Introduction to Logic
Relational Proofs

Michael Genesereth
Computer Science Department
Stanford University

Logical Entailment

A set of premises Δ logically entails a conclusion φ ($\Delta \models \varphi$) if and only if every interpretation that satisfies Δ also satisfies φ.

Determining Logical Entailment

$$\{m \Rightarrow p \vee q, p \Rightarrow q\} \models m \Rightarrow q?$$

m	p	q	$m \Rightarrow p \vee q$	$p \Rightarrow q$	$m \Rightarrow q$
1	1	1	1	1	1
1	1	0	1	0	0
1	0	1	1	1	1
1	0	0	0	1	0
0	1	1	1	1	1
0	1	0	1	0	1
0	0	1	1	1	1
0	0	0	1	1	1

Determining Logical Entailment

$$\{p(a) \vee p(b), \forall x.(p(x) \Rightarrow q(x))\} \models \exists x.q(x)?$$

$p(a)$	$p(b)$	$q(a)$	$q(b)$	$p(a) \vee p(b)$	$\forall x.(p(x) \Rightarrow q(x))$	$\exists x.q(x)$
1	1	1	1	1	1	1
1	1	1	0	1	0	1
1	1	0	1	1	0	1
1	1	0	0	1	0	0
1	0	1	1	1	1	1
1	0	1	0	1	1	1
1	0	0	1	1	0	1
1	0	0	0	1	0	0
0	1	1	1	1	1	1
0	1	1	0	1	0	1
0	1	0	1	1	1	1
0	1	0	0	1	0	0
0	0	1	1	0	1	1
0	0	1	0	0	1	1
0	0	0	1	0	1	1
0	0	0	0	0	1	0

Analysis

Object constants: n
Binary relation constants: k
Factoids in Herbrand Base: $k*n^2$
Interpretations: 2^{k*n^2}

Object constants: 4
Binary relation constants: 4
Factoids in Herbrand Base: 64
Interpretations: 2^{64} = 18,446,744,073,709,551,616

Good News

Good News: If Δ logically entails φ, then there is a finite proof of φ from Δ. And vice versa.

Good News: If Δ logically entails φ, it is possible to find such a proof in finite time.

More Good News: Such proofs are often *much* smaller than the corresponding truth tables.

Fitch System for Relational Logic

Logical Rules of Inference

Negation Introduction
Negation Elimination

And Introduction
And Elimination

Or Introduction
Or Elimination

Assumption
Implication Elimination
Implication Introduction

Biconditional Introduction
Biconditional Elimination

New Rules of Inference

Universal Elimination Domain Closure Universal Introduction	*Universal Reasoning*
Existential Introduction Existential Elimination	*Existential Reasoning*

Logical Entailment and Provability

A set of premises Δ *logically entails* a conclusion φ ($\Delta \models \varphi$) if and only if every interpretation that satisfies Δ also satisfies φ.

If there exists a proof of a sentence ϕ from a set Δ of premises using the rules of inference in R, we say that ϕ is *provable* from Δ using R (written $\Delta \vdash_R \phi$).

Soundness and Completeness

A proof system is *sound* if and only if every provable conclusion is logically entailed.

$$\text{If } \Delta \vdash \phi, \text{ then } \Delta \vDash \phi.$$

A proof system is *complete* if and only if every logical conclusion is provable.

$$\text{If } \Delta \vDash \phi, \text{ then } \Delta \vdash \phi.$$

Fitch

Theorem: Fitch is sound and complete for **Relational Logic**.

$$\Delta \models \varphi \text{ if and only if } \Delta \vdash_{\text{Fitch}} \phi.$$

Upshot: The truth table method and the proof method succeed in exactly the same cases!

Universal Elimination

Universal Elimination (UE)

$$\frac{\forall \nu . \phi}{\phi_{\nu \Leftarrow \tau}}$$

where τ is ground

NB: $\phi_{\nu \Leftarrow \tau}$ is an instance of ϕ with *all* occurrences of ν replaced by τ.

UE Examples

Premise:

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

Conclusions:

$hates(jane, jill)$ $\quad x \Leftarrow jill$

$hates(jane, jane)$ $\quad x \Leftarrow jane$

Non-Conclusions:

$hates(jane, y)$ $\quad x \Leftarrow y$ Wrong!
Must be ground.

