Logic in Computer Science III

Lesson 6, automated theorem proving

Lesson 6, automated theorem proving

Recall:

Theorem (Herbrand's Theorem)

Let A be a prenex sentence, let

$$He(A) := \exists x_1 \ldots \exists x_n \psi(x_1, \ldots, x_n).$$

Then A is logically valid iff there exist terms t_{ij} such that

$$\bigvee_{j=1}^{m} \psi(t_{1j}, \ldots, t_{nj})$$

is a propositional tautology.

Lesson 6, automated theorem proving

Recall:

Theorem (Herbrand's Theorem)

Let A be a prenex sentence, let

$$He(A) := \exists x_1 \ldots \exists x_n \psi(x_1, \ldots, x_n).$$

Then A is logically valid iff there exist terms t_{ij} such that

$$\bigvee_{j=1}^{m} \psi(t_{1j}, \ldots, t_{nj})$$

is a propositional tautology.

This is reduces the task of proving a theorem to

1. finding suitable terms, and
2. proving a propositional tautology.

We want to use Resolution to prove the Herbrand disjunction. But

- searching randomly, or systematically, for terms and then trying to prove the disjunction in Resolution

is not a good strategy, because we will generate a lot of useless terms.

There is a better approach:

- look for terms that enable us to do resolution steps.

Resolution in first order logic

Example

The pair of clauses

$$A \vee P(f(x), y), \quad B \vee \neg P(z, g(u))$$

cannot be resolved.

Resolution in first order logic

Example

The pair of clauses

$$A \vee P(f(x), y), \quad B \vee \neg P(z, g(u))$$

cannot be resolved.

But we can substitute $y := g(u)$ *and* $z := f(x)$ *and get*

$$\frac{A \vee P(f(x), g(u)), \quad B \vee \neg P(f(x), g(u))}{A \vee B}.$$

Resolution in first order logic

Example

The pair of clauses

$$A \vee P(f(x), y), \quad B \vee \neg P(z, g(u))$$

cannot be resolved.

But we can substitute $y := g(u)$ *and* $z := f(x)$ *and get*

$$\frac{A \vee P(f(x), g(u)), \quad B \vee \neg P(f(x), g(u))}{A \vee B}.$$

Caveat: *We must substitute* $y := g(u)$ *also in* A *and* $z := f(x)$ *also in* B*!*

Resolution in first order logic

Example

The pair of clauses

$$A \vee P(f(x), y), \quad B \vee \neg P(z, g(u))$$

cannot be resolved.

But we can substitute $y := g(u)$ *and* $z := f(x)$ *and get*

$$\frac{A \vee P(f(x), g(u)), \quad B \vee \neg P(f(x), g(u))}{A \vee B}.$$

Caveat: *We must substitute* $y := g(u)$ *also in* A *and* $z := f(x)$ *also in* B*!*

By successive substitutions we eventually obtain the term needed.

unification of terms

- a substitution is a mapping $\sigma : \mathit{Variables} \rightarrow \mathit{Terms}$
- for a term t, $\sigma(t)$ denotes the term obtained by substitution σ
- σ is a unifier of a pair of terms s, t if $\sigma(s) = \sigma(t)$
- σ is a most general unifier (MGU) of a pair of terms s, t if $\sigma(s) = \sigma(t)$ and for every unifier τ there exists ρ such that $\tau = \rho\sigma$.

examples

- $\sigma = \{y \mapsto g(u), z \mapsto f(x)\}$ is an MGU of $P(f(x), y)$ and $P(z, g(u))$.
- x and $f(x)$ cannot be unified.
- $f(s_1, \ldots, s_n)$ and $g(t_1, \ldots, t_n)$ cannot be unified if $f \neq g$.

Theorem

If there exists a unifier, then there exists an MGU.

Theorem

There exists an algorithm that either finds an MGU, or outputs "NO unifier". The algorithm runs in polynomial time in the input and output size. (The output may be exponentially larger than the input.)

See Buss's chapter.

Robinson's first-order resolution

John Alan Robinson, 1965

Robinson's first-order resolution

John Alan Robinson, 1965

We use two rules

1. given a clause C, derive $\sigma(C)$ for some substitution σ,
2. propositional Resolution rule.

Robinson's first-order resolution

John Alan Robinson, 1965

We use two rules

1. given a clause C, derive $\sigma(C)$ for some substitution σ,
2. propositional Resolution rule.

