Workshop on Inference from Text: Natural Logic

Larry Moss, Indiana University

NASSLLI, June 21–22, 2010

Can we reason in language? Two views

If we were to devise a logic of ordinary language
for direct use on sentences as they come,
we would have to complicate our rules of inference
in sundry unilluminating ways.

W. V. O. Quine, Word and Object

1		John is a man	Hyp
2		Any woman is a mystery to any man	Hyp
3	Jane	Jane is a woman	Hyp
4		Any woman is a mystery to any man	R, 2
5		Jane is a mystery to any man	Any Elim, 4
6		John is a man	R, 1
7		Jane is a mystery to John	Any Elim, 6
8		Any woman is a mystery to John	Any intro, 3, 7

1		Thistledew is a house	Hyp
2		Any man who owns a house should paint that house	Hyp
3	Joe	Joe is a man	Any which elim, 2
4		If any man owns a house, then that man should paint that house	Hyp
5		If Joe owns a house, then Joe should paint that house	Any elim, 3, 4
6		If Joe owns Thistledew, then Joe should paint Thistledew	special a elim, 3, 4
7		If any man owns Thistledew, that man should paint that Thistledew	Any intro
8		Any man who owns Thistledew should paint that Thistledew	Any which intro

This talk deals with new logical systems tuned to natural language

- The raison d'être of logic is the study of inference in language.
- However, modern logic was developed in connection with the foundations of mathematics.
- So we have a mismatch, leading to
  — neglect of language in the first place
  — use of first-order logic and no other tools
- First-order logic is both too big and too small:
  — cannot handle many interesting phenomena
  — is undecidable

Program

Show that significant parts of natural language inference can be carried out in decidable logical systems.

Whenever possible, to obtain complete axiomatizations, because the resulting logical systems are likely to be interesting.

To be completely mathematical and hence to work using all tools and to make connections to fields like
complexity theory, (finite) model theory,
decidable fragments of first-order logic, and algebraic logic.

Natural Logic: parallel studies

I won't have much to say on these, but you can ask me about them

- History of logic: reconstruction of original ideas
- Philosophy of language: proof-theoretic semantics
- Philosophy of logic: why variables?
- Cognitive science: models of human reasoning
- Linguistic semantics:
  Are deep structures necessary, or can we just
  use surface forms?
  And is a complete logic a semantics?
- Computational linguistics/artificial intelligence:
  this connection is the overall topic of our workshop.

All men are mortal.
Socrates is a man.

Socrates is mortal.

Some men are mortal.
Socrates is a man.

Socrates is mortal.

All frogs are reptiles.
All reptiles are animals.

All frogs are animals.

All frogs are reptiles.
All frogs are animals.

All reptiles are animals.

All sagatricians are maltnomans.
All sagatricians are aikims.

All maltnomans are aikims.

The first point is that there is an exact definition of *validity* for arguments.
The second point here is that the *form* is as important, even more important, than the particular words.

$$\frac{\begin{array}{c}\textit{All } X \textit{ are } Y.\\ \textit{All } Y \textit{ are } Z.\\ \textit{All } W \textit{ are } X.\end{array}}{\textit{All } W \textit{ are } Z.}$$

So valid arguments can have more than two *premises*.

Our plan is to deal only with sentences containing all, some, and no.

Probably the key point of logic is that there is a distinction between

syntax and semantics.

The idea is that syntax is the raw symbols.

Syntax is usually painful: think of computer programming.
I'll try to avoid it as much as possible to concentrate on the ideas.

The semantics is where we get the meaning.
So in our examples, we need some context or model to give a meaning.
In our examples, the syntax will start with some variables p, q, n, n_1, ...
The our sentences are expressions of the form
All p *are* q, *Some* p *are* q, and *No* p *are* q

To say whether

All sagatricians (s) are maltnomans (m).

is true or not needs a model.
This is given by a few things:

First, a set U called the universe.
Second, for the words sagatrician and maltnoman,
we need sets $[\![\textit{sagatrician}]\!] \subseteq U$ and $[\![\textit{maltnomanan}]\!] \subseteq U$.

Given all of this, we say that

All s are m

is true in the model if $[\![S]\!] \subseteq [\![M]\!]$.
Otherwise, *All s are m* is false in the model.

What should we say about *Some s are m* and *No s are m*?

Syntax: *All p are q*, *Some p are q*, *No p are q*

Semantics: A model $\mathcal{M}$ is a set M,
and for each variable p we have an interpretation $[\![p]\!] \subseteq M$.

$$\mathcal{M} \models \textit{All p are q} \qquad \text{iff} \qquad [\![p]\!] \subseteq [\![q]\!]$$

The symbols $\mathcal{M} \models S$ is read as $\mathcal{M}$satisfies S.

A statement like $\mathcal{M} \models \textit{All p are q}$ could also be read as

All p are q is true in $\mathcal{M}$

Is All X are Y true or not?

Is All X are Y true or not?

Without a model (also called a context), the question makes no sense.

So let's take an example model, and ask whether our sentence is true in that model or not.

Let $U = \{1, 2, 3, 4, 5\}$
Let $[\![X]\!] = \{1, 2, 4\}$.
Let $[\![Y]\!] = \{3, 4\}$.

Is All X are Y true or not?

Without a model (also called a context), the question makes no sense.

So let's take an example model, and ask whether our sentence is true in that model or not.

Let $U = \{1, 2, 3, 4, 5\}$
Let $[\![X]\!] = \{1, 2, 4\}$.
Let $[\![Y]\!] = \{3, 4\}$.

In this model, All X are Y is false!

But if we change the model by re-setting $[\![Y]\!]$ to $\{1, 2, 3, 4\}$, then our sentence is true.

One fine point on the definition is that if $[\![X]\!]$ is the empty set $\emptyset$, then our sentence All X are Y is *true*!
So in this room now,

All people in the room over 7 feet tall are standing

is (on this definition) true.

This strange point will lead us to various issues over the next few weeks.

For now, it might be best to say that it's true because there are *no exceptions*.

But we again admit that the semantics of All that we are giving is not what most people would agree to in cases where $[\![X]\!] = \emptyset$.
(There is also another quirk that we'll see soon.)

At this point, we know how to give the semantics of single sentences.
We say that a sentence S follows from sentences $A_1, \ldots, A_n$ if every model that makes all of the As true also makes S true.
We write this as

$$A_1, \ldots, A_n \models S$$

and we also say that the A's semantically imply S.

To argue that $A_1, \ldots, A_n \models S$ we need some reasoning.
Usually, we do this in English and in an informal way,
just as one would do ordinary reasoning.

But to argue that $A_1, \ldots, A_n \not\models S$ we can produce a counterexample.

In all of this work, the main thing is that we have a rigorous definition.

A small note on notation

We use letters like Γ (Greek letter Gamma) for sets of sentences.

And then we would write $\Gamma \models S$ to mean
that every model of all the sentences in Γ is also a model of φ.

However, if Γ is a set that we have listed out, say
$\Gamma = \{A_1, A_2, \ldots, A_{104}\}$.
then usually we would write $\Gamma \models S$ as

$$A_1, A_2, \ldots, A_{104} \models S$$

rather than as

$$\{A_1, A_2, \ldots, A_{104}\} \models S.$$

That is, we drop the set braces on the left of the $\models$ symbol.
We do this to make things a little more readable.

$$\overbrace{A_1, A_2, \ldots, A_n}^{\text{premises}} \quad \models \quad \overbrace{S}^{\text{conclusion}}$$

The intuition is that

$$A_1, A_2, \ldots, A_n \models S$$

means that

any circumstance in which the premises $A_1, A_2, \ldots, A_n$ are all true is also a circumstance in which the conclusion S is true

$$\frac{\textit{All frogs are reptiles.} \quad \textit{All frogs are animals.}}{\textit{All reptiles are animals.}}$$

We can take $U = \{1, 2, 3, 4, 5, 6\}$.
$[\![F]\!] = \{1, 2\}$,
$[\![R]\!] = \{1, 2, 3, 4\}$.
$[\![A]\!] = \{1, 2, 4, 5, 6\}$.
In this context, the assumptions are true but the conculsion is false.
So the argument is invalid.

All frogs are reptiles, *All frogs are animals*, $\not\models$ *All reptiles are animals.*

Note the difference between syntax and semantics.
$\models$ is intended to mean follows by general-purpose reasoning.

We can check whether our definitions match with our intuitions.
In the case of our very simple fragment, this mostly is right.

The main exception is that people usually wouldn't say
All X are Y
in a context where they know that there are no X.

A secondary point is that a computer should be able to decide whether

$$A_1, \ldots, A_n \models S$$

or not.

The entailment problem should be decidable.

Another way to make these points:
the definitions and theory should be "tight" enough so that the decision can be made *without semantics* (!),
by only looking at the *form* of the argument.

Hoch means every.

Hoch verengan Ha'DIbah jajlo' Qa'
Hoch jajlo'Qa' Qa'Hom.

Hoch verengan Qa'Hom.

Let Γ be a set of sentences $\{A_1, \ldots, A_n\}$.
A proof tree over Γ is a tree following properties:

1. The leaves are either labeled with sentences from Γ, or with sentences of the form *All* X *are* X.
2. The interior leaves match one of the rules of our system (see the next slide).

The trees are drawn with the root at the bottom and the leaves at the top.

If there is a proof tree over Γ whose root is labeled S,
we write $\Gamma \vdash S$.

We say that S is provable from Γ in our system.

The rules for building trees

$$\frac{}{\textit{All } p \textit{ are } p} \qquad\qquad \frac{\textit{All } p \textit{ are } n \quad \textit{All } n \textit{ are } q}{\textit{All } p \textit{ are } q}$$

Here is an example: Let Γ be the set

$$\{\textit{All A are B}, \textit{All Q are A}, \textit{All B are D}, \textit{All C are D}, \textit{All A are Q}\}$$

Let S be *All Q are D*. Here is a proof tree showing that $\Gamma \vdash S$:

$$\dfrac{\textit{All Q are A} \qquad \dfrac{\textit{All A are B} \quad \textit{All B are D}}{\textit{All A are D}}}{\textit{All Q are D}}$$

All of the leaves belong to Γ.
Note also that some elements of Γ are not used as leaves.

This is permitted according to our definition.
The proof tree above shows that $\Gamma \vdash S$.

What are we doing here?

The idea is that proof trees are our model of basic reasoning using the words all, some, no.

A proof tree is like a caricature of a real proof.