UE Examples

Premise:

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

Conclusions:

$hates(jane,jane)$	$x \Leftarrow jane$
$hates(jill,jill)$	$x \Leftarrow jill$

Non-Conclusions:

$hates(jane,x)$	$x \Leftarrow jane$	Wrong!
$hates(x,jane)$	$x \Leftarrow jane$	Wrong!
$hates(x,x)$	$x \Leftarrow jane$	Wrong!

Must be ground.

UE Examples

Premise:

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

Conclusions:

$$\exists y. hates(jane,y) \qquad x \Leftarrow jane$$

UE Examples

Premise:

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

Conclusion:

$$\forall y.hates(jane,y) \qquad x \Leftarrow jane$$

Subsequent Conclusion:

$$hates(jane,jill) \qquad y \Leftarrow jill$$

$$hates(jane,jane) \qquad y \Leftarrow jane$$

Domain Closure

Domain Closure

$$\frac{\begin{array}{c}\phi[\sigma_1] \\ \ldots \\ \phi[\sigma_n]\end{array}}{\forall \upsilon . \phi[\upsilon]}$$

} *every object constant*

DC Example

$$
\frac{\begin{array}{l} likes(abby,cody) \\ likes(bess,cody) \\ likes(cody,cody) \\ likes(dana,cody) \end{array}}{\forall x.likes(x,cody)}
$$

DC Example

$$likes(abby, abby)$$
$$likes(bess, bess)$$
$$likes(cody, cody)$$
$$likes(dana, dana)$$

$$\forall x.likes(x,x)$$

DC Example

$$likes(abby,cody) \Rightarrow likes(cody,abby)$$
$$likes(bess,cody) \Rightarrow likes(cody,bess)$$
$$likes(cody,cody) \Rightarrow likes(cody,cody)$$
$$likes(dana,cody) \Rightarrow likes(cody,dana)$$

$$\forall x.(likes(x,cody) \Rightarrow likes(cody,x))$$

DC Example

$\exists y.likes(abby,y)$

$\exists y.likes(bess,y)$

$\exists y.likes(cody,y)$

$\exists y.likes(dana,y)$

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

Universal Introduction

Proof

1.	$\forall y.(likes(cody,y) \Rightarrow happy(y))$	Premise
2.	$\forall y.likes(cody,y)$	Premise
3.	$likes(cody,abby) \Rightarrow happy(abby)$	UE: 1
4.	$likes(cody,abby)$	UE: 2
5.	$happy(abby)$	IE: 4, 3
6.	$likes(cody,bess) \Rightarrow happy(bess)$	UE: 1
7.	$likes(cody,bess)$	UE: 2
8.	$happy(bess)$	IE: 6, 7
9.	$likes(cody,cody) \Rightarrow happy(cody)$	UE: 1
10.	$likes(cody,cody)$	UE: 2
11.	$happy(cody)$	IE: 9, 10
12.	$likes(cody,dana) \Rightarrow happy(dana)$	UE: 1
13.	$likes(cody,dana)$	UE: 2
14.	$happy(dana)$	IE: 12, 13
15.	$\forall y.happy(y)$	DC: 5, 8, 11, 14

Example

1.	$\forall y.(likes(cody, y) \Rightarrow happy(y))$	Premise
2.	$\forall y.likes(cody, y)$	Premise
3.	$likes(cody, c) \Rightarrow happy(c)$	UE: 1
4.	$likes(cody, c)$	UE: 2
5.	$happy(c)$	IE: 4, 3
6.	$\forall y.happy(y)$	UI: 5

Reasoning About *Arbitrary Objects*

If we can prove a property about an *arbitrary object*, then it must be true of all objects.

Common type of mathematical reasoning:

Let c be an arbitrary object.
We can prove that a particular property is true of c.
Therefore, the property is true of everything.

Placeholders

A *placeholder* is a new type of symbol that *stands for* an arbitrary object constant but *is not itself* an object constant. Spelled the same as object constants.

Placeholders must be disjoint from object constants.

Object Constants: *abby*, *bess*, *cody*, *dana*
Placeholder: *c*
Sometimes written in brackets: [*c*]

Placeholders are used only within the Fitch procedure, *never* used outside of the procedure.

Universal Introduction (UI)

ϕ

$\forall \upsilon . \phi_{\tau \Leftarrow \nu}$

where τ is a placeholder
not used in any *active* assumption

NB: $\phi_{\tau \Leftarrow \nu}$ is an instance of ϕ with *all* occurrences of τ replaced by ν.

