Skolemization, Most General Unifiers, First-Order Resolution

Torsten Hahmann

CSC 384, University of Toronto

March 07, 2011
Skolemization

Conversion of sentences FOL to CNF requires skolemization.

Skolemization: remove existential quantifiers by introducing new function symbols.

How: For each existentially quantified variable introduce a n-place function where n is the number of previously appearing universal quantifiers.

Special case: introducing constants (trivial functions: no previous universal quantifier).
Every philosopher writes at least one book.
\[\forall x [\text{Philo}(x) \rightarrow \exists y [\text{Book}(y) \land \text{Write}(x, y)]]\]
Skolemization - Example 1

- Every philosopher writes at least one book.
  \( \forall x [\text{Philo}(x) \rightarrow \exists y [\text{Book}(y) \land \text{Write}(x, y)]] \)

- Eliminate Implication:
  \( \forall x [\neg \text{Philo}(x) \lor \exists y [\text{Book}(y) \land \text{Write}(x, y)]] \)
Every philosopher writes at least one book.
\[ \forall x[\text{Philo}(x) \rightarrow \exists y[\text{Book}(y) \land \text{Write}(x, y)]] \]

Eliminate Implication:
\[ \forall x[\neg \text{Philo}(x) \lor \exists y[\text{Book}(y) \land \text{Write}(x, y)]] \]

Skolemize: substitute \( y \) by \( g(x) \)
\[ \forall x[\neg \text{Philo}(x) \lor \text{Book}(g(x)) \land \text{Write}(x, g(x))]] \]
Skolemization - Example 2

- All students of a philosopher read one of their teacher’s books.
  \[ \forall x \forall y [\text{Philo}(x) \land \text{StudentOf}(y, x) \rightarrow \exists z [\text{Book}(z) \land \text{Write}(x, z) \land \text{Read}(y, z)]] \]
Skolemization - Example 2

All students of a philosopher read one of their teacher’s books.
\(\forall x \forall y [\text{Philo}(x) \land \text{StudentOf}(y, x) \rightarrow \exists z [\text{Book}(z) \land \text{Write}(x, z) \land \text{Read}(y, z)]]\)

Eliminate Implication:
\(\forall x \forall y [\neg \text{Philo}(x) \lor \neg \text{StudentOf}(y, x) \lor \exists z [\text{Book}(z) \land \text{Write}(x, z) \land \text{Read}(y, z)]]\)
Skolemization - Example 2

- All students of a philosopher read one of their teacher’s books.
  \[\forall x \forall y [\text{Philo}(x) \land \text{StudentOf}(y, x) \rightarrow \exists z [\text{Book}(z) \land \text{Write}(x, z) \land \text{Read}(y, z)]]\]

- Eliminate Implication:
  \[\forall x \forall y [\neg \text{Philo}(x) \lor \neg \text{StudentOf}(y, x) \lor \exists z [\text{Book}(z) \land \text{Write}(x, z) \land \text{Read}(y, z)]]\]

- Skolemize: substitute \(z\) by \(h(x, y)\)
  \[\forall x \forall y [\neg \text{Philo}(x) \lor \neg \text{StudentOf}(y, x) \lor \exists z [\text{Book}(h(x, y)) \land \text{Write}(x, h(x, y)) \land \text{Read}(y, h(x, y))]]\]
There exists a philosopher with students.
\[ \exists x \exists y [\text{Philo}(x) \land \text{StudentOf}(y, x)] \]
There exists a philosopher with students.
\[ \exists x \exists y [\text{Philo}(x) \land \text{StudentOf}(y, x)] \]

Skolemize: substitute \( x \) by \( a \) and \( y \) by \( b \)
\[ \text{Philo}(a) \land \text{StudentOf}(b, a) \]
Most General Unifier

Least specialized unification of two clauses.

We can compute the MGU using the disagreement set $D_k = \{e_1, e_2\}$: the pair of expressions where two clauses first disagree.

