CSE 541 - Logic in Computer Science
Solutions for Selected Problems on Skolemization, Unification, and Resolution

Prenex form.

A possible prenex form of

$$\neg\exists x\, ((\forall y\, \forall z\, P(y,z)) \wedge \neg P(x,z))$$

is

$$\forall x \exists y \exists z\, (\neg P(y,z) \vee P(x,z)).$$

Logical equivalence.

The two sentences $\forall x\, \exists y\, (P(x) \wedge Q(y))$ and $\exists y\, \forall x\, (P(x) \wedge Q(y))$

are equivalent, as the the following proof shows:

$$\begin{aligned}
&\forall x\, \exists y\, (P(x) \wedge Q(y)) \\
&\quad \sim \quad \forall x\, (P(x) \wedge \exists y\, Q(y)) \\
&\quad \sim \quad \forall x\, P(x) \wedge \exists y\, Q(y) \\
&\quad \sim \quad \exists y\, (\forall x\, P(x) \wedge Q(y)) \\
&\quad \sim \quad \exists y\, \forall x\, (P(x) \wedge Q(y))
\end{aligned}$$

Logical equivalence.

The two sentences $\forall x\, \exists y\, P(x,y)$ and $\exists y\, \forall x\, P(y,x)$ are not equivalent. Consider a model $\mathcal{M}$ with the set of (negative and nonnegative) integers as universe, where $P^{\mathcal{M}}$ is the less-than relation. The first sentence (which asserts that every integer is less than some other integer) is true in this model, but the second sentence (which states that there is a smallest integer) is false.

Logical equivalence.

Consider $\forall x\, \exists x\, P(x,x)$ and $\exists x\, \forall x\, P(x,x)$. Since $\forall x\, \exists x\, P(x,x)$ is logically equivalent to $\exists x\, P(x,x)$, whereas $\exists x\, \forall x\, P(x,x)$ is equivalent to $\forall x\, P(x,x)$, the two formulas are not equivalent.

Logical consequence.

The sentence $\exists x\, (P(x) \wedge R(x))$ is *not* a logical consequence of $\exists x\, (P(x) \wedge Q(x))$ and $\exists x\, (Q(x) \wedge R(x))$.

For instance, consider a model $\mathcal{M}$ with domain $\{a, b\}$, where $P^{\mathcal{M}} = \{a\}$, $Q^{\mathcal{M}} = \{a, b\}$, and $R^{\mathcal{M}} = \{b\}$. Then $\exists x\ (P(x) \wedge Q(x))$ is true in $\mathcal{M}$ as $P(x)$ and $Q(x)$ both evaluate to true if a is assigned to x. Similarly, $\exists x\ (Q(x) \wedge R(x))$ is true in $\mathcal{M}$ as Q and R both evaluate to true if b is assigned to x. But $\exists x\ (P(x) \wedge R(x))$ is not true in $\mathcal{M}$, as there is no assignment to x for which both $P(x)$ and $R(x)$ evaluate to true at the same time.

Skolemization.

We skolemize various sentences.

1. $\exists x\ \forall y\ \exists z\ (P(x, y) \wedge P(y, z) \to P(x, z))$
   Solution: $\forall y\ (P(c, y) \wedge P(y, f(y)) \to P(c, f(y)))$
2. $\forall x\ \forall y\ (P(x, y) \to \exists z\ (P(x, z) \to P(z, y)))$
   Solution: $\forall x\ \forall y\ (P(x, y) \to (P(x, f(x, y)) \to P(f(x, y), y)))$
3. $\forall x\ \exists x\ P(x, x)$
   Solution: $\forall x\, P(f(x), f(x))$
4. $\exists x\ \forall x\ P(x, x)$
   Solution: $\forall x\, P(x, x)$

Prenex form and Skolemization.

