Reasoning About the Sizes of Sets

Larry Moss
Indiana University

North American Summer School on
Logic, Language, and Information
June 23-27, 2025

Program

To put inference at the center of semantics.

Show the aspects of natural language inference
that can be modeled at all
can be modeled using decidable logical systems.

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

To make connections to proof-theoretic semantics,
and to psycholinguistic studies about human reasoning.

What are the simplest kinds of quantity reasoning?

Our candidate would be combinations of

All x are y
Some x are y
No x are y

There are at least as many x as y
There are more x than y
There are at most as many x as y
There are fewer x than y
There are as many x as y
Most x are y

What are the simplest kinds of quantity reasoning?

Our candidate would be combinations of

All x are y
Some x are y
No x are y

There are at least as many x as y
There are more x than y
There are at most as many x as y
There are fewer x than y
There are as many x as y
Most x are y

To add to the complexity, we could add the ability to use non-x or even to take unions and intersections.

How can we do logic with these?

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

$\mathcal{M} \models$ *All p are q*	iff	$[\![p]\!] \subseteq [\![q]\!]$
$\mathcal{M} \models$ *Some p are q*	iff	$[\![p]\!] \cap [\![q]\!] \neq \emptyset$
$\mathcal{M} \models$ *No p are q*	iff	$[\![p]\!] \cap [\![q]\!] = \emptyset$
$\mathcal{M} \models$ *There are at least as many p as q*	iff	$\lvert[\![p]\!]\rvert \geq \lvert[\![q]\!]\rvert$
$\mathcal{M} \models$ *There are more p than q*	iff	$\lvert[\![p]\!]\rvert > \lvert[\![q]\!]\rvert$
$\mathcal{M} \models$ *There are at most as many p as q*	iff	$\lvert[\![p]\!]\rvert \leq \lvert[\![q]\!]\rvert$
$\mathcal{M} \models$ *There are fewer p than q*	iff	$\lvert[\![p]\!]\rvert < \lvert[\![q]\!]\rvert$
$\mathcal{M} \models$ *There are as many p as q*	iff	$\lvert[\![p]\!]\rvert = \lvert[\![q]\!]\rvert$
$\mathcal{M} \models$ *Most p are q*	iff	$\lvert[\![p]\!] \cap [\![q]\!]\rvert > \frac{1}{2}\lvert[\![p]\!]\rvert$

I know that “most” people in this room will object to the last definition.

All + Some + "there are at least as many" + "there are more than"

There are at least as many x as y is written $\exists^{\geq}(x, y)$
There are more x than y is written $\mathit{More}(x, y)$

$$\frac{}{\forall(p,p)}\ (\textsc{axiom}) \qquad \frac{\forall(n,p) \quad \forall(p,q)}{\forall(n,q)}\ (\textsc{barbara})$$

$$\frac{\exists(p,q)}{\exists(p,p)}\ (\textsc{some}) \qquad \frac{\exists(q,p)}{\exists(p,q)}\ (\textsc{conversion})$$

$$\frac{\exists(p,n) \quad \forall(n,q)}{\exists(p,q)}\ (\textsc{darii}) \qquad \frac{\forall(p,q) \quad \exists^{\geq}(p,q)}{\forall(q,p)}\ (\textsc{card-mix})$$

$$\frac{\forall(p,q)}{\exists^{\geq}(q,p)}\ (\textsc{subset-size}) \qquad \frac{\exists^{\geq}(n,p) \quad \exists^{\geq}(p,q)}{\exists^{\geq}(n,q)}\ (\textsc{card-trans})$$

$$\frac{\exists(p,p) \quad \exists^{\geq}(q,p)}{\exists(q,q)}\ (\textsc{card-}\exists) \qquad \frac{\mathit{More}(p,q)}{\exists^{\geq}(p,q)}\ (\textsc{more-at least})$$

$$\frac{\mathit{More}(n,p) \quad \exists^{\geq}(p,q)}{\mathit{More}(n,q)}\ (\textsc{more-left}) \qquad \frac{\exists^{\geq}(n,p) \quad \mathit{More}(p,q)}{\mathit{More}(n,q)}\ (\textsc{more-right})$$

$$\frac{\exists^{\geq}(p,q) \quad \mathit{More}(q,p)}{\varphi}\ (\mathrm{X})$$

All + Some + "there are at least as many" + "there are more than"

