Natural Logic: Visions, Results, Plans

Larry Moss

A presentation to Thomas Icard's NASSLLI Course

June 20, 2012

Doing logic in 2012 means living in two worlds

My talk today will be an exploration of this tension.

At times I will be unashamedly anachronistic,
letting the voices of the past ricochet off the future.

Overall questions

- What is the current relation of logic and language?
- What could/should it be?
- What does logic for NL look like when it is done with a minimum of translation?
- Can we re-work semantics in the light of computational linguistics?
- What does any of this have to do with other courses at this NASSLLI?

We want to account for natural language inferences such as

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

We want to account for natural language inferences such as

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

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:

$$\begin{array}{ll} & \text{Common sense} + \text{Frege's favorite food was chimichangas} \\ \models & \text{Frege ate chimichangas at least once} \end{array}$$

or a via a formal deduction:

$$\begin{array}{ll} & \text{Common sense} + \text{Frege's favorite food was chimichangas} \\ \vdash & \text{Frege ate chimichangas at least once} \end{array}$$

- To carry our this program, it would be advisable to take as expressive a logical system as possible.
- First-order logic (FOL) is a good starting point, but for many phenomena we'll need to go further.
- Being more expressive, FOL is vastly superior to traditional (term) logic.
- Various properties of FOL are interest in this discussion, but only secondarily so.

And anyways, what choice to we really have?

One can easily object to the whole enterprise of using FOL in connection with NL inference, on the grounds that FOL cannot handle

- vague words
- intentions of speakers
- missing words and phrases
- poetic language

$\vdots$

In other words, FOL is too small for the job.

FOL is also too big!

The point is that for "everyday inference", a small fragment of FOL should be sufficient.

Also, there is a long tradition in linguistics of dissatisfaction with models which are "Turing complete "
and in favor of ones with much less expressive power.

This actually was decisive in syntax: the Peters-Ritchie Theorem.

FOL is also too big!

The point is that for "everyday inference", a small fragment of FOL should be sufficient.

Also, there is a long tradition in linguistics of dissatisfaction with models which are "Turing complete "
and in favor of ones with much less expressive power.

This actually was decisive in syntax: the Peters-Ritchie Theorem.

You decide

Consider three activites:

- A mathematics: prove the Pythagorean Theorem $a^2 + b^2 = c^2$.
- B syntax: parse John feared his mother saw him at her house.
- C semantics: tell whether the text of The Yellow Rose of Texas entails that

  Some African-American man once missed a (specific) girl.

Where would you put semantics?

A. mathematics B. syntax

Theres a yellow girl in Texas
That I'm going down to see;
No other darkies know her,
No darkey, only me;
She cried so when I left her
That it like to broke my heart,
And if I only find her,
We never more will part.

What does undecidability have to do with it?

Theorem (Church 1936)

There is no algorithm, which given
a finite set Γ of sentences in FOL and another sentence φ,
decides whether or not $\Gamma \models \varphi$.

The same goes for the proof-theoretic notion $\Gamma \vdash \varphi$,
since this comes to the same thing,
by the Completeness Theorem of FOL.

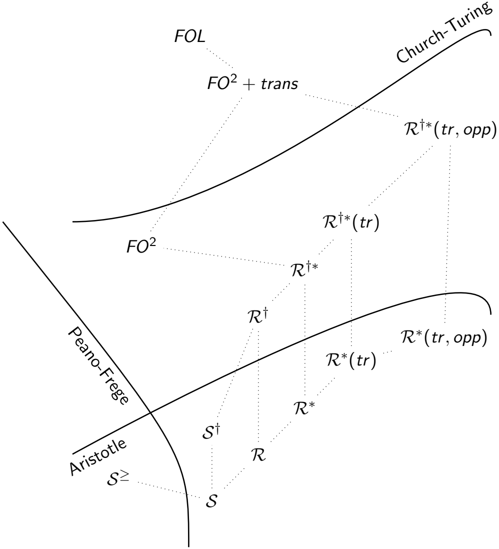
PROGRAM

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

Raise the question of how much semantics can be done in decidable fragments.

To axiomatize as much as possible,
because the resulting logical systems are likely to be interesting.

