CPSC420-200 Final Exam (5/8/2005, Mon)[1]

Last name: _______________, First name: _______________

Time: **3:30pm–5:30pm (2 hours)**, Total Points: **100**

Subject	Score
First-order logic	/45
Uncertainty	/15
Probabilistic reasoning	/26
Learning	/30
Natural language processing	/9
Total	/125

- The total adds up to 125, but your final score will be $\min(raw\_score, 100)$. For example, if you solve all problems and get 85 points total, then your score will be $\min(85, 100) = 85$. If you get 112 points total, then your score will be $\min(112, 100) = 100$. So, if you have time, solve all problems!
- Be as **succinct** (i.e., brief) as possible.
- Read the questions carefully to see what kind of answer is expected (*explain blah* **in terms of ...** *blah*).
- Total of __11__ pages, including this cover and the Appendix at the end. **Before starting, count the pages and see if you have all __11__.**
- Problems are on **both** sides of each sheet.
- This is a closed-book, closed-note exam.
- You may rip off the last page (Appendix) to view it while solving the logic problems.

[1] Instructor: Yoonsuck Choe.

1 First-Order Logic (45pts)

Question 1 (5 pts): What is a limitation of propositional logic that first-order logic overcomes?

Question 2 (5 pts): What is the purpose of doing unification (and substitution) in automated inference for first-order logic?

Question 3 (6 pts): Are the following pairs of predicates unifiable? Answer Y or N to each, and give the most general unifier and the unified expression if unifiable. Your answer must be exact. (Note: P, Q, R, S are predicates, u, v, w, x, y, z are variables, A, B, C, D are constants, and $f(\cdot), g(\cdot)$ are functions.)

1. $Q(A, x, f(x))$ and $Q(y, y, z)$.

2. $Q(x, x, B)$ and $Q(u, f(v, w), w)$.

Question 4 (6 pts): Convert the following statement into standard form (find the prenex normal form , turn into CNF, and then skolemize). Show all your work.

$$\exists x, [\neg(\exists y, P(x, y)) \rightarrow (\exists z, (Q(z) \rightarrow R(x)))]$$

Question 5 (10 pts): Show that $R(A)$ is a logical consequence of the following. Use **resolution**.

1. $P(x) \vee Q(x, y) \vee R(y)$
2. $\neg P(A)$
3. $\neg Q(w, z) \vee R(w)$

Question 6 (8 pts): Resolution is domain (independent, dependent): **circle one** (2pts). Why is this both an advantage and also a disadvantage? Briefly explain in the two cases **in terms of dependence/independence**.

(1) **Advantage** (3pts):

(2) **Disadvantage** (3pts):

Question 7 (5 pts): Briefly explain how question answering can be done using a resolution theorem prover.

2 Uncertainty (15pts)

Question 8 (5 pts): Why is it difficult to apply first-order logic in domains with uncertainty? Briefly explain.

Question 9 (10 pts): Consider the two random variables $Cause$ and $Phenomenon$, and answer the two questions.

1. $P(Cause|Phenomenon)$ is (easier, harder) to calculate directly than $P(Phenomenon|Cause)$ (**circle one**). (5pts)
2. Explain why is this observation important in recognizing the usefulness of the Bayes rule? (5pts)

3 Probabilistic Reasoning (26pts)

Question 10 (6 pts): Given the belief network below, calculate the probability of the event: Both John and Mary called when a burglary actually happened and the alarm turned on while there was no earthquake. You don't need to reduce it to a single value – just leave the formula as is.

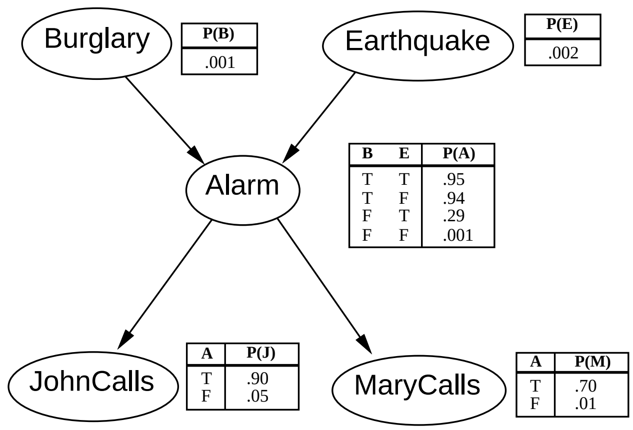

Figure 1: **Belief Network.**

Question 11 (6 pts): Why is ordering of the nodes important when constructing a belief network?