It can be examined (and even constructed) by a person or computer
who has no understanding of anything but the rules!

There are several hopes about this work:
⋆ The whole thing will "scale up" to include many more words. (This would call on linguistic semantics to provide the correct notion of context.)
⋆ The formal relation $\vdash$ should have something to do with $\models$ (logic)
⋆ The proof system $\vdash$ should have something to do with actual human reasoning (psychology)
⋆ A computer should be able to work with $\vdash$ without understanding anything.

A computer could check whether a purported tree actually satisfies our definition, even if it didn't "understand" *All*.
So one important question is: what is the relation between

$$\Gamma \vdash A \quad \text{and} \quad \Gamma \models A \quad ?$$

Soundness

If $\Gamma \vdash S$, then $\Gamma \models S$.

This means that proof trees do not lead us astray:
if $\Gamma \vdash S$, then in any context where the sentences of Γ all hold, S too must hold.
Our proof system will not lead us to believe that bogus syllogisms are in fact valid.

Here is the basic idea of why the Soundness Lemma holds.
The two most basic facts about $\subseteq$ are:

1. $X \subseteq X$ for all sets X.
2. For all sets X, Y, and Z: if $X \subseteq Y$ and $Y \subseteq Z$, then $X \subseteq Z$.

(Probably the third would be that $\emptyset \subseteq X$ for all X.)

Soundness Sketch, Continued

Let's go back to our example proof tree.

$$\dfrac{\textit{All } Q \textit{ are } A \qquad \dfrac{\textit{All } A \textit{ are } B \qquad \textit{All } B \textit{ are } D}{\textit{All } A \textit{ are } D}}{\textit{All } Q \textit{ are } D}$$

Take any model, say $\mathcal{M}$.
Assume that in $\mathcal{M}$, $[\![A]\!] \subseteq [\![B]\!]$, etc.
We have to show that in this same model $\mathcal{M}$, $[\![Q]\!] \subseteq [\![D]\!]$.

The idea is to use our proof tree and specialize it to $\mathcal{M}$:

$$\dfrac{[\![Q]\!] \subseteq [\![A]\!] \qquad \dfrac{[\![A]\!] \subseteq [\![B]\!] \qquad [\![B]\!] \subseteq [\![D]\!]}{[\![A]\!] \subseteq [\![D]\!]}}{[\![Q]\!] \subseteq [\![D]\!]}$$

And then going downward mirrors *intuitively valid reasoning in the model*.

Since the model $\mathcal{M}$ was arbitrary (had no special features),
the conclusion $\Gamma \models S$ holds.

$$\Gamma = \left\{ \begin{array}{l} \text{All A are B,} \\ \text{All A are C,} \\ \text{All B are C,} \\ \text{All C are B,} \\ \text{All C are D,} \\ \text{All B are E,} \\ \text{All D are G,} \\ \text{All F are G,} \\ \text{All G are F} \end{array} \right\}$$

We see that $\Gamma \vdash$ All B are G.

Do you think that $\Gamma \vdash$ *All D are E*?

Is there an algorithm to tell yes or no?

Definition

A preorder is a pair $(P, \leq)$,
where P is a set
and $\leq$ is a relation on it with the following properties:

REFLEXIVE $p \leq p$

TRANSITIVE If $p \leq q$ and $q \leq r$, then $p \leq r$.

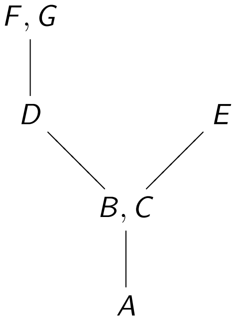
We need not have the following property:

ANTI-SYMMETRIC if $p \leq q$ and $q \leq p$, then $p = q$.

An anti-symmetric preorder is a partially ordered set (poset).

A picture of a preorder

$$\Gamma = \left\{ \begin{array}{l} \text{All A are B}, \\ \text{All A are C}, \\ \text{All B are C}, \\ \text{All C are B}, \\ \text{All C are D}, \\ \text{All B are E}, \\ \text{All D are G}, \\ \text{All F are G}, \\ \text{All G are F} \end{array} \right\}$$

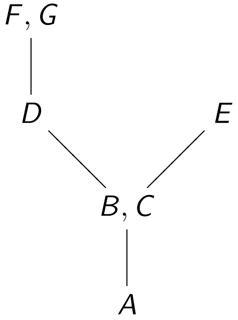


The set P here is $\{A, \ldots, G\}$.
The order $\leq$ is given by

$$X \leq Y \quad \text{iff} \quad \Gamma \vdash \text{All } X \text{ are } Y.$$

Downsets in preorders

In a preorder, $\downarrow p = \{x : x \leq p\}$.



$\downarrow F = \downarrow G = \{A, \ldots, G\}$

$\downarrow D = \{A, B, C, D\}, \downarrow E = \{A, B, C, E\}$

$\downarrow B = \{A, B, C\} = \downarrow C$

$\downarrow A = \{A\}$

$\downarrow$ is monotone: if $p \leq q$, then $\downarrow p \ \subseteq \downarrow q$.

Suppose that $\Gamma \models$ *All X are Y*.
Let M be the set of variables.
Define $A \leq B$ to mean that $\Gamma \vdash$ *All A are B*.
Check that this is reflexive and transitive, using the logic.
The semantics is via downsets:

$$[\![A]\!] \quad = \quad \downarrow A \quad = \quad \{B : B \leq A\}$$

By transitivity, $\mathcal{M} \models \Gamma$.

Suppose that $\Gamma \models$ *All X are Y*.
Let M be the set of variables.
Define $A \leq B$ to mean that $\Gamma \vdash$ *All A are B*.
Check that this is reflexive and transitive, using the logic.
The semantics is via downsets:

$$[\![A]\!] \quad = \quad \downarrow A \quad = \quad \{B : B \leq A\}$$

By transitivity, $\mathcal{M} \models \Gamma$.
In more detail, suppose Γ contains All C are D.
Then if $\mathrm{W} \leq C$, then also $\mathrm{W} \leq D$.

Suppose that $\Gamma \models$ *All X are Y*.
Let M be the set of variables.
Define $A \leq B$ to mean that $\Gamma \vdash$ *All A are B*.
Check that this is reflexive and transitive, using the logic.
The semantics is via downsets:

$$[\![A]\!] \quad = \quad \downarrow A \quad = \quad \{B : B \leq A\}$$

By transitivity, $\mathcal{M} \models \Gamma$.
So $[\![X]\!] \subseteq [\![Y]\!]$.
But by reflexivity $X \in [\![X]\!]$.
And so $X \in [\![X]\!]$; this means that $X \leq Y$.

An example of how the proof works

All A are B
All A are C
All B are C
All C are B
All C are D
All B are E
All D are G
All F are G
All G are F

$[\![A]\!] = \{A\}$
$[\![B]\!] = \{A, B, C\}$
$[\![C]\!] = \{A, B, C\}$
$[\![D]\!] = \{A, B, C, D\}$
$[\![E]\!] = \{A, B, C, E\}$
$[\![F]\!] = \{A, \ldots, G\}$
$[\![G]\!] = \{A, \ldots, G\}$

Getting back to the question of whether or not $\Gamma \vdash$ *All D are E*.

Since $[\![D]\!]$ is not a subset of $[\![E]\!]$, $\Gamma \not\models$ *All D are E*.
By soundness, $\Gamma \nvdash$ *All D are E*.

Syllogistic Logic of *All* and *Some*

Syntax: *All p are q*, *Some p are q*

Semantics: A model $\mathcal{M}$ is a set M,
and for each noun p we have an interpretation $[\![p]\!] \subseteq M$.

$$\begin{array}{lll} \mathcal{M} \models \textit{All p are q} & \text{iff} & [\![p]\!] \subseteq [\![q]\!] \\ \mathcal{M} \models \textit{Some p are q} & \text{iff} & [\![p]\!] \cap [\![q]\!] \neq \emptyset \end{array}$$

Proof system:

$$\frac{}{\textit{All p are p}} \qquad \frac{\textit{All p are n} \quad \textit{All n are q}}{\textit{All p are q}}$$

$$\frac{\textit{Some p are q}}{\textit{Some q are p}} \qquad \frac{\textit{Some p are q}}{\textit{Some p are p}} \qquad \frac{\textit{All q are n} \quad \textit{Some p are q}}{\textit{Some p are n}}$$

Some n are n, *All n are p*, *All n are q* $\vdash$ *Some p are q*.

The proof tree is

$$\dfrac{\textit{All n are q} \qquad \dfrac{\dfrac{\textit{All n are p} \quad \textit{Some n are n}}{\textit{Some n are p}}}{\textit{Some p are n}}}{\textit{Some p are q}}$$

Beyond first-order logic: cardinality

Read $\exists^{\geq}(X, Y)$ as "there are at least as many Xs as Ys".

$$\frac{\textit{All } Y \textit{ are } X}{\exists^{\geq}(X, Y)} \qquad \frac{\exists^{\geq}(X, Y) \quad \exists^{\geq}(Y, Z)}{\exists^{\geq}(X, Z)}$$

$$\frac{\textit{All } Y \textit{ are } X \quad \exists^{\geq}(Y, X)}{\textit{All } X \textit{ are } Y}$$

$$\frac{\textit{Some } Y \textit{ are } Y \quad \exists^{\geq}(X, Y)}{\textit{Some } X \textit{ are } X} \qquad \frac{\textit{No } Y \textit{ are } Y}{\exists^{\geq}(X, Y)}$$

The point here is that by working with a weak basic system, we can go beyond the expressive power of first-order logic.

The languages $\mathcal{S}$ and $\mathcal{S}^\dagger$ add noun-level negation

Let us add complemented atoms $\overline{p}$ on top of
the language of All and Some,
with interpretation via set complement: $[\![\overline{p}]\!] = M \setminus [\![p]\!]$.