REPEAT UNTIL no more disagreement $\rightarrow$ found MGU.

IF either $e_1$ or $e_2$ is a variable $V$ and the other is some term (or a variable) $t$, then choose $V = t$ as substitution.

Then substitute to obtain $S_{k+1}$ and find disagreement set $D_{k+1}$.

ELSE unification is not possible.
Find the MGU of \( p(f(a), g(X)) \) and \( p(Y, Y) \):

\[
S_0 = \{ p(f(a), g(X)) ; p(Y, Y) \}
\]
Find the MGU of $p(f(a), g(X))$ and $p(Y, Y)$:

- $S_0 = \{ p(f(a), g(X)) ; p(Y, Y) \}$
- $D_0 = \{ f(a), Y \}$
- $\sigma = \{ Y = f(a) \}$

no unification possible!
Find the MGU of \( p(f(a), g(X)) \) and \( p(Y, Y) \):

- \( S_0 = \{ p(f(a), g(X)) ; p(Y, Y) \} \)
- \( D_0 = \{ f(a), Y \} \)
- \( \sigma = \{ Y = f(a) \} \)
- \( S_1 = \{ p(f(a), g(X)) ; p(f(a), f(a)) \} \)

no unification possible!
Find the MGU of \( p(f(a), g(X)) \) and \( p(Y, Y) \):

- \( S_0 = \{ p(f(a), g(X)) ; p(Y, Y) \} \)
- \( D_0 = \{ f(a), Y \} \)
- \( \sigma = \{ Y = f(a) \} \)
- \( S_1 = \{ p(f(a), g(X)) ; p(f(a), f(a)) \} \)
- \( D_1 = \{ g(X), f(a) \} \)

\( \text{no unification possible!} \)
Find the MGU of \( p(f(a), g(X)) \) and \( p(Y, Y) \):

- \( S_0 = \{ p(f(a), g(X)), p(Y, Y) \} \)
- \( D_0 = \{ f(a), Y \} \)
- \( \sigma = \{ Y = f(a) \} \)
- \( S_1 = \{ p(f(a), g(X)), p(f(a), f(a)) \} \)
- \( D_1 = \{ g(X), f(a) \} \)
- no unification possible!
MGU - Example 2

- $S_0 = \{ p(a, X, h(g(Z))) ; p(Z, h(Y), h(Y)) \}$
MGU - Example 2

- \( S_0 = \{ p(a, X, h(g(Z))) ; p(Z, h(Y), h(Y)) \} \)
- \( D_0 = \{ a, Z \} \)
MGU - Example 2

- $S_0 = \{p(a, X, h(g(Z))) ; p(Z, h(Y), h(Y))\}$
- $D_0 = \{a, Z\}$
- $\sigma = \{Z = a\}$
MGU - Example 2

- $S_0 = \{ p(a, X, h(g(Z))) ; p(Z, h(Y), h(Y)) \}$
- $D_0 = \{ a, Z \}$
- $\sigma = \{ Z = a \}$
- $S_1 = \{ p(a, X, h(g(a))) ; p(a, h(Y), h(Y)) \}$
MGU - Example 2

\[ S_0 = \{ p(a, X, h(g(Z))) ; p(Z, h(Y), h(Y)) \} \]
\[ D_0 = \{ a, Z \} \]
\[ \sigma = \{ Z = a \} \]
\[ S_1 = \{ p(a, X, h(g(a))) ; p(a, h(Y), h(Y)) \} \]
\[ D_1 = \{ X, h(Y) \} \]
MGU - Example 2

- \( S_0 = \{ p(a, X, h(g(Z))) ; p(Z, h(Y), h(Y)) \} \)
- \( D_0 = \{ a, Z \} \)
- \( \sigma = \{ Z = a \} \)
- \( S_1 = \{ p(a, X, h(g(a))) ; p(a, h(Y), h(Y)) \} \)
- \( D_1 = \{ X, h(Y) \} \)
- \( \sigma = \{ Z = a, X = h(Y) \} \)
MGU - Example 2

