

Monotonicity Inference in Theory and Practice

Larry Moss

NASSLLI'2025

June 23-25, 2025

Entertainment to initiate this topic: a 3 minute video on monotonicity

<https://www.youtube.com/watch?v=-zBPHuBGZAE&feature=youtu.be>

This was part of a competition for non-technical videos on mathematics for the general public.

Monotonicity is a big part of the Natural Logic Program

Show that significant parts of **inference** in natural language can be done automatically, using surface forms as much as possible.

To work in decidable, and “light” logical systems.

To make connections with all relevant fields of logic.

To re-work semantics in light of computational linguistics and cognitive science.

Let's give up on grammar, at least on the very strict forms of grammar which we have seen so far!

Monotonicity is a big part of the Natural Logic Program

Show that significant parts of **inference** in natural language can be done automatically, using surface forms as much as possible.

To work in decidable, and “light” logical systems.

To make connections with all relevant fields of logic.

To re-work semantics in light of computational linguistics and cognitive science.

In another direction

Since 2018, neural machine learners have been able to do some forms of inference at human level, or better.

Thus, we have a watershed moment in the history of logic.

This calls for a response.

- (1) some[↑] dog[↑] hit[↑] some[↑] cat[↑]
- (2) some[↑] dog[↑] kissed[↓] no[↑] cat[↓]
- (3) most[↑] dog⁼ hit[↓] no[↑] cat[↓]
- (4) no[↑] dog[↓] hit[↑] no[↓] cat[↑]
- (5) at-most two[↑] dog[↓] chased[↑] at-most three[↓] cats[↑]

knowledge base for nouns, transitive verbs,
determiners, and numbers

dog \leq animal
 cat \leq animal
 poodle \leq dog
 ragdoll \leq cat
 ⋮
 ⋮

kiss \leq touch
 hit \leq touch
 thrash \leq hit
 hit vigorously \leq hit
 hit lightly \leq hit
 ⋮

every \leq most
 most \leq some

 one \leq two
 two \leq three
 three \leq four
 no \leq at-most one

(6) some[↑] dog[↑] hit[↑] some[↑] cat[↑]

(7) some[↑] dog[↑] kissed[↓] no[↑] cat[↓]

(8) most[↑] dog⁼ hit[↓] no[↑] cat[↓]

⇒ some dog thrashed no ragdoll

(9) no[↑] dog[↓] hit[↑] no[↓] cat[↑]

⇒ at most one poodle touched no animal

(10) at-most two[↑] dog[↓] chased[↑] at-most three[↓] cats[↑]

knowledge base for nouns, transitive verbs,
determiners, and numbers

dog \leq animal

cat \leq animal

poodle \leq dog

ragdoll \leq cat

⋮

⋮

kiss \leq touch

hit \leq touch

thrash \leq hit

hit vigorously \leq hit

hit lightly \leq hit

⋮

every \leq most

most \leq some

one \leq two

two \leq three

three \leq four

no \leq at-most one

Put the arrow on **chased** in

No dog (chased no cat)

Put the arrow on **chased** in

No dog (chased no cat)

= Every dog (chased some cat)

Put the arrow on **chased** in

No dog (chased no cat)

= Every dog (chased some cat)

Moss 2012: **no**[↑] **dog**[↓] **chased**[↓] **no**[↓] **cat**[↑].

The arrow on **chased**[↓] is wrong.

An algebraic expression like

$$z - (v + w)$$

is **increasing** in z , and **decreasing** in v and w .

If we assume

▶ $z_1 \leq z_2$

▶ $v_2 \leq v_1$

▶ $w_2 \leq w_1$

Then we are entitled to conclude

$$z_1 - (v_1 + w_1) \leq z_2 - (v_2 + w_2)$$

We had

$$z - (v + w)$$

We would write

$$\frac{v \downarrow \quad w \downarrow \quad z \uparrow}{(z - (v + w)) \uparrow}$$

The responsible parties here are the facts that

- $+$: $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ is increasing (monotone) in both arguments
- $-$: $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ is increasing in the first argument
and decreasing (antitone) in the second argument

An algebraic expression like

$$z - (v + |w|)$$

is **increasing** in z , and **decreasing** in v , and
there's nothing we can say about w .

