Computational Logic — Lecture 7

Herbrand Method

Michael Genesereth — Spring 2002

Propositional Interpretations

p	q	r	
0	0	0	$\{\}$
0	0	1	$\{r\}$
0	1	0	$\{q\}$
0	0	1	$\{q,r\}$
1	0	0	$\{p\}$
1	0	1	$\{p,r\}$
1	1	0	$\{p,q\}$
1	1	1	$\{p,q,r\}$

For a language with n constants, there are 2^n interpretations.

Relational Interpretations

$$\{\}$$

$$\{p(a)\}$$

$$\{q(a)\}$$

$$\longrightarrow \{p(a), q(a)\}$$

$$\{p(b)\}$$

$$\{q(b)\}$$

$$\{p(b), q(b)\}$$

$$\{p(a), p(b)\}$$

$$\{p(a), p(b), q(a)\}$$

$$\{p(a), p(b), q(b)\}$$

$$\{p(a), p(b), q(a), q(b)\}$$

$$\ldots$$

Infinitely many interpretations.

Logical Entailment

A set of premises logically entails a conclusion if and only if every interpretation that satisfies the premise also satisfies conclusion.

In the case of Propositional Logic, the number of interpretations is finite, and so it is possible to check logical entailment directly in finite time.

In the case of Relational Logic, the number of interpretations is infinite, and so a direct check of logical entailment is not feasible.

Good News

Given any set of sentences, there is a specially defined subset of interpretations called *Herbrand interpretations*.

Under *certain conditions*, checking just the Herbrand interpretations suffices to determine logical entailment.

Checking just the Herbrand interpretations is less work than checking all interpretations.

HHHHerbrand

The *Herbrand universe* for a set of sentences in Relational Logic (with at least one object constant) is the set of all ground terms that can be formed from just the constants used in those sentences. If there are no object constants, then we add an arbitrary object constant, say a.

The *Herbrand base* for a set of sentences is the set of all ground atomic sentences that can be formed using just the constants in the Herbrand universe.

A *Herbrand interpretation* for a set of sentences is any subset of the Herbrand base for those sentences.

Example

Sentences

$$\forall x.(r(a,x) \Rightarrow r(x,b))$$
$$\forall x.\forall y.\forall z.(r(x,y) \wedge r(x,y) \Rightarrow r(x,z))$$

Herbrand Universe (constants used in sentences only)

$$\{a,b\}$$

Herbrand Base

$$\{r(a,a),\ r(a,b),\ r(b,a),\ r(b,b)\}$$

Herbrand Interpretations

$\{\}$
$\{r(a,a)\}$
$\{r(a,b)\}$
$\{r(b,a)\}$
$\{r(b,b)\}$
$\{r(a,a), r(a,b)\}$
$\{r(a,a), r(b,a)\}$
$\{r(a,a), r(b,b)\}$
$\{r(a,b), r(b,a)\}$
$\{r(a,b), r(b,b)\}$
$\{r(b,a), r(b,b)\}$
$\{r(a,a), r(a,b), r(b,a)\}$
$\{r(a,a), r(a,b), r(b,b)\}$
$\{r(a,a), r(b,a), r(b,b)\}$
$\{r(a,b), r(b,a), r(b,b)\}$
$\{r(a,a), r(a,b), r(b,a), r(b,b)\}$

16 Herbrand interpretations in all. Note: $16<\infty$.

Herbrand Theorem

Herbrand Theorem: A set of quantifier-free sentences has a model if and only if it has a Herbrand model.

Proof. Assume the set of sentences contains at least one object constant. If a set of quantifier-free sentences is satisfiable, then there is an interpretation that satisfies it. Take the intersection of this interpretation with the Herbrand base. By definition, this is a Herbrand interpretation. Moreover, it is easy to see that it is a model. If the sentences are ground, it must agree with the original interpretation on all of the sentences, since they are all ground and mention only the constants common to both interpretations. If the sentences contain variables, the instances must all be true, including those in which the variables are replaced by elements in the Herbrand universe.