- $S_0 = \{ p(a, X, h(g(Z))) ; p(Z, h(Y), h(Y)) \}$
- $D_0 = \{ a, Z \}$
- $\sigma = \{ Z = a \}$
- $S_1 = \{ p(a, X, h(g(a))) ; p(a, h(Y), h(Y)) \}$
- $D_1 = \{ X, h(Y) \}$
- $\sigma = \{ Z = a, X = h(Y) \}$
- $S_2 = \{ p(a, h(Y), h(g(a))) ; p(a, h(Y), h(Y)) \}$
MGU - Example 2

- $S_0 = \{ p(a, X, h(g(Z))) ; p(Z, h(Y), h(Y)) \}$
- $D_0 = \{ a, Z \}$
- $\sigma = \{ Z = a \}$
- $S_1 = \{ p(a, X, h(g(a))) ; p(a, h(Y), h(Y)) \}$
- $D_1 = \{ X, h(Y) \}$
- $\sigma = \{ Z = a, X = h(Y) \}$
- $S_2 = \{ p(a, h(Y), h(g(a))) ; p(a, h(Y), h(Y)) \}$
- $D_2 = \{ g(a), Y \}$
MGU - Example 2

- $S_0 = \{ p(a, X, h(g(Z))) ; p(Z, h(Y), h(Y)) \}$
- $D_0 = \{ a, Z \}$
- $\sigma = \{ Z = a \}$
- $S_1 = \{ p(a, X, h(g(a))) ; p(a, h(Y), h(Y)) \}$
- $D_1 = \{ X, h(Y) \}$
- $\sigma = \{ Z = a, X = h(Y) \}$
- $S_2 = \{ p(a, h(Y), h(g(a))) ; p(a, h(Y), h(Y)) \}$
- $D_2 = \{ g(a), Y \}$
- $\sigma = \{ Z = a, X = h(Y), Y = g(a) \}$

No disagreement $\Rightarrow \sigma = \{ Z = a, X = h(Y), Y = g(a) \}$ is MGU
MGU - Example 2

- $S_0 = \{p(a, X, h(g(Z))) ; p(Z, h(Y), h(Y))\}$
- $D_0 = \{a, Z\}$
- $\sigma = \{Z = a\}$
- $S_1 = \{p(a, X, h(g(a))) ; p(a, h(Y), h(Y))\}$
- $D_1 = \{X, h(Y)\}$
- $\sigma = \{Z = a, X = h(Y)\}$
- $S_2 = \{p(a, h(Y), h(g(a))) ; p(a, h(Y), h(Y))\}$
- $D_2 = \{g(a), Y\}$
- $\sigma = \{Z = a, X = h(Y), Y = g(a)\}$
- $S_3 = \{p(a, h(g(a)), h(g(a))) ; p(a, h(g(a)), h(g(a)))\}$

$\Rightarrow \sigma = \{Z = a, X = h(Y), Y = g(a)\}$ is MGU
MGU - Example 2

\( S_0 = \{ p(a, X, h(g(Z))) ; \ p(Z, h(Y), h(Y)) \} \)
\( D_0 = \{ a, Z \} \)
\( \sigma = \{ Z = a \} \)
\( S_1 = \{ p(a, X, h(g(a))) ; \ p(a, h(Y), h(Y)) \} \)
\( D_1 = \{ X, h(Y) \} \)
\( \sigma = \{ Z = a, X = h(Y) \} \)
\( S_2 = \{ p(a, h(Y), h(g(a))) ; \ p(a, h(Y), h(Y)) \} \)
\( D_2 = \{ g(a), Y \} \)
\( \sigma = \{ Z = a, X = h(Y), Y = g(a) \} \)
\( S_3 = \{ p(a, h(g(a)), h(g(a))) ; \ p(a, h(g(a)), h(g(a))) \} \)
\( \text{No disagreement} \)
\( \Rightarrow \sigma = \{ Z = a, X = h(Y), Y = g(a) \} \) is MGU
$S_0 = \{p(X, X) ; p(Y, f(Y))\}$
MGU - Example 3

