time.

$(S(x) \wedge P(x)) \Leftrightarrow (S(x) \wedge Sps(x, y) \wedge A(y))$

(b) [**7 pts**] Convert the logical sentence from (a) into conjunctive normal form.

$(S(x) \land P(x)) \Leftrightarrow (S(x) \land Sps(x, y) \land A(y))$

$(\neg(S(x) \land P(x)) \lor (S(x) \land Sps(x, y) \land A(y))) \land ((S(x) \land P(x)) \lor \neg (S(x) \land Sps(x, y) \land A(y)))$

$(\neg S(x) \lor \neg P(x) \lor (S(x) \land Sps(x, y) \land A(y))) \land ((S(x) \land P(x)) \lor \neg S(x) \lor \neg Sps(x, y) \lor \neg A(y))$

$(\neg S(x) \lor \neg P(x) \lor (S(x) \land Sps(x, y) \land A(y))) \land ((S(x) \land P(x)) \lor \neg S(x) \lor \neg Sps(x, y) \lor \neg A(y))$

$(\neg S(x) \lor \neg P(x) \lor Sps(x, y)) \land (\neg S(x) \lor \neg P(x) \lor A(y)) \land (\neg S(x) \lor P(x) \lor \neg Sps(x, y) \lor \neg A(y))$

$\neg S(x) \lor P(x) \lor \neg Sps(x, y) \lor \neg A(y)$

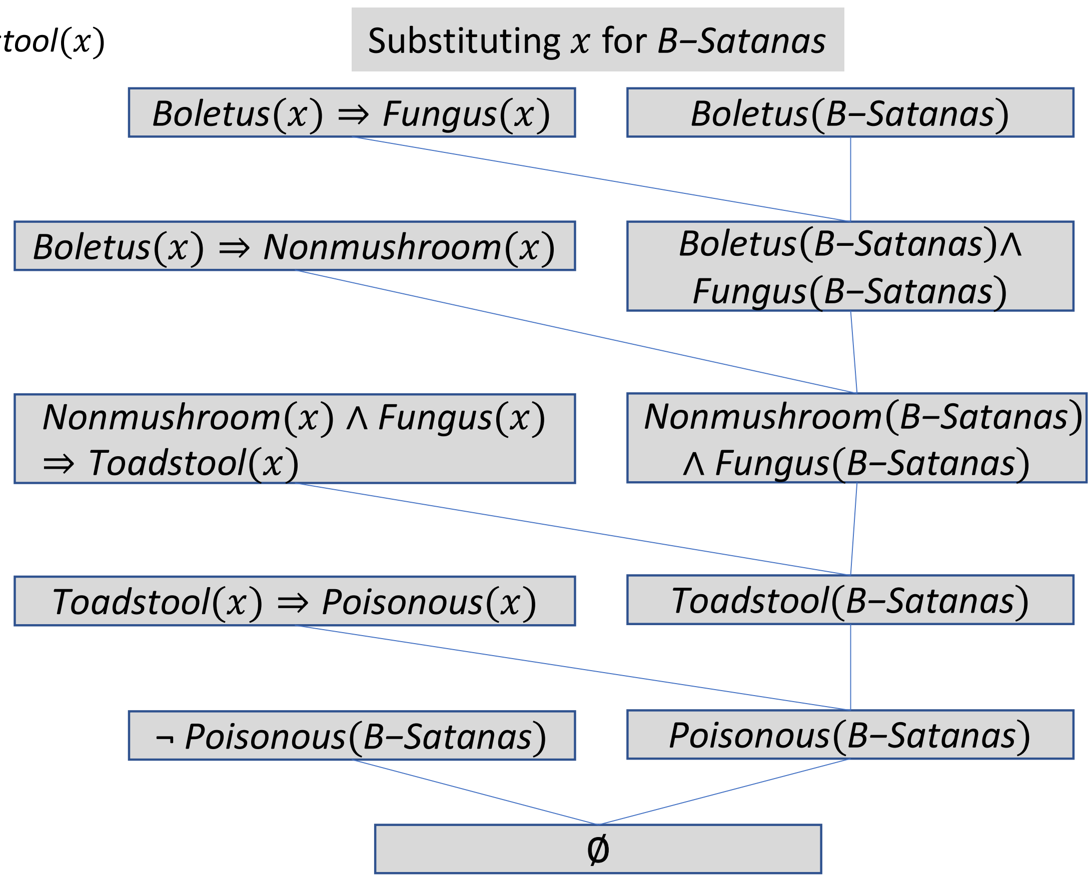
4. **[10 pts]** *Inference in First-Order Logic*