If there is no object constant, then create a tautology involving a new constant (say a) and add to the set. This does not change the satisfiability of the sentences but satisfies proof above. QED

Example

Sentences

$$r(a,b) \Rightarrow r(b,b)$$
$$r(a,b) \vee r(b,b)$$

Model

$$\{r(a,b), r(b,b), r(a,c), r(b,c)\}$$

Herbrand Base

$$\{r(a,a), r(a,b), r(b,a), r(b,b)\}$$

Herbrand Model

$$\{r(a,b), r(b,b)\}$$

Example

Sentences

$$r(a,x)$$
$$r(x,y) \Rightarrow r(y,\mathrm{x})$$

Model

$$\{r(a,a), r(a,b), r(a,c), r(b,a), r(c,a)\}$$

Herbrand Base

$$\{r(a,a)\}$$

Herbrand Model

$$\{r(a,a)\}$$

Herbrand Method

Definition: Add negation of conclusion to the premises to form the *satisfaction set*. Loop over Herbrand interpretations. Cross out each interpretation that does *not* satisfy the sentences in the satisfaction set. If all Herbrand interpretations are crossed out, by the Herbrand Theorem, the set is unsatisfiable.

Sound and Complete: Negating the conclusion leads to a contradiction; therefore, the premises logically entail the conclusion.

Termination: Since there are only finitely many Herbrand interpretations, the process halts.

Special Cases

,, *Ground Relational Logic*

no variables, no functions, no quantifiers

Universal Relational Logic

no functions, no quantifiers

free variables implicitly universally quantified

Existential Relational Logic

no functions

Functional Relational Logic

no quantifiers

Example

Premises:

$$p(a) \Rightarrow q(a)$$

$$p(b) \Rightarrow q(b)$$

$$p(a) \vee p(b)$$

Conclusion:

$$q(a) \vee q(b)$$

Satisfaction Set:

$$p(a) \Rightarrow q(a)$$

$$p(b) \Rightarrow q(b)$$

$$p(a) \vee p(b)$$

$$\neg(q(a) \vee q(b))$$

$\{\}$

$\{q(a)\}$

$\{q(b)\}$

$\{q(a), q(b)\}$

$\{p(a)\}$

$\{p(a), q(a)\}$

$\{p(a), q(b)\}$

$\{p(a), q(a), q(b)\}$

$\{p(b)\}$

$\{p(b), q(a)\}$

$\{p(b), q(b)\}$

$\{p(b), q(a), q(b)\}$

$\{p(a), p(b)\}$

$\{p(a), p(b), q(a)\}$

$\{p(a), p(b), q(b)\}$

$\{p(a), p(b), q(a), q(b)\}$

Special Cases

Ground Relational Logic
no variables, no functions, no quantifiers

Universal Relational Logic
no functions, no quantifiers
free variables implicitly universally quantified

Existential Relational Logic
no functions

Functional Relational Logic
no quantifiers

Example

Premises:

$p(x) \Rightarrow q(x)$

$p(a) \vee q(a)$

Conclusion:

$q(a)$

Satisfaction Set:

$p(x) \Rightarrow q(x)$

$p(a) \vee q(a)$

$\neg q(a)$

Herbrand Interpretations:

$\{\}$

$\{p(a)\}$

$\{q(a)\}$

$\{p(a), q(a)\}$

Problem

Premises

$p(x) \Rightarrow q(x)$

$p(x) \vee q(x)$

Conclusion

$q(x)$

What is the Satisfaction set?

$\neg q(x)$ No. This says q is false for all args.

$\neg \forall x.q(x)$ Yes, but this has an explicit quantifier.

Universal and Existential Sentences

The negation of a sentence φ in Universal Logic is the universally quantified sentence $\neg\forall x_1 \ldots \forall x_n . \varphi$, where $x_1, \ldots, x_n$ are the free variables in φ.

Negation distributed over universal quantification by flipping the universal quantifier to an existential quantifier.

$$\begin{array}{c} \neg\forall x_1 \ldots \forall x_n . \varphi \\ \downarrow \\ \exists x_1 \ldots \exists x_n . \neg\varphi \end{array}$$

If a sentence is purely universal, then this distribution leads to a purely existential sentence.

Skolemization

The *Skolemization* of a purely existential sentence is the sentence obtained by dropping the existential quantifiers and replacing all variables systematically by brand new constants.

Example:

$$\exists x.\neg q(x)$$
$$\downarrow$$
$$\neg q(c)$$

Example:

$$\exists x.\exists y.(p(x,y) \wedge q(x,b))$$
$$\downarrow$$
$$p(c,d) \wedge q(c,b)$$

Significance

Skolemization Theorem: A set of sentences in Relational Logic is satisfiable if and only if its Skolemization is satisfiable.