We convert the following formula to a set of clauses so that satisfiability is preserved:

$$\neg(\forall x \exists y\, P(x, y) \to (\forall y \exists z\, \neg Q(x, z) \wedge \forall y \neg \forall z\, R(y, z))).$$

First we rename bound variables so that different quantifiers bind different variables and no variable has both free and bound occurrences:

$$\neg(\forall u \exists v\, P(u, v) \to (\forall y \exists z\, \neg Q(x, z) \wedge \forall s \neg \forall t\, R(s, t))).$$

Next observe that this formula is satisfiable if, and only if, its existential closure is satisfiable:

$$\exists x[\neg(\forall u \exists v\, P(u, v) \to (\forall y \exists z\, \neg Q(x, z) \wedge \forall s \neg \forall t\, R(s, t)))].$$

Conversion to prenex form takes several steps; one intermediate formula is

$$\exists x[\forall u \exists v\, P(u, v) \wedge (\exists y \forall z\, Q(x, z) \vee \exists s \forall t\, R(s, t))].$$

A possible prenex formula is

$$\exists x \forall u \exists v \exists y \forall z \exists s \forall t \, (P(u, v) \wedge (Q(x, z) \vee R(s, t)).$$

Skolemization yields a universal formula,

$$\forall u \forall z \forall t \, (P(u, f_v(u)) \wedge (Q(c_x, z) \vee R(f_s(u, z), t))),$$

where c_x and f_s are Skolem symbols. (Other universal sentences can also be obtained from the given initial formula.) The corresponding clauses are $P(u, f_v(u))$ and $Q(c_x, z) \vee R(f_s(u, z), t))$.

Substitution.

Let σ_1 be the substitution $[x \mapsto y, y \mapsto z, z \mapsto x]$, σ_2 the substitution $[x \mapsto y, y \mapsto z, z \mapsto y]$, and σ_3 the substitution $[x \mapsto x + y, y \mapsto y + z, z \mapsto x + z]$.

Since $\sigma_2 = \sigma_1[x \mapsto y]$, the substitution σ_1 is more general than σ_2. We also have $\sigma_3 = \sigma_1[x \mapsto x + z, y \mapsto x + y, z \mapsto y + z]$ so that σ_1 is more general than σ_3. (But neither σ_2 nor σ_3 is more general than σ_1.)

We also have

$$\begin{aligned}
\sigma_1\sigma_2 &= [x \mapsto z, z \mapsto y] \\
\sigma_2\sigma_2 &= [x \mapsto z] \\
\sigma_2\sigma_3 &= [x \mapsto y + z, y \mapsto x + z, z \mapsto y + z] \\
\sigma_1\sigma_2\sigma_3 &= [x \mapsto x + z, y \mapsto y + z, z \mapsto y + z]
\end{aligned}$$

Unification.

The unification problem $\{x =^? f(y, g(y)), g(f(z, a)) =^? g(y)\}$ is solvable. The derivation,

$$\begin{array}{ll}
\multicolumn{2}{l}{x =^? f(y, g(y)), g(f(z, a)) =^? g(y)} \\
\Rightarrow_{\text{DECOMPOSE}} & x =^? f(y, g(y)), f(z, a) =^? y \\
\Rightarrow_{\text{ORIENT}} & x =^? f(y, g(y)), y =^? f(z, a) \\
\Rightarrow_{\text{ELIMINATE}} & x =^? f(f(z, a), g(f(z, a))), y =^? f(z, a)
\end{array}$$

yields a most general unifier, $[x \mapsto f(f(z, a), g(f(z, a))), y \mapsto f(z, a)]$. Another unifer, but not a most general one, is $[x \mapsto f(f(a, a), g(f(a, a))), y \mapsto f(a, a)]$.

Unification.

The unification problem

$$f(x, g(a, y)) =^? f(h(y), g(y, a)), g(x, h(y)) =^? g(z, z)$$

where x, y, and z are the only variables (and all other symbols denote functions or constants), is solvable. The derivation,

$$\begin{array}{ll} \multicolumn{2}{l}{f(x, g(a, y)) =^? f(h(y), g(y, a)), g(x, h(y)) =^? g(z, z)} \\ \Rightarrow_{\text{DECOMPOSE}} & x =^? h(y), g(a, y) =^? g(y, a), g(x, h(y)) =^? g(z, z) \\ \Rightarrow_{\text{DECOMPOSE}} & x =^? h(y), a =^? y, y =^? a, g(x, h(y)) =^? g(z, z) \\ \Rightarrow_{\text{ELIMINATE}} & x =^? h(a), a =^? a, y =^? a, g(x, h(a)) =^? g(z, z) \\ \Rightarrow_{\text{DELETE}} & x =^? h(a), y =^? a, g(x, h(a)) =^? g(z, z) \\ \Rightarrow_{\text{ELIMINATE}} & x =^? h(a), y =^? a, g(h(a), h(a)) =^? g(z, z) \\ \Rightarrow_{\text{DECOMPOSE}} & x =^? h(a), y =^? a, h(a) =^? z \\ \Rightarrow_{\text{ORIENT}} & x =^? h(a), y =^? a, z =^? h(a) \end{array}$$

yields a most general unifier,

$$[x \mapsto h(a), y \mapsto a, z \mapsto a].$$

Unification.