There are at least as many x as y is written $\exists^{\geq}(x, y)$
There are more x than y is written $More(x, y)$

$$\frac{}{\forall(p,p)}\ (\text{AXIOM}) \qquad \frac{\forall(n,p) \quad \forall(p,q)}{\forall(n,q)}\ (\text{BARBARA})$$

$$\frac{\exists(p,q)}{\exists(p,p)}\ (\text{SOME}) \qquad \frac{\exists(q,p)}{\exists(p,q)}\ (\text{CONVERSION})$$

$$\frac{\exists(p,n) \quad \forall(n,q)}{\exists(p,q)}\ (\text{DARII}) \qquad \frac{\forall(p,q) \quad \exists^{\geq}(p,q)}{\forall(q,p)}\ (\text{CARD-MIX})$$

$$\frac{\forall(p,q)}{\exists^{\geq}(q,p)}\ (\text{SUBSET-SIZE}) \qquad \frac{\exists^{\geq}(n,p) \quad \exists^{\geq}(p,q)}{\exists^{\geq}(n,q)}\ (\text{CARD-TRANS})$$

$$\frac{\exists(p,p) \quad \exists^{\geq}(q,p)}{\exists(q,q)}\ (\text{CARD-}\exists) \qquad \frac{More(p,q)}{\exists^{\geq}(p,q)}\ (\text{MORE-AT LEAST})$$

$$\frac{More(n,p) \quad \exists^{\geq}(p,q)}{More(n,q)}\ (\text{MORE-LEFT}) \qquad \frac{\exists^{\geq}(n,p) \quad More(p,q)}{More(n,q)}\ (\text{MORE-RIGHT})$$

$$\frac{\exists^{\geq}(p,q) \quad More(q,p)}{\varphi}\ (\text{X})$$

Soundness/Completeness Theorem

$\Gamma \models \varphi$ iff $\Gamma \vdash \varphi$.
Moreover, there's an easy algorithm to tell whether or not $\Gamma \vdash \varphi$

Rules of inference using complemented variables $\overline{p}$

We use the classical semantics $[\![\overline{p}]\!] = M \setminus [\![p]\!]$.

$$\frac{\forall(p, \overline{p})}{\forall(p, q)}\ (\textsc{zero}) \qquad \frac{\forall(\overline{p}, p)}{\forall(q, p)}\ (\textsc{one})$$

$$\frac{\forall(q, p) \quad \exists(p, \overline{q})}{More(p, q)}\ (\textsc{more}) \qquad \frac{More(p, q)}{\exists(p, \overline{q})}\ (\textsc{more-some})$$

$$\frac{More(q, p)}{More(\overline{p}, \overline{q})}\ (\textsc{more-anti}) \qquad \frac{\forall(p, q)}{\forall(\overline{q}, \overline{p})}\ (\textsc{anti})$$

$$\frac{\exists^{\geq}(p, q)}{\exists^{\geq}(\overline{q}, \overline{p})}\ (\textsc{card-anti}) \qquad \frac{\exists(p, p) \quad \exists^{\geq}(q, \overline{q})}{\exists(q, q)}\ (\textsc{int})$$

$$\frac{\exists^{\geq}(p, \overline{p}) \quad \exists^{\geq}(\overline{q}, q)}{\exists^{\geq}(p, q)}\ (\textsc{half}) \qquad \frac{More(p, \overline{p}) \quad \exists^{\geq}(\overline{q}, q)}{More(p, q)}\ (\textsc{strict half})$$

$$\frac{\exists^{\geq}(p, \overline{p}) \quad \exists^{\geq}(q, \overline{q}) \quad \exists(\overline{p}, \overline{q})}{\exists(p, q)}\ (\textsc{maj}) \qquad \frac{\exists(p, q) \quad \forall(q, \overline{q})}{\varphi}\ (\textsc{X})$$

$$\frac{\exists^{\geq}(p,\overline{p}) \quad \exists^{\geq}(q,\overline{q}) \quad \exists(\overline{p},\overline{q})}{\exists(p,q)}\ (\textsc{maj})$$

It is typical for these logical systems that we get strange rules!

$$\frac{\exists^{\geq}(p,\overline{p}) \quad \exists^{\geq}(q,\overline{q}) \quad \exists(\overline{p},\overline{q})}{\exists(p,q)}\ (\textsc{maj})$$

It is typical for these logical systems that we get strange rules!

Here is a related principle:

$$\frac{\exists^{\geq}(p,\overline{p}) \quad \exists^{\geq}(q,\overline{q}) \quad \forall(p,\overline{q})}{\forall(\overline{p},q)}\ (\textsc{maj-contrapositive})$$

This one actually can be proved:

$$\cfrac{\cfrac{\cfrac{\exists^{\geq}(p,\overline{p}) \quad \exists^{\geq}(q,\overline{q})}{\exists^{\geq}(p,\overline{q})}\ (\textsc{half}) \qquad \forall(p,\overline{q})}{\forall(\overline{q},p)}\ (\textsc{card-mix})}{\forall(p,\overline{q})}\ (\textsc{anti})$$

However, our logic doesn't have reductio ad absurdum, and so the derivation above doesn't show that (MAJ) can be proved from the rest of the rules.

