# CPSC 420-500 Homework #2
# SOLUTION

Total: 130 pts

Yoonsuck Choe

November 1, 2007

## 1 First-Order Logic

**Important:** In this section, assume that $w, x, y, z$ are variables; $A, B, C, D$ are constants; and $f(\cdot), g(\cdot), h(\cdot)$ are functions; and $P(\cdot), Q(\cdot), R(\cdot)$ are predicates.

### 1.1 Standard Forms

To do automatic theorem proving in first-order logic, you need to go through three steps to convert your initial first-order logic expression into a standard form. These are:

1. Prenex normal form,
2. Conjunctive normal form, and
3. Skolemization.

**Question 1 (12 pts):** Convert to prenex normal form (4 points each):

1. $\forall x, \neg(\exists y, \neg P(x,y))$
2. $\neg \forall x\, (P(x) \vee \neg(\exists y, \neg Q(x,y)))$
3. $\neg \forall x\, (\exists y, Q(x,y) \rightarrow \neg P(x))$

**SOLUTION:**

1. $\forall x, \neg(\exists y, \neg P(x,y))$
   $\forall x, (\forall y, \neg\neg P(x,y))$
   $\forall x, (\forall y, P(x,y))$
   $\forall x, \forall y, P(x,y)$
2. $\neg \forall x\, (P(x) \vee \neg(\exists y, \neg Q(x,y)))$
   $\exists x \neg\, (P(x) \vee \neg(\exists y, \neg Q(x,y)))$
   $\exists x\, (\neg P(x) \wedge \neg\neg(\exists y, \neg Q(x,y)))$
   $\exists x\, (\neg P(x) \wedge (\exists y, \neg Q(x,y)))$
   $\exists x, \exists y, \neg P(x) \wedge \neg Q(x,y)$

3. $\neg\forall x\,(\exists y, Q(x,y) \rightarrow \neg P(x))$
   $\exists x \neg\,(\neg(\exists y, Q(x,y)) \vee \neg P(x))$
   $\exists x\,(\neg\neg(\exists y, Q(x,y)) \wedge \neg\neg P(x))$
   $\exists x\,(\exists y, Q(x,y) \wedge P(x))$
   $\exists x, \exists y, Q(x,y) \wedge P(x)$

**Question 2 (20 pts):** Skolemize the expressions (4 points each):

1. $\exists x P(x)$
2. $\forall x \exists y P(x,y)$
3. $\exists x, \exists y, \forall z P(x,y) \wedge Q(y,z)$
4. $\forall x, \exists y, \exists z P(x,y) \wedge Q(y,z)$
5. $\forall x, \forall y, \exists z P(x,y) \wedge Q(y,z)$

**SOLUTION:**

1. $P(A)$
2. $P(x, f(x))$
3. $P(A,B) \wedge Q(B,z)$
4. $P(x, f(x)) \wedge Q(f(x), g(x))$
   The question was ambiguous, so you will get full points for this one if your answer was reasonable. The qeustion was supposed to be $\forall x, \exists y, \exists z P(x,y) \wedge Q(y,z)$
5. $P(x,y) \wedge Q(y, f(x,y))$

**Question 3 (9 pts):** Convert the following into a standard form:

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

**SOLUTION:**

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

$$\forall x, [\neg\neg P(x) \vee \neg(\exists y, Q(x,y))]$$

$$\forall x, [P(x) \vee (\forall y, \neg Q(x,y))]$$

$$\forall x, \forall y, [P(x) \vee \neg Q(x,y)]$$

$$P(x) \vee \neg Q(x,y)$$

## 1.2 Substitution and Unification

**Question 1 (9 pts):** Apply the following substitutions to the expressions (3 point each);

1. Apply $\{x/f(A)\}$ to $P(x, y) \vee Q(x)$.
2. Apply $\{x/A, y/f(z)\}$ to $P(x, y) \vee Q(x)$.
3. Apply $\{y/x\}$ to $P(x, y) \vee Q(x)$.