UI Example

Object Constants: *jane*, ...
Placehoders: *c*, ...

Premise:

$$hates(c,jane)$$

Conclusion:

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

UI Example

Object Constants: *jane*, ...

Placehoders: *c*, ...

Premise:

$$hates(c,jane) \Rightarrow hates(jane,c)$$

Conclusion:

$$\forall x.(hates(x,jane) \Rightarrow hates(jane,x)) \qquad c \Leftarrow x$$

Example

Premises:

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

$$\forall x.p(x)$$

Goal:

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

Proof

1. $\forall x.(p(x) \Rightarrow q(x))$	Premise
2. $\forall x.p(x)$	Premise
3. $p(c) \Rightarrow q(c)$	UE: 1
4. $p(c)$	UE: 2
5. $q(c)$	IE: 4, 3
6. $\forall x.q(x)$	UI: 5

Problem

Premises:

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

$$\forall x.(q(x) \Rightarrow r(x))$$

Goal:

$$\forall x.(p(x) \Rightarrow r(x))$$

Proof

1.	$\forall x.(p(x) \Rightarrow q(x))$	Premise
2.	$\forall x.(q(x) \Rightarrow r(x))$	Premise
3.	$p(c) \Rightarrow q(c)$	UE: 1
4.	$q(c) \Rightarrow r(c)$	UE: 2
5.	$\mid p(c)$	Assumption
6.	$\mid q(c)$	IE: 5, 3
7.	$\mid r(c)$	IE: 6, 4
8.	$p(c) \Rightarrow r(c)$	II: 5, 7
9.	$\forall x.(p(x) \Rightarrow r(x))$	UI: 8

Lovers

Everybody loves somebody. Everybody loves a lover.
Show that everybody loves everybody.

Premises:

$$\forall y.\exists z.loves(y,z)$$

$$\forall x.\forall y.(\exists z.loves(y,z) \Rightarrow loves(x,y))$$

Conclusion:

$$\forall x.\forall y.loves(x,y)$$

Proof

Everybody loves somebody. Everybody loves a lover. Show that everybody loves everybody.

1.	$\forall y.\exists z.loves(y,z)$	Premise
2.	$\forall x.\forall y.(\exists z.loves(y,z) \Rightarrow loves(x,y))$	Premise
3.	$\exists z.loves(d,z)$	UE: 1
4.	$\forall y.(\exists z.loves(y,z) \Rightarrow loves(c,y))$	UE: 2
5.	$\exists z.loves(d,z) \Rightarrow loves(c,d)$	UE: 4
6.	$loves(c,d)$	IE: 5, 3
7.	$\forall y.loves(c,y)$	UI: 6
8.	$\forall x.\forall y.loves(x,y)$	UI: 7

Universal Introduction (UI)

ϕ

$\forall \upsilon . \phi_{\tau \Leftarrow \nu}$

where τ is a placeholder

not used in any *active* assumption

NB: $\phi_{\tau \Leftarrow \nu}$ is an instance of ϕ with *all* occurrences of τ replaced by ν.

Bad, Bad, Bad "Proof"

1.	$\forall x.(p(x) \Rightarrow q(x))$	Premise	
2.	$p(a)$	Premise	
3.	$p(c) \Rightarrow q(c)$	UE: 1	
4.	$\vert\ p(c)$	Assumption	
5.	$\vert\ q(c)$	IE: 5, 3	
6.	$\vert\ \forall y.q(y)$	UI: 5	NO!!!
7.	$p(c) \Rightarrow \forall y.q(y)$	II: 4, 6	
8.	$\forall x.(p(x) \Rightarrow \forall y.q(y))$	UI: 7	
9.	$p(a) \Rightarrow \forall y.q(y)$	UE: 8	
10	$\forall y.q(y)$	IE: 9, 2	Wrong.

Universal Introduction (UI)

ϕ

$\forall \upsilon . \phi_{\tau \Leftarrow \nu}$

where τ is a placeholder

not used in any *active* assumption

NB: $\phi_{\tau \Leftarrow \nu}$ is an instance of ϕ with *all* occurrences of τ replaced by ν.

Name Conflict Not Cool

1.	$\forall y.(likes(cody,y) \Rightarrow happy(y))$	Premise	
2.	$likes(cody,abby)$	Premise	
3.	$likes(cody,abby) \Rightarrow happy(abby)$	UE: 1	
4.	$happy(abby)$	IE: 3, 2	
5.	$\forall y.happy(y)$	UI: 4	Wrong.

