Automatic Reasoning (Section 8.3)

Automatic Reasoning

- Can reasoning be automated? Yes, for some logics, including first-order logic. We could try to automate *natural deduction*, but there are many proof rules that can be applied in many ways.
- We'll introduce a single proof rule, called resolution that can be applied automatically by a computer. Using resolution requires that we try to prove that a wff is unsatisfiable. But we know that a wff is valid iff its negation is unsatisfiable, so we will work with negations of the wff.

Resolution for Propositional Logic

Just to get started, we will look at resolution for Propositional Logic.

- The Resolution Rule for Propositional Logic is: $\frac{p \vee A, \neg p \vee B}{(A - p) \vee (B - \neg p)}$ where $A - p$ denotes A with all occurrences of p removed and $B - \neg p$ denotes B with all occurrences of $\neg p$ removed.
- *Special case:* $\frac{p, \neg p}{[\,]}$ (where [ ] denotes false).
- *Special case:* $\frac{p \to q, p}{q}$ also known as $\frac{\neg p \vee q, p}{q}$ which is Modus Ponens.

Steps in a resolution proof

To prove a wff is valid:

1. negate the wff
2. convert it to CNF
3. write down the fundamental disjunctions as premises
4. use resolution to find a false statement

Example

- Prove $(p \to q) \land (q \to r) \to (p \to r)$ is a tautology.
- *Solution:* Negate the wff and transform it into CNF: $(\neg p \lor q) \land (\neg q \lor r) \land p \land \neg r$.

Proof:

1.	$\neg p \lor q$	P
2.	$\neg q \lor r$	P
3.	p	P
4.	$\neg r$	P
5.	q	1,3,R
6.	r	2,5,R
7.	[ ]	4,6,R

QED

Another Example

- Prove
  $(A \vee B) \wedge (A \to C \wedge D) \wedge (B \to E \wedge F) \to (C \wedge D) \vee (E \wedge F)$
  is a tautology.
- Negate and transform into CNF to get: $(A \vee B) \wedge (\neg A \vee C) \wedge (\neg A \vee D) \wedge (\neg B \vee E) \wedge (\neg B \vee F) \wedge (\neg C \vee \neg D) \wedge (\neg E \vee \neg F)$.
- Proof on next page.

Proof

Proof:

1.	$A \vee B$	P
2.	$\neg A \vee C$	P
3.	$\neg A \vee D$	P
4.	$\neg B \vee E$	P
5.	$\neg B \vee F$	P
6.	$\neg C \vee \neg D$	P
7.	$\neg E \vee \neg F$	P

Proof

Proof:

1. $A \vee B$	P
2. $\neg A \vee C$	P
3. $\neg A \vee D$	P
4. $\neg B \vee E$	P
5. $\neg B \vee F$	P
6. $\neg C \vee \neg D$	P
7. $\neg E \vee \neg F$	P
8. $\neg A \vee \neg C$	3,6,R
9. $\neg B \vee \neg F$	4,7,R
10. $\neg B \vee \neg B$	5,9,R
11. $\neg A \vee \neg A$	2,8,R
12. B	1,11,R
13. [ ]	10,12,R

QED

Resolution for FOL

Example (to show the idea): Suppose some proof contains the following two lines:

$$
\begin{array}{lll}
1. & p(a,x) \vee q(x,y,a) & \text{P} \\
2. & \neg p(x,b) \vee r(x,y,b) & \text{P}
\end{array}
$$

First we need to change variable names so all clauses have distinct variables. So, line 2 becomes:

$$
3.\ \neg p(v,b) \vee r(v,w,b) \quad \text{P}
$$

Resolution for FOL (continued)

- Unify (i.e., match) $p(a,x)$ on line 1 and $\neg p(v,b)$ on line 3 to get a substitution (i.e., a set of bindings) $\theta = \{v/a, x/b\}$
- Apply θ to clauses on lines 1 and 3.
  - $p(a,b) \vee q(b,y,a)$
  - $\neg p(a,b) \vee r(a,w,b)$
- Apply resolution to the transformed wffs.

$$4.\ \ q(b,y,a) \vee r(a,w,b) \quad 1,3,\text{R},\ \theta = \{v/a, x/b\}$$

First-order resolution

So, three major topics need to be examined:

1. the clausal form of wffs
2. unification of atoms
3. the inference rule

Clausal Form

- A clause is a disjunction of literals. A clausal form is the universal closure of a conjunction of clauses.
- *Example:* The clausal form:
  - $\forall x \forall y \, (p(x) \wedge (q(x, y) \vee \neg r(x)) \wedge (\neg p(x) \vee s(y)))$
  - is represented by the set:
  - $\{p(x), q(x, y) \vee \neg r(x), \neg p(x) \vee s(y)\}$
- Not every wff is equivalent to a clausal form. For example, the wff $\exists x \; p(x)$ is not equivalent to a clausal form. But we have a result from Skolem that is sufficient for resolution. (See next slide.)