**SOLUTION:**

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

**Question 2 (16 pts):** For each of the following, (1) find the unifier, and (2) show the unified expression. For example, given $P(A)$ and $P(x)$, the unifier would be $\{x/A\}$, and the unified expression $P(A)$. If the pair of expressions is not unifiable, indicate so. (4 points each):

1. $P(x, f(B))$ and $P(A, f(y))$
2. $P(x, f(A))$ and $P(y, y)$
3. $P(x, f(y), y)$ and $P(A, f(g(w)), g(A))$
4. $P(A, f(y), y, A)$ and $P(x, f(g(x)), g(B), w)$

**SOLUTION:**

1. $\{x/A, y/B\}, P(A, f(B))$
2. $\{x/f(A), y/f(A)\}, P(f(A), f(A))$
   Note that initially you have $\{x/y\}$, but the $y$ on the right eventually gets replaced by $f(A)$.
3. $\{x/A, y/g(A), w/A\}, P(A, f(g(A)), g(A))$
4. $\{x/A, y/g(A), ...\}$ Cannot proceed any further since you cannot unify $g(A)$ and $g(B)$, or more precisely, you cannot unify $A$ and $B$.

**Question 3 (20 pts):** Show that $R(v)$ is a logical consequence of the following. Use **resolution**. Turn into a normal form as necessary.

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

**SOLUTION:**

1. $\forall x, \forall y, (\neg P(x) \rightarrow (Q(x,y) \vee R(y)))$
   $\forall x, \forall y, (\neg\neg P(x) \vee (Q(x,y) \vee R(y)))$
   $\forall x, \forall y, (P(x) \vee Q(x,y) \vee R(y))$
   $P(x) \vee Q(x,y) \vee R(y)$
2. $\exists x, \neg P(x)$
   $\neg P(A)$
3. $\forall w, \forall z, (\neg Q(w,z) \vee R(w))$
   $\neg Q(w,z) \vee R(w)$
4. Negated conclusion: $\neg R(v)$
5. $Q(A,y) \vee R(y)$ (1 and 2), $\sigma = \{x/A\}$
6. $Q(A,y)$ (4 and 5), $\sigma = \{v/y\}$
7. $\neg Q(w,z)$ (3 and 4), $\sigma = \{v/w\}$
8. False (6 and 7), $\sigma = \{w/A, y/z\}$

# 2 Uncertainty and Probabilistic Reasoning

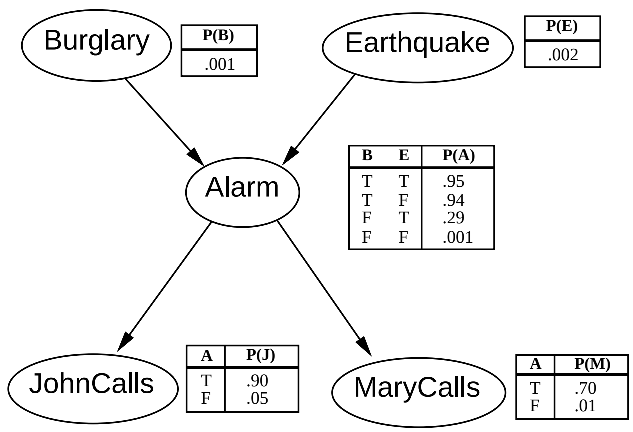


Figure 1: **Belief Network.** See problem 1.

