Common Mistakes

- Unification
- Factors
- English to FOL
- Normal forms
- Resolution
- Probability
- Partition of sample space
- $P(D|T)$ vs. $P(T|D)$
- Belief network

Unification

- These are constants: $a$, $b$, $c$, ...
- Be very careful to distinguish between the two.

Unification (cont’d)

- These are variables: $x$, $y$, $z$, ...
- Substitution notation: $\{variable/term\}$, replace $variable$ with $term$.

Unification (cont’d)

- Variable is also a term, but the other way around is not always true (constant or function is not a variable)! Thus, $\{variable_1/variable_2\}$ is allowed, because $variable_2$ is a term, as well as a variable.

Unification (cont’d)

- You cannot replace a constant:
  Unify $P(a)$ and $P(x)$: $\{a/x\}$ is NOT allowed because $a$ is not a variable!

- You cannot replace a function, whether or not it contains a variable:
  - Unify $P(f(a))$ and $P(x)$: $\{f(a)/x\}$ is NOT allowed because $f(a)$ is not a variable!
  - Unify $P(f(x))$ and $P(y)$: $\{f(x)/y\}$ is NOT allowed because $f(x)$ is not a variable, even though it contains a variable $x$!

Unification (cont’d)

- When finding a unifier by going from left to right, the current substitution should be applied to the previous one. For example, unify $C_1 = P(f(x), x)$ and $C_2 = P(z, a)$
  - After first step, unifier is $\sigma_1 = \{z/f(x)\}$, and the resulting predicates are $C_1\sigma_1 = P(f(x), x)$ and $C_2\sigma_1 = P(f(x), a)$.
  - After second step, unifier is $\sigma_2 = \{x/a\}$, and thus the full unifier is
    $$\sigma = \sigma_1 \circ \sigma_2 = \{z/f(x)\sigma_2, x/a\} = \{z/f(x)\{x/a\}, x/a\} = \{z/f(a), x/a\}.$$

1  2  3  4
Unification (cont’d)

- Substitution \( \{ x / f(x) \} \) is a valid substitution, but it **cannot** be used in a unifier.

In the unification algorithm, you need to find variable \( v \) and term \( t \) in the disagreement set where \( t \) does not contain \( v \)!

- Example: unify \( P(a, x, f(g(y))) \) and \( P(z, f(z), f(u)) \):
  answer: \( \{ z/a, x/f(a), u/g(y) \} \).

Factors

- **Definition:** If **two or more** literals of a clause \( C \) (with the same sign) have a most general unifier \( \sigma \), then \( C\sigma \) is called a **Factor** of \( C \).

- Thus, if you have \( P(a) \lor P(b) \lor P(x) \), it has a factor \( P(a) \lor P(b) \) where the most general unifier can be either \( \{ x/a \} \) or \( \{ x/b \} \).

English to FOL: Common Patterns

5. Beware of different uses of **Any**:

   From WordNet (r) 1.7 [wn]:

   **any**

   adj 1: one or some or every or all without specification; "give me any peaches you don't want"; "not any milk is left"; "any child would know that"; "pick any card"; "any day now"; "cars can be rented at almost any airport"; "at twilight or any other time"; "beyond any doubt"; "need any help we can get"; "give me whatever peaches you don’t want"; "no milk whatsoever is left" [syn: {any(a)}, {whatev er}, {whatsoever}]

   (a) Any bird can fly:
   \( \forall x (Bird(x) \rightarrow Fly(x)) \)

   (b) Any student who has any book is smart:
   \( \forall x \left( \text{student}(x) \land \exists y \left( \text{have}(x, y) \land \text{book}(y) \right) \rightarrow \text{smart}(x) \right) \)
Normal Forms

- Prenex normal form:
  - Move all negations inside:
    Change $\neg \forall x(M(x))$ to $\exists x(\neg M(x))$.
    Change $\neg \exists x(M(x))$ to $\forall x(\neg M(x))$.