Skolem's Algorithm

Every wff can be associated with a clausal form in which the two are either both satisfiable or both unsatisfiable.

1. Construct the prenex CNF for the wff.
2. Replace all occurrences of each free variable by a new constant.
3. Eliminate the existential quantifiers by Skolem's Rule:
   - If $\exists x\ W(x)$ is not inside the scope of a universal quantifier, then replace $\exists x\ W(x)$ by $W(c)$ for a new constant c.
   - If $\exists x\ W(x)$ is inside the scope of $\forall x_1 \ldots x_n$, then replace $\exists x\ W(x)$ by $W(f(x_1, \ldots, x_n))$ for a new function symbol f.

Skolem examples

Find a clausal form for each wff:

1. $\exists x \, A(x)$

Skolem examples

Find a clausal form for each wff:

1. $\exists x\ A(x)$	1. $A(c)$
2. $\forall x \exists y\ B(x, y)$	2.

Skolem examples

Find a clausal form for each wff:

1. $\exists x\ A(x)$	1. $A(c)$
2. $\forall x \exists y\ B(x, y)$	2.$\forall x\ B(x, f(x))$
3. $\forall x \forall y \exists z\ C(x, y, z)$	3.

Skolem examples

Find a clausal form for each wff:

1. $\exists x \; A(x)$	1. $A(c)$
2. $\forall x \exists y \; B(x, y)$	2.$\forall x \; B(x, f(x))$
3. $\forall x \forall y \exists z \; C(x, y, z)$	3.$\forall x \forall y \; C(x, y, g(x, y))$
4. $\forall x \exists y \exists z \; D(x, y, z)$	4.

Skolem examples

Find a clausal form for each wff:

1. $\exists x\ A(x)$	1. $A(c)$
2. $\forall x \exists y\ B(x, y)$	2.$\forall x\ B(x, f(x))$
3. $\forall x \forall y \exists z\ C(x, y, z)$	3.$\forall x \forall y\ C(x, y, g(x, y))$
4. $\forall x \exists y \exists z\ D(x, y, z)$	4.$\forall x\ D(x, f(x), g(x))$

Clausal form example

Find a clausal form for the wff:

- $\exists y \forall x\; p(x,y) \rightarrow \exists z\; (q(x) \wedge r(z))$

$\exists y \forall x\; p(x,y) \rightarrow \exists z\; (q(w) \wedge r(z))$	(renamed)
$\equiv \neg \exists y \forall x\; p(x,y) \vee \exists z\; (q(w) \wedge r(z))$	(removed $\rightarrow$)
$\equiv \forall y \exists x\; \neg p(x,y) \vee \exists z\; (q(w) \wedge r(z))$	(moved $\neg$ inside)
$\equiv \forall y (\exists x\; \neg p(x,y) \vee \exists z\; (q(w) \wedge r(z)))$	(moved $\forall y$ out)
$\equiv \forall y \exists x \exists z\; (\neg p(x,y) \vee (q(w) \wedge r(z)))$	(moved $\exists x$, $\exists z$ ou
$\equiv \forall y \exists x \exists z\; ((\neg p(x,y) \vee q(w)) \wedge (\neg p(x,y) \vee r(z)))$	(constructed CNF

- (replace free variable w by constant c)

$$\forall y \exists x \exists z\; ((\neg p(x,y) \vee q(c)) \wedge (\neg p(x,y) \vee r(z)))$$

- Apply Skolem's Rule to eliminate $\exists x$ and $\exists z$.

$$\forall y ((\neg p(f(y),y) \vee q(c)) \wedge (\neg p(f(y),y) \vee r(g(y))))$$

- Giving us two clauses:
  $\{\neg p(f(y),y) \vee q(c), \neg p(f(y),y) \vee r(g(y))\}$

Substitutions and Unification

- A binding of a variable x to a term t is denoted x/t and it means replace x by t.
- *Applying a binding to an expression:* If x/t is a binding and E is an expression, then $E(x/t)$ denotes the expression obtained from E by replacing all free occurrences of x with t.
- *Examples:*
  - $p(x, y, z)(x/y) = p(y, y, z)$
  - $p(x, y, z)(y/f(z)) = p(x, f(z), z)$
- A substitution is a finite set of bindings with distinct numerators. (We use lowercase greek letters for substitutions.)
- *Examples:*
  - $\theta = \{x/y, y/f(z)\}$
  - $\sigma = \{y/f(a), z/b\}$