To ask how much of language could have been done if the traditional logicians
had the mathematical tools to go further than they were able to.



first-order logic

FO^2 + "R is trans"

2 variable FO logic

$\dagger$ 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

Syntax: Start with a collection of unary atoms (for nouns).
Then the sentences are the expressions

$$\textit{All } p \textit{ are } q$$

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

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

Proof system is based on the following rules:

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

If Γ is a set of sentences, we write $\mathcal{M} \models \Gamma$ if for all $\varphi \in \Gamma$, $\mathcal{M} \models \varphi$.

$\Gamma \models \varphi$ means that every $\mathcal{M} \models \Gamma$ also has $\mathcal{M} \models \varphi$.

A proof tree over Γ is a finite tree $\mathcal{T}$
whose nodes are labeled with sentences,
and each node is either an element of Γ,
or comes from its parent(s) by an application of one of the rules.

$\Gamma \vdash \varphi$ means that there is a proof tree $\mathcal{T}$ for over Γ
whose root is labeled φ.

The simplest completeness theorem in logic

If $\Gamma \models$ *All p are q*, then $\Gamma \vdash$ *All p are q*

Suppose that $\Gamma \models$ *All p are q*.

Build a model $\mathcal{M}$, taking M to be the set of variables.

Define $u \leq v$ to mean that $\Gamma \vdash$ *All u are v*.
The semantics is $[\![u]\!] = \downarrow u$.
Then $\mathcal{M} \models \Gamma$.
Hence for the p and q in our statement, $[\![p]\!] \subseteq [\![q]\!]$.

But by reflexivity, $p \in [\![p]\!]$.
And so $p \in [\![q]\!]$; this means that $p \leq q$.

But this is exactly what we want:
$\Gamma \vdash$ *All p are q*.

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

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

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

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

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

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

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

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

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.

A rude interruption

Robert van Rooij: from an email message of July, 2009
quoted with permission

I also like the idea (as a semanticist) of having a variable free semantics,
and a natural logic, and this seems to be what
the traditional logicians were (very slowly)
developing before they were
so rudely interrupted by Frege, Peano, Russell and others.
...
i agree that proofs, and computability, should play a bigger part in semantics (theories of meaning).

Actually I am also interested in semantics/pragmatics where bounded rationality plays an important part.
This is the move many economists are now taking in game theory.

I hope, one day, to connect both of these research trends
(bounded rationality in game theory, and thus pragmatics),
and natural logic, with emphasis on monotonicity and so on.

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

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

- The logical systems that one would get from looking at inference involving surface sentences would contain many copies of similar-looking rules. Presenting things in this way would miss a lot of generalizations.

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

- The systems would contain 'rules' that are more like complex deduction patterns that need to be framed as rules only because one lacks the machinery to break them down into more manageable sub-deductions.

  Moreover, those complex rules would be unilluminating.

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

- The systems would lack variables,
  and thus they would be tedious and inelegant.

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

- Turning to the standard topic of quantifier-scope ambiguities, it would be impossible to handle inferences among sentences exhibiting this phenomenon in an elegant way.

The next language uses "see" as a variable 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 language $\mathcal{R}^\dagger$ has complemented variables $\overline{p}$ on top of $\mathcal{R}$.

Theorem

There are no purely syllogistic logical systems complete for $\mathcal{R}$. However, there is a logical system **R** which uses reductio ad absurdum

$$\dfrac{\cdots[\bar{\varphi}] \quad \cdots[\varphi]\cdots \atop \vdots \atop \bot}{\psi}\ \mathsf{RAA}$$

and which is complete.

Theorem

There are no purely syllogistic logical systems complete for $\mathcal{R}$. However, there is a logical system **R** which uses reductio ad absurdum

$$\dfrac{\begin{array}{c}\cdots[\bar{\varphi}]\ \ \cdots[\varphi]\cdots \\ \vdots \\ \bot\end{array}}{\psi}\ \mathsf{RAA}$$

and which is complete.