If we assume

▶ $z_1 \leq z_2$

▶ $v_2 \leq v_1$

▶ $w_2 = w_1$

Then we are entitled to conclude

$$z_1 - (v_1 + w_1) \leq z_2 - (v_2 + w_2)$$

We had

$$z - (v + |w|)$$

We would write

$$\frac{v \downarrow \quad w = \quad z \uparrow}{(z - (v + w)) \uparrow}$$

We just saw

$$\frac{v \downarrow \quad w = \quad z \uparrow}{(z - (v + w)) \uparrow}$$

and it follows that

$$\frac{v \downarrow \quad w = \quad z =}{(z - (v + w)) \uparrow}$$

We have strengthened the hypotheses and thus weakened the statement:

If we assume

- ▶ $z_1 = z_2$
- ▶ $v_2 \leq v_1$
- ▶ $w_2 = w_1$

Then we are once again entitled to conclude

$$z_1 - (v_1 + w_1) \leq z_2 - (v_2 + w_2)$$

... because algebra is first-order; we need higher-order stuff

Some[↑] dog[↑] and[↑] no[↑] cat[↓] sleeps⁼

Abstract some words:

$s(x, y, z) = \llbracket \text{Some } x \text{ and no } y \text{ } z \rrbracket$ in some underlying model

Our intuitions tell us that
 s is decreasing in x , decreasing in y , and
there's nothing we can say about z .

If we assume

- ▶ $x_1 \leq x_2$
- ▶ $y_2 \leq y_1$
- ▶ $z_1 = z_2$

Then we are entitled to conclude

$$\llbracket \text{Some } x_1 \text{ and no } y_1 \text{ } z_1 \rrbracket \leq \llbracket \text{Some } x_2 \text{ and no } y_2 \text{ } z_2 \rrbracket$$

... because algebra is first-order; we need higher-order stuff

Some[↑] dog[↑] and[↑] no[↑] cat[↓] sleeps⁼

Abstract some words:

$s(x, y, z) = \llbracket \text{Some } x \text{ and no } y \text{ } z \rrbracket$ in some underlying model

Our intuitions tell us that
 s is decreasing in x , decreasing in y , and
there's nothing we can say about z .

If we assume

- ▶ $x_1 \leq x_2$, but where?
- ▶ $y_2 \leq y_1$, but where?
- ▶ $z_1 = z_2$, but where? And what about the Dets?

Because the model is hypothetical,

Some x_1 and no y_1 z_1 **implies** Some x_2 and no y_2 z_2

Observe the 3 *np* types (!) and the $+/-/\cdot$ notation

item	category	type
Fido, Felix	NP	e
cat, dog	N	$n = pr = e \rightarrow t$
swim, run	IV = $S \backslash NP$	pr
chase, see, hit, kiss	TV = IV / NP	$e \rightarrow pr$
every	DET = NP / N	$pr \bar{\rightarrow} np^+$
some	NP / N	$pr \overset{+}{\rightarrow} np^+$
no	NP / N	$pr \bar{\rightarrow} np^-$
most	NP / N	$pr \dot{\rightarrow} np^+$
who	$(N \backslash N) / (S / NP)$	$(np^+ \overset{+}{\rightarrow} t) \overset{+}{\rightarrow} (pr \overset{+}{\rightarrow} pr)$
didn't	IV / IV TV / TV	$pr \bar{\rightarrow} pr$ $(e \rightarrow pr) \bar{\rightarrow} (e \rightarrow pr)$
and	$X / (X \backslash X)$	$x \overset{+}{\rightarrow} (x \overset{+}{\rightarrow} x)$
one, two, three	NUM	num
more than	DET / NUM	$num \bar{\rightarrow} (pr \overset{+}{\rightarrow} np^+)$
less than	DET / NUM	$num \overset{+}{\rightarrow} (pr \bar{\rightarrow} np^-)$
if ... then ...	$(S \backslash S) / S$	$t \bar{\rightarrow} (t \overset{+}{\rightarrow} t)$

Markings vs. Polarities

Markings are what words have in the lexicon:

$$\text{every} : pr \xrightarrow{-} np^{+}$$

We indicate markings by $+$, $-$, and \cdot .