So we have

$$\left.\begin{array}{l}
\mathcal{S}\quad\left\{\begin{array}{l}
\textit{All } p \textit{ are } q \\
\textit{Some } p \textit{ are } q \\
\textit{All } p \textit{ are } \overline{q} \equiv \textit{No } p \textit{ are } q \\
\textit{Some } p \textit{ are } \overline{q} \equiv \textit{Some } p \textit{ aren't } q
\end{array}\right. \\
\\
\qquad\quad \textit{Some non-p are non-q}
\end{array}\right\}\quad \mathcal{S}^\dagger$$

$$\frac{}{\textit{All } p \textit{ are } p} \qquad \frac{\textit{Some } p \textit{ are } q}{\textit{Some } p \textit{ are } p} \qquad \frac{\textit{Some } p \textit{ are } q}{\textit{Some } q \textit{ are } p}$$

$$\frac{\textit{All } p \textit{ are } n \quad \textit{All } n \textit{ are } q}{\textit{All } p \textit{ are } q} \qquad \frac{\textit{All } n \textit{ are } p \quad \textit{Some } n \textit{ are } q}{\textit{Some } p \textit{ are } q}$$

$$\frac{\textit{All } q \textit{ are } \overline{q}}{\textit{All } q \textit{ are } p}\ \textit{Zero} \qquad \frac{\textit{All } \overline{q} \textit{ are } q}{\textit{All } p \textit{ are } q}\ \textit{One}$$

$$\frac{\textit{All } p \textit{ are } \overline{q}}{\textit{All } q \textit{ are } \overline{p}}\ \textit{Antitone} \qquad \frac{\textit{Some } p \textit{ are } \overline{p}}{\varphi}\ \textit{Ex falso quodlibet}$$

The system uses

$$\frac{\textit{Some } p \textit{ are } \overline{p}}{\varphi}\ \textit{Ex falso quodlibet}$$

and this is prima facie weaker than reductio ad absurdum.

One of the logical issues in this work is to determine exactly where various principles are needed.

Completeness via representation of orthoposets

Definition

An orthoposet is a tuple $(P, \leq, 0, ')$ such that

- POSET: $\leq$ is a reflexive, transitive, and antisymmetric relation on the set P.
- ZERO: $0 \leq p$ for all $p \in P$.
- ANTITONE: If $x \leq y$, then $y' \leq x'$.
- INVOLUTIVE: $x'' = x$.
- INCONSISTENCY: If $x \leq y$ and $x \leq y'$, then $x = 0$.

A Key Point

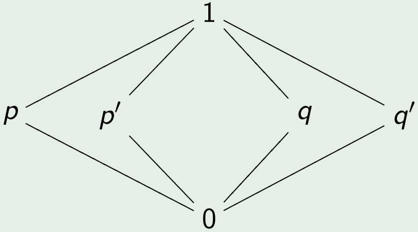
Orthoposets need not have a meet or join operation.

Orthoposets: two examples

Example

For all sets X we have an orthoposet $(\mathcal{P}(X), \subseteq, \emptyset, ')$, where $a' = X \setminus a$ for all subsets a of X.

Example



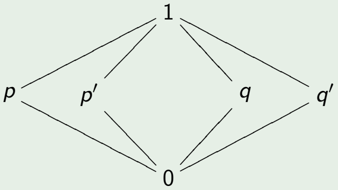
$(x')' = x$, $0' = 1$, $1' = 0$.

Orthoposets: two examples

Example

For all sets X we have an orthoposet $(\mathcal{P}(X), \subseteq, \emptyset, ')$, where $a' = X \setminus a$ for all subsets a of X.

Example



$(x')' = x$, $0' = 1$, $1' = 0$.

The idea

$$\frac{\text{boolean algebra}}{\text{propositional logic}} = \frac{\text{orthoposet}}{\text{logic of All, Some and }'}$$

The details concerning completeness are somewhat different, and the whole thing would take about 10 minutes.

Let Γ be any set of sentences in the fragment.
Let $\mathcal{V}$ be the set of variables.
We already know the preorder $\leq$:

$$X \leq \mathrm{Y} \quad \text{iff} \quad \Gamma \vdash \textit{All } X \textit{ are } Y.$$

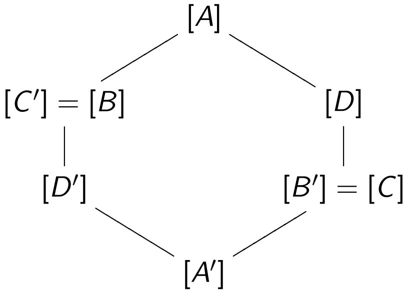
(so Some plays no role)
We have an induced equivalence relation $\equiv$,
and we take $\mathcal{V}_\Gamma$ to be the quotient $\mathcal{V}/\!\equiv$.
If there is some X such that $X \leq X'$, then set 0 to be $[X]$.
We finally define $[X]' = [X']$.
If there is no X such that $X \leq X'$, we add fresh elements 0 and 1 to $\mathcal{V}/\!\equiv$.
It is not hard to check that we have an orthoposet $\mathcal{V}_\Gamma$.

Orthoposets from logic, concretely

Let $\Gamma =$
$\{$All B are A, All B' are A, All C' are B, All C are B', All C are $D\}$.
Then

$$\begin{array}{ll} [A] = \{A\} & [A'] = \{A'\} \\ [B] = \{B, C'\} & [B'] = \{B', C\} \\ [C] = \{B', C\} & [C'] = \{B, C'\} \\ [D] = \{D\} & [D'] = \{D'\} \end{array}$$

Here is a picture of the orthoposet $\mathcal{V}_\Gamma$:



A point of a orthoposet $P = (P, \leq, 0, ')$ is a subset $\mathcal{S} \subseteq P$ with the following properties:

UP-CLOSED If $p \in S$ and $p \leq q$, then $q \in \mathcal{S}$.

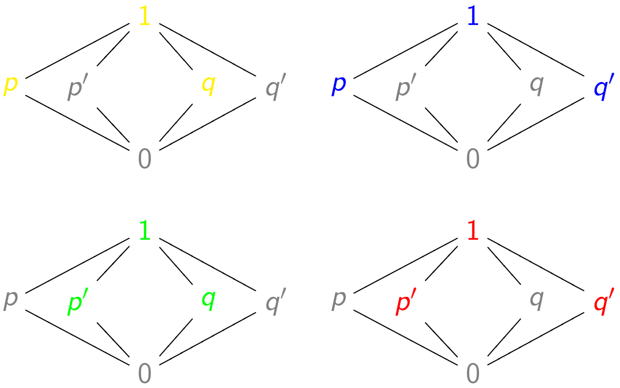
COMPLETE For all p, either $p \in S$ or $p' \in \mathcal{S}$.

PAIRWISE COMPATIBLE For all $p, q \in \mathcal{S}$, $p \not\leq q'$.

Points are sets

Look back at the Chinese lantern.

There are four points here: the sets marked •, •, •, and •:



What are the points?

$$
\begin{tikzcd}
 & {[A]} \arrow[dl, dash] \arrow[dr, dash] & \\
{[C'] = [B]} \arrow[d, dash] & & {[D]} \arrow[d, dash] \\
{[D']} \arrow[dr, dash] & & {[B'] = [C]} \arrow[dl, dash] \\
 & {[A']} &
\end{tikzcd}
$$

What are the points?

There are three points.

$$\begin{tikzcd}[arrows=dash]
 & {[A]} \arrow[dl] \arrow[dr] & \\
{[C']=[B]} \arrow[d] & & {[D]} \arrow[d] \\
{[D']} \arrow[dr] & & {[B']=[C]} \arrow[dl] \\
 & {[A']} &
\end{tikzcd}$$

$\mathcal{S}=\{[D'],[B],[A]\}$, $\mathcal{T}=\{[B'],[D],[A]\}$, $\mathcal{U}=\{[B],[D],[A]\}$.

Let $X = \{1, 2, 3\}$, and let $\mathcal{P}(X)$ be the power set orthoposet.
Then $\mathcal{S}$ is a point, where

$$\mathcal{S} \quad = \quad \{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}.$$

It is easy to check that the points on this $\mathcal{P}(X)$ are exactly $\mathcal{S}$ as above and the three principal ultrafilters.
$\mathcal{S}$ shows that a point of a boolean algebra need not be a filter.

The Extension Lemma for pairwise consistent sets

Lemma

Let $\mathcal{S} \subseteq P$ be pairwise consistent: $(\forall p, q \in \mathcal{S}) p \not\leq q'$.
Then for all $x \in P$, either $\mathcal{S} \cup \{x\}$ or $\mathcal{S} \cup \{x'\}$ is again pairwise consistent.

Proof.

Suppose not. Then x and x' figure in to problems both times.
There is some $p \in \mathcal{S}$ such that $p \leq x'$.
There is some $q \in \mathcal{S}$ such that $q \leq x'' = x$.
And now: $q \leq x \leq p'$. Ooops! $\square$

The Extension Lemma for pairwise consistent sets

Lemma

Let $\mathcal{S} \subseteq P$ be pairwise consistent: $(\forall p, q \in \mathcal{S}) p \not\leq q'$.
Then for all $x \in P$, either $\mathcal{S} \cup \{x\}$ or $\mathcal{S} \cup \{x'\}$ is again pairwise consistent.

Lemma

If $p \not\leq q$, then $\{p, q'\}$ is pairwise consistent.
Thus there is a point $\mathcal{S}$ containing p but not q.

Pairwise consistent sets extend to points

Lemma

For a subset S_0 of an orthoposet $P = (P, \leq, 0, ')$, the following are equivalent:

1. *S_0 is a subset of a point $\mathcal{S}$ in P.*
2. *S_0 is pairwise compatible.*

Proof.

Clearly (1) $\Longrightarrow$ (2).
For the more important direction, use Zorn's Lemma to get $\mathcal{S} \supseteq S_0$ which is pairwise compatible, and maximal with this property.
For all p, either p or p' belongs to $\mathcal{S}$. [By maximality.]
Check easily that $\mathcal{S}$ is up-closed:
If $p \in \mathcal{S}$, $p \leq q$, but $q \notin S$, then $q' \in \mathcal{S}$.
And now $p \leq q = (q')'$, so $\mathcal{S}$ is not pairwise compatible. $\square$

Let $P = (P, \leq, ')$ be an orthoposet.
Let $points(P)$ be the set of points of P.
We have an orthoposet

$$(\mathcal{P}(points(P)), \subseteq, \emptyset, ')$$

Let $m : P \to \mathcal{P}(points(P))$ be given by

$$m(p) \quad = \quad \{\mathcal{S} : p \in \mathcal{S}\}.$$

Theorem

m is a strict morphism of orthoposets:
$m(0) = \emptyset$,
$m(p') = (m(p))'$,
and $p \leq q$ *iff* $m(p) \subseteq m(q)$.