The unification problem $\{x_1 =^? f(x_2), x_2 =^? f(x_3), g(x_4) =^? x_3, g(x_1) =^? x_4\}$ is not solvable: after applying several orientation and elimination steps to the given set, one obtains a unification problem to which the occurs-check rule applies.

Ground resolution.

We use ground resolution to show that the set of clauses

$$\{P \vee \neg Q, P \vee R, \neg Q \vee R, \neg P \vee Q, Q \vee \neg R, \neg P \vee \neg R\}$$

is unsatisfiable. Here is one possible derivation of a contradiction:

$P \vee \neg Q$	given	(1)
$P \vee R$	given	(2)
$\neg Q \vee R$	given	(3)
$\neg P \vee Q$	given	(4)

$Q \vee \neg R$	given	(5)
$\neg P \vee \neg R$	given	(6)
$P \vee Q$	RES 2,5	(7)
$\neg P \vee \neg Q$	RES 3,6	(8)
$P \vee P$	RES 1,7	(9)
P	FACT 9	(10)
$\neg P \vee \neg P$	RES 4,8	(11)
$\neg P$	FACT 11	(12)
$\bot$	RES 10,12	(13)

Ground resolution.

Let N be the set containing the following (ground) clauses:

$$\neg P \vee Q \vee R \tag{14}$$
$$P \vee \neg R \tag{15}$$
$$Q \vee \neg R \tag{16}$$
$$P \vee R \vee \neg S \tag{17}$$
$$\neg P \vee T \tag{18}$$
$$\neg Q \vee R \vee T \tag{19}$$
$$Q \vee R \vee S \vee T \tag{20}$$
$$\neg Q \vee \neg T \tag{21}$$
$$P \vee S \vee \neg T \tag{22}$$

We derive new clauses by resolution:

$P \vee Q \vee R \vee S \vee S$	[7 and 9]	(23)
$P \vee \neg Q \vee R \vee S$	[6 and 9]	(24)
$\neg Q \vee \neg Q \vee R$	[6 and 8]	(25)
$\neg P \vee \neg Q$	[5 and 8]	(26)
$P \vee P \vee Q \vee R \vee R$	[10 and 4]	(27)
$P \vee \neg Q \vee \neg Q$	[12 and 2]	(28)
$\neg P \vee Q \vee Q$	[1 and 3]	(29)
$P \vee \vee P \vee Q \vee Q$	[14 and 3]	(30)
$\neg P \vee \neg P$	[16 and 13]	(31)
$P \vee P \vee P$	[17 and 15]	(32)
$\bot$	[18 and 19]	(33)

Since a contradiction has been derived the initial set N is unsatisfiable.

Ground resolution.

We derive a contradiction from the following clauses using resolution:

$$\begin{array}{lll} P_{1,1} \vee P_{1,2} & P_{2,1} \vee P_{2,2} & P_{3,1} \vee P_{3,2} \\ \neg P_{1,1} \vee \neg P_{2,1} & \neg P_{1,2} \vee \neg P_{2,2} & \neg P_{1,1} \vee \neg P_{3,1} \\ \neg P_{1,2} \vee \neg P_{3,2} & \neg P_{2,1} \vee \neg P_{3,1} & \neg P_{2,2} \vee \neg P_{3,2} \end{array}$$

In each inference the maximal literals in each premise were resolved, where maximality is determined by the following ordering: using the following order on literals:

$$\neg P_{3,2} \succ P_{3,2} \succ \neg P_{3,1} \succ P_{3,1} \succ \neg P_{2,2} \succ \cdots \succ \neg P_{1,1} \succ P_{1,1}.$$