Polarities arrows \uparrow , \downarrow and $=$

indicate “in sentence” monotonicity/antitonicity/neither assertions.

They are indicating the same things, but markings are permanent features of the lexical items.

Polarities are what we want to calculate in this work.

The rules

Don't worry, I explain what all this is about

Rules

$$\frac{u^d : x \xrightarrow{m} y \quad v^{md} : x}{(uv)^d : y} >$$

$$\frac{u^{md} : x}{u^d : (x \xrightarrow{m} y) \xrightarrow{+} y} \text{ T}$$

$$\frac{u^{md} : u \xrightarrow{n} y \quad v^{md} : y \xrightarrow{n} z}{(uv)^d : (x \xrightarrow{mn} z)} \text{ B}$$

$$\frac{u^{md} : e \rightarrow b}{u^d : np^m \xrightarrow{+} b} \text{ K}$$

$$\frac{u^d : x \xrightarrow{m} y}{u^d : x \xrightarrow{\cdot} y} \text{ M}$$

$$\frac{u^= : x \xrightarrow{m} y}{u^d : x \xrightarrow{m} y} \text{ W}$$

	↑	↓	=
+	↑	↓	=
-	↓	↑	=
·	=	=	=

d can be ↑ or ↓ or =.

m can be +, -, or ·.

There are two operations that combine these notation,
and we'll see one of them soon.

What one does with the polarization rules

One looks for a tree with the following features:

- ▶ The leaves must have a type from the lexicon.
- ▶ The leaves may be polarized \uparrow , \downarrow , or $=$; it's our choice.
- ▶ The type at the bottom must be t .
The polarity at the bottom must be \uparrow .
- ▶ At each non-leaf node, one of the rules of the system must be matched.

The goal

If we take the sentence at the root of the tree,
then all polarity arrows at the leaves
should match our intuitions.

Example

$$\begin{array}{c}
 \frac{\text{no}^\uparrow : pr \rightarrow np^- \quad \text{dog}^\downarrow : pr}{\text{no dog}^\uparrow : pr \rightarrow t} > \quad \frac{\frac{\text{chased}^\uparrow : e \rightarrow pr}{\text{chased}^\downarrow : np^- \rightarrow pr} \text{K} \quad \frac{\text{no}^\downarrow : pr \rightarrow np^- \quad \text{cat}^\uparrow : pr}{\text{no cat}^\downarrow : np^-} >}{\text{chased no cat}^\downarrow : pr} > \\
 \text{no dog chased no cat}^\uparrow : t
 \end{array}$$

Rules

$$\begin{array}{ccc}
 \frac{u^d : x \xrightarrow{m} y \quad v^{md} : x}{(uv)^d : y} > & \frac{u^{md} : x}{u^d : (x \xrightarrow{m} y) \xrightarrow{+} y} \text{T} & \frac{u^{md} : u \xrightarrow{n} y \quad v^{md} : y \xrightarrow{n} z}{(uv)^d : (x \xrightarrow{mn} z)} \text{B} \\
 \\
 \frac{u^{md} : e \rightarrow b}{u^d : np^m \xrightarrow{+} b} \text{K} & \frac{u^d : x \xrightarrow{m} y}{u^d : x \xrightarrow{\cdot} y} \text{M} & \frac{u^= : x \xrightarrow{m} y}{u^d : x \xrightarrow{m} y} \text{W}
 \end{array}$$