We use unification

1. to obtain complementary literals for applying Resolution,
2. to obtain same literals in a clause in order to make it shorter; this is called factoring.

Robinson's first-order resolution

John Alan Robinson, 1965

We use two rules

1. given a clause C, derive $\sigma(C)$ for some substitution σ,
2. propositional Resolution rule.

We use unification

1. to obtain complementary literals for applying Resolution,
2. to obtain same literals in a clause in order to make it shorter; this is called factoring.

Strategy: *use only MGUs as substitutions.*
However, we may need to substitute different variables in order to enable unification.

Example (of substituting variables)

Consider clauses

$$Q(x, z) \vee P(x), \qquad R(x) \vee \neg P(f(x)).$$

We cannot unify x *with* $f(x)$*, but we can first apply substitution* $x \to y$ *to get*

$$Q(x, z) \vee P(x), \qquad R(y) \vee \neg P(f(y)).$$

and then we can unify x *and* $f(y)$ *and get*

$$Q(f(y), z) \vee R(y).$$

Example (of substituting variables)

Consider clauses

$$Q(x,z) \vee P(x), \qquad R(x) \vee \neg P(f(x)).$$

We cannot unify x *with* $f(x)$*, but we can first apply substitution* $x \to y$ *to get*

$$Q(x,z) \vee P(x), \qquad R(y) \vee \neg P(f(y)).$$

and then we can unify x *and* $f(y)$ *and get*

$$Q(f(y),z) \vee R(y).$$

Example (of factoring)

$$\dfrac{\dfrac{Q(z,x) \vee P(x) \vee P(f(z))}{Q(z,f(z)) \vee P(f(z)) \vee P(f(z))}\ \textit{unification of } x \textit{ and } f(z)}{Q(z,f(z)) \vee P(f(z))}\ \textit{contraction}$$

Example (of substituting variables)

Consider clauses

$$Q(x,z) \vee P(x), \qquad R(x) \vee \neg P(f(x)).$$

We cannot unify x *with* $f(x)$, *but we can first apply substitution* $x \to y$ *to get*

$$Q(x,z) \vee P(x), \qquad R(y) \vee \neg P(f(y)).$$

and then we can unify x *and* $f(y)$ *and get*

$$Q(f(y),z) \vee R(y).$$

Example (of factoring)

$$\cfrac{\cfrac{Q(z,x) \vee P(x) \vee P(f(z))}{Q(z,f(z)) \vee P(f(z)) \vee P(f(z))}\ \textit{unification of } x \textit{ and } f(z)}{Q(z,f(z)) \vee P(f(z))}\ \textit{contraction}$$

Exercise

Define a general rule that would join factoring, unification, and resolution into one step.

how to prove a sentence Φ

1. put Φ in to prenex form with the matrix[1] in the DNF form,

[1] =the formula without the quantifier prefix

how to prove a sentence Φ

1. put Φ in to prenex form with the matrix[1] in the DNF form,
2. construct $He(\Phi)$, and let ψ be the matrix of $He(\Phi)$,

[1] =the formula without the quantifier prefix

how to prove a sentence Φ

1. put Φ in to prenex form with the matrix[1] in the DNF form,
2. construct $He(\Phi)$, and let ψ be the matrix of $He(\Phi)$,
3. transform $\neg\psi$ into a CNF $C_1 \wedge \cdots \wedge C_n$,

[1] =the formula without the quantifier prefix

how to prove a sentence Φ

1. put Φ in to prenex form with the matrix[1] in the DNF form,
2. construct $He(\Phi)$, and let ψ be the matrix of $He(\Phi)$,
3. transform $\neg\psi$ into a CNF $C_1 \wedge \cdots \wedge C_n$,
4. use Robinson's Resolution to derive the empty clause from $\{C_1, \ldots, C_n\}$.

[1]=the formula without the quantifier prefix

how to prove a sentence Φ

1. put Φ in to prenex form with the matrix[1] in the DNF form,
2. construct $He(\Phi)$, and let ψ be the matrix of $He(\Phi)$,
3. transform $\neg\psi$ into a CNF $C_1 \wedge \cdots \wedge C_n$,
4. use Robinson's Resolution to derive the empty clause from $\{C_1, \ldots, C_n\}$.

Exercise

Prove that every sentence can be proved using this (nondeterministic) procedure.

