Computational Logic Lecture 8

Relational Proofs

Michael Genesereth Spring 2002

Logical Entailment

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

Herbrand Summary

Herbrand Method works for Basic Logic and Universal Logic, but there can be *many* interpretations.

Herbrand Method does not work for Existential Logic.

Herbrand Method works for Functional Logic, but infinitely many interpretations

Solution: Use formal proofs!

Formal Proofs

A *formal proof* of φ from Δ is a sequence of sentences terminating in φ in which each item is either:

1. a premise (a member of Δ)
2. an instance of an axiom schema
3. the result of applying a rule of inference to earlier items in the sequence.

Old Rules of Inference

Modus Ponens (MP)

$$\frac{\varphi \Rightarrow \psi \qquad \varphi}{\psi}$$

Modus Tolens (MT)

$$\frac{\varphi \Rightarrow \psi \qquad \neg\psi}{\neg\varphi}$$

And Introduction (AI)

$$\frac{\varphi \qquad \psi}{\varphi \wedge \psi}$$

And Elimination (AE)

$$\frac{\varphi \wedge \psi}{\varphi \qquad \psi}$$

Idea for Universal Instantiation

$$\frac{\forall \nu . \varphi}{\varphi[\nu \leftarrow \tau]}$$

Warning: This is not quite right.

Examples

$\forall y.hates(jane,y)$

$hates(jane,jill)$	$y \leftarrow jill$
$hates(jane,mother(jane))$	$y \leftarrow mother(jane)$
$hates(jane,y)$	$y \leftarrow y$
$hates(jane,z)$	$y \leftarrow z$

$\forall x.\exists y.hates(x,y)$

$\exists y.hates(jane,y)$	$y \leftarrow jane$	
$\exists y.hates(y,y)$	$y \leftarrow y$	*Wrong!!*

Inappropriateness

A term τ is *inappropriate* for a variable ν in φ if and only if τ contains a variable μ *and* there is some free occurrence of ν in φ that lies in the scope of a quantifier of μ.

mother(*x*) is inappropriate for *y* in $\exists x.$*hates*(*x*,*y*).

Substitutability

A term τ is *substitutable* for ν in φ if and only if it is not inappropriate with ν in φ.

Some texts say "x is *free* for y in φ" instead of "x is substitutable for y in φ".

$mother(jane)$ is free for y in $hates(jane,y)$.
$mother(x)$ is free for y in $hates(jane,y)$.
$mother(x)$ is free for y in $\exists z.hates(z,y)$.
$mother(x)$ is *not* free for y in $\exists x.hates(x,y)$.
$mother(x)$ is free for y in $(\forall x.\forall y.loves(x,y) \land \exists z.hates(z,y))$.

Universal Instantiation

$$\frac{\forall \nu . \varphi}{\varphi[\nu \leftarrow \tau]}$$

where τ is free for ν in φ

Existential Instantiation I

$$\frac{\exists \nu . \varphi}{\varphi[\nu \leftarrow \sigma]}$$

where φ contains no free variables

where σ is a new object constant

Examples

$\exists y.p(y)$

$p(c)$

$\exists y.y*y=0$

$1*1=0$ *Wrong!*

$\exists y.y*y=x$

$c*c=x$ *Wrong!*

$c*c=4$

$c*c=6$

Existential Instantiation II

$$\frac{\exists \nu . \varphi}{\varphi[\nu \leftarrow \pi(\tau_1, ..., \tau_n)]}$$

where $\tau_1, ..., \tau_n$ are free in φ

where π is a new function constant

Examples

$\exists y . y * y = x$

$f(x) * f(x) = x$

$f(4) * f(4) = 4$

$f(6) * f(6) = 6$

$\exists y . y * y = x$

$sqrt(x) * sqrt(x) = x$

$log(x) * log(x) = x$ *Wrong!*

Formal Proofs

A *formal proof* of φ from Δ is a sequence of sentences terminating in φ in which each item is either:

1. a premise (a member of Δ)
2. an instance of an axiom schema
3. the result of applying a rule of inference to earlier items in the sequence.

Example

Everybody loves somebody. Everybody loves a lover. Show that Jack loves Jill.

1.	$\forall x.\exists y.loves(x,y)$	Premise
2.	$\forall u.\forall v.\forall w.(loves(v,w) \Rightarrow loves(u,v))$	Premise
3.	$\exists y.loves(jill,y)$	$\mathrm{UI}:1$
4.	$loves(jill, f(jill))$	$\mathrm{EI}:3$
5.	$\forall v.\forall w.(loves(v,w) \Rightarrow loves(jack,v))$	$\mathrm{UI}:2$
6.	$\forall w.(loves(jill,w) \Rightarrow loves(jack, jill))$	$\mathrm{UI}:5$
7.	$loves(jill, f(jill)) \Rightarrow loves(jack, jill)$	$\mathrm{UI}:6$
8.	$loves(jack, jill)$	$\mathrm{MP}:7,4$

Harry and Ralph

Every horse can outrun every dog. Some greyhounds can outrun every rabbit. Harry is a horse. Ralph is a rabbit. Can Harry outrun Ralph?