Theorem

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

A complete system **R** for $\mathcal{R}$

On top of the system $\mathcal{S}$, one rule is missing, and so is *RAA*

$$\textit{All } X^{\color{red}\downarrow} \textit{ (don't) see all } Y^{\color{red}\downarrow}$$
$$\textit{Some } X^{\color{red}\uparrow} \textit{ (don't) see all } Y^{\color{red}\downarrow}$$
$$\textit{All } X^{\color{red}\downarrow} \textit{ (don't) see some } Y^{\color{red}\uparrow}$$
$$\textit{Some } X^{\color{red}\uparrow} \textit{ (don't) see some } Y^{\color{red}\uparrow}$$

$$\frac{\textit{All X aren't X}}{\textit{All X see all Y}}$$

$$\frac{\textit{All X (don't) see all Z} \quad \textit{Some Y are Z}}{\textit{All X (don't) see some Y}}$$

$$\frac{\textit{All Z (don't) see all Y} \quad \textit{Some X are Z}}{\textit{Some X (don't) see all Y}}$$

$$\frac{\textit{Some X don't see some Y} \quad \textit{All X see all Y}}{\textit{No X are X}}$$

$$\frac{\textit{Some X (don't) see some Y}}{\textit{Some Y is a Y}}$$

Example of a proof in this system

What do you think?

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

Example of a proof in this system

What do you think?

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

The conclusion does indeed follow.
We should have a formal proof.

Example of a proof in this system

What do you think?

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

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\color{red}\textit{Some } X \textit{ see no } Y}{\exists X} \qquad \textit{All } X \textit{ see some } Z}{\textit{Some } X \textit{ see some } Z}}{\exists Z} \qquad \textit{All } Z \textit{ see some } Y}{\textit{Some } Z \textit{ see some } Y}}{\exists Y} \qquad \textit{All } X \textit{ see all } Y}{\textit{All } X \textit{ see some } Y} \qquad {\color{red}\textit{Some } X \textit{ see no } Y}}{\textit{Some } X \textit{ aren't } X}$$

$\exists X$ abbreviates *Some* X *are* X

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

This shows that

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

I'm not at all committed to the particular semantics here, and in fact there are good reasons not to like it.

One can change it: see Pawel Garbacz, "A System of Syllogistic for Cooperative Conversation", 2012.

One can also tune the systems to experimental results in cognitive science, as you saw on Monday.

What do you think about this one?

$$\frac{\textit{All armadillos are mammals}}{\textit{All who fear all who respect all arm's fear all who respect all mammals}}$$

It follows, using an interesting antitonicity principle:

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

It follows, using an interesting antitonicity principle

$$\cfrac{\cfrac{\textit{All armadillos are mammals}}{\textit{All who respect all mammals respect all armadillos}}}{\textit{All who fear all who respect all arm's fear all who respect all mammals}}$$

$\mathcal{R}^*$ allows subject noun phrases to contain relative clauses of the form

who see all p	*who see some p*
who don't see all p	*who don't see some p*

$\mathcal{R}^{*\dagger}$ has full negation on nouns.

$$\frac{\textit{All p are q}}{\textit{All (see all q) (see all p)}} \qquad \frac{\textit{All p are q}}{\textit{All (see some p) (see some q)}}$$

$$\frac{\textit{Some p are q}}{\textit{All (see all p) (see some q)}} \qquad \frac{\textit{Some p see some q}}{\textit{Some q are q}}$$

$$\frac{\textit{All p aren't p}}{\textit{All q see all p}} \qquad \frac{\textit{All p aren't p}}{\textit{All (see all q) see all p}}$$

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

Every giraffe is taller than every gnu
Some gnu is taller than every lion
<u>Some lion is taller than some zebra</u>
Every giraffe is taller than some zebra

$$\frac{\forall(p, \exists(q, r))}{\forall(\exists(p, r), \exists(q, r))} \text{ (tr1)} \qquad \frac{\forall(p, \forall(q, r))}{\forall(\exists(p, r), \forall(q, r))} \text{ (tr2)}$$

$$\frac{\exists(p, \forall(q, r))}{\forall(\forall(p, r), \forall(q, r))} \text{ (tr3)} \qquad \frac{\exists(p, \exists(q, r))}{\forall(\forall(p, r), \exists(q, r))} \text{ (tr4)}$$

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

$$\dfrac{\dfrac{\forall(\mathsf{giraffe}, \forall(\mathsf{gnu}, \mathsf{taller})) \quad \exists(\mathsf{gnu}, \forall(\mathsf{lion}, \mathsf{taller}))}{\forall(\mathsf{giraffe}, \forall(\mathsf{lion}, \mathsf{taller}))}\ (\rho 1) \quad \exists(\mathsf{lion}, \exists(\mathsf{zebra}), \mathsf{taller})}{\forall(\mathsf{giraffe}, \exists(\mathsf{zebra}, \mathsf{taller}))}\ (\rho 2)$$

Decidable logic beyond the Aristotle boundary

Having relative clauses + negation on nouns
leads to systems beyond the Aristotle boundary.