(In fact, an exhaustive search shows that (MAJ) cannot be proved from the rest of the system.)

The full logic: adding reductio ad absurdum simplifies the system but it complicates the proof search

$$\frac{}{\forall(p,p)}\ \text{(axiom)} \qquad \frac{\forall(n,p)\quad \forall(p,q)}{\forall(n,q)}\ \text{(Barbara)} \qquad \frac{\exists(p,q)}{\exists(p,p)}\ \text{(some)}$$

$$\frac{\exists(q,p)}{\exists(p,q)}\ \text{(conversion)} \qquad \frac{\forall(p,q)}{\forall(\overline{q},\overline{p})}\ \text{(anti)} \qquad \frac{\forall(p,\overline{p})}{\forall(p,q)}\ \text{(zero)}$$

$$\frac{\exists(p,n)\quad \forall(n,q)}{\exists(p,q)}\ \text{(Darii)} \qquad \frac{\forall(\overline{p},p)}{\forall(q,p)}\ \text{(one)} \qquad \frac{\forall(p,q)}{\exists^{\geq}(q,p)}\ \text{(subset-size)}$$

$$\frac{\exists^{\geq}(p,q)}{\exists^{\geq}(\overline{q},\overline{p})}\ \text{(card-mon)} \qquad \frac{\exists^{\geq}(p,q)}{\exists^{\geq}(\overline{q},\overline{p})}\ \text{(card-anti)} \qquad \frac{\forall(p,q)\quad \exists^{\geq}(p,q)}{\forall(q,p)}\ \text{(card-mix)}$$

$$\frac{\exists(p,p)\quad \exists^{\geq}(p,q)}{\exists(q,q)}\ \text{(card-}\exists\text{)} \qquad \frac{\forall(q,p)\quad \exists(p,\overline{q})}{\mathit{More}(p,q)}\ \text{(more)} \qquad \frac{\mathit{More}(p,q)}{\exists(p,\overline{q})}\ \text{(more-some)}$$

$$\frac{\mathit{More}(p,q)}{\exists^{\geq}(p,q)}\ \text{(more-at least)} \qquad \frac{\mathit{More}(n,p)\quad \exists^{\geq}(p,q)}{\mathit{More}(n,q)}\ \text{(more-left)} \qquad \frac{\mathit{More}(q,p)}{\mathit{More}(\overline{p},\overline{q})}\ \text{(more-anti)}$$

$$\frac{\exists(p,p)\quad \exists^{\geq}(q,\overline{q})}{\exists(q,q)}\ \text{(int)} \qquad \frac{\exists^{\geq}(p,\overline{p})\quad \exists^{\geq}(\overline{q},q)}{\exists^{\geq}(p,q)}\ \text{(half)} \qquad \frac{\mathit{More}(p,\overline{p})\quad \exists^{\geq}(\overline{q},q)}{\mathit{More}(p,q)}\ \text{(strict half)}$$

$$\frac{\exists^{\geq}(p,\overline{p})\quad \exists^{\geq}(q,\overline{q})\quad \exists(\overline{p},\overline{q})}{\exists(p,q)}\ \text{(maj)}$$

$$\frac{\exists(p,q)\quad \forall(p,\overline{q})}{\varphi}\ \text{(X)} \qquad \frac{\mathit{More}(p,q)\quad \exists^{\geq}(q,p)}{\varphi}\ \text{(X)}$$

The logic has been implemented in Sage.
Similar ones have been implemented in Haskell and we're putting them on the web these days.

For example, one may enter

```
assumptions= ['All non-a are b',
'There are more c than non-b',
'There are more non-c than non-b',
'There are at least as many non-d as d',
'There are at least as many c as non-c',
'There are at least as many non-d as non-a']
conclusion = 'All a are non-c'
follows(assumptions,conclusion)
```

We get

```
The conclusion does not follow
Here is a counter-model.
We take the universe of the model to be
{0, 1, 2, 3, 4, 5}
noun semantics
a      {2, 3}
b      {0, 1, 4, 5}
c      {0, 2, 3}
d      {}
```

So it gives the semantics of `a`, `b`, `c`, and `d` as subsets of $\{0, \ldots, 5\}$. Notice that the assumptions are true in the model, but the conclusion is false.