Question 12 (8 pts): Bayesian updating makes belief update very efficient as new evidence comes in. What is the major **domain-specific property** that allows such an efficient update? Explain in terms of two of the following conditional probabilities: $P(Cavity|Catch)$, $P(Catch|Cavity)$, $P(Catch|Cavity, ToothAche)$, and $P(Cavity|Catch, ToothAche)$.

Question 13 (6 pts): Explain why intercausal inference works, e.g., $P(Alarm|Burglary) >> P(Alarm|Burglary, Earthquake)$.

4 Learning (30pts)

Question 14 (6 pts): What are the appropriate tasks suitable for (1) supervised learning, (2) unsupervised learning, and (3) reinforcement learning? Give one example each and briefly explain.

Question 15 (9 pts): (1) What is the limitation of perceptrons and (2) how does a geometrical interpretation of the connection weights and bias help in clearly showing the limitation? Explain using the XOR function as an example. (3) How does backpropagation solve the limitation in perceptrons?

Question 16 (6 pts): Decision tree learning falls under the broader category of inductive learning. (1) What is the **particular** inductive bias in decision tree learning algorithm and (2) why does it make the resulting tree better than a tree that explicitly represents all example input-output pairs?

Question 17 (9 pts): In decision tree learning, suppose the initial set of input examples contain 14 positive and 13 negative examples. After testing with an attribute A which can take on three different values a_1, a_2, and a_3, the set it partitioned as follows:

- $A = a_1$: 8 positive and 0 negative examples.
- $A = a_2$: 4 positive and 8 negative examples.
- $A = a_3$: 2 positive and 16 negative examples.

What is the information gain? (You don't need to reduce the answer to a single number.) **Hint:** It may help to draw a tree with attribute A at the root, and then try to figure it out.

5 Natural Language Processing (9pts)

Question 18 (9 pts): What is the relationship between/among the (a) serial nature of vocal communication, (b) effeciency of language, and (c) ambiguity of language?

Appendix

Note: There is no exam question on this page.

Logic:

- $P \vee Q = Q \vee P$,
  $P \wedge Q = Q \wedge P$ (commutative)
- $(P \vee Q) \vee H = P \vee (Q \vee H)$,
  $(P \wedge Q) \wedge H = P \wedge (Q \wedge H)$,(associative)
- $P \vee (Q \wedge H) = (P \vee Q) \wedge (P \vee H)$,
  $P \wedge (Q \vee H) = (P \wedge Q) \vee (P \wedge H)$ (distributive)
- $P \vee \mathbf{F} = P, P \wedge \mathbf{F} = \mathbf{F}$ (**F**: False)
- $P \vee \mathbf{T} = \mathbf{T}$
  $P \wedge \mathbf{T} = P$ (**T**: True)
- $P \vee \neg P = \mathbf{T}$
  $P \wedge \neg P = \mathbf{F}$
- $\neg(P \vee Q) = \neg P \wedge \neg Q$,
  $\neg(P \wedge Q) = \neg P \vee \neg Q$ (DeMorgan's law)
- $P \rightarrow Q = \neg Q \rightarrow \neg P$ (contrapositive)
- $P \rightarrow Q = \neg P \vee Q$
- $(Qx, F(x)) \vee G = Qx, (F(x) \vee G)$
  $(Qx, F(x)) \wedge G = Qx, (F(x) \wedge G)$
- $\neg(\forall x, F(x)) = \exists x, (\neg F(x))$
  $\neg(\exists x, F(x)) = \forall x, (\neg F(x))$
- $(\forall x, F(x)) \wedge (\forall x, G(x)) = \forall x, (F(x) \wedge G(x))$
  $(\exists x, F(x)) \vee (\exists x, G(x)) = \exists x, (F(x) \vee G(x))$
- $(Q_1 x, F(x)) \vee (Q_2 x, H(x)) = Q_1 x, Q_2 z, (F(x) \vee H(z))$
  $(Q_1 x, F(x)) \wedge (Q_2 x, H(x)) = Q_1 x, Q_2 z, (F(x) \wedge H(z))$

Probability (for Boolean random variables):

1. $P(A, B) = P(A \wedge B) = P(A = \mathbf{True} \wedge B = \mathbf{True})$
2. $P(A|B) = \frac{P(A,B)}{P(B)}$
3. $P(A, B) = P(A|B)P(B)$
4. Bayes rule: $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$
5. Condition for independence: $P(A, B) = P(A)P(B)$
6. $P(A, B|C) = P(\underbrace{A, B}|C)$, **not** $P(A, \underbrace{B|C})$