It is possible to formulate logical systems with
restricted notions of variables
and yet stay inside the Turing boundary.

Example: $\forall(c, d) \vdash \forall(\exists(c, r), \exists(d, r))$

if all watches are expensive items, then everyone who owns a watch owns an expensive item

$$\dfrac{\dfrac{[\exists(c, r)(x)]^2 \qquad \dfrac{[r(x, y)]^1 \qquad \dfrac{[c(y)]^1 \quad \forall(c, d)}{d(y)}\ \forall E}{\exists(d, r)(x)}\ \exists I}{\exists(d, r)(x)}\ \exists E^1}{\forall(\exists(c, r), \exists(d, r))}\ \forall I^2$$

Example: $\forall(c, d) \vdash \forall(\exists(c, r), \exists(d, r))$

IF ALL WATCHES ARE EXPENSIVE ITEMS, THEN EVERYONE WHO OWNS A WATCH OWNS AN EXPENSIVE ITEM

1		$\forall(c, d)$	hyp
2	x	$\exists(c, r)(x)$	hyp
3		$c(y)$	$\exists E$, 2
4		$r(x, y)$	$\exists E$, 2
5		$d(y)$	$\forall E$, 1, 3
6		$\exists(d, r)(x)$	$\exists I$, 4, 5
7		$\forall(\exists(c, r), \exists(d, r))$	$\forall I$, 1–6

Theorem

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

If Γ is consistent, then Γ has a model of size at most 2^{2n}, where n is the number of set terms in Γ.

WHY WAS FITCH'S 1973 PAPER FORGOTTEN?

NATURAL DEDUCTION RULES FOR ENGLISH, *Phil. Studies*, 24:2 (1973), 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

Montague's "English as a Formal Language" in 1970,
"The Proper Treatment of Quantification in Ordinary English" in 1973.

We extend our language $\mathcal{L}$ to a language $\mathcal{L}(adj)$ by taking a basic set $\mathbf{A}$ of comparative adjective phrases in the base.

In the semantics, we would require
that for an adjective $a \in \mathbf{A}$, $[\![a]\!]$ must be a transitive relation
(in every model $\mathcal{M}$)

We add a rule:

$$\frac{a(t_1, t_2) \quad a(t_2, t_3)}{a(t_1, t_3)}\ \text{trans}$$

This rule is added for all $a \in \mathbf{A}$.

Example of the transitivity rule

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

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

Transitivity should not be treated as a meaning postulate, since this renders the logic undecidable.

Transitivity should not be treated as a meaning postulate, since this renders the logic undecidable.

Instead, it is a proof rule.

(I have not conclusively that this cannot be so, but there are results that strongly suggest it.)

This is an important result: it shows that the decidability requirement for natural logics has a bite.

It suggests that we'll have to re-think the semantic enterprise in interesting ways.

My statements of the objections, again

- The logical systems that one would get
  from looking at inference involving surface sentences
  would contain many copies of similar-looking rules.
  Presenting things in this way would miss a lot of
  generalizations.

- The systems would contain 'rules' that
  are more like complex deduction patterns
  that need to be framed as rules
  only because one lacks the machinery to break them down
  into more manageable sub-deductions.
  Moreover, those complex rules would be unilluminating.

- The systems would lack variables,
  and thus they would be tedious and inelegant.

- Turning to the standard topic of quantifier-scope ambiguities,
  it would be impossible to handle inferences among sentences
  exhibiting this phenomenon in an elegant way.

- The many similar rules can be succinctly grouped into meta-rules.
  It has already been suggested that this is a good thing to do for monotonicity.

- It's true that sometimes the systems have some complex "rules".
  Perhaps this could be turned into an advantage,
  but aiming for a theory of "shallow inference" in language.

- The systems can have variables in some form.

- Quantifier-scope ambiguities can be handled.

More to do

- Implement the existing systems
- Combine with treatments of conversation, vagueness, anaphora, conditionals, abduction, ...
- Logic beyond grammar
- Raise the question of a proof theory/syntax interface
- Ask the question of whether a (complete or incomplete) logical system is a semantics
- Further develop logical systems for use with RTE

Joining the perspectives of semantics, complexity theory, proof theory, and computational linguistics will allow us to ask and answer interesting questions.