Example

$$\frac{\text{no}^{\uparrow} : pr \rightarrow np^- \quad \text{dog}^{\downarrow} : pr}{\text{no dog}^{\uparrow} : pr \rightarrow t} > \quad \frac{\text{no}^{\downarrow} : pr \rightarrow np^- \quad \text{cat}^{\uparrow} : pr}{\text{no cat}^{\downarrow} : pr \rightarrow t} >$$

item	type
cat, dog	$pr = e \rightarrow t$
no	$pr \rightarrow np^- = (e \rightarrow t) \rightarrow np^-$

van Benthem's rule

$$\frac{u^d : x \xrightarrow{m} y \quad v^{md} : x}{(uv)^d : y} >$$

	\uparrow	\downarrow	$=$
$+$	\uparrow	\downarrow	$=$
$-$	\downarrow	\uparrow	$=$
\cdot	$=$	$=$	$=$

In the left application, we have $d = \uparrow$ and $m = -$, so that $md = \downarrow$.

In the right application, we have $d = \downarrow$ and $m = -$, so that $md = \uparrow$.

The arrow in the conclusion is inherited from the function above, not the argument.

van Benthem's groundbreaking contributions

Work on the logical side of monotonicity/polarity stems from the seminal contributions of Johan van Benthem.

His central insight

Polarities can be computed from derivation trees in categorial grammar (CG).

References

Essays in Logical Semantics, 1986

Language in Action: Categories, Lambdas,
and Dynamic Logic, 1991

"A Brief History of Natural Logic", 2008

see also the 1991 PhD dissertation by Victor Sánchez Valencia
Studies on Natural Logic and Categorial Grammar

The K rule in this example

Example

$$\frac{\text{chased}^{\uparrow} : e \rightarrow (e \rightarrow t)}{\text{chased}^{\downarrow} : np^{\neg} \rightarrow^+ (e \rightarrow t)} \text{ K}$$

item	type
chased	$e \rightarrow pr = e \rightarrow (e \rightarrow t)$

K

$$\frac{u^{md} : e \rightarrow b}{u^d : np^m \rightarrow^+ b} \text{ K}$$

	\uparrow	\downarrow	$=$
$+$	\uparrow	\downarrow	$=$
$-$	\downarrow	\uparrow	$=$
\cdot	$=$	$=$	$=$

b must be a boolean type:

t is boolean, and if β is boolean, so is $\sigma \rightarrow \beta$.

We take $m = -$ and $d = \downarrow$.

Some dog chased no cat

Example

$$\begin{array}{c}
 \frac{\text{some}^\uparrow : pr \xrightarrow{+} np^+ \quad \text{dog}^\uparrow : pr}{\text{some dog}^\uparrow : pr \xrightarrow{+} t} > \quad \frac{\frac{\text{chased}^\downarrow : e \xrightarrow{+} pr}{\text{chased}^\uparrow : np^- \xrightarrow{+} pr} \text{K} \quad \frac{\text{no}^\uparrow : pr \xrightarrow{-} np^- \quad \text{cat}^\downarrow : pr}{\text{no cat}^\uparrow : np^-} >}{\text{chased no cat}^\uparrow : pr} > \\
 \text{some dog chased no cat}^\uparrow : t
 \end{array}$$

Rules

$$\begin{array}{ccc}
 \frac{u^d : x \xrightarrow{m} y \quad v^{md} : x}{(uv)^d : y} > & \frac{u^{md} : x}{u^d : (x \xrightarrow{m} y) \xrightarrow{+} y} \text{T} & \frac{u^{md} : u \xrightarrow{n} y \quad v^{md} : y \xrightarrow{n} z}{(uv)^d : (x \xrightarrow{mn} z)} \text{B} \\
 \\
 \frac{u^{md} : e \rightarrow b}{u^d : np^m \xrightarrow{+} b} \text{K} & \frac{u^d : x \xrightarrow{m} y}{u^d : x \xrightarrow{\cdot} y} \text{M} & \frac{u^= : x \xrightarrow{m} y}{u^d : x \xrightarrow{m} y} \text{W}
 \end{array}$$

	↑	↓	=
+	↑	↓	=
-	↓	↑