Let $P = (P, \leq, ')$ be an orthoposet.
Let $points(P)$ be the set of points of P.
We have an orthoposet

$$(\mathcal{P}(points(P)), \subseteq, \emptyset, ')$$

Let $m : P \to \mathcal{P}(points(P))$ be given by

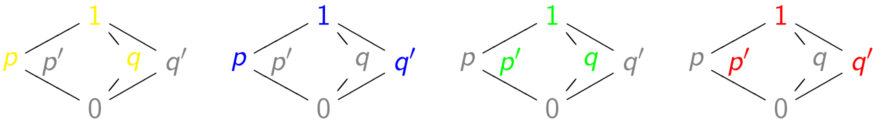
$$m(p) \quad = \quad \{\mathcal{S} : p \in \mathcal{S}\}.$$

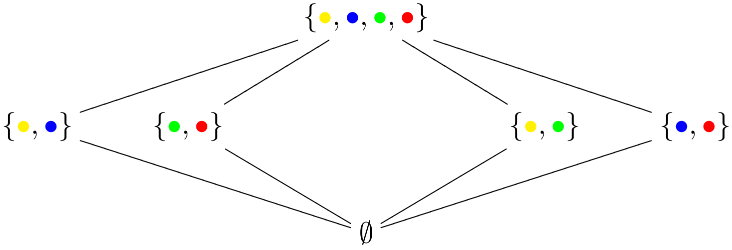
Theorem

m is a strict morphism of orthoposets:
$m(0) = \emptyset$,
$m(p') = (m(p))'$,
and $p \leq q$ iff $m(p) \subseteq m(q)$.

Corollary

Every orthoposet is isomorphic to a sub-orthoposet of a power set orthoposet.


1
p
p′
q
q′
0
1
p
p′
q
q′
0
1
p
p′
q
q′
0
1
p
p′
q
q′
0


∅

Another one

$\mathcal{S} = \{[D'], [B], [A]\}$, $\mathcal{T} = \{[B'], [D], [A]\}$, $\mathcal{U} = \{[B], [D], [A]\}$.

$$
\begin{tikzcd}[arrows=dash]
 & {[A]} \arrow[dl] \arrow[dr] & \\
{[C'] = [B]} \arrow[d] & & {[D]} \arrow[d] \\
{[D']} \arrow[dr] & & {[B'] = [C]} \arrow[dl] \\
 & {[A']} &
\end{tikzcd}
$$

$$
\begin{array}{ll}
m([A]) = \{\mathcal{S}, \mathcal{T}, \mathcal{U}\} & m([A']) = \emptyset \\
m([B]) = \{\mathcal{S}, \mathcal{U}\} & m([B']) = \{\mathcal{T}\} \\
m([D]) = \{\mathcal{T}, \mathcal{U}\} & m([D']) = \{\mathcal{S}\}
\end{array}
$$

This pretty much solves our earlier problem of getting a model of Γ where $[\![B]\!] \not\subseteq [\![D]\!]$.

But how?

Sources the Representation Theorem

N. Zierler and M. Schlessinger

Boolean embeddings of orthomodular sets and quantum logic.
Duke Mathematical Journal 32 (1965), 251–262.

F. Katrnoška

On the representation of orthocomplemented posets.
Comment. Math. Univ. Carolinae 23 (1982), 489–498.

C. S. Calude, P. H. Hertling, K. Svozil

Embedding quantum universes into classical ones.
Foundations of Physics, 29, 3 (1999), 349-379.

Lemma

Let Γ be consistent in $\mathcal{L}(\text{all}, \text{some}, ')$.
There is a canonical model $\mathcal{M} = (M, [\![\ ]\!])$ such that

1. $\mathcal{M} \models \Gamma$.
2. *If $\mathcal{M} \models$ All X are Y, then $\Gamma \vdash$ All X are Y.*

Proof.

Let $\mathcal{V}_\Gamma$ be the syntactic orthoposet for Γ. Let $M = \mathit{points}(\mathcal{V}_\Gamma)$.
The interpretation $[\![\ ]\!] : \mathcal{V} \to \mathcal{P}(M)$ is given by

$$\mathcal{V} \xrightarrow{\ n\ } \mathcal{V}_\Gamma \xrightarrow{\ m\ } \mathcal{P}(\mathit{points}(\mathcal{V}_\Gamma)) = \mathcal{P}(M)$$

Key point If Γ contains Some U are V, need a point including $\{[\mathsf{U}], [\mathsf{V}]\}$.
If none exists, then wlog $\mathsf{U} \leq \mathsf{V}'$. But then Γ is inconsistent. $\square$

Lemma

Let Γ be consistent in $\vdash$.
There is a canonical model $\mathcal{M} = (M, [\![\ ]\!])$ such that

1. $\mathcal{M} \models \Gamma$.
2. *If $\mathcal{M} \models$ All X are Y, then $\Gamma \vdash$ All X are Y.*

This gives half of completeness:

If $\Gamma \models$ All X are Y,
then $\mathcal{M} \models$ All X are Y,
and so $\Gamma \vdash$ All X are Y.

For Some sentences, we need a little more.

If Γ is consistent and $\Gamma \models$ Some X are Y, then $\Gamma \vdash$ Some X are Y

Lemma (Ian Pratt-Hartmann 2007)

There is some existential sentence in Γ, say Some A are B, such that

$$\Gamma_{all} \cup \{\textit{Some A are B}\} \models \textit{Some X are Y}.$$

If Γ is consistent and $\Gamma \models$ Some X are Y, then $\Gamma \vdash$ Some X are Y

Fix A and B as in the lemma.
Consider the model $\mathcal{M} = \mathcal{M}(\mathcal{V}_{\Gamma_{all}})$ of points on $\mathcal{V}_{\Gamma_{all}}$. $\mathcal{M} \models \Gamma_{all}$.
Consider $\{[A], [B], [X']\}$.
If this set were a subset of a point $\mathcal{S}$, then consider $\{\mathcal{S}\}$ as a one-point submodel of $\mathcal{M}$.
In the submodel, $\Gamma_{all} \cup \{$Some A are $B\}$ would hold, and yet Some X are Y would fail, since $[\![X]\!] = \emptyset$.
Therefore $\{[A], [B], [X']\}$ is not pairwise compatible.
There are six cases:

$$\begin{array}{ll} A \leq A' & A \leq B' \\ A \leq X & B \leq B' \\ B \leq X & X' \leq X \end{array}$$

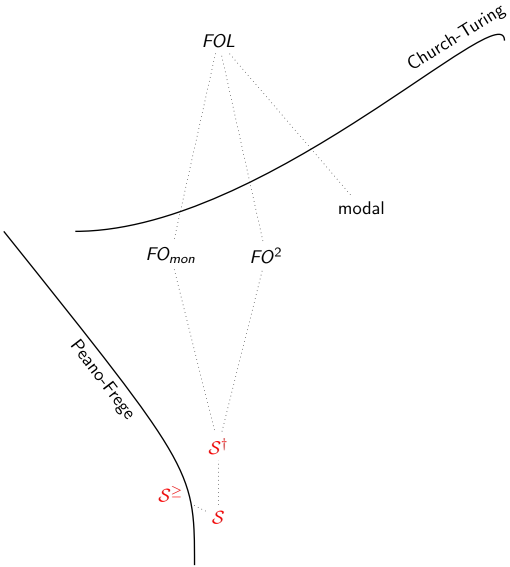
Only two are significant.

Next, consider $\{A, B, Y'\}$.
The same analysis gives two other cases: $A \leq Y$ and $B \leq Y$.
Putting these together with the other two gives four pairs.
The case when $A \leq X$ and $B \leq Y$ is representative:

$$
\dfrac{\begin{array}{c}\vdots \\ \textit{All } B \textit{ are } Y\end{array} \qquad \dfrac{\dfrac{\begin{array}{c}\vdots \\ \textit{All } A \textit{ are } X\end{array} \quad \textit{Some } B \textit{ are } A}{\textit{Some } B \textit{ are } X}}{\textit{Some } X \textit{ are } B}}{\textit{Some } X \textit{ are } Y}
$$

The other cases are similar. This completes the proof.

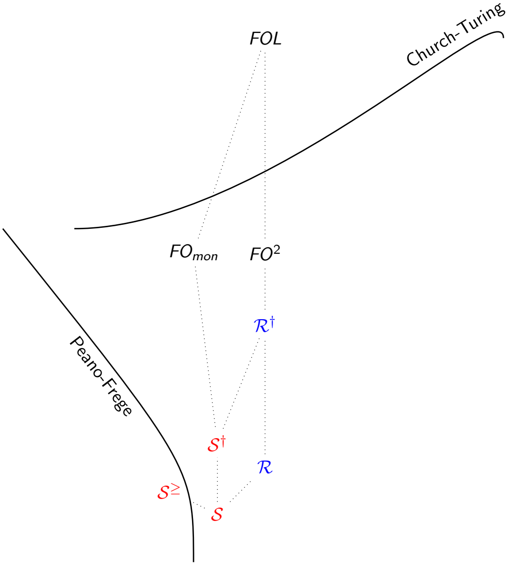
Picture



monadic FOL
2 variable fragment

† adds full N-negation
We have discussed these

How about verbs?



† adds full N-negation

We have discussed these
relational syllogistic
next

The next language uses "see" or r as variables for transitive verbs.

All p are q	*All p aren't q* $\equiv$ *No p are q*
Some p are q	*Some p aren't q*
All p see all q	*All p don't see all q* $\equiv$ *No p sees any q*
All p see some q	*All p don't see some q* $\equiv$ *No p sees all q*
Some p see all q	*Some p don't see any q*
Some p see some q	*Some p don't see some q*

The interpretation is the natural one, using the subject wide scope readings in the ambiguous cases.

This is $\mathcal{R}$.
(The first system of its kind was Nishihara, Morita, Iwata 1990.)

The language $\mathcal{R}^{\dagger}$ has complemented atoms $\overline{p}$ on top of $\mathcal{R}$.