Apply a Substitution

- If θ is a substitution and E is an expression, then $E\theta$ denotes the expression obtained from E by simultaneously applying the bindings in θ to E.
- *Example:* If $\theta = \{x/y, y/f(z)\}$, then
  - $p(x, y, z)\theta = p(x, y, z)\{x/y, y/f(z)\} = p(y, f(z), z)$
- If S is a set of expressions, then $S\theta$ is the set of expressions obtained from S by applying θ to each expression of S
  - *Example:* If $\theta = \{x/y, y/f(z)\}$ and $S = \{p(x, y), q(y, g(z))\}$, then $S\theta = \{p(x, y)\theta, q(y, g(z))\theta\} = \{p(y, f(z)), q(f(z), g(z))\}$

Composing Substitutions

- If θ and σ are two substitutions, then the composition of them $\theta\sigma$ is applied to an expression E by $E(\theta\sigma) = (E\theta)\sigma$.
- We can calculate $\theta\sigma$ by applying it to an atom that contains all the numerator variables of the two substitutions. Then collect the bindings that result from the application.
- *Example:* Let $\theta = \{x/y, y/f(z)\}$ and $\sigma = \{y/f(a), z/b\}$. Since the numerator variables are x, y, z we calculate:
    - $p(x, y, z)(\theta\sigma) = (p(x, y, z)\theta)\sigma = p(y, f(z), z)\sigma = p(f(a), f(b), b)$
    - So $\theta\sigma = \{x/f(a), y/f(b), z/b\}$.

Unifiers

- A unifier of a set S of expressions is a substitution θ such that $S\theta$ is a singleton set.
- *Example:* $\{x/a\}$ is a unifier of $\{p(x), p(a)\}$ because $\{p(x), p(a)\}\{x/a\}$ is $\{p(a)\}$.
- *Example Quiz:* What are some unifiers of $S = \{p(x, a), p(y, z)\}$?

Unifiers

- A unifier of a set S of expressions is a substitution θ such that $S\theta$ is a singleton set.
- *Example:* $\{x/a\}$ is a unifier of $\{p(x), p(a)\}$ because $\{p(x), p(a)\}\{x/a\}$ is $\{p(a)\}$.
- *Example Quiz:* What are some unifiers of $S = \{p(x, a), p(y, z)\}$?
- *Answer:* Since a is a constant, any unifier must include the binding z/a. A unifier might also include x/y or y/x. So two unifiers of S are $\{x/y, z/a\}$ and $\{y/x, z/a\}$ because:
  - $S\{x/y, z/a\} = \{p(y, a)\}$ and $S\{y/x, z/a\} = \{p(x, a)\}$.
- Notice also that $\{x/t, y/t, z/a\}$ is a unifier of S for any term t because:
  - $S\{x/t, y/t, z/a\} = \{p(t, a)\}$.

Robinson's Unification Algorithm

For a finite set S of atoms, find whether S has an mgu.

1. $k := 0; \theta_0 := \epsilon$; go to Step 2.
2. If $S\theta_k$ is a singleton then stop with mgu θ_k. Otherwise construct D_k (set of terms in leftmost position of disagreement); go to Step 3.
3. If D_k has a variable v and a term t such that v *does not occur in* t then $\theta_{k+1} := \theta_k\{v/t\}; k := k+1$; go to Step 2. Otherwse stop (S is not unifiable).

Unification Example

Trace the algorithm for $S = \{p(x, h(x, g(y)), y), p(x, h(a, z), b\}$.

1. $k := 0; \theta_0 := \epsilon$.
2. $S\theta_0 = S\epsilon = \{p(x, h(x, g(y)), y), p(x, h(a, z), b\}$ is not a singleton; $D_0 = \{x, a\}$.
3. $\theta_1 := \theta_0\{x/a\} = \{x/a\}; k := 1$
4. (2) $S\theta_1 = \{p(a, h(a, g(y)), y), p(a, h(a, z), b)\}$ is not a singleton; $D_1 = \{g(y), z\}$.
5. (3) $\theta_2 := \theta_1\{z/g(y)\} = \{x/a, z/g(y)\}; k := 2$.
6. (2) $S\theta_2 = \{p(a, h(a, g(y)), y), p(a, h(a, g(y)), b)\}$ is not a singleton; $D_2 = \{y, b\}$.
7. (3) $\theta_3 := \theta_2\{y/b\} = \{x/a, z/g(b), y/b\}; k := 3$.
8. (2) $S\theta_3 = \{p(a, h(a, g(b)), b)\}$ is a singleton; stop with mgu $\theta_3 = \{x/a, z/g(b), y/b\}$.