**Question 1 (10 pts):** Given the Belief network as shown in figure 1, calculate the two joint probability values and answer the question. Note that in this section $P(\cdot)$ denotes the probability of the event. (5 points each):

1. $P(MaryCalls \wedge JohnCalls \wedge Alarm \wedge \neg Earthquake \wedge Burglary)$
2. $P(\neg MaryCalls \wedge \neg JohnCalls \wedge \neg Alarm \wedge \neg Earthquake \wedge Burglary)$

**SOLUTION:**

1. $P(MaryCalls \wedge JohnCalls \wedge Alarm \wedge \neg Earthquake \wedge Burglary)$
   $= P(M|A)P(J|A)P(A|\neg E, B)P(\neg E)P(B)$
   $= 0.7 \times 0.9 \times 0.94 \times (1 - 0.002) \times 0.001$

2. $P(\neg MaryCalls \wedge \neg JohnCalls \wedge \neg Alarm \wedge \neg Earthquake \wedge Burglary)$
$= P(\neg M|\neg A)P(\neg J|\neg A)P(\neg A|\neg E, B)P(\neg E)P(B)$
$= (1 - 0.01) \times (1 - 0.05) \times (1 - 0.94) \times (1 - 0.002) \times 0.001$

**Question 2 (5 pts):** Why do belief networks give a much more compact representation of the joint probability distribution, compared to a full joint probability table?

**SOLUTION:** Because the joint probability can be represented as a product of conditional probabilities where each node is conditioned on only a small number of other nodes (their parents) in the network due to the conditional independence structure of the domain, and thus much smaller probability tables for each node.

# 3 Learning

## 3.1 Decision Tree Learning

Consider the following set of examples where you are trying to make a decision whether to take a course or not.

| Example# | Usefulness in life | Toughness | Fun | Decision (Take course?) |
|---|---|---|---|---|
| 1 | Stellar | Okay | Medium | Yes |
| 2 | Not bad | Very tough | High | Yes |
| 3 | Zero | Light | Medium | No |
| 4 | Not bad | Light | Low | No |
| 5 | Not bad | Light | High | Yes |
| 6 | Not bad | Very tough | Low | No |
| 7 | Zero | Okay | Medium | No |
| 8 | Stellar | Very tough | Low | Yes |
| 9 | Stellar | Okay | Medium | Yes |
| 10 | Zero | Light | High | No |

**Question 1 (12 pts):** For each of the three attributes above, draw a decision tree rooted at that attribute with a **single depth**. See slide06, page 12, $(a)$ and $(b)$ which show some examples. (4 points each)

**SOLUTION:**

1. Usefulness
   Stellar: (+: 1, 8, 9) (-: None)
   Not bad: (+: 2, 5) (-: 4, 6)
   Zero: (+: None) (-: 3, 7, 10)
2. Toughness
   Okay: (+: 1, 9) (-: 7)
   Very tough: (+: 2, 8) (-: 6)
   Light: (+: 5) (-: 3, 4, 10)
3. Fun
   Low: (+: 8) (-: 4, 6)
   Medium: (+: 1, 9) (-: 3, 7)
   High: (+: 2, 5) (-: 10)

**Question 2 (12 pts):** Calculate the information gain for each of the three attributes. (4 points each)

**SOLUTION:**

1. $Ent(1/2, 1/2) - (3/10 \times Ent(3/3, 0/3) + 4/10 \times Ent(1/2, 1/2) + 3/10 \times Ent(0/3, 3/3))$
   $= 1 - 0.4 = 0.6$
   Note: Let's define $Ent(x, y) = -x \log_2 x - y \log_2 y$.
2. $Ent(1/2, 1/2) - (3/10 \times Ent(2/3, 1/3) + 3/10 \times Ent(2/3, 1/3) + 4/10 \times Ent(1/4, 3/4))$
   $= 1 - 0.87549 = 0.12451$
3. $Ent(1/2, 1/2) - (3/10 \times Ent(1/3, 2/3) + 4/10 \times Ent(2/2, 2/2) + 3/10 \times Ent(2/3, 1/3))$
   $= 1 - 0.95098 = 0.049022$

**Question 3 (5 pts):** If you are supposed to choose from the three attributes for the first test in a decision tree construction, which one would you choose and why? **Do not** base your answer on your personal preference.

**SOLUTION:** Usefulness, since it results in the highest information gain.