[1]=the formula without the quantifier prefix

example

$$\exists x(\neg P(0) \vee (P(x) \wedge \neg P(Sx)) \vee P(S^n 0))$$

example

$$\exists x(\neg P(0) \vee (P(x) \wedge \neg P(Sx)) \vee P(S^n 0))$$

Herbrand's theorem gives us:

$$\begin{aligned}
&\neg P(0) \vee \\
&(P(0) \wedge \neg P(S0)) \vee (P(S0) \wedge \neg P(SS0)) \vee (P(SS0) \wedge \neg P(SSS0)) \vee \ldots \\
&\hfill \vee P(S^n(0))
\end{aligned}$$

example

$$\exists x(\neg P(0) \vee (P(x) \wedge \neg P(Sx)) \vee P(S^n 0))$$

Herbrand's theorem gives us:

$$\begin{aligned}
&\neg P(0)\vee \\
&(P(0) \wedge \neg P(S0)) \vee (P(S0) \wedge \neg P(SS0)) \vee (P(SS0) \wedge \neg P(SSS0)) \vee \dots \\
&\hfill \vee P(S^n(0))
\end{aligned}$$

Claim. Every Herbrand's disjunction must contain all terms $0, S0, SS0, \dots, S^n 0$.

Proof.

Suppose it does not contain $S^i 0$ for $0 < i < n$. Define a truth assignment by

- $P(S^j 0) \mapsto$ true, for $j < i$,
- $P(S^j 0) \mapsto$ false, for $j > i$.

Then all disjuncts are falsified. $\square$

example, contd.

But first-order resolution is more efficient. Suppose $n = 2^k$.

[2] we use $\dots \to \dots$ instead of $\neg \dots \vee \dots$, resolution becomes transitivity of $\to$

example, contd.

But first-order resolution is more efficient. Suppose $n = 2^k$.

1. $P(x) \to P(S(x))$ (initial clause)[2]
2. $P(Sx) \to P(SS(x))$ (substitution $x \mapsto Sx$)
3. $P(x) \to P(SS(x))$ (resolution of 1 and 2)
4. $P(SSx) \to P(SSSS(x))$ (substitution $x \mapsto SSx$)
5. $P(x) \to P(SSSS(x))$ (resolution of 3 and 4)

   ...

2k+1. $P(x) \to P(S^{2^k}(x))$

2k+2. $P(0) \to P(S^{2^k}(0))$ (substitution $x \mapsto 0$)

2k+3. $P(0)$ (initial clause)

2k+4. $P(S^{2^k}(0))$ (resolution of $2k+2$ and $2k+3$)

2k+5. $\neg P(S^{2^k}(0))$ (initial clause)

2k+6. $\perp$ (resolution of $2k+4$ and $2k+5$)

[2] we use $\cdots \to \ldots$ instead of $\neg \cdots \vee \ldots$, resolution becomes transitivity of $\to$

example, contd.

But where is unification?

example, contd.

But where is unification?

An alternative view:

1. $P(x) \to P(S(x))$ (initial clause)
2. $P(y) \to P(S(y))$ (substitution $x \mapsto y$)
3. $P(Sx) \to P(SS(x))$ (unification of $S(x)$ and y)
4. $P(x) \to P(SS(x))$ (resolution of 1 and 3)
5. $\ldots$

an important improvement of efficiency

Theorem

Let $A_1, \ldots, A_n$ be prenex sentences. Then

$$\vdash A_1 \vee \cdots \vee A_n \quad \Leftrightarrow \quad \vdash He(A_1) \vee \cdots \vee He(A_n)$$

Proof.

- **exercise** $\square$

an important improvement of efficiency

Theorem

Let $A_1, \ldots, A_n$ be prenex sentences. Then

$$\vdash A_1 \vee \cdots \vee A_n \quad \Leftrightarrow \quad \vdash He(A_1) \vee \cdots \vee He(A_n)$$

Proof.