Reasoning Tip for Universal Reasoning

If you have some universal sentences and you want to prove a universal sentence, use placeholders to eliminate the universals, prove a specific conclusion, then generalize.

Proof

$$\forall x.(p(x) \Rightarrow q(x)), \forall x.(q(x) \Rightarrow r(x)) \models \forall x.(p(x) \Rightarrow r(x))$$

1.	$\forall x.(p(x) \Rightarrow q(x))$	Premise
2.	$\forall x.(q(x) \Rightarrow r(x))$	Premise
3.	$p(c) \Rightarrow q(c)$	UE: 1
4.	$q(c) \Rightarrow r(c)$	UE: 2
5.	$\quad\mid p(c)$	Assumption
6.	$\quad\mid q(c)$	IE: 5, 3
7.	$\quad\mid r(c)$	IE: 6, 4
8.	$p(c) \Rightarrow r(c)$	II: 5, 7
9.	$\forall x.(p(x) \Rightarrow r(x))$	UI: 8

Existential Introduction

Existential Introduction (EI)

$$\frac{\phi}{\exists \nu . \phi_{\tau \leftarrow \nu}}$$

where τ is a constant

NB: $\phi_{\tau \leftarrow \nu}$ is an instance of ϕ with *0 or more* occurrences of τ replaced by ν.

EI Examples

Premise:

$$hates(jill, jill)$$

Conclusions:

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

$$\exists x.hates(jill, x)$$

$$\exists x.hates(x, jill)$$

Two Applications:

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

EI Examples

Premise:

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

Non-Conclusion:

$\exists y.\forall x.hates(x,y)$ Wrong. Constants only!

Existential Elimination

EE Example

Premises:

$$\exists x.hates(jane,x)$$

$$\forall x.(hates(jane,x) \Rightarrow mean(jane))$$

Conclusion:

$$mean(jane)$$

EE Example

Metatheorem: $\forall \nu.(\phi \Rightarrow \psi)$ is equivalent to $(\exists \nu.\phi \Rightarrow \psi)$ so long as ψ is free of ν.

Example:

$$\forall x.(hates(jane,x) \Rightarrow mean(jane))$$

is equivalent to

$$(\exists x.hates(jane,x) \Rightarrow mean(jane))$$

Proof Reminder:

$\forall \nu.(\phi \Rightarrow \psi)$ is equivalent to $\forall \nu.(\neg\phi \vee \psi)$,
which is equivalent to $(\forall \nu.\neg\phi \vee \psi)$ since ψ free of ν,
which is equivalent to $(\neg\exists \nu.\phi \vee \psi)$,
which is equivalent to $(\exists \nu.\phi \Rightarrow \psi)$.

EE Example

Premises:

$$\exists x.hates(jane,x)$$

$$\forall x.(hates(jane,x) \Rightarrow mean(jane))$$

Equivalent Premises:

$$\exists x.hates(jane,x)$$

$$\exists x.hates(jane,x) \Rightarrow mean(jane)$$

Conclusion (by Implication Elimination):

$$mean(jane)$$

Existential Elimination (EE)

$\exists\nu.\phi$

$\forall\nu.(\phi \Rightarrow \psi)$

ψ

where ν does not occur free in ψ

EE Example

Premises:

$$\exists x. hates(jane, x)$$

$$\forall x. (hates(jane, x) \Rightarrow mean(jane))$$

Conclusion:

$$mean(jane)$$

Existential Elimination (EE)

$\exists \nu.\phi$

$\forall \nu.(\phi \Rightarrow \psi)$

ψ

where ν does not occur ***free*** in ψ

EE Example

Premises:

$$\exists x.hates(jane,x)$$

$$\forall x.(hates(jane,x) \Rightarrow \forall x.hates(x,jane))$$

Conclusion:

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

Or Elimination

$$\frac{\begin{array}{l}\phi \vee \psi \\ \phi \Rightarrow \chi \\ \psi \Rightarrow \chi\end{array}}{\chi}$$

EE = OE on Steroids

EE

$$\frac{\exists \nu.\phi \qquad \forall \nu.(\phi \Rightarrow \psi)}{\psi} \quad \longrightarrow \quad \frac{\phi_1 \vee \ldots \vee \phi_n \qquad (\phi_1 \Rightarrow \psi) \wedge \ldots \wedge (\phi_n \Rightarrow \psi)}{\psi}$$

$\updownarrow$

$$\frac{\phi_1 \vee \ldots \vee \phi_n \qquad \phi_1 \Rightarrow \psi \qquad \ldots \qquad \phi_n \Rightarrow \psi}{\psi}$$

$\longleftarrow$

OE

$$\frac{\phi \vee \psi \qquad \phi \Rightarrow \chi \qquad \psi \Rightarrow \chi}{\chi}$$

Intuition Analogous to Universal Reasoning

Suppose we know $\exists \nu . \phi(\nu)$.