(This is also known as "ordered resolution.") Factoring has been systematically applied to eliminate multiple occurrences of the same literal from a clause, and for simplicity only clauses without multiple occurrences of the same literal are listed. The first nine clauses are given.

$$\begin{array}{rlr} P_{1,1} \vee P_{1,2} & \text{[given]} & (1) \\ P_{2,1} \vee P_{2,2} & \text{[given]} & (2) \\ P_{3,1} \vee P_{3,2} & \text{[given]} & (3) \\ \neg P_{1,1} \vee \neg P_{2,1} & \text{[given]} & (4) \\ \neg P_{1,2} \vee \neg P_{2,2} & \text{[given]} & (5) \\ \neg P_{1,1} \vee \neg P_{3,1} & \text{[given]} & (6) \\ \neg P_{1,2} \vee \neg P_{3,2} & \text{[given]} & (7) \\ \neg P_{2,1} \vee \neg P_{3,1} & \text{[given]} & (8) \\ \neg P_{2,2} \vee \neg P_{3,2} & \text{[given]} & (9) \\ \neg P_{2,2} \vee P_{3,1} & 3 \,\&\, 9 & (10) \\ \neg P_{1,2} \vee P_{3,1} & 3 \,\&\, 7 & (11) \\ \neg P_{2,1} \vee \neg P_{2,2} & 10 \,\&\, 8 & (12) \\ \neg P_{1,1} \vee \neg P_{2,2} & 10 \,\&\, 6 & (13) \\ \neg P_{1,2} \vee \neg P_{2,1} & 11 \,\&\, 8 & (14) \\ \neg P_{1,1} \vee \neg P_{1,2} & 11 \,\&\, 6 & (15) \\ \neg P_{1,2} \vee P_{2,1} & 2 \,\&\, 5 & (16) \\ \neg P_{1,1} \vee P_{2,1} & 2 \,\&\, 13 & (17) \end{array}$$

$$\neg P_{1,2} \quad \text{16 \& 14, plus factoring} \tag{18}$$
$$\neg P_{1,1} \lor \neg P_{1,2} \quad \text{16 \& 4} \tag{19}$$
$$\neg P_{1,1} \quad \text{17 \& 4, plus factoring} \tag{20}$$
$$P_{1,1} \quad \text{1 \& 18} \tag{21}$$
$$\bot \quad \text{21 \& 20} \tag{22}$$

Instantiation of clauses.

Consider the following clauses,

$$\neg R(x, x) \tag{1}$$
$$\neg R(x, y) \lor R(f(x), y) \tag{2}$$
$$R(x, f(x)) \tag{3}$$

Suitable instantiation yields a set of ground clauses,

$$\neg R(f(a), f(a)) \tag{1'}$$
$$\neg R(a, f(a)) \lor R(f(a), f(a)) \tag{2'}$$
$$R(a, f(a)) \tag{3'}$$

that is unsatisfiable, as one can obtain a contradiction by two steps of resolution. Hence, the initial set of clauses is also unsatisfiable.

Resolution.

Consider the following clauses:

$$\neg R(x, y) \lor \neg R(y, x)$$
$$R(fx, fx)$$

We apply resolution to the first clause and a renamed version (renaming x to x') of the second clause, using most general unifier $\sigma = [x \mapsto fx', y \mapsto fx']$, to obtain

$$\neg R(fx', fx').$$

From the (original) second clause and the new clause we obtain a contradiction by applying resolution with most general unifier $\sigma = [x \mapsto x']$. The initial set of clauses is therefore not satisfiable.

Resolution.

We use resolution to show that the set of two clauses,

$$\neg R(x,y) \vee \neg R(y,x)$$
$$R(ffx, fy)$$

is unsatisfiable. After renaming x to x' and y to y' in the second clause, we apply resolution to the two given clauses to obtain

$$\neg R(fy', ffx')$$

by using the unifier $\sigma = [x \mapsto ffx', y \mapsto fy']$. From the second clause and this new clause we get a contradiction by applying resolution with unifier $\sigma = [y \mapsto fx', y' \mapsto fx]$. The initial set of clauses is therefore not satisfiable.

Resolution.