- **exercise** $\square$

Corollary

Let $A_1, \ldots, A_n, B$ be prenex sentences. Then

$$A_1, \ldots, A_n \vdash B \quad \Leftrightarrow \quad \vdash Sk(A_1) \wedge \cdots \wedge Sk(A_n) \rightarrow He(B).$$

modification of the procedure

for derivations from axioms $A_1, \ldots, A_n \vdash B$

1. Skolemize $A_1, \ldots, A_n$
2. put the matrices of $Sk(A_1), \ldots, Sk(A_n)$ into CNF forms
3. Herbrandize B
4. put the negation of the matrix of $He(B)$ into a CNF form
5. — the rest is the same

[3] chang-lee-examples.pdf

modification of the procedure

for derivations from axioms $A_1, \ldots, A_n \vdash B$

1. Skolemize $A_1, \ldots, A_n$
2. put the matrices of $Sk(A_1), \ldots, Sk(A_n)$ into CNF forms
3. Herbrandize B
4. put the negation of the matrix of $He(B)$ into a CNF form
5. — the rest is the same

See examples in *Symbolic Logic and Mechanical Theorem Proving* by C.-L. Chang, R. C.-T. Lee.[3]

[3] chang-lee-examples.pdf

the completeness theorem from Herbrand's theorem

- We had assumed that the sequent calculus was complete, when we proved Herbrand's theorem.
- Now we will prove it without this assumption.
- This gives us
  1. model-theoretical proof of Herbrand's theorem
  2. and completeness of the sequent calculus w.r.t. prenex sentences
- We can prove in the sequent calculus that every sentence is equivalent to a prenex sentence, hence we get completeness for the sequent calculus for all sentences.

Theorem

Let A be a prenex sentence and let ψ be the matrix of $He(A)$. Suppose that for no m and no terms t_{ij},

$$\bigvee_{j=1}^{m} \psi(t_{1j}, \ldots, t_{nj})$$

is a propositional tautology. Then there exists a model M such that

$$M \models \neg A.$$

Theorem

Let A be a prenex sentence and let ψ be the matrix of $He(A)$. Suppose that for no m and no terms t_{ij},

$$\bigvee_{j=1}^{m} \psi(t_{1j}, \ldots, t_{nj})$$

is a propositional tautology. Then there exists a model M such that

$$M \models \neg A.$$

We will actually prove

$$M \models Sk(prenex(\neg A)).$$

We know that $\vdash Sk(B) \rightarrow B$ for prenex formulas.

We will assume that the sequent calculus is complete for propositional logic.

Proof.

Let A be given, and let $\psi(x_1, \ldots, x_n)$ be the matrix of $He(A)$.

Proof.

Let A be given, and let $\psi(x_1, \ldots, x_n)$ be the matrix of $He(A)$.

Then $Sk(prenex(\neg A)) = \forall x_1 \ldots \forall x_n \neg\psi(x_1, \ldots, x_n)$.

Proof.

Let A be given, and let $\psi(x_1, \ldots, x_n)$ be the matrix of $He(A)$.

Then $Sk(prenex(\neg A)) = \forall x_1 \ldots \forall x_n \neg\psi(x_1, \ldots, x_n)$.

If there is no Herbrand disjunction witnessing the validity of A, we need a model M such that

$$M \models \neg\psi[a_1, \ldots, a_n] \quad \text{for all } a_1, \ldots, a_n \in M.$$

Proof.

Let A be given, and let $\psi(x_1, \ldots, x_n)$ be the matrix of $He(A)$.

Then $Sk(prenex(\neg A)) = \forall x_1 \ldots \forall x_n \neg\psi(x_1, \ldots, x_n)$.

If there is no Herbrand disjunction witnessing the validity of A, we need a model M such that

$$M \models \neg\psi[a_1, \ldots, a_n] \quad \text{for all } a_1, \ldots, a_n \in M.$$

Construction of the model

- the universe of M: all terms
- a constant c is interpreted as c
- a function symbol f is interpreted as the function

  $t_1, \ldots, t_k \mapsto f(t_1, \ldots, t_k)$
- interpretation of predicates and relations: we need to assign truth values to atomic formulas $R(t_1, \ldots, t_l)$ so that all propositions $\psi(t_1, \ldots, t_k)$ are evaluated false.

See next page:

interpretation of predicates and relations

- Since no Herbrand disjunction is a tautology, we have for every m and every finite set of terms $t_{i,j}$ an assignment that falsifies all $\psi(t_{1,j}, \ldots, t_{n,j})$, $j = 1, \ldots, m$.

interpretation of predicates and relations

- Since no Herbrand disjunction is a tautology, we have for every m and every finite set of terms $t_{i,j}$ an assignment that falsifies all $\psi(t_{1,j}, \ldots, t_{n,j})$, $j = 1, \ldots, m$.
- By the compactness of the propositional calculus, we have an assignment a to all atomic formulas such that a falsifies all propositions $\psi(t_1, \ldots, t_n)$.
- So we define the relations using a as follows $M \models R(t_1, \ldots, t_l)$ iff $a : R(t_1, \ldots, t_l) \mapsto \top$.