- $S_0 = \{p(X, X) ; p(Y, f(Y))\}$
- $D_0 = \{X, Y\}$

no unification possible!
$S_0 = \{ p(X, X) ; p(Y, f(Y)) \}$

$D_0 = \{ X, Y \}$

$\sigma = \{ X = Y \}$
MGU - Example 3

- $S_0 = \{ p(X, X) ; p(Y, f(Y)) \}$
- $D_0 = \{ X, Y \}$
- $\sigma = \{ X = Y \}$
- $S_1 = \{ p(Y, Y) ; p(Y, f(Y)) \}$

no unification possible!
MGU - Example 3

- $S_0 = \{p(X, X) ; p(Y, f(Y))\}$
- $D_0 = \{X, Y\}$
- $\sigma = \{X = Y\}$
- $S_1 = \{p(Y, Y) ; p(Y, f(Y))\}$
- $D_1 = \{Y, f(Y)\}$
MGU - Example 3

- $S_0 = \{p(X, X) ; p(Y, f(Y))\}$
- $D_0 = \{X, Y\}$
- $\sigma = \{X = Y\}$
- $S_1 = \{p(Y, Y) ; p(Y, f(Y))\}$
- $D_1 = \{Y, f(Y)\}$
- no unification possible!
Given the following sentences, answer the question ‘What is connected to the Galbraith building?’ using resolution with answer extraction:

Connected is a binary symmetric relation.

An object $X$ is part of another object $Y$ iff everything $X$ is connected to, $Y$ is also connected to.

Room GB221 is part of Galbraith building.

Room GB221 is connected to itself.
(a) Represent these sentences in first order logic.

Connected is a binary symmetric relation.

An object $X$ is part of another object $Y$ iff everything $X$ is connected to, $Y$ is also connected to.

Room GB221 is part of Galbraith building.

Room GB221 is connected to itself.
Full example problem - Representation in FOL

- Connected is a symmetric relation.
  \[(\forall X, Y) \text{connected}(X, Y) \supset \text{connected}(Y, X)\]
Full example problem - Representation in FOL

- An object $X$ is part of another object $Y$ iff everything $X$ is connected to, $Y$ is also connected to.

$(\forall X, Y) \ (part(X, Y) \equiv ((\forall Z) \ connected(Z, X) \supset connected(Z, Y)))$
Room GB221 is part of Galbraith building.

part(gb221, galbraith)
Full example problem - Representation in FOL

- Room GB221 is connected to itself.
  \(\text{connected}(gb221, gb221)\)
(b) Convert the formulas to clausal form. Indicate any Skolem functions or constants used.

\((\forall X, Y)\) \text{connected}(X, Y) \supset \text{connected}(Y, X)

\((\forall X, Y)\) \text{part}(X, Y) \equiv (\forall Z) \text{connected}(Z, X) \supset \text{connected}(Z, Y)

\text{part}(gb221, galbraith)

\text{connected}(gb221, gb221)
Full example problem - CNF Conversion

\[ (\forall X, Y) \text{connected}(X, Y) \supset \text{connected}(Y, X) \]
\[ [\neg \text{connected}(X, Y), \text{connected}(Y, X)] \]
Full example problem - CNF Conversion

\[
(\forall X, Y) \text{part}(X, Y) \equiv (\forall Z) \text{connected}(Z, X) \supset \text{connected}(Z, Y)
\]

\[\rightarrow: [\neg \text{part}(X, Y), \neg \text{connected}(Z, X), \text{connected}(Z, Y)]\]

\[\leftarrow: [\text{part}(X, Y), \text{connected}(f(X, Y), X)]\]