All p are q	$\forall(p, q)$
Some p are q	$\exists(p, q)$
All p r all q	$\forall(p, \forall(q, r))$
All p r some q	$\forall(p, \exists(q, r))$
Some p r all q	$\exists(p, \forall(q, r))$
Some p r some q	$\exists(p, \exists(q, r))$
No p are q	$\forall(p, \overline{q})$
Some p aren't q	$\exists(p, \overline{q})$
All p don't r all q $\equiv$	
No p r any q	$\forall(p, \forall(q, \overline{r}))$
All p don't r some q $\equiv$	
No p r all q	$\forall(p, \exists(q, \overline{r}))$
Some p don't r any q	$\exists(p, \forall(q, \overline{r}))$
Some p don't r some q	$\exists(p, \exists(q, \overline{r}))$

All p are q	$\forall(p, q)$
Some p are q	$\exists(p, q)$
All p r all q	$\forall(p, \forall(q, r))$
All p r some q	$\forall(p, \exists(q, r))$
Some p r all q	$\exists(p, \forall(q, r))$
Some p r some q	$\exists(p, \exists(q, r))$
No p are q	$\forall(p, \overline{q})$
Some p aren't q	$\exists(p, \overline{q})$
No p r any q	$\forall(p, \forall(q, \overline{r}))$
No p r all q	$\forall(p, \exists(q, \overline{r}))$
Some p don't r any q	$\exists(p, \forall(q, \overline{r}))$
Some p don't r some q	$\exists(p, \exists(q, \overline{r}))$

set terms c

positive	p	$\forall(p, r)$	$\exists(p, r)$
negative	$\overline{p}$	$\exists(p, \overline{r})$	$\forall(p, \overline{r})$

$\forall(p, r)$	those who r all p
$\exists(p, r)$	those who r some p
$\forall(p, \overline{r})$	those who fail-to-r all $p \approx$ those who r no p
$\exists(p, \overline{r})$	those who fail-to-r some $p \approx$ those who don't r some p

$$\left.\begin{array}{ll}
\textit{All p are q} & \forall(p,q) \\
\textit{Some p are q} & \exists(p,q) \\
\textit{All p r all q} & \forall(p,\forall(q,r)) \\
\textit{All p r some q} & \forall(p,\exists(q,r)) \\
\textit{Some p r all q} & \exists(p,\forall(q,r)) \\
\textit{Some p r some q} & \exists(p,\exists(q,r)) \\
\textit{No p are q} & \forall(p,\overline{q}) \\
\textit{Some p aren't q} & \exists(p,\overline{q}) \\
\textit{No p sees any q} & \forall(p,\forall(q,\overline{r})) \\
\textit{No p sees all q} & \forall(p,\exists(q,\overline{r})) \\
\textit{Some p don't r any q} & \exists(p,\forall(q,\overline{r})) \\
\textit{Some p don't r some q} & \exists(p,\exists(q,\overline{r}))
\end{array}\right\} \begin{array}{l}\text{simplifies to} \\ \\ \forall(p,c) \quad \exists(p,c)\end{array}$$

set terms c

positive	p	$\forall(p,r)$	$\exists(p,r)$
negative	$\overline{p}$	$\exists(p,\overline{r})$	$\forall(p,\overline{r})$

We start with one collection of unary atoms (for nouns) and another of binary atoms (for transitive verbs).

expression	variables	syntax
unary atom	$p,\ q$	
binary atom	r	
positive set term	c^+	$p \mid \exists(p,r) \mid \forall(p,r)$
set term	c, d	$p \mid \exists(p,r) \mid \forall(p,r) \mid$
		$\overline{p} \mid \exists(p,\overline{r}) \mid \forall(p,\overline{r})$
$\mathcal{R}$ sentence	φ	$\forall(p,c) \mid \exists(p,c)$
$\mathcal{R}^\dagger$ sentence	φ	$\forall(p,c) \mid \exists(p,c) \mid \forall(\overline{p},c) \mid \exists(\overline{p},c)$

We need one last concept, syntactic negation:

expression	syntax	negation
positive set term c	p	$\overline{p}$
	$\overline{p}$	p
	$\exists(p, r)$	$\forall(p, \overline{r})$
	$\forall(p, r)$	$\exists(p, \overline{r})$
	$\exists(p, \overline{r})$	$\forall(p, r)$
	$\forall(p, \overline{r})$	$\exists(p, r)$
$\mathcal{R}$ sentence φ	$\forall(p, c)$	$\exists(p, \overline{c})$
	$\exists(p, c)$	$\forall(p, \overline{c})$

Note that $\overline{\overline{p}} = p$, $\overline{\overline{c}} = c$ and $\overline{\overline{\varphi}} = \varphi$.

THEOREM

There are no finite syllogistic logical systems which are sound and complete for $\mathcal{R}$.

However, there is a logical system (presented below) which uses reductio ad absurdum

$$\dfrac{\begin{array}{c}[\varphi] \\ \vdots \\ \exists(p, \overline{p})\end{array}}{\overline{\varphi}}\ \mathsf{RAA}$$

and which is complete.

Theorem

There are no finite syllogistic logical systems which are sound and complete for $\mathcal{R}$.

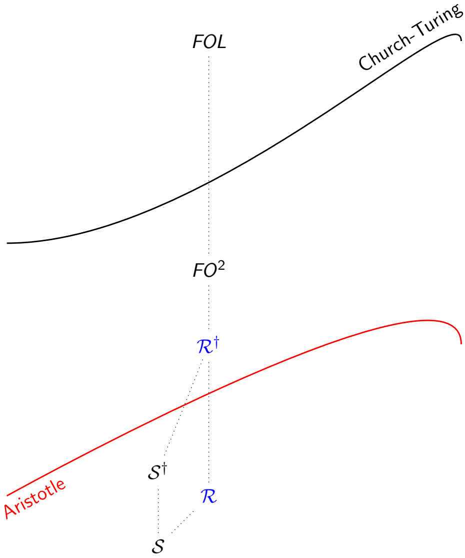
However, there is a logical system (presented below) which uses reductio ad absurdum

$$\frac{\begin{array}{c}[\varphi] \\ \vdots \\ \exists(p, \overline{p})\end{array}}{\overline{\varphi}}\ \mathsf{RAA}$$

and which is complete.

Theorem

There are no finite, sound and complete syllogistic logical systems for $\mathcal{R}^\dagger$, even ones which allow *RAA*.



† adds full N-negation

relational syllogistic

There are no finite, sound, and complete purely syllogistic logics $\vdash_{\mathcal{X}}$ for $\mathcal{R}$

Suppose towards a contradiction that $\mathcal{X}$ did it.
We allow rules with arbitrarily many premises.
Fix $n \in \mathsf{N}$ greater than the number of premises in any rule in $\mathcal{X}$.
Let $Y_1, \ldots, Y_n$ be distinct variables .
Let Γ be the following set of $\mathcal{R}$-formulas:

All Y_i *see some* Y_{i+1}	$(1 \leq i < n)$
All Y_1 *see all* Y_n	
All Y_i *are* Y_i	$(1 \leq i < n)$
All Y_i *aren't* Y_j	$(1 \leq i \neq j \leq n)$

There are no finite, sound, and complete purely syllogistic logics $\vdash_{\mathcal{X}}$ for $\mathcal{R}$

Suppose towards a contradiction that $\mathcal{X}$ did it.
We allow rules with arbitrarily many premises.
Fix $n \in \mathsf{N}$ greater than the number of premises in any rule in $\mathcal{X}$.
Let $Y_1, \ldots, Y_n$ be distinct variables .
Let Γ be the following set of $\mathcal{R}$-formulas:

All Y_i *see some* Y_{i+1}	$(1 \leq i < n)$
All Y_1 *see all* Y_n	
All Y_i *are* Y_i	$(1 \leq i < n)$
All Y_i *aren't* Y_j	$(1 \leq i \neq j \leq n)$

Observe that $\Gamma \models$ *All* Y_1 *see some* Y_n,
but this sentence is not in Γ.

$$\Gamma \quad = \quad \begin{cases} \textit{All } Y_i \textit{ see some } Y_{i+1} & (1 \leq i < n) \\ \textit{All } Y_1 \textit{ see all } Y_n & \\ \textit{All } Y_i \textit{ are } Y_i & (1 \leq i < n) \\ \textit{All } Y_i \textit{ aren't } Y_j & (1 \leq i \neq j \leq n) \end{cases}$$

For $1 \leq i < n$, let $\Delta_i = \Gamma \setminus \{\textit{All } Y_i \textit{ see some } Y_{i+1}\}$.

Claim If $\varphi \in \mathcal{R}$ and $\Delta_i \models \varphi$, then $\varphi \in \Gamma$.

$$\Gamma \quad = \quad \begin{cases} \textit{All } Y_i \textit{ see some } Y_{i+1} & (1 \leq i < n) \\ \textit{All } Y_1 \textit{ see all } Y_n & \\ \textit{All } Y_i \textit{ are } Y_i & (1 \leq i < n) \\ \textit{All } Y_i \textit{ aren't } Y_j & (1 \leq i \neq j \leq n) \end{cases}$$

For $1 \leq i < n$, let $\Delta_i = \Gamma \setminus \{\textit{All } Y_i \textit{ see some } Y_{i+1}\}$.

Claim If $\varphi \in \mathcal{R}$ and $\Delta_i \models \varphi$, then $\varphi \in \Gamma$.
It follows this claim that $\Gamma \not\vdash_{\mathcal{X}} \gamma$.

Claim If $\varphi \in \mathcal{R}$ and $\Delta_i \models \varphi$, then $\varphi \in \Gamma$.

We show by induction on proof trees using $\mathcal{X}$ that all deductions from Γ must have an element of Γ on the root.

No rule of $\mathcal{X}$ has more than $n - 1$ premises.
By induction hypothesis, the sentences just above the root are contained in Γ.
So by the claim the root is in Γ.

Therefore the logic is not complete.

$$\begin{array}{ll} \textit{All } Y_i \textit{ see some } Y_{i+1} & (1 \leq i < n) \\ \textit{All } Y_1 \textit{ see all } Y_n & \\ \textit{All } Y_i \textit{ are } Y_i & (1 \leq i < n) \\ \textit{All } Y_i \textit{ aren't } Y_j & (1 \leq i \neq j \leq n) \end{array}$$

For $1 \leq i < n$, let $\Delta_i = \Gamma \setminus \{\textit{All } Y_i \textit{ see some } Y_{i+1}\}$.
Proof sketch We consider every sentence in the languge $\mathcal{R}$.
We check is either in Γ or is falsified in some model of Δ_i.