- Conjunctive normal form:
  - First, remove all implications ($\rightarrow$).
  - Then, move all negations on parentheses inside.
  - If the connectives are the same, you can parenthesize them arbitrarily (this includes removing the parentheses as well):
    \[
    (P(x) \lor Q(y)) \lor R(z) = P(x) \lor Q(y) \lor R(z) = P(x) \lor (Q(y) \lor R(z)) = \ldots
    \]
    So, $(P(x) \lor Q(y)) \lor R(z)$ is already a clause!

Normal Forms (cont'd)

- $P(x) \lor Q(y) \lor R(z)$ is both in conjunctive normal form and also at the same time in disjunctive normal form. How come?:
  - CNF $\left( \underbrace{\ldots \lor \ldots} \right) \land \left( \underbrace{\ldots \lor \ldots} \right) \land \left( \underbrace{\ldots \lor \ldots} \right)$: it is a single clause
  - DNF $\left( \underbrace{\ldots \land \ldots} \right) \lor \left( \underbrace{\ldots \land \ldots} \right) \lor \left( \underbrace{\ldots \land \ldots} \right)$: it has three terms

Resolution

- You MUST add the negated conclusion to the list of axioms before beginning.

- You are done when you derive False.

Probability

- When $X$ and $Y$ are boolean random variables,
  \[
  P(A, B) = P(A \land B)
  \]

- When $X$ and $Y$ are boolean random variables,
  \[
  P(A \land B) = P(A = \text{True} \land B = \text{True})
  \]
  \[
  P(\neg A \land \neg B) = P(A = \text{False} \land B = \text{False})
  \]

- Conditional probabilities:
  \[
  P(A|B, C) = \frac{P(A, B, C)}{P(B, C)}
  \]
  \[
  P(A, B|C) = \frac{P(A, B, C)}{P(C)}
  \]
Total Probability

\[ P(T) = \sum_{i=1}^{n} P(T \land X_i) = \sum_{i=1}^{n} P(T \mid X_i) P(X_i) \]

If \( \sum_{i=1}^{n} P(X_i) = 1 \),

\[ P(T) = \sum_{i=1}^{n} P(T \land X_i) = \sum_{i=1}^{n} P(T \mid X_i) P(X_i) \]

Belief Network: Representing Joint Prob. Dist. (cont’d)

\[ P(X_1 = x_1, \ldots, X_n = x_n) = P(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid Parent(X_i)) \]

\( Parent(X_i) \) refers to the event when \( X_j = x_j \) for parents \( X_j \) of \( X_i \).

Imagine a case where each node in the example has a T or F assignment. \( X_j \) will then be either T or F for all \( j \).

The belief network fully defines a joint probability distribution!

\[ P(D \mid T) \text{ vs. } P(T \mid D) \]

- \( P(T \mid D) = 0.9 \)
  - A test is 90% accurate
  - a test returns positive 90% of the time when tested on those with the disease
  - if you have the disease, the chance of your test results returning positive is 90%.

- \( P(D \mid T) = 0.9 \)
  - if you tested positive on a test, the chance of you having the disease is 90%.

Calculating Probability of a Joint Event

\[ \text{Calculate the probability of the event that the alarm (A) has sounded but neither a burglary (¬B) nor an earthquake (¬E) occurred, and both John (J) and Mary (M) call:} \]

\[ P(J \land M \land A \land ¬B \land ¬E) = P(J \mid Parent(J)) P(M \mid Parent(M)) P(A \mid Parent(A)) P(¬B) P(¬E) \]

\[ = P(J \mid A) P(M \mid A) P(A \mid ¬B \land ¬E) P(¬B) P(¬E) \]

\textbf{Note that} \( \text{Parents}(A) \) \textbf{in this case is not} \( B \land E \) \textbf{but} \( ¬B \land ¬E \) \textbf{because} \( Parent(X_i) \) \textbf{refers to the event when} \( X_j = x_j \) \textbf{for parents} \( X_j \) \textbf{of} \( X_i \).
Calculating Probability of a Joint Event (cont’d)

\[
P(J \mid A) P(M \mid A) P(\neg B \land \neg E) P(\neg B) P(\neg E)
\]
\[
= 0.9 \times 0.7 \times 0.001 \times 0.999 \times 0.998 = 0.00062
\]