Here is an example of a derivation found by our implementation. We ask whether the putative conclusion below really follows:

All non-x are x
Some non-y are z
____
There are more x than y

```
Here is a formal proof in our system:
1 All non-x are x     Assumption
2 All y are x     One 1
3 All non-x are x     Assumption
4 All non-y are x     One 3
5 Some non-y are z     Assumption
6 Some non-y are non-y     Some 5
7 Some non-y are x     Darii 4 6
8 Some x are non-y     Conversion 7
9 There are more x than y     More 2 8
```

$$\frac{}{\forall(p,p)}\ \text{(axiom)} \qquad \frac{\forall(n,p) \quad \forall(p,q)}{\forall(n,q)}\ \text{(Barbara)} \qquad \frac{\exists(p,q)}{\exists(p,p)}\ \text{(some)}$$

$$\frac{\exists(q,p)}{\exists(p,q)}\ \text{(conversion)} \qquad \frac{\forall(p,q)}{\forall(\overline{q},\overline{p})}\ \text{(anti)} \qquad \frac{\forall(p,\overline{p})}{\forall(p,q)}\ \text{(zero)}$$

$$\frac{\exists(p,n) \quad \forall(n,q)}{\exists(p,q)}\ \text{(Darii)} \qquad \frac{\forall(\overline{p},p)}{\forall(q,p)}\ \text{(one)} \qquad \frac{\forall(p,q)}{\exists^{\geq}(q,p)}\ \text{(subset-size)}$$

$$\frac{\exists^{\geq}(p,q)}{\exists^{\geq}(\overline{q},\overline{p})}\ \text{(card-mon)} \qquad \frac{\exists(p,p) \quad \exists^{\geq}(p,q)}{\exists(q,q)}\ \text{(card-}\exists\text{)} \qquad \frac{\mathit{More}(p,q)}{\exists(p,\overline{q})}\ \text{(more-some)}$$

$$\frac{\mathit{More}(p,q)}{\exists^{\geq}(p,q)}\ \text{(more-at least)} \qquad \frac{\mathit{More}(n,p) \quad \exists^{\geq}(p,q)}{\mathit{More}(n,q)}\ \text{(more-left)}$$

$$\frac{\exists^{\geq}(p,\overline{p})}{\exists(p,p)}\ \text{(NON-EMPTY)} \qquad \frac{\forall(\overline{p},p)}{\mathit{More}(p,\overline{p})}\ \text{(NON-EMPTY-MORE)}$$

$$\frac{\mathit{More}(q,p)}{\exists^{\geq}(\overline{p},x)}\ \text{(WEAK-MORE-ANTI)} \qquad \frac{\exists^{\geq}(x,p) \quad \exists^{\geq}(x,\overline{p})}{\exists^{\geq}(x,q)}\ \text{(UP)}$$

$$\frac{\exists(p,q) \quad \forall(p,\overline{q})}{\varphi}\ \text{(X)} \qquad \frac{\mathit{More}(p,q) \quad \exists^{\geq}(q,p)}{\varphi}\ \text{(X)}$$

$$\frac{}{\text{All } X \text{ are } X} \qquad \frac{\text{All } X \text{ are } Y \quad \text{All } Y \text{ are } Z}{\text{All } X \text{ are } Z}$$

$$\frac{\text{Some } X \text{ are } Y}{\text{Some } Y \text{ are } X} \qquad \frac{\text{Some } X \text{ are } Y}{\text{Some } X \text{ are } X} \qquad \frac{\text{Some } X \text{ are } Y \quad \text{All } Y \text{ are } Z}{\text{Some } X \text{ are } Z}$$

Can you think of any valid laws that add Most X are Y on top of All X are Y and Some X are Y?

$$\frac{}{\text{All } X \text{ are } X} \qquad \frac{\text{All } X \text{ are } Y \quad \text{All } Y \text{ are } Z}{\text{All } X \text{ are } Z}$$

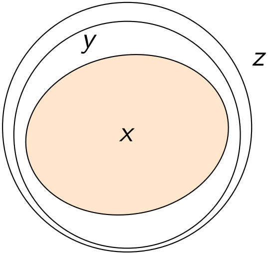
$$\frac{\text{Some } X \text{ are } Y}{\text{Some } Y \text{ are } X} \qquad \frac{\text{Some } X \text{ are } Y}{\text{Some } X \text{ are } X} \qquad \frac{\text{Some } X \text{ are } Y \quad \text{All } Y \text{ are } Z}{\text{Some } X \text{ are } Z}$$

$$\frac{\text{Most } X \text{ are } Y}{\text{Some } X \text{ are } Y}\, m_1 \qquad \frac{\text{Some } X \text{ are } X}{\text{Most } X \text{ are } X}\, m_2 \qquad \frac{\text{Most } X \text{ are } Y \quad \text{All } Y \text{ are } Z}{\text{Most } X \text{ are } Z}\, m_3$$

$$\frac{\text{Most } X \text{ are } Z \quad \text{All } X \text{ are } Y \quad \text{All } Y \text{ are } X}{\text{Most } Y \text{ are } Z}\, m_4$$

$$\frac{\text{All } Y \text{ are } X \quad \text{All } X \text{ are } Z \quad \text{Most } Z \text{ are } Y}{\text{Most } X \text{ are } Y}\, m_5$$

$$\frac{X_1 \triangleright_{A,B} Y_1 \quad Y_1 \triangleright_{B,A} X_2 \quad \cdots \quad X_n \triangleright_{A,B} Y_n \quad Y_n \triangleright_{B,A} X_1}{\text{Some } A \text{ are } B}\, \triangleright$$



If most z are x,
then a fortiori, most y are x, too.