\[\text{[part}(X, Y), \neg \text{connected}(g(X, Y), Y)]\]
Full example problem - CNF Conversion

\[ (\forall X, Y) \text{connected}(X, Y) \supset \text{connected}(Y, X) \]

\[ \neg \text{connected}(X, Y), \neg \text{connected}(Y, X) \]

\[ (\forall X, Y) \text{part}(X, Y) \equiv (\forall Z) \text{connected}(Z, X) \supset \text{connected}(Z, Y) \rightarrow \]

\[ \neg \text{part}(X, Y), \neg \text{connected}(Z, X), \text{connected}(Z, Y) \]

\[ \neg \text{part}(X, Y), \neg \text{connected}(g(gb221, gb221), gb221) \]

\[ \text{part}(gb221, galbraith) \]

\[ \text{part}(gb221, galbraith) ]
Full example problem - CNF Conversion

▶ $\forall X, Y \quad \text{connected}(X, Y) \supset \text{connected}(Y, X)$

$\neg \text{connected}(X, Y), \neg \text{connected}(Y, X)$

▶ $\forall X, Y \quad \text{part}(X, Y) \equiv \forall Z \quad \text{connected}(Z, X) \supset \text{connected}(Z, Y)$

$\to: [\neg \text{part}(X, Y), \neg \text{connected}(g(X, Y), X), \text{connected}(g(X, Y), Y)]$  

$\leftarrow: [\text{part}(X, Y), \text{connected}(f(X, Y), X)]$

▶ $\text{part}(gb221, gb221)$

$[\text{connected}(gb221, gb221)]$

$\forall X, Y \quad \text{connected}(X, Y)$
(c) Convert the negation of the statement ‘What is connected to the Galbraith building?’ to clause form (using an answer literal).

- FOL: $(\exists X) \text{connected}(\text{galbraith}, X)$
(c) Convert the negation of the statement ‘What is connected to the Galbraith building?’ to clause form (using an answer literal).

- FOL: $(\exists X)\ connected(galbraith, X)$
- negate goal!! $(\neg \exists X)\ connected(galbraith, X)$
(c) Convert the negation of the statement ‘What is connected to the Galbraith building?’ to clause form (using an answer literal).

- FOL: \((\exists X) \, \text{connected}(\text{galbraith}, X)\)
- negate goal!! \((\neg \exists X) \, \text{connected}(\text{galbraith}, X)\)
- CNF with answer literal: \((\forall X) \, \neg \text{connected}(\text{galbraith}, X)\)
  \[\neg \text{connected}(\text{galbraith}, X), \text{ans}(X)\]
(d) Answer the question using resolution and answer extraction. Use the notation developed in class: every new clause must be labeled by the resolution step that was used to generate it. For example, a clause labeled $R[4c, 1d]x = a, y = f(b)$ means that it was generated by resolving literal $c$ of clause 4 against literal $d$ of clause 1, using the MGU $x = a, y = f(b)$.

Our clauses:

1. $[\neg connected(R, S), connected(S, R)]$
2. $[\neg part(T, U), \neg connected(V, T), connected(V, U)]$
3. $[part(W, X), connected(f(W, X), W)]$
4. $[part(Y, Z), \neg connected(g(Y, Z), Z)]$
5. $[part(gb221, galbraith)]$
6. $[connected(gb221, gb221)]$
7. $[\neg connected(galbraith, A), ans(A)]$
Full example problem - Resolution

1. \([-\text{connected}(R, S), \text{connected}(S, R)]\)
2. \([-\text{part}(T, U), -\text{connected}(V, T), \text{connected}(V, U)]\)
3. \([\text{part}(W, X), \text{connected}(f(W, X), W)]\)
4. \([\text{part}(Y, Z), -\text{connected}(g(Y, Z), Z)]\)
5. \([\text{part}(\text{gb}221, \text{galbraith})]\)
6. \([\text{connected}(\text{gb}221, \text{gb}221)]\)
7. \([-\text{connected}(\text{galbraith}, A), \text{ans}(A)]\)
Full example problem - Resolution