Consider the set of three clauses,

$$\neg R(x,y) \vee \neg R(y,z) \vee R(x,z)$$
$$\neg R(fx, fffx)$$
$$R(x, fx)$$

We rename x to x' in the second clause and apply resolution with most general unifier $\sigma = [x \mapsto fx', z \mapsto fffx']$ to the renamed clause and the first clause, to obtain

$$\neg R(fx', y) \vee \neg R(y, fffx').$$

Applying resolution to the third and fourth clause we get

$$\neg R(ffx', fffx')$$

using the most general unifier $[x \mapsto fx', y \mapsto ffx']$.

From the third and fifth clause we obtain a contradiction by resolution via most general unifier $[x \mapsto ffx']$. The initial set of clauses is therefore not satisfiable.

Resolution.

We use resolution and factoring to show that the following set of clauses is unsatisfiable:

$$\neg P(x,y) \vee \neg P(y,x) \vee \neg P(x,a)$$
$$P(x,a) \vee P(x, f(x))$$
$$P(x,a) \vee P(f(x), x)$$

where a is a constant and x and y are variables. Here is one possible derivation of a contradiction:

$$\begin{array}{rll}
\neg P(x,y) \vee \neg P(y,x) \vee \neg P(x,a) & \text{given} & (1) \\
P(x,a) \vee P(x,f(x)) & \text{given} & (2) \\
P(x,a) \vee P(f(x),x) & \text{given} & (3) \\
\neg P(x,x) \vee \neg P(x,a) & \text{FACT 1 } [y \mapsto x] & (4) \\
\neg P(a,a) & \text{FACT 4 } [x \mapsto a] & (5) \\
P(a,f(a)) & \text{RES 2,5 } [x \mapsto a] & (6) \\
P(f(a),a) & \text{RES 3,5 } [x \mapsto a] & (7) \\
\neg P(f(a),y) \vee \neg P(y,f(a)) & \text{RES 1,7 } [x \mapsto f(a)] & (8) \\
\neg P(a,f(a)) & \text{RES 7,8 } [y \mapsto a] & (9) \\
\bot & \text{RES 6,9} & (10)
\end{array}$$

Resolution.

We use resolution to determine whether

$$\eta : \forall x \exists y \forall z [R(f(x),y) \vee R(y,f(z))]$$

is a logical consequence of

$$\phi : \forall x \exists y [R(x,f(y)) \rightarrow R(y,f(x))]$$

and

$$\psi : \exists x \forall y \exists z [\neg R(x,f(y)) \rightarrow \neg R(y,f(z))].$$

First note that η is a logical consequence of ϕ and ψ if, and only if, the implication $\phi \wedge \psi \rightarrow \eta$ is valid. The latter problem is equivalent to determining whether $\phi \wedge \psi \wedge \neg\eta$ is unsatisfiable.

We next skolemize ϕ, ψ, and $\neg\eta$ to obtain universal sentences,

$$\begin{aligned}
\phi' &: \forall x [R(x,f(g(x))) \rightarrow R(g(x),f(x))] \\
\psi' &: \forall y [\neg R(c,f(y)) \rightarrow \neg R(y,f(h(y)))] \\
\eta' &: \forall y \neg [R(f(d),y) \vee R(y,f(i(y)))]
\end{aligned}$$

where c, d, g, h, and i denote Skolem functions. The formula $\phi \wedge \psi \wedge \neg\eta$ is unsatisfiable if, and only if, $\phi' \wedge \psi' \wedge \eta'$ is unsatisfiable. The latter formula is unsatisfiable if, and only if, the following set of clauses S is unsatisfiable:

$$\begin{aligned} \neg R(x, f(g(x))) \lor R(g(x), f(x)) \\ R(c, f(y)) \lor \neg R(y, f(h(y))) \\ \neg R(f(d), y) \\ \neg R(y, f(i(y))) \end{aligned}$$

Each clause in S contains a negative literal. In general, if both premises of a resolution inference contain a negative literal, so does the conclusion; and, similarly, if factoring is applied to a clause with a negative literal, the conclusion also contains a negative literal. Thus, we can only derive clauses with negative literals from S (by resolution and factoring), but not the empty clause (a contradiction). We may conclude that S is satisfiable and, hence, η is not a logical consequence of ϕ and ψ.