Suppose a knowledge base consists of the following five first-order definite clauses:

$$Nonmushroom(x) \wedge Fungus(x) \Rightarrow Toadstool(x)$$
$$Toadstool(x) \Rightarrow Poisonous(x)$$
$$Boletus(x) \Rightarrow Fungus(x)$$
$$Boletus(x) \Rightarrow Nonmushroom(x)$$
$$Boletus(B\text{–}Satanas)$$

In the above, x is a variable while *B-Satanas* is a constant. Use forward chaining to prove $Poisonous(B\text{–}Satanas)$ by constructing a generated proof tree to the right.

(You may draw a tree with substitutions, if any, shown beneath nodes, or a tree after all substitutions have been carried out.)



5. **[12 pts]** *Inference Using Full Joint Distributions*

Consider three Boolean random variables A, B, and C. We abbreviate the propositions $A = \textit{true}$, $B = \textit{true}$, $C = \textit{true}$ as a, b, c, respectively, and $A = \textit{false}$, $B = \textit{false}$, $C = \textit{false}$ as $\neg a$, $\neg b$, $\neg c$, respectively. Given below is a table showing their joint probability distributions.

	b		$\neg b$	
	c	$\neg c$	c	$\neg c$
a	0.1	0.15	0.1	0.2
$\neg a$	0.3	0.05	0.05	0.05

(a) [**4 pts**] Compute the conditional probability distribution $\boldsymbol{P}(A \mid \neg c) = \langle P(a \mid \neg c), P(\neg a \mid \neg c)\rangle$.
(To normalize a vector, say, $\langle x, y\rangle$, you may write $\alpha\langle x, y\rangle$, which represents $\langle \frac{x}{x+y}, \frac{y}{x+y}\rangle$.)

$\boldsymbol{P}(\neg c)$ = 0.15 + 0.05 + 0.2 + 0.05 = 0.45
$\boldsymbol{P}(A \mid \neg c) = \langle (0.15 + 0.2)/0.45, (0.05 + 0.05)/0.45\rangle = \langle 0.7778, 0.2222\rangle$

(b) [**4 pts**] Compute the conditional probability distribution $\boldsymbol{P}(B \mid \neg c) = \langle P(b \mid \neg c), P(\neg b \mid \neg c)\rangle$.

$\boldsymbol{P}(\neg c)$ = 0.15 + 0.05 + 0.2 + 0.05 = 0.45

$\boldsymbol{P}(B \mid \neg c) = \langle \frac{0.15+0.05}{0.45}, \frac{0.2+0.05}{0.45} \rangle = \langle 0.4444, 0.5556 \rangle$

(c) [**4 pts**] Are A and B conditionally independent given C? Explain why or why not.

$\boldsymbol{P}(A$, B | C) == $\boldsymbol{P}(A|$ C) * $\boldsymbol{P}(B|$ C) ?