1. $[-\text{connected}(R, S), \text{connected}(S, R)]$
2. $[-\text{part}(T, U), -\text{connected}(V, T), \text{connected}(V, U)]$
3. $[\text{part}(W, X), \text{connected}(f(W, X), W)]$
4. $[\text{part}(Y, Z), -\text{connected}(g(Y, Z), Z)]$
5. $[\text{part}(gb221, \text{galbraith})]$
6. $[\text{connected}(gb221, gb221)]$
7. $[-\text{connected}(\text{galbraith}, A), \text{ans}(A)]$
8. $R[7a, 1b] \{ S = \text{galbraith}, R = U \}$
   $[-\text{connected}(A, \text{galbraith}), \text{ans}(A)]$
1. \([-\text{connected}(R, S), \text{connected}(S, R)]\]
2. \([-\text{part}(T, U), \neg\text{connected}(V, T), \text{connected}(V, U)]\]
3. \(\text{part}(W, X), \text{connected}(f(W, X), W)\]
4. \(\text{part}(Y, Z), \neg\text{connected}(g(Y, Z), Z)\]
5. \(\text{part}(gb221, \text{galbraith})\]
6. \(\text{connected}(gb221, gb221)\]
7. \([-\text{connected}(\text{galbraith}, A), \text{ans}(A)\]

8. R[7a, 1b] \{S = \text{galbraith}, R = U\}
   \([-\text{connected}(A, \text{galbraith}), \text{ans}(A)\]

9. R[8a, 2c] \{V = A, U = \text{galbraith}\}
   \([-\text{part}(T, \text{galbraith}), \neg\text{connected}(A, T), \text{ans}(A)\]
Full example problem - Resolution

1. \([-\text{connected}(R, S), \text{connected}(S, R)]\)
2. \([-\text{part}(T, U), -\text{connected}(V, T), \text{connected}(V, U)]\)
3. \([\text{part}(W, X), \text{connected}(f(W, X), W)]\)
4. \([\text{part}(Y, Z), -\text{connected}(g(Y, Z), Z)]\)
5. \([\text{part}(gb221, galbraith)]\)
6. \([\text{connected}(gb221, gb221)]\)
7. \([-\text{connected}(galbraith, A), \text{ans}(A)]\)

8. R[7a, 1b] \(\{S = galbraith, R = U\}\)

9. R[8a, 2c] \(\{V = A, U = galbraith\}\)

10. R[9a, 5] \(\{T = gb221\}\)
Full example problem - Resolution

1. $[\neg \text{connected}(R, S), \text{connected}(S, R)]$
2. $[\neg \text{part}(T, U), \neg \text{connected}(V, T), \text{connected}(V, U)]$
3. $[\text{part}(W, X), \text{connected}(f(W, X), W)]$
4. $[\text{part}(Y, Z), \neg \text{connected}(g(Y, Z), Z)]$
5. $[\text{part}(gb221, \text{galbraith})]$
6. $[\text{connected}(gb221, gb221)]$
7. $[\neg \text{connected}(\text{galbraith}, A), \text{ans}(A)]$

8. $R[7a, 1b] \{ S = \text{galbraith}, R = U \}$
   $[\neg \text{connected}(A, \text{galbraith}), \text{ans}(A)]$

9. $R[8a, 2c] \{ V = A, U = \text{galbraith} \}$
   $[\neg \text{part}(T, \text{galbraith}), \neg \text{connected}(A, T), \text{ans}(A)]$

10. $R[9a, 5] \{ T = gb221 \}$
    $[\neg \text{connected}(A, gb221), \text{ans}(A)]$

11. $R[10a, 6] \{ A = gb221 \}$
    $[\text{ans}(gb221)]$