All Y_i *are* Y_j	*All* Y_i *aren't* Y_j $\equiv$ *No* X *are* Y_j
Some Y_i *are* Y_j	*Some* Y_i *aren't* Y_j
All Y_i *see all* Y_j	*All* Y_i *don't see all* Y_j $\equiv$ *No* X *sees any* Y_j
All Y_i *see some* Y_j	*All* Y_i *don't see some* Y_j $\equiv$ *No* X *sees all* Y_j
Some Y_i *see all* Y_j	*Some* Y_i *don't see any* Y_j
Some Y_i *see some* Y_j	*Some* Y_i *don't see some* Y_j

The proof goes on

$$y_1 \to y_2 \to \cdots y_j \to y_{j+1} \to y_{j+2} \to \cdots \qquad \to y_i \qquad y_{i+1} \to \cdots \qquad y_n$$

This structure satisfies Γ_i and makes a few false:

All Y_i *are* Y_j $\surd$	*No* Y_j *are* Y_k $\surd$ for $j = k$
Some Y_i *are* Y_j $\surd$	*Some* Y_i *aren't* Y_j
All Y_i *r all* Y_j	*All* Y_i *don't r all* $Y_j \equiv$ *No* Y_i *sees any* Y_j
All Y_i *r some* Y_j	*All* Y_i *don't r some* $Y_j \equiv$ *No* Y_i *sees all* Y_j
Some Y_i *r all* Y_j	*Some* Y_i *don't r any* Y_j
Some Y_i *r some* Y_j	*Some* Y_i *don't r some* Y_j

The empty structure satisfies Γ_i and makes a few more false:

All Y_j *are* Y_k $\surd$

Some Y_j *are* Y_k $\surd$

All Y_j *r all* Y_k

All Y_j *r some* Y_k

Some Y_j *r all* Y_k $\surd$

Some Y_j *r some* Y_k $\surd$

No Y_j *are* Y_k $\surd$ for $j = k$

Some Y_j *aren't* Y_k

All Y_j *don't r all* $Y_k \equiv$ *No* Y_i *sees any* Y_j

All Y_j *don't r some* $Y_k \equiv$ *No* Y_i *sees all* Y_j

Some Y_j *don't r any* Y_k $\surd$

Some Y_j *don't r some* Y_k $\surd$

p and q range over unary atoms,
c over set terms, and t over binary atoms or their negations.

$$\frac{\exists(p,q) \qquad \forall(q,c)}{\exists(p,c)} \qquad \frac{\forall(p,q) \qquad \forall(q,c)}{\forall(p,c)}$$

$$\frac{\forall(p,q) \qquad \exists(p,c)}{\exists(q,c)} \qquad \frac{}{\forall(p,p)} \qquad \frac{\exists(p,c)}{\exists(p,p)}$$

$$\frac{\forall(q,\bar{c}) \qquad \exists(p,c)}{\exists(p,\bar{q})} \qquad \frac{\forall(p,\bar{p})}{\forall(p,c)} \qquad \frac{\exists(p,\exists(q,t))}{\exists(q,q)}$$

$$\frac{\forall(p,\forall(n,t)) \qquad \exists(q,n)}{\forall(p,\exists(q,t))} \qquad \frac{\exists(p,\exists(q,t)) \qquad \forall(q,n)}{\exists(p,\exists(n,t))}$$

$$\frac{\forall(p,\exists(q,t)) \qquad \forall(q,n)}{\forall(p,\exists(n,t))} \qquad \begin{array}{c}[\varphi] \\ \vdots \\ \dfrac{\exists(p,\overline{p})}{\overline{\varphi}}\ \mathrm{RAA}\end{array}$$

Most are monotonicty principles

$$\begin{array}{ll} \exists(p^{\uparrow}, q^{\uparrow}) & \forall(p^{\downarrow}, q^{\uparrow}) \\ \exists(p^{\uparrow}, \forall(q^{\downarrow}, t)) & \exists(p^{\uparrow}, \exists(q^{\uparrow}, t)) \\ \forall(p^{\downarrow}, \forall(q^{\downarrow}, t)) & \forall(p^{\downarrow}, \exists(q^{\uparrow}, t)) \end{array}$$

Plus also

$$\frac{}{\forall(p, p)} \qquad \frac{\exists(p, c)}{\exists(p, p)} \qquad \frac{\forall(p, \bar{p})}{\forall(p, c)} \qquad \frac{\exists(p, \exists(q, t))}{\exists(q, q)}$$

$$\frac{\forall(q, \bar{c}) \qquad \exists(p, c)}{\exists(p, \bar{q})}\ (\star) \qquad \frac{\forall(p, \forall(n, t)) \qquad \exists(q, n)}{\forall(p, \exists(q, t))}$$

Of these, $(\star)$ is the most interesting.

Most are monotonicty principles

$$\begin{array}{ll} \exists(p^{\uparrow}, q^{\uparrow}) & \forall(p^{\downarrow}, q^{\uparrow}) \\ \exists(p^{\uparrow}, \forall(q^{\downarrow}, t)) & \exists(p^{\uparrow}, \exists(q^{\uparrow}, t)) \\ \forall(p^{\downarrow}, \forall(q^{\downarrow}, t)) & \forall(p^{\downarrow}, \exists(q^{\uparrow}, t)) \end{array}$$

Tomorrow I'll talk about monotonicity
and its relation to categorial grammar,
generalized quantifiers, and other areas.

Relevant papers:

van Benthem (2007) and earlier
Sanchez Valencia (1991)
van Eijck (2007)
Zamansky, Francez, and Winter (2006)

Example of a proof in the system for $\mathcal{R}^\dagger$

What do you think? Sound or unsound?

$$\begin{array}{ll} & \textit{All } X \textit{ see all } Y, \textit{All } X \textit{ see some } Z, \textit{All } Z \textit{ see some } Y \\ \models & \textit{All } X \textit{ see some } Y \end{array}$$

Example of a proof in the system for $\mathcal{R}^\dagger$

What do you think? Sound or unsound?

$$\begin{array}{ll} & \textit{All } X \textit{ see all } Y, \textit{All } X \textit{ see some } Z, \textit{All } Z \textit{ see some } Y \\ \models & \textit{All } X \textit{ see some } Y \end{array}$$

The conclusion does indeed follow:
take cases as to whether or not there are Z.

We should have a formal proof.

$$\textit{All } X \textit{ see all } Y, \textit{All } X \textit{ see some } Z, \textit{All } Z \textit{ see some } Y$$
$$\models \quad \textit{All } X \textit{ see some } Y$$

$$\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\color{red}\textit{Some } X \textit{ see no } Y}{\textit{Some } X \textit{ are } X} \quad \textit{All } X \textit{ see some } Z}{\textit{Some } X \textit{ see some } Z}}{\textit{Some } Z \textit{ are } Z} \quad \textit{All } Z \textit{ see some } Y}{\textit{Some } Z \textit{ see some } Y}}{\textit{Some } Y \textit{ are } Y} \quad \textit{All } X \textit{ see all } Y}{\textit{All } X \textit{ see some } Y} \quad {\color{red}\textit{Some } X \textit{ see no } Y}}{\textit{Some } X \textit{ aren't } X}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\textit{Some } X \textit{ see no } Y]}{\textit{Some } X \textit{ are } X} \quad \textit{All } X \textit{ see some } Z}{\textit{Some } X \textit{ see some } Z}}{\textit{Some } Z \textit{ are } Z} \quad \textit{All } Z \textit{ see some } Y}{\textit{Some } Z \textit{ see some } Y}}{\textit{Some } Y \textit{ are } Y} \quad \textit{All } X \textit{ see all } Y}{\textit{All } X \textit{ see some } Y} \quad [\textit{Some } X \textit{ see no } Y]}{\textit{Some } X \textit{ aren't } X}}{\textit{All } X \textit{ see some } Y} \; RAA$$

This shows that

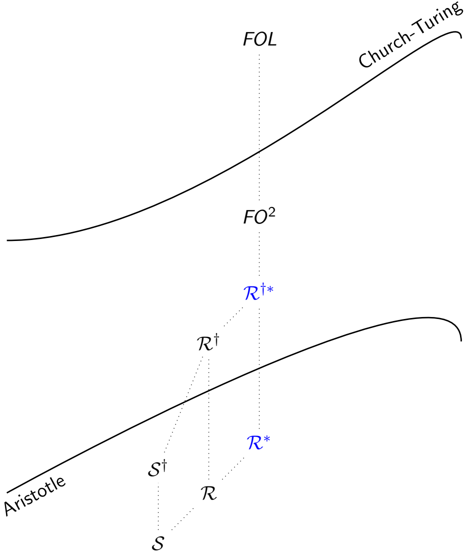
$$\textit{All } X \textit{ see all } Y, \textit{All } X \textit{ see some } Z, \textit{All } Z \textit{ see some } Y \vdash \textit{All } X \textit{ see some } Y$$

Again, $\mathcal{R}$ has no pure syllogistic proof system.
But it has an indirect system (one using RAA).

With a lot more work, one can show that $\mathcal{R}^\dagger$
doesn't even have an indirect system!

The arguments are reminiscent of arguments in finite model theory,
but without the boolean connectives there are many differences.

Next: relative clauses



† adds full N-negation

add relative clauses
= relativized quantifiers

What do you think about this one?

$$\frac{\textit{All skunks are mammals}}{\textit{All who fear all who respect all skunks fear all who respect all mammals}}$$

It follows, using an interesting antitonicity principle:

$$\frac{\textit{All skunks are mammals}}{\textit{All who respect all mammals respect all skunks}}$$

Inference with relative clauses

It follows, using an interesting antitonicity principle:

$$\dfrac{\dfrac{\textit{All skunks are mammals}}{\textit{All who respect all mammals respect all skunks}}}{\textit{All who fear all who respect all skunks fear all who respect all mammals}}$$

$\mathcal{R}^*$ allows sentential subjects to be noun phrases containing subject relative clauses.

who r all p	*who r some p*
who don't r all p	*who don't r any p*

---

expression	syntax	
$\mathcal{R}^*$ sentence	$\forall(d^+, c)$	$\exists(d^+, c)$
$\mathcal{R}^{\dagger *}$ sentence	$\forall(d, c)$	$\exists(d, c)$

d^+ is a positive set term, and c is an arbitrary set term.

$$\frac{\forall(p,q)}{\forall(\forall(q,r),\forall(p,r))} \quad \frac{\forall(p,q)}{\forall(\exists(p,r),\exists(q,r))} \quad \frac{\exists(p,q)}{\forall(\forall(p,r),\exists(q,r))}$$

These rules are based on McAllester and Givan (1992).

The remaining rules for $\mathcal{R}^*$ are generalizations of the $\mathcal{R}$ rules to the bigger syntax.