The last infinite batch of rules

$$\frac{X_1 \triangleright_{A,B} Y_1 \quad Y_1 \triangleright_{B,A} X_2 \quad \cdots \quad X_n \triangleright_{A,B} Y_n \quad Y_n \triangleright_{B,A} X_1}{\text{Some } A \text{ are } B} \triangleright$$

Examples:

$$\frac{\text{Most } Z \text{ are } X \quad \text{Most } Z \text{ are } Y}{\text{Some } X \text{ are } Y} \triangleright$$

You call this an inference rule?!

From

Most X are B', All A' are A, Most Y are A', All B' are B, All X are Y
Most Y are A'', All A'' are A, Most X are B'', All B'' are B, All A'' are X

infer

Some A are B.

It means either the 5 sentences below

Most X are B',
All A' are A,
Most Y are A',
All B' are B,
All X are Y,

$$\begin{tikzcd}
 & & Y \arrow[dl] & & \\
A & A' \arrow[l, hook'] & & B' \arrow[r, hook] & B \\
 & & X \arrow[uu, hook'] \arrow[ur] & &
\end{tikzcd}$$

or else the 5 sentences below

Most X are B',
All A' are A,
Most Y are A',
All B' are B,
All B' are Y.

$$\begin{tikzcd}
 & & Y \arrow[dl] & & \\
A & A' \arrow[l, hook'] & & B' \arrow[ul, hook'] \arrow[r, hook] & B \\
 & & X \arrow[ur] & &
\end{tikzcd}$$

In these, A' and B' are arbitrary.

Theorem (Jörg Endrullis & LM (2015))

The logical system for this language is complete.

Theorem

Infinitely many axioms are needed in the system.

Theorem

The decision problem for the consequence relation

$$\Gamma \vdash \varphi$$

is in polynomial time.

With Tri Lai (combinatorist, then a graduate student)
we showed that

- Most X are Y
- boolean connectives, especially negation

has a very simple proof system and is algorithmically manageable.

Propositional logic with most X are Y

We start with a collection of unary atoms X, Y, Z, ...
We then form atomic sentences of the form Most X are Y.
We form sentences from atomic sentences using the boolean connectives of propositional logic:

negation ($\neg$)

conjunction ($\wedge$)

disjunction ($\vee$),

implication ($\rightarrow$)

So as just one example of a sentence, we would have

$$(\text{Most } X \text{ are } Y \wedge \neg\text{Most } X \text{ are } Z) \vee \text{Most } Y \text{ are } X \text{ .}$$

We call this logic $\mathcal{L}(most)$.

A model of this fragment is a structure $\mathcal{M} = (M, [\![\ ]\!])$ consisting of a finite set M together with interpretations $[\![X]\!] \subseteq M$ of each one-place relation symbol X.

We then interpret our sentences in a model as follows

$$\mathcal{M} \models \text{Most } X \text{ are } Y \quad \text{iff} \quad |[\![X]\!] \cap [\![Y]\!]| > \tfrac{1}{2}|[\![X]\!]| \,.$$

An example of the kind of question we are interested in

Let

$$\Gamma \quad = \quad \left\{ \begin{array}{rr} \text{Most } X \text{ are } Y & \text{Most } Z \text{ are } Y \\ \text{Most } Y \text{ are } X & \text{Most } Y \text{ are } W \\ \text{Most } X \text{ are } Z & \neg\text{Most } W \text{ are } Y \\ \neg\text{Most } Z \text{ are } X & \text{Most } Z \text{ are } W \\ \text{Most } Y \text{ are } Z & \end{array} \right\}$$

Is it true or not that

$$\Gamma \models \text{Most } W \text{ are } Z \text{ ?}$$

An example of the kind of question we are interested in

Let

$$\Gamma \quad = \quad \left\{ \begin{array}{rr} \text{Most } X \text{ are } Y & \text{Most } Z \text{ are } Y \\ \text{Most } Y \text{ are } X & \text{Most } Y \text{ are } W \\ \text{Most } X \text{ are } Z & \neg\text{Most } W \text{ are } Y \\ \neg\text{Most } Z \text{ are } X & \text{Most } Z \text{ are } W \\ \text{Most } Y \text{ are } Z & \end{array} \right\}$$

Is it true or not that

$$\Gamma \models \text{Most } W \text{ are } Z\ ?$$

Can we take the graph below

$$\begin{tikzcd} X \arrow[r, leftrightarrow] \arrow[d] & Y \arrow[d] \arrow[dl, leftrightarrow] \\ Z \arrow[r] & W \end{tikzcd}$$

and turn the nodes into sets in some appropriate way?