A Failed Example

- Apply the algorithm to $S = \{p(x), p(f(x))\}$.

A Failed Example

- Apply the algorithm to $S = \{p(x), p(f(x))\}$.
- S is not unifiable because in the set $D_0 = \{x, f(x))\}$ x *occurs* in $f(x)$.

Resolution Rule

Given the following two clauses:

- $L_1 \vee \ldots \vee L_k \vee C$ and $\neg M_1 \vee \ldots \vee \neg M_n \vee D$

where L_i and M_i are atoms and C and D are disjunctions of other literals. Assume also that:

1. The clauses have distinct sets of variable names (rename if necessary).
2. θ is the mgu of $\{L_1, \ldots, L_k, M_1, \ldots, M_n\}$.
3. $N = L_1\theta$, where $\{L_1, \ldots, L_k, M_1, \ldots, M_n\}\theta = \{N\}$.

Then we have:

- $\frac{L_1 \vee \ldots \vee L_k \vee C, \neg M_1 \vee \ldots \vee \neg M_n \vee D}{(C\theta - N) \vee (D\theta - \neg N)}$

Example

Given the two clauses in the following proof segment:

k.	$p(a,y) \vee p(a,z) \vee q(x,y,z)$	P
k+1.	$\neg p(w,f(b)) \vee r(w,v,g(w))$	P

- These two clauses have the form $L_1 \vee L_2 \vee C$ and $\neg M_1 \vee D$. The two clauses have distinct sets of variables.
- The set of atoms $\{p(a,y), p(a,z), p(w,f(b))\}$ has mgu $\theta = \{w/a, y/f(b), z/f(b)\}$.
- Notice that $\{p(a,y), p(a,z), p(w,f(b))\}\theta = \{p(a,f(b))\}$
- So, the resolution rule can be applied to the two clauses to obtain the resolvant:

k+2. $q(x,f(b),f(b)) \vee r(a,v,g(a))$ $\quad k, k+1, R, \{w/a, y/f(b), z/f(b)$

Whole Process

- *To prove a wff is valid:* negate the wff and convert it to clausal form; write down the clauses as premises; use resolution to find a false statement.
- *Example::* Use resolution to prove the following wff is valid:
  - $\exists x\ (\forall y\ p(x,y) \vee \forall z\ q(x,z) \rightarrow \forall y\ (p(x,y) \vee q(x,y)))$
- *Solution:* Negate the wff, giving you:
  - $\neg\exists x\ (\forall y\ p(x,y) \vee \forall z\ q(x,z) \rightarrow \forall y\ (p(x,y) \vee q(x,y)))$.
- Then convert it to clausal form:

Convert to clausal form

Convert: $\neg\exists x\ (\forall y\ p(x,y) \vee \forall z\ q(x,z) \rightarrow \forall y\ (p(x,y) \vee q(x,y)))$

$\neg\exists x\ (\forall y\ p(x,y) \vee \forall z\ q(x,z) \rightarrow \forall w\ (p(x,w) \vee q(x,w)))$
(renamed variables)
$\forall x\ ((\forall y\ p(x,y) \vee \forall z\ q(x,z)) \wedge \exists w\ (\neg p(x,w) \wedge \neg q(x,w)))$
(removed $\rightarrow$ and moved $\neg$ right)
$\forall x \exists w\ ((\forall y\ p(x,y) \vee \forall z\ q(x,z)) \wedge \neg p(x,w) \wedge \neg q(x,w))$
(moved $\exists w$ left)
$\forall x \exists w \forall y \forall z\ ((p(x,y) \vee q(x,z)) \wedge \neg p(x,w) \wedge \neg q(x,w))$
(moved $\forall y$ and $\forall z$ left)
$\forall x \forall y \forall z\ ((p(x,y) \vee q(x,z)) \wedge \neg p(x,f(x)) \wedge \neg q(x,f(x)))$
(removed $\exists w$ by Skolem)

So, the set of clauses is:
$\{p(x,y) \vee q(x,z), \neg p(x,f(x)), \neg q(x,f(x))\}$

Resolution steps

Now do a resolution proof by making each of the three clauses a premise (rename to get distinct variable names).

1.	$p(x, y) \vee q(x, z)$	P
2.	$\neg p(u, f(u))$	P
3.	$\neg q(w, f(w))$	P
4.	$q(u, z)$	1,2,R,{x/u,y/f(u)}
5.	[ ]	3,4,R,{w/u,z/f(u)}

QED