CPSC 625-600 Final Exam (11/7/2007, Wed)[1]

Last name: _______________, First name: _______________

Time: **12:40pm–1:30pm (2 hours)**, Total Points: **100**

Subject	Score
FOL	/30
Uncertainty and Probabilistic reasoning	/30
Learning	/40
Total	/100

- You may use a calculator.
- You may use the back of the sheet, but please **prominently mark** on the front in such a case.
- 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*).
- Solve all problems.
- Total of _7_ pages, including this cover and the Appendix at the end. **Before starting, count the pages and see if you have all _7_.**
- 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

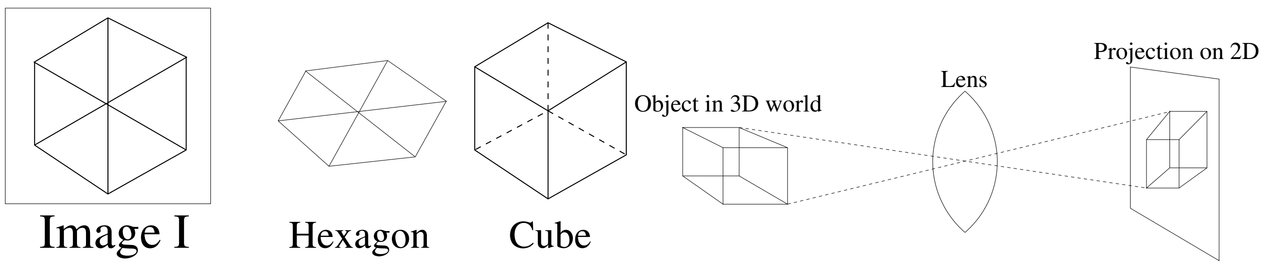
Question 1 (10 pts): (1) Resolution is domain (independent, dependent): **circle one**. (2) Why is this both an advantage and also a disadvantage?

Question 2 (20 pts): Show that $\exists x(D(x) \wedge C(x))$ is a logical consequence of the following, using **resolution**.

1. $\neg E(x) \vee V(x) \vee S(x, f(x))$
2. $\neg E(x) \vee V(x) \vee C(f(x))$
3. $E(a)$
4. $D(a)$
5. $\neg S(a, y) \vee D(y)$
6. $\neg D(x) \vee \neg V(x)$

2 Uncertainty and Probabilistic Reasoning

Question 3 (15 pts): Consider the vision problem below. Given an image I, you want to know what was the object in the environment that gave rise to the image. Let's say there are 2D hexagons and 3D cubes.



(1) Why is $P(Image|Object)$ easier to obtain than $P(Object|Image)$? Explain in terms of Cause and Effect. (2) If cubes and hexgons are encountered with an equal probability (i.e., $P(Object = Cube) = P(Object = Hexagon)$), what would be the relationship between $P(Object = Cube|Image = I)$ and $P(Object = Hexagon|Image = I)$? Choose among $>, <, =$ and explain why. Use the Bayes rule.

Question 4 (15 pts): Explain why you can calculate the joint probability of events in a belief network as $P(X_1 = v_1, ..., X_n = v_n) = \prod_{i=1}^{N} P(X_i = v_i | \forall X_j = Parent(X_i), X_j = v_j)$. Show your derivation, and state your assumptions. (Hint: Repeatedly apply axiom 3 in the probability cheat sheet in the end.)

3 Learning

Question 5 (14 pts): In decision tree learning, the **Choose-Attribute** function uses the information gain measure to select the root node for a subtree during the construction process. (1) How is this process related to the inductive bias in decision tree learning and (2) how does information gain relate to the degree of uncertainty before and after a particular attribute has been tested?

Question 6 (10 pts): If you have a random variable $Weather \in \{Sunny, Cloudy, Rainy\}$, and the probability of each is given $P(Weather = Sunny)$, $P(Weather = Cloudy)$, and $P(Weather = Rainy)$, how can you calculate the entropy of $Weather$? (1) Write the formula (**don't** use the summation symbol $\sum$), and (2) explain under what condition the entropy would be maximum, in terms of the three probabilities (hint: think of entropy as average uncertainty, and think about when that uncertainty would be maximum).

Question 7 (16 pts): (1) What does a single perceptron unit represent, with its threshold and connection weights, in geometric space? (2) Why does this geometric interpretation give insights on the kind of problems learnable or not by perceptrons?

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** $P(A, \underbrace{B|C})$