We say that $\mathcal{M} \models \Gamma$ if $\mathcal{M} \models \psi$ for all $\psi \in \Gamma$.
The main semantic definition is:

$$\Gamma \models \varphi \text{ if for all (finite) models } \mathcal{M}, \text{ if } \mathcal{M} \models \Gamma, \text{ then } \mathcal{M} \models \varphi.$$

This relation $\Gamma \vdash \varphi$ between finite sets of sentences and single sentences is called the consequence relation of the logic.

Axioms all propositional tautologies

(Most X are Y) $\rightarrow$ ((Most X are X) $\wedge$ (Most Y are Y))

((Most X_1 are X_2) $\wedge$ (Most X_2 are X_3) $\wedge \cdots \wedge$ (Most X_n are X_1))
$\rightarrow$ ((Most X_2 are X_1) $\vee$ (Most X_3 are X_2) $\vee \cdots \vee$ (Most X_1 are X_n))

Rule from $\varphi \rightarrow \psi$ and φ, infer ψ (Modus Ponens)

The heart of the completeness argument

A $\frac{1}{2}$-representation of a finite simple graph $(G, \to)$ is a family of finite sets A_g for $g \in G$ with the following property:

$g \to h$ if and only if "more than half of the A_g are A_h".

That is,

$$g \to h \quad \text{iff} \quad |A_g \cap A_h| > \frac{1}{2} \cdot |A_g|.$$

THE HEART OF THE COMPLETENESS ARGUMENT

A $\frac{1}{2}$-representation of a finite simple graph $(G, \to)$ is a family of finite sets A_g for $g \in G$ with the following property:

$g \to h$ if and only if "more than half of the A_g are A_h".

That is,

$$g \to h \quad \text{iff} \quad |A_g \cap A_h| > \frac{1}{2} \cdot |A_g|.$$

A two-way edge in a graph is an edge $g \to h$ such that also $h \to g$.
A one-way edge in a graph is an edge $g \to h$ such that $h \not\to g$.

If G has a $\frac{1}{2}$-representation,
and there is a one-way edge from g to h,
then $|A_h| > |A_g|$.

Thus G cannot have one-way cycles: there are no paths

$$g_1 \to g_2 \to \cdots \to g_n = g_1$$

such that $g_{i+1} \not\to g_i$. (There may be cycles with two-way edges.)

The heart of the completeness argument

A $\frac{1}{2}$-representation of a finite simple graph $(G, \to)$ is a family of finite sets A_g for $g \in G$ with the following property:

$g \to h$ if and only if "more than half of the A_g are A_h".

That is,

$$g \to h \quad \text{iff} \quad |A_g \cap A_h| > \frac{1}{2} \cdot |A_g|.$$

So our logic question before may be refomulated:
Does the graph below have a $\frac{1}{2}$-representation?

$$\begin{tikzcd}
X \arrow[r, leftrightarrow] \arrow[d] & Y \arrow[d] \arrow[dl, leftrightarrow] \\
Z \arrow[r] & W
\end{tikzcd}$$

Theorem (Tri Lai, Jörg Endrullis, and LM 2013), conjectured by Chloe Urbanski

Every graph without two-way cycles has a $\frac{1}{2}$-representation.

Illustration

$$\begin{tikzcd}
X \arrow[r, leftrightarrow] \arrow[d] & Y \arrow[d] \arrow[dl, leftrightarrow] \\
Z \arrow[r] & W
\end{tikzcd}$$

Illustration

$$\begin{tikzcd}
X \arrow[r, leftrightarrow] \arrow[d] & Y \arrow[d] \\
Z \arrow[ur, leftrightarrow] \arrow[r] & W
\end{tikzcd}$$

We begin with four subsets of $\{1, \ldots, 16\}$ each of size 8, with the property that distinct sets have intersections of size 4:

$$\begin{array}{rcl}
A_X & = & \{1,2,3,4,5,6,7,8\} \\
A_Y & = & \{1,2,3,4,9,10,11,12\} \\
A_Z & = & \{1,2,5,6,9,10,13,14\} \\
A_W & = & \{1,3,5,7,9,11,13,15\}
\end{array}$$

For $i \neq j$, we write $A_i \sqcap A_j$ for the private intersection:

$$A_i \sqcap A_j \quad = \quad (A_i \cap A_j) \setminus \bigcup_{k \neq i,j} A_k$$

For $i \neq j$, $A_i \sqcap A_j$ has size 1.

$$\begin{tikzcd}
X \arrow[r, leftrightarrow] \arrow[d] & Y \arrow[d] \arrow[dl, leftrightarrow] \\
Z \arrow[r] & W
\end{tikzcd}$$

For $i \neq j$, $A_i \sqcap A_j$ has size 1.
For example, $A_X \sqcap A_Z = \{6\}$.