We hypothesize an object c and assume $\phi(c)$.

We try to prove ψ.

If does not contain c, then it is true for *any* such c.

We know that *there is some* ν from $\exists \nu . \phi(\nu)$.

So we can conclude ψ.

But we are still in the subproof.

Intuition Analogous to Universal Reasoning

Suppose we know $\exists \nu . \phi(\nu)$.

We hypothesize an object c and assume $\phi(c)$.

We try to prove ψ.

If does not contain c, then it is true for *any* such c.

We know that *there is some* ν from $\exists \nu . \phi(\nu)$.

So we can conclude ψ.

But we are still in the subproof.

So, we exit the subproof with $(\phi(c) \Rightarrow \psi)$.

We apply Universal Introduction to get $\forall \nu . (\phi(\nu) \Rightarrow \psi)$

Then we apply EE to get ψ outside the subproof.

Reasoning Tip for Existential Reasoning

If you have an existential sentence,
(1) assume the scope with a placeholder instead of variable,
(2) prove some conclusion,
(3) exit the assumption with an implication,
(4) generalize with Universal Introduction, and
(5) use Existential Elimination to derive the conclusion.

Proof

$$\exists y.\forall x.likes(x,y) \models \forall x.\exists y.likes(x,y)$$

	1.	$\exists y.\forall x.likes(x,y)$	Premise
(1)	2.	$\forall x.likes(x,d)$	Assumption
	3.	$likes(c,d)$	UE: 2
(2)	4.	$\exists y.likes(c,y)$	EI: 3
(3)	5.	$\forall x.likes(x,d) \Rightarrow \exists y.likes(c,y)$	II: 2, 4
(4)	6.	$\forall y.(\forall x.likes(x,y) \Rightarrow \exists y.likes(c,y))$	UI: 5
(5)	7.	$\exists y.likes(c,y)$	EE: 1, 6
	8.	$\forall x.\exists y.likes(x,y)$	UI: 7

Useful Result

1.	$\forall x.(p(x) \Rightarrow q(a))$	Premise
2.	$\mid \exists x.p(x)$	Assumption
3.	$\mid q(a)$	EE: 2, 1
4.	$\exists x.p(x) \Rightarrow q(a)$	II: 2, 3

Another Useful Result

1.	$\exists x.p(x) \Rightarrow q(a)$	Premise
2.	$\quad \mid p(c)$	Assumption
3.	$\quad \mid \exists x.p(x)$	EI: 2
4.	$\quad \mid q(a)$	IE: 1, 3
5.	$p(c) \Rightarrow q(a)$	II: 2, 4
6.	$\forall x.(p(x) \Rightarrow q(a))$	UI: 5

Fitch Online System

Course Website

http://logica.stanford.edu

Proof

1.	$\forall x.(p(x) \Rightarrow q(x))$	Premise
2.	$\forall x.p(x)$	Premise
3.	$p(c) \Rightarrow q(c)$	UE: 1
4.	$p(c)$	UE: 2
5.	$q(c)$	IE: 4, 3
6.	$\forall x.q(x)$	UI: 5

Proof

$$\forall x.(p(x) \Rightarrow q(x)), \forall x.(q(x) \Rightarrow r(x)) \models \forall x.(p(x) \Rightarrow r(x))$$

1.	$\forall x.(p(x) \Rightarrow q(x))$	Premise
2.	$\forall x.(q(x) \Rightarrow r(x))$	Premise
3.	$p(c) \Rightarrow q(c)$	UE: 1
4.	$q(c) \Rightarrow r(c)$	UE: 2
5.	$\vert\ p(c)$	Assumption
6.	$\vert\ q(c)$	IE: 5, 3
7.	$\vert\ r(c)$	IE: 6, 4
8.	$p(c) \Rightarrow r(c)$	II: 5, 7
9.	$\forall x.(p(x) \Rightarrow r(x))$	UI: 8