A modification of the Herbrand Method can be used!

Modified Herbrand Method

Original Definition: Add negation of conclusion to the premises to form the *satisfaction set*… Loop over Herbrand interpretations…

New Definition: Negate the conclusion. Add its Skolemization to the premises to form the *satisfaction set*… Loop over Herbrand interpretations…

Example

Premises:

$$p(x) \Rightarrow q(x)$$
$$p(x) \vee q(x)$$

Conclusion:

$$q(x)$$

Negation:

$$\neg\forall x.q(x)$$

Existentialization:

$$\exists x.\neg q(x)$$

Satisfaction Set:

$$p(x) \Rightarrow q(x)$$
$$p(x) \vee q(x)$$
$$\neg q(c)$$

Skolemization:

$$\neg q(c)$$

Example

Premises:

$r(a,a)$

$r(b,b)$

Conclusion:

$r(x,x)$

Negation:

$\neg\forall x.r(x,x)$

Existentialization:

$\exists x.\neg r(x,x)$

Skolemization:

$\neg r(c,c)$

Satisfaction Set:

$r(a,a)$

$r(b,b)$

$\neg r(c,c)$

Special Cases

Ground Relational Logic

no variables, no functions, no quantifiers

Universal Relational Logic

no functions, no quantifiers

free variables implicitly universally quantified

,, *Existential Relational Logic*

no functions

Functional Relational Logic

no quantifiers

Herbrand Theorem Does Not Apply

Sentences

$$r(a,a)$$
$$r(b,b)$$
$$\exists x.\neg r(x,x)$$

Model

$$\{r(a,a), r(b,b), r(a,c), r(b,c)\}$$

Herbrand Base

$$\{r(a,a), r(a,b), r(b,a), r(b,b)\}$$

Intersection

$$\{r(a,a), r(b,b)\}$$

Not a model and there is no Herbrand Model

Good News (Sort of…)

Solution: Skolemize the sentences to get rid of explicit quantifiers.

Small hitch #1 is that we are now dealing with existential sentences but not purely existential sentences; so we need to extend the technique we just looked at. See upcoming lecture.

Small hitch #2 is that there are some difficulties using the Modified Herbrand Method with Functional Logic. See next few slides.

Special Cases

Ground Relational Logic
no variables, no functions, no quantifiers

Universal Relational Logic
no functions, no quantifiers
free variables implicitly universally quantified

Existential Relational Logic
no functions

,, *Functional Relational Logic*
no quantifiers

Herbrand Universe for Functional Logic

Old Definition: The *Herbrand universe* for a set of sentences in Relational Logic is the set of all ground terms that can be formed from just the constants used in those sentences.

In the absence of function constants, this is exactly the set of constants in the set of sentences.

$$\{a, b\}$$

In the presence of function constants, all ground functional terms are included.

$$\{a, b, f(a), f(b), g(a), g(b), f(f(a)), f(f(b)), f(g(a)), f(g(b)), \ldots\}$$

Herbrand Theorem for Functional Logic

Herbrand Theorem: A set of quantifier-free sentences has a model if and only if it has a Herbrand model.

No quantifiers; so we are okay.

The Modified Herbrand Method works. Hooray!!

Sad Theorem

The size of the Herbrand universe for a functional language is infinite.

Upshot: Checking the Herbrand interpretations for a language to determine logical entailment is not feasible in finite time.

General Problem

The number of Herbrand models can be very large.

n object constants
m k-ary relation constants
Number of k-ary tuples: n^k
Number of k-ary relations: $2^\wedge n^k$
Number of Herbrand Interpretations: $(2^\wedge n^k)^\wedge m$

10 object constants
3 2-ary relation constants
Number of k-ary tuples: 100
Number of k-ary relations: 2^{100}
Number of Herbrand Interpretations: 2^{300}

Summary

Herbrand Method *works* for Ground Logic.

Modified Herbrand Method *works* for Universal Logic.

Herbrand Method *does not work* for Existential Logic. These cases *can be handled* by Skolemizing to form sentences in Functional Logic and then using Modified Herbrand Method.

Modified Herbrand Method *works* for Functional Logic, but there are infinitely many interpretations.

In any case, the number of Herbrand interpretations can be very large.

Solution: Use formal proofs!