We replace each point x by three copies of itself, $3x - 2$, $3x - 1$, and $3x$.

$$\begin{aligned}
A_X &= \{1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24\} \\
A_Y &= \{1,2,3,4,5,6,7,8,9,10,11,12,25,26,27,28,29,30,31,32,33,34,35,36\} \\
A_Z &= \{1,2,3,4,5,6,13,14,15,16,17,18,25,26,27,28,29,30,37,38,39,40,41,42\} \\
A_W &= \{1,2,3,7,8,9,13,14,15,19,20,21,25,26,27,31,32,33,37,38,39,43,44,45\}
\end{aligned}$$

$$\begin{tikzcd} X \arrow[r, leftrightarrow] \arrow[d] & Y \arrow[d] \arrow[dl, leftrightarrow] \\ Z \arrow[r] & W \end{tikzcd}$$

We then take three fresh points, 49, 50, and 51, add them to all sets A_i.
Then add one new point to A_Y, two new points to A_Z, and three to A_W.

$$\begin{aligned}
A_X &= \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 49, 50, 51\} \\
A_Y &= \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 49, 50, 51, 52\} \\
A_Z &= \{1, 2, 3, 4, 5, 6, 13, 14, 15, 16, 17, 18, 25, 26, 27, 28, 29, 30, 37, 38, 39, 40, 41, 42, 49, 50, 51, 53, 54\} \\
A_W &= \{1, 2, 3, 7, 8, 9, 13, 14, 15, 19, 20, 21, 25, 26, 27, 31, 32, 33, 37, 38, 39, 43, 44, 45, 49, 50, 51, 55, 56, 57\}
\end{aligned}$$

Now $|A_X| = 27$, $|A_Y| = 28$, $|A_Z| = 29$, and $|A_W| = 30$.
For $i \neq j$, $|A_i \cap A_j| = 15$, and $|A_i \sqcap A_j| = 3$.

$$\begin{tikzcd} X \arrow[r, leftrightarrow] \arrow[d] & Y \arrow[d] \arrow[dl, leftrightarrow] \\ Z \arrow[r] & W \end{tikzcd}$$

$$\begin{aligned} A_X &= \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 49, 50, 51\} \\ A_Y &= \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 49, 50, 51, 52\} \\ A_Z &= \{1, 2, 3, 4, 5, 6, 13, 14, 15, 16, 17, 18, 25, 26, 27, 28, 29, 30, 37, 38, 39, 40, 41, 42, 49, 50, 51, 53, 54\} \\ A_W &= \{1, 2, 3, 7, 8, 9, 13, 14, 15, 19, 20, 21, 25, 26, 27, 31, 32, 33, 37, 38, 39, 43, 44, 45, 49, 50, 51, 55, 56, 57\} \end{aligned}$$

We have already arranged that $A_X \to A_Y$ and $A_Y \to A_X$.

Here is how we arrange that $A_X \to A_Z$ and $A_Z \not\to A_X$.

Take the "private intersection" $A_X \sqcap A_Z = \{16, 17, 18\}$.
Remove 16 from A_X and A_Z, and return it as two fresh points $58 \in A_X$ and $59 \in A_Z$.
The point is that now $|A_X \cap A_Z| = 14$, and $\frac{14}{29} < \frac{1}{2} < \frac{14}{27}$.

$$\begin{tikzcd} X \arrow[r, leftrightarrow] \arrow[d] & Y \arrow[d] \arrow[dl, leftrightarrow] \\ Z \arrow[r] & W \end{tikzcd}$$

Similar tricks arrange all of our other requirements.