In a variant of this language which
admits iterated relative clauses, we would just have

$$\forall(s, m) \vdash \forall(\forall(\forall(s, r), f), \forall(\forall(m, r), f),$$

$$\cfrac{\cfrac{\forall(s, m)}{\forall(\forall(m, r), \forall(s, r))}}{\forall(\forall(\forall(s, r), f), \forall(\forall(m, r), f))}$$

$\mathcal{R}^{\dagger}$ and $\mathcal{R}^{\dagger *}$ lie beyond the Aristotle boundary,
due to full negation on nouns.

It is possible to formulate a logical system with
a restricted notion of variables,
prove completeness,
and yet stay inside the Turing boundary.

It's a fairly involved definition, so we might skip the details
and instead look at examples.

DETAILS ON THE PROOF SYSTEM FOR $\mathcal{R}^{\dagger *}$

Expression	Variables	Syntax
unary atom	p, q	
binary atom	s	
constant	j, k	
unary literal	l	$p \mid \bar{p}$
binary literal	r	$s \mid \bar{s}$
set term	b, c, d	$l \mid \exists(c, r) \mid \forall(c, r)$
sentence	φ, ψ	$\forall(c, d) \mid \exists(c, d) \mid c(j) \mid r(j, k)$

Think of the constants as proper names: John, Mary, etc.
the unary atoms as predicates like boys or girls,
the binary atoms by transitive verbs such as likes and sees.

Recursion allows us to embed set terms, and so we have set terms like

$$\exists(\forall(\forall(b, \bar{s}), h), a)$$

which may be taken to symbolize
a verb phrase such as
admires someone who hates everyone who does not see any boy.

We should note that the relative clauses which can be obtained in this way are all "subject relatives", never "object relatives".

The language is too poor to express predicates like
$\lambda x.$all boys see x.

Proof system: general sentences

General sentences in this fragment are what usually are called formulas.
We prefer to change the standard terminology to make the point that here, sentences are not built from formulas by quantification.
Sentences in our sense do not have variable occurrences.
But general sentences do allow variables.

Expression	Variables	Syntax
individual variable	x, y	
individual term	t, u	$x \mid j$
general sentence	α	$\varphi \mid c(t) \mid r(t, u) \mid \perp$

It will turn out that for this fragment, only two variables are needed.
We don't need general sentences of the form $r(j, x)$ or $r(x, j)$.

Proof system: half of the rules

$$\frac{c(t) \quad \forall(c,d)}{d(t)}\ \forall E \qquad \frac{c(u) \quad \forall(c,r)(t)}{r(t,u)}\ \forall E$$

$$\frac{c(t) \quad d(t)}{\exists(c,d)}\ \exists I \qquad \frac{r(t,u) \quad c(u)}{\exists(c,r)(t)}\ \exists I$$

Proof system: the second half of the rules

$$\dfrac{\begin{array}{c}[c(x)]\\ \vdots\\ d(x)\end{array}}{\forall(c,d)}\ \forall I \qquad\qquad \dfrac{\begin{array}{c}[c(x)]\\ \vdots\\ r(t,x)\end{array}}{\forall(c,r)(t)}\ \forall I$$

$$\dfrac{\exists(c,d) \qquad \begin{array}{c}[c(x)]\quad [d(x)]\\ \vdots\\ \alpha\end{array}}{\alpha}\ \exists E \qquad\qquad \dfrac{\exists(c,r)(t) \qquad \begin{array}{c}[c(x)]\quad [r(t,x)]\\ \vdots\\ \alpha\end{array}}{\alpha}\ \exists E$$

$$\dfrac{\alpha \quad \overline{\alpha}}{\bot}\ \bot I \qquad\qquad \dfrac{\begin{array}{c}[\overline{\varphi}]\\ \vdots\\ \bot\end{array}}{\varphi}\ RAA$$

Proof system: side conditions

$$\dfrac{\begin{array}{c}[c(x)]\\ \vdots \\ d(x)\end{array}}{\forall(c,d)}\ \forall I \qquad\qquad \dfrac{\begin{array}{c}[c(x)]\\ \vdots \\ r(t,x)\end{array}}{\forall(c,r)(t)}\ \forall I$$

$$\dfrac{\exists(c,d) \qquad \begin{array}{c}[c(x)]\quad [d(x)]\\ \vdots \\ \alpha\end{array}}{\alpha}\ \exists E \qquad\qquad \dfrac{\exists(c,r)(t) \qquad \begin{array}{c}[c(x)]\quad [r(t,x)]\\ \vdots \\ \alpha\end{array}}{\alpha}\ \exists E$$

In $(\forall I)$, x must not occur free in any uncanceled hypothesis.

In $(\exists E)$, the variable x must not occur free in the conclusion α or in any uncanceled hypothesis in the subderivation of α.

In contrast to usual first-order natural deduction systems, there are no side conditions on the rules $(\forall E)$ and $(\exists I)$.

Example of a proof in the system

From all keys are old items, infer everyone who owns a key owns an old item

$$\dfrac{\dfrac{[\exists(key, own)(x)]^2 \qquad \dfrac{[own(x,y)]^1 \qquad \dfrac{[key(y)]^1 \quad \forall(key, old\text{–}item)}{old\text{–}item(y)}\ \forall E}{\exists(old\text{–}item, own)(x)}\ \exists I}{\exists(old\text{–}item, own)(x)}\ \exists E^1}{\forall(\exists(key, own), \exists(old\text{–}item, own))}\ \forall I^2$$

Example of a proof in the system

From all keys are old items, infer everyone who owns a key owns an old item

1	$\forall(\mathit{key}, \mathit{old\text{–}item})$	hyp
2	$\quad \exists(\mathit{key}, \mathit{own})(x)$	hyp
3	$\quad \mathit{key}(y)$	$\exists E$, 2
4	$\quad \mathit{own}(x, y)$	$\exists E$, 2
5	$\quad \mathit{old\text{–}item}(y)$	$\forall E$, 1, 3
6	$\quad \exists(\mathit{old\text{–}item}, \mathit{own})(x)$	$\exists I$, 4, 5
7	$\forall(\exists(\mathit{key}, \mathit{own}), \exists(\mathit{old\text{–}item}, \mathit{own}))$	$\forall I$, 1–6

Frederic Fitch, 1973

Natural deduction rules for English, *Phil. Studies*, 24:2, 89–104.

1		John is a man	Hyp
2		Any woman is a mystery to any man	Hyp
3	Jane	Jane is a woman	Hyp
4		Any woman is a mystery to any man	R, 2
5		Jane is a mystery to any man	Any Elim, 4
6		John is a man	R, 1
7		Jane is a mystery to John	Any Elim, 6
8		Any woman is a mystery to John	Any intro, 3, 7

A word on completeness/decidability/complexity of the logics for $\mathcal{R}^{\dagger}$ and $\mathcal{R}^{\dagger *}$

For these logics, one can prove completeness by a Henkin-style argument.

The easiest way to prove decidability would be via the

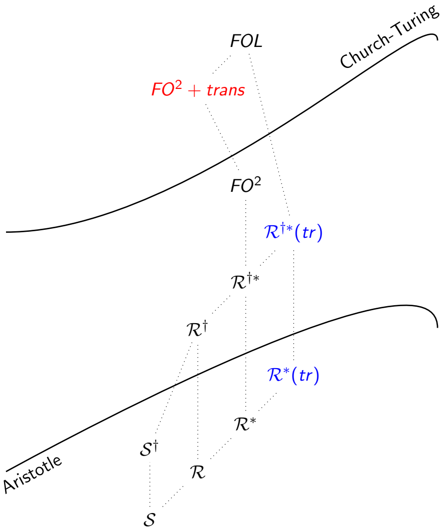
- finite model property: use filtration from modal logic
- embedding into FO^2
- embedding into boolean modal logic (better complexity)
- results on resolution in Pratt-Hartmann 2004 (better complexity)

Also, there is a lower bound using K + universal modality.

The upshot: the validity problem is complete for exponential time.

Next: comparative adjectives

used for inferences involving phrases like bigger than some kitten



Grädel, Otto, Rosen 1999

!!

† adds full N-negation

* adds relative clauses

tr adds comparatives, requiring transitivity

Every giraffe is taller than every gnu
Some gnu is taller than every lion
Some lion is taller than some zebra
Every giraffe is taller than some zebra

We extend $\mathcal{R}^*$ to a language $\mathcal{R}^*(tr)$ by taking a set **A** of comparative adjective phrases in the base.

In the semantics, we would require of a model that for $a \in \mathbf{A}$, $[\![a]\!]$ must be a transitive relation. (At the end of the talk we'll see irreflexivity.)

Every giraffe is taller than every gnu
Some gnu is taller than every lion
Some lion is taller than some zebra

Every giraffe is taller than some zebra

$$\frac{\forall(p, \exists(q, r))}{\forall(\exists(p, r), \exists(q, r))} \qquad \frac{\forall(p, \forall(q, r))}{\forall(\exists(p, r), \forall(q, r))}$$

$$\frac{\exists(p, \forall(q, r))}{\forall(\forall(p, r), \forall(q, r))} \qquad \frac{\exists(p, \exists(q, r))}{\forall(\forall(p, r), \exists(q, r))}$$

$$\begin{array}{l}\text{Every giraffe is taller than every gnu}\\ \text{Some gnu is taller than every lion}\\ \underline{\text{Some lion is taller than some zebra}}\\ \text{Every giraffe is taller than some zebra}\end{array}$$

$$\dfrac{\dfrac{\forall(\mathsf{gir}, \forall(\mathsf{gnu}, \mathsf{taller})) \quad \exists(\mathsf{gnu}, \forall(\mathsf{lion}, \mathsf{taller}))}{\forall(\mathsf{gir}, \forall(\mathsf{lion}, \mathsf{taller}))} \qquad \exists(\mathsf{lion}, \exists(\mathsf{zebra}, \mathsf{taller}))}{\forall(\mathsf{giraffe}, \exists(\mathsf{zebra}, \mathsf{taller}))}$$

We begin with the logical system for $\mathcal{R}^{\dagger *}$,
and then we add a rule:

$$\frac{a(x,y) \quad a(y,z)}{a(x,z)}\ \textsf{trans}$$

This rule is added for all $a \in \mathbf{A}$, and all x, y, z.

This gives a language $\mathcal{R}^{\dagger *}(tr)$.