Thus we get

$$M \models \neg\psi[a_1, \ldots, a_n] \quad \text{for all } a_1, \ldots, a_n \in M.$$

$\square$

We proved

Theorem

A prenex sentence A is logically valid iff there exists a tautological Herbrand disjunction for A.

It remains to prove:

Lemma

If there exists a tautological Herbrand disjunction for A, then A is provable in the sequent calculus.

Proof.

Let the disjunction be given.

- for every term consider all maximal subterms that start with some Herbrand function symbol and replace them by a free variable; the same terms by the same variable, different by different
- the substitution preserves provability in the propositional calculus, so assuming that the propositional part of the sequent calculus is complete, we get a proof of the sequent with the substituted terms
- now we introduce quantifiers as follows
  1. whenever it is possible to contract, we contract, otherwise
  2. whenever it is possible to introduce $\exists$ we do it, otherwise
  3. we introduce $\forall$

It remains to prove the following claim: $\square$

Claim The procedure only stops when the sequent becomes A.

Proof.

We need to show that we can introduce $\forall$ if no contraction or $\exists$-introduction is possible and the sequent is not A yet.

Claim The procedure only stops when the sequent becomes A.

Proof.

We need to show that we can introduce $\forall$ if no contraction or $\exists$-introduction is possible and the sequent is not A yet.

In such a situation, in each formula B, different from A, we can only introduce $\forall$ provided that the free b variable does not occur elsewhere. So we need to show that there exist at least one B whose b does not occur elsewhere.

Claim The procedure only stops when the sequent becomes A.

Proof.

We need to show that we can introduce $\forall$ if no contraction or $\exists$-introduction is possible and the sequent is not A yet.

In such a situation, in each formula B, different from A, we can only introduce $\forall$ provided that the free b variable does not occur elsewhere. So we need to show that there exist at least one B whose b does not occur elsewhere.

We take B whose b corresponds to the most complex term t. For some Herbrand function h, B and t have the form

$$t := h(\dots s_i \dots) \quad \text{and} \quad B := \Box\psi(\dots s_i \dots b \dots),$$

where $\Box$ is some prefix of quantifiers. Clearly

1. since t cannot occur as a proper subterm in any term at this stage, so does b
2. t cannot be equal to a term in any other formula, because t encodes terms s_i, hence the formula would have to be equal, but we have contracted equal formulas.

$\Box$

We get the completeness of other proof system by simulating the sequent calculus.

complexity issues

We know that PHP requires exponential size Resolution proofs, while it has polynomial size proofs in the standard proof systems such as the sequent and Frege calculi.

complexity issues

We know that PHP requires exponential size Resolution proofs, while it has polynomial size proofs in the standard proof systems such as the sequent and Frege calculi.

The comparison of Herbrand's theorem with the sequent calculus with cuts, or Hilbert style calculi is much worse.

complexity issues

We know that PHP requires exponential size Resolution proofs, while it has polynomial size proofs in the standard proof systems such as the sequent and Frege calculi.

The comparison of Herbrand's theorem with the sequent calculus with cuts, or Hilbert style calculi is much worse.

Theorem

There exists a sequence of logically valid sentences ϕ_n such that

1. *ϕ_n have polynomial size proofs in the sequent calculus with cuts, and Hilbert style calculi, but*
2. *every cut-free proof, or Herbrand disjunction for ϕ_n has size at least*

$$2^{2^{2^{\cdots^{2}}}} \Big\} \; n \text{ times.}$$

complexity issues

We know that PHP requires exponential size Resolution proofs, while it has polynomial size proofs in the standard proof systems such as the sequent and Frege calculi.

The comparison of Herbrand's theorem with the sequent calculus with cuts, or Hilbert style calculi is much worse.

Theorem

There exists a sequence of logically valid sentences ϕ_n such that

1. *ϕ_n have polynomial size proofs in the sequent calculus with cuts, and Hilbert style calculi, but*
2. *every cut-free proof, or Herbrand disjunction for ϕ_n has size at least*

$$\left.2^{2^{2^{\cdots^{2}}}}\right\} n \text{ times.}$$

This also applies to Robinson's first-order resolution, maybe, with one 2 in the stack less.