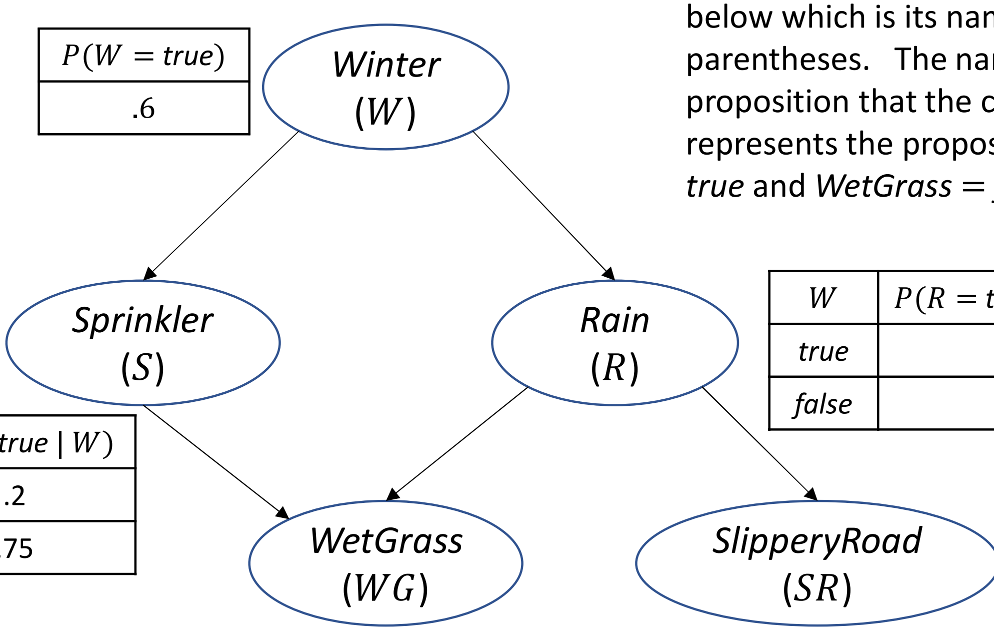
$\boldsymbol{P}(A$, B | C) = P(C|A,B) P(A,B) = $\langle 0.1, 0.15 \rangle$

$\boldsymbol{P}(A|$ C) * $\boldsymbol{P}(B|$ C) = $\langle 0.3457, 0.1235 \rangle$

As the two equations are not equivalent, the probabilities are not conditionally independent.

6. **[12 pts]** *Bayesian networks*

Show on the left is a Bayesian network with five nodes and their conditional probability tables (CPTs). Within each node is a Boolean random variable, below which is its name abbreviation in upper case letter(s) inside a pair of parentheses. The name abbreviation in lower case letter(s) represents the proposition that the corresponding variable is true, and if preceded with $\neg$, represents the proposition that the variable is false. For example, *Winter* $=$ *true* and *WetGrass* $=$ *false* are abbreviated as *w* and $\neg$*wg*, respectively.



$P(W = true)$
.6

W	$P(S = true \mid W)$
true	.2
false	.75

W	$P(R = true \mid W)$
true	.8
false	.1

S	R	$P(WG = true \mid S, R)$
true	*true*	.95
true	*false*	.9
false	*true*	.8
false	*false*	0

R	$P(SR = true \mid R)$
true	.7
false	.05

(a) [**3 pts**] Evaluate the joint probability below.

$$P(\neg w \land s \land \neg r \land wg) =$$

(1-0.6) * (1-0.2) * (1-0.1) * 0.9 = 0.2592

(b) [**3 pts**] Consider a query for the probability distribution $\boldsymbol{P}(R \mid sr) = \langle P(r \mid sr),\ P(\neg r \mid sr)\rangle$.

The evidence variable is ____sr__________.

The query variable is ______R_________.

Variables that are irrelevant to the evaluation include ___W, S, WG___.

(c) [**6 pts**] Compute the probability distribution $\boldsymbol{P}(R \mid sr) = \langle P(r \mid sr),\ P(\neg r \ \mid sr)\rangle$.

$P(r$

$$P(r \mid sr) = \frac{\text{P(sr|r)P(r)}}{\text{P(sr)}} = \frac{(0.7)(0.8*(1\text{-}.9))}{(0.7*(1\text{-}0.05))} = \frac{(0.7)(0.72)}{(0.665)} = \frac{(0.7)(0.72)}{(0.665)} = 0.7579$$