Example of the transitivity rule

$$\frac{\text{Every sweet fruit is bigger than every kumquat}}{\text{Every fruit bigger than some sweet fruit is bigger than every kumquat}}$$

$$\dfrac{\dfrac{[\exists(\mathsf{sw}, \mathsf{bigger})(x)]^3 \qquad \dfrac{\dfrac{[\mathsf{bigger}(x,y)]^2 \qquad \dfrac{[\mathsf{kq}(z)]^1 \qquad \dfrac{[\mathsf{sw}(y)]^2 \quad \forall(\mathsf{sw}, \forall(\mathsf{kq}, \mathsf{bigger}))}{\forall(\mathsf{kq}, \mathsf{bigger})(y)}\ \forall E}{\mathsf{bigger}(y,z)}\ \forall E}{\mathsf{bigger}(x,z)}\ {\color{red}\mathsf{trans}}}{\forall(\mathsf{kq}, \mathsf{bigger})(x)}\ \forall I^1}{\forall(\mathsf{kq}, \mathsf{bigger})(x)}\ \exists E^2}{\forall(\exists(\mathsf{sw}, \mathsf{bigger}), \forall(\mathsf{kq}, \mathsf{bigger}))}\ \forall I^3$$

We want to account for inferences such as

$$\frac{\text{Frege's favorite food was sushi}}{\text{Frege ate sushi at least once}}$$

We want to account for inferences such as

$$\frac{\text{Frege's favorite food was sushi}}{\text{Frege ate sushi at least once}}$$

The hypothesis and conclusion would be
rendered in some logical system or other.
There would be a background theory ($\approx$ common sense),
and then the inference would be modeled either as a semantic fact:

$$\text{Common sense} + \text{Frege's favorite food was sushi} \models \text{Frege ate sushi at least once}$$

or a via a formal deduction:

$$\text{Common sense} + \text{Frege's favorite food was sushi} \vdash \text{Frege ate sushi at least once}$$

Either way, it's all in one and the same language.

Transitivity should not be treated as a meaning postulate,
since even stating it would seem to render the logic undecidable.

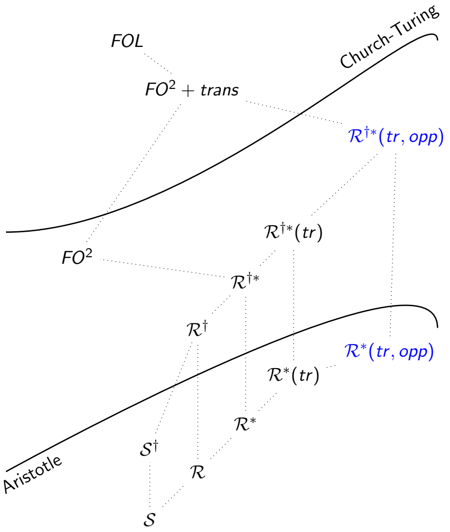
Instead, it is a proof rule:

$$\frac{a(x,y) \quad a(y,z)}{a(x,z)}\ \text{trans}$$

(I have not proved that one can't formulate a decidable logic which can directly express transitivity using variables and also cover the sentences we've seen.
But there are results that suggest it.)

Next: relational converses

used for inferences relating bigger and smaller



† adds full N-negation
* adds relative clauses

opp adds opposites
of comparative adjectives

Converses of transitive relations

On top of all the other syllogistic systems we have seen

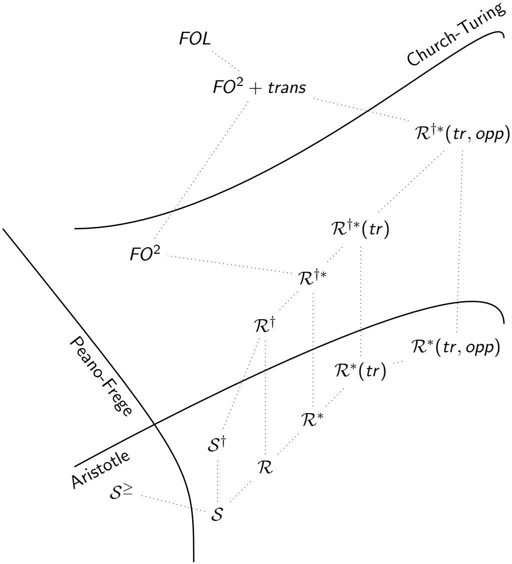
$$\frac{\forall(p,\forall(q,t))}{\forall(q,\forall(p,t^{-1}))} \qquad \frac{\exists(p,\forall(q,t))}{\forall(q,\exists(p,t^{-1}))}\ (\text{scope}) \qquad \frac{\forall(p,\exists(q,r^{-1}))}{\forall(\forall(q,r),\forall(p,r))}$$

$$\frac{\exists(\exists(p,r^{-1}),\exists(q,r))}{\exists(p,\exists(q,r))} \qquad \frac{\exists(\forall(p,r),\forall(q,r^{-1}))}{\forall(p,\forall(q,r^{-1}))} \qquad \frac{\exists(\forall(p,r),\exists(q,r^{-1}))}{\exists(q,\forall(p,r^{-1}))}$$

$$\frac{\forall(p,\exists(q,r)) \quad \forall(\exists(p,r^{-1}),\exists(n,r))}{\forall(p,\exists(n,r))}\ (\star) \qquad \frac{\forall(p,\exists(q,r)) \quad \forall(\exists(p,r^{-1}),\forall(n,r))}{\forall(p,\forall(n,r))}$$

(scope): if some p is bigger than all q,
then all q are smaller than some p or other.

($\star$): if every dog is bigger than some hedgehog,
and everything smaller than some dog is bigger than some cat,
then every dog is bigger than some cat.



first-order logic

FO^2 + "R is trans"

2 variable FO logic

† adds full N-negation

$\mathcal{R}^*(tr)$ + opposites

$\mathcal{R}^*$ + (transitive) comparative adjs

$\mathcal{R}$ + relative clauses

$\mathcal{S}$ + full N-negation

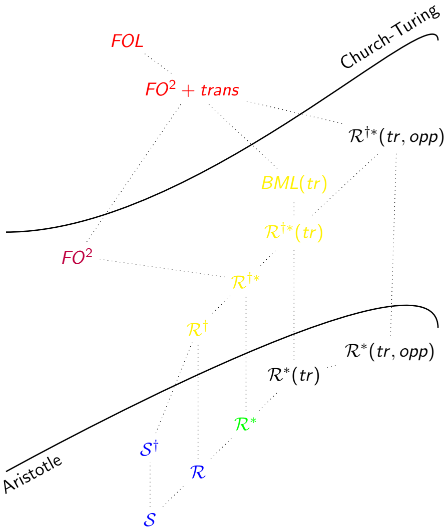
$\mathcal{R}$ = relational syllogistic

$\mathcal{S}^{\geq}$ adds $|p| \geq |q|$

$\mathcal{S}$: all/some/no p are q

Complexity

(mostly) best possible results on the validity problem



undecidable
Church 1936
Grädel, Otto, Rosen 1999

in co-NEXPTIME
EXPTIME
Lutz & Sattler 2001

Co-NEXPTIME
Grädel, Kolaitis, Vardi '97
EXPTIME
Pratt-Hartmann 2004

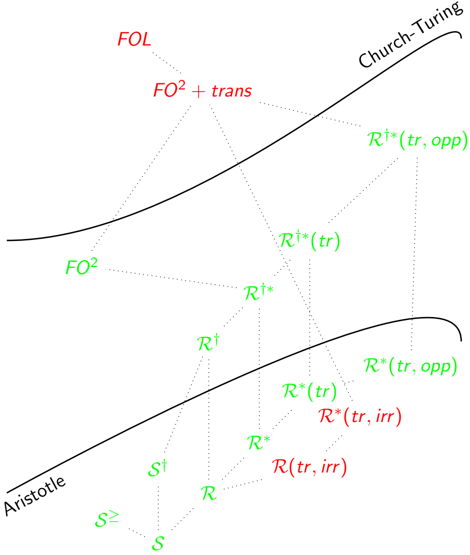
lower bounds also open

Co-NP
McAllester & Givan 1992

NLOGSPACE

$\mathcal{S}$	NLOGSPACE	lower bound via reachability problem for directed graphs
$\mathcal{S}^{\dagger}$	NLOGSPACE	upper bound via $2\mathrm{SAT}$
$\mathcal{R}$	NLOGSPACE	upper bound takes special work based on the proof system
$\mathcal{R}^{\dagger}$	EXPTIME	lower bound via K^U, Hemaspaandra 1996
$\mathcal{R}^{*\dagger}$	EXPTIME	upper bound by Pratt-Hartmann 2004
$BML(tr)$	EXPTIME	Boolean modal logic on transitive models Lutz and Sattler 2001
$\mathcal{R}^{*}$	Co-NPTIME	essentially in McAllester and Givan 1992
FO^2	NEXPTIME	Grädel, Kolaitis, and Vardi 1997

The finite model property: Yes$^{\downarrow}$ and No$^{\uparrow}$



filtration of a
Henkin model

Mortimer 1975

irr means that
comparative adjectives
must have irreflexive
interpretations.
$\forall(p, \exists(p, r)) + \exists p$

Some p are q
Some p are not q
Some q are not p
Every p is smaller than some q
Every q is smaller than some p

$$\begin{tikzcd}
 & p_0 \arrow[r] \arrow[ddr] & p_1 \arrow[r] \arrow[ddr] & \cdots & p_n \arrow[r] \arrow[ddr] & p_{n+1} & \cdots \\
\langle p, q \rangle \arrow[ur] \arrow[dr] & & & & & & \\
 & q_0 \arrow[r] \arrow[uur] & q_1 \arrow[r] \arrow[uur] & \cdots & q_n \arrow[r] \arrow[uur] & q_{n+1} & \cdots
\end{tikzcd}$$

The relation in the model is the transitive closure of the arrows.

Natural logic: what I hope to have gotten across

Program

Show that significant parts of natural language inference can be carried out in decidable logical systems.

Whenever possible, to obtain complete axiomatizations, because the resulting logical systems are likely to be interesting.

To be completely mathematical and hence to work using all tools and to make connections to fields like complexity theory, (finite) model theory, decidable fragments of first-order logic, and algebraic logic.

- We must ask whether a complete proof system is a semantics.
- We should not be afraid of doing logic beyond logic.
- Joining the perspectives of semantics, complexity theory, proof theory, cognitive science, and computational linguistics should allow us to ask interesting questions and answer them.