We get

$$\begin{aligned}
A_X &= \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 17, 18, 22, 23, 24, 49, 50, 51, 58, 60, 61, 62\} \\
A_Y &= \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 49, 50, 51, 52\} \\
A_Z &= \{1, 2, 3, 4, 5, 6, 13, 14, 15, 17, 18, 25, 26, 27, 28, 29, 30, 37, 38, 39, 40, 41, 42, 49, 50, 51, 53, 54, 59\} \\
A_W &= \{1, 2, 3, 7, 8, 9, 13, 14, 15, 25, 26, 27, 31, 32, 33, 37, 38, 39, 43, 44, 45, 49, 50, 51, 55, 56, 57, 63, 64, 65\}
\end{aligned}$$

This is a $\frac{1}{2}$-representation of our graph G.

$$\begin{tikzcd}
X \arrow[r, leftrightarrow] \arrow[d] & Y \arrow[d] \arrow[dl, leftrightarrow] \\
Z \arrow[r] & W
\end{tikzcd}$$

$$\begin{aligned}
A_X &= \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 17, 18, 22, 23, 24, 49, 50, 51, 58, 60, 61, 62\} \\
A_Y &= \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 49, 50, 51, 52\} \\
A_Z &= \{1, 2, 3, 4, 5, 6, 13, 14, 15, 17, 18, 25, 26, 27, 28, 29, 30, 37, 38, 39, 40, 41, 42, 49, 50, 51, 53, 54, 59\} \\
A_W &= \{1, 2, 3, 7, 8, 9, 13, 14, 15, 25, 26, 27, 31, 32, 33, 37, 38, 39, 43, 44, 45, 49, 50, 51, 55, 56, 57, 63, 64, 65\}
\end{aligned}$$

Recall our set

$$\Gamma = \left\{ \begin{array}{rr}
\text{Most } X \text{ are } Y & \text{Most } Z \text{ are } Y \\
\text{Most } Y \text{ are } X & \text{Most } Y \text{ are } W \\
\text{Most } X \text{ are } Z & \neg\text{Most } W \text{ are } Y \\
\neg\text{Most } Z \text{ are } X & \text{Most } Z \text{ are } W \\
\text{Most } Y \text{ are } Z &
\end{array} \right\}$$

We have built a model to see that

$$\Gamma \not\models \text{Most } W \text{ are } Z$$

Theorem (Tri Lai, Jörg Endrullis, and LM 2013)

Every graph without one-way cycles has a $\frac{1}{2}$-representation.

Theorem

The logical system for this language is complete.

Recall that the main axiom was

$$\begin{array}{ll} & (\text{Most } X_1 \text{ are } X_2 \wedge \text{Most } X_2 \text{ are } X_3 \wedge \cdots \wedge \text{Most } X_n \text{ are } X_1) \\ \rightarrow & (\text{Most } X_2 \text{ are } X_1 \vee \text{Most } X_3 \text{ are } X_2 \vee \cdots \vee \text{Most } X_1 \text{ are } X_n) \end{array}$$

corresponding to the absence of one-way cycles in $\frac{1}{2}$-graphs.

Theorem

The satisfiability problem for the logic is NP-complete.

Add $\exists^{\geq}(X, Y)$ and *More*(X, Y) on top of this logic.

Her rules include

$$\frac{\text{Most } y \text{ are } x \quad \text{There are at least as many } y \text{ are } x}{\text{Most } x \text{ are } y}$$

$$\frac{\text{Most } y \text{ are } x \quad \neg\text{Most } x \text{ are } y}{\text{There are at more } y \text{ than } x}$$

Add All x are y to the last fragment.

The complete system should be all the rules that I mentioned, plus:

$$\frac{\text{Most } x \text{ are } y \quad \text{All } y \text{ are } z}{\text{Most } x \text{ are } z}$$

What I would like to know

- Get a such complete logic for

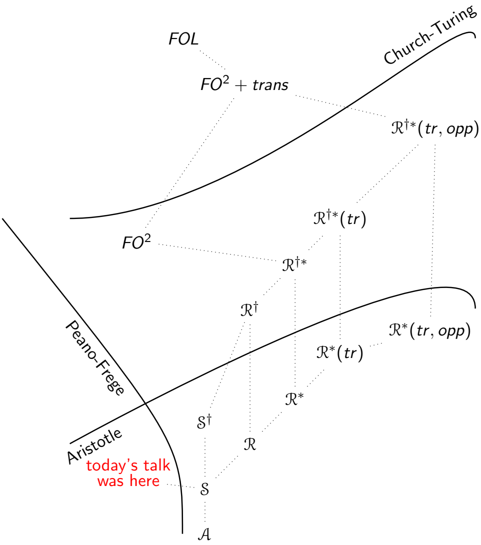
All X are Y	Some X are Y	Most X are Y
No X are Y	$\exists^{\geq}(X, Y)$	

and sentential $\wedge$, $\vee$, and $\neg$.

- Alternatively, prove that there is no such logic.
- Investigate the algorithmic properties of the logic.

I also want to mention the very important
complexity results on
"Deciding Boolean Algebra with Presburger Arithmetic"
Viktor Kuncak and collaborators.

Although not motivated by linguistics or by logic,
this line of work seems relevant to my topic today.



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

$\mathcal{A}$: all p are q

Now what do you think?

Quine, from *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.

This is something we'll talk about throughout the week.