Harry and Ralph (continued)

Harry and Ralph (continued)

Standard Axiom Schemata

All generalizations* of the following:

II: $\varphi \Rightarrow (\psi \Rightarrow \varphi)$

ID: $(\varphi \Rightarrow (\psi \Rightarrow \chi)) \Rightarrow ((\varphi \Rightarrow \psi) \Rightarrow (\varphi \Rightarrow \chi))$

CR: $(\neg\psi \Rightarrow \varphi) \Rightarrow ((\neg\psi \Rightarrow \neg\varphi) \Rightarrow \psi)$

$(\psi \Rightarrow \varphi) \Rightarrow ((\psi \Rightarrow \neg\varphi) \Rightarrow \neg\psi)$

EQ: $(\varphi \Leftrightarrow \psi) \Rightarrow (\varphi \Rightarrow \psi)$

$(\varphi \Leftrightarrow \psi) \Rightarrow (\psi \Rightarrow \varphi)$

$(\varphi \Rightarrow \psi) \Rightarrow ((\psi \Rightarrow \varphi) \Rightarrow (\varphi \Leftrightarrow \psi))$

OQ: $(\varphi \Leftarrow \psi) \Leftrightarrow (\psi \Rightarrow \varphi)$

$(\varphi \vee \psi) \Leftrightarrow (\neg\varphi \Rightarrow \psi)$

$(\varphi \wedge \psi) \Leftrightarrow \neg(\neg\varphi \vee \neg\psi)$

Standard Axiom Schemata (continued)

UD: $\forall \nu.(\varphi \Rightarrow \psi) \Rightarrow (\forall \nu.\varphi \Rightarrow \forall \nu.\psi)$

UG: $\varphi \Rightarrow \forall \nu.\varphi$
where ν is not free in φ

UI: $\forall \nu.\varphi \Rightarrow \varphi[\nu \leftarrow \tau]$
where τ is free for ν in φ

ED: $\exists \nu.\varphi \Leftrightarrow \neg \forall \nu.\neg \varphi$

*A generalization of φ is $\forall \nu_1 \ldots \forall \nu_k.\varphi$, for variables $\nu_1, \ldots, \nu_k$.

Provability

A sentence φ is *provable* from a set of sentences Δ if and only if there is a finite formal proof of φ from Δ using only Modus Ponens and the standard axiom schemata.

Soundness Theorem: If φ is provable from Δ, then Δ logically entails φ.

Completeness Theorem (Godel): If Δ logically entails φ, then φ is provable from Δ.

Decidability

A *class* of questions is *decidable* if and only if there is a procedure such that, when given as input any question in the class, the procedure halts and says *yes* if the answer is positive and *no* if the answer is negative.

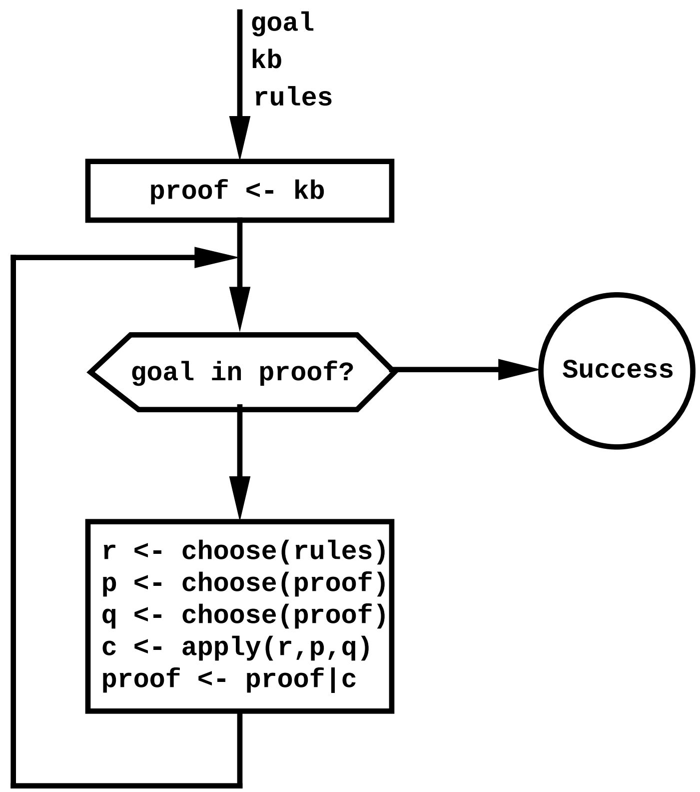
Example: For any natural number *n*, determining whether *n* is prime.

Semidecidability

A *class* of questions is semidecidable if and only if there is a procedure that halts and says *yes* if the answer is positive.

Obvious Fact: If a class of questions is decidable, it is semidecidable.

Semidecidability of Logical Entailment



Decidability Not Proved

Note that we have *not* shown that logical entailment for Relational Logic is decidable.

The procedure may not halt.

$p(x) \Rightarrow p(f(x))$

$p(f(f(a)))$

$p(f(b))$?

We cannot just run procedure on negated sentence because that may not be logically implied either!

$p(x) \Rightarrow p(f(x))$

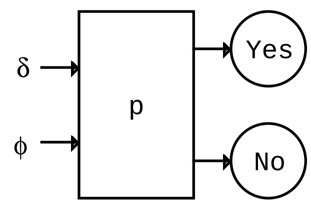
$p(f(f(a)))$

$\neg p(f(b))$?

Undecidability of Logical Entailment

Metatheorem: Logical Entailment for Relational Logic is *not* decidable.

Proof: Suppose there is a procedure *p* that decides the question of logical entailment. Its inputs are Δ and φ.

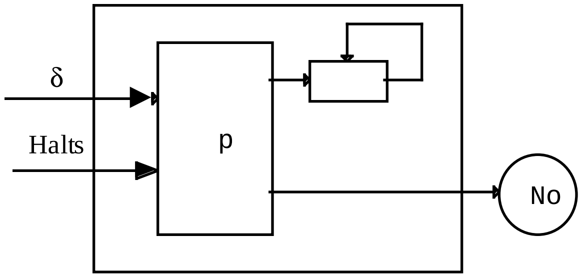


We can encode the behavior of a machine and its inputs as sentences and ask whether the machine halts as a conclusion.

What happens if we give this description and question to *p*? It says *yes*.

Undecidability (continued)

It is possible to construct a larger machine *p'* that enters an infinite loop if *p* says *yes* and halts if *p* says *no*.



We can encode a description of this machine as a set of sentences and ask whether the machine halts as a conclusion.

What happens if we give this description and question to *p*? If *p* says *yes*, then *p'* runs forever, contradicting the hypothesis that *p* computes correctly. If *p* says *no*, then *p'* halts, once again contradicting answer from *p*. QED

Closure

The *closure* S^* of a set S of sentences is the set of all sentences logically entailed by S.

$$S^* = \{\varphi \mid S \models \varphi\}$$

Set of Sentences:

$p(a)$
$p(x) \Rightarrow p(f(x))$

Closure:

$p(a)$
$p(f(a))$
$p(f(f(a)))$
$p(a) \wedge p(f(a))$
$p(x) \Rightarrow p(f(x))$
…

Theories

A theory is a set of sentences closed under logical entailment, i.e. T is a theory if and only if $T^*=T$.

A theory T is *finitely axiomatizable* if and only if there is a *finite* set Δ of sentences such that $T=\Delta^*$.

A theory T is *complete* if and only, for all φ, either $\varphi \in T$ or $\neg\varphi \in T$.

Note: Not every theory is complete. Consider the theory consisting of all consequences of $p(a,b)$. Does this include $p(b,a)$? Does it include $\neg p(b,a)$?

Relationships on Theories

Decidable

Semidecidable

Finitely Axiomatizable

Arithmetization of Logical Entailment

The theory of arithmetic is the set of all sentences true of the natural numbers, 0, 1, +, *, and <.

Fact: It is possible to assign numbers to sentences such that
(1) Every sentence φ is assigned a unique number n_φ.
(2) The question of logical entailment $\Delta \models \varphi$ can be expressed as a numerical condition $r(n_\Delta, n_\varphi)$.

Conclusion: The theory of arithmetic is not decidable.

Incompleteness Theorem

Metatheorem (Godel): If Δ is a finite subset of the theory of arithmetic, then Δ^* is not complete.

Variant: Arithmetic is not finitely axiomatizable.

Proof: If there were a finite axiomatization, then the theory would be decidable. However, arithmetic is not decidable. Therefore, there is no finite axiomatization.

Summary

Logical Entailment for Relational Logic is semidecidable.

Logical Entailment for Relational Logic is *not* decidable.

Arithmetic is not finitely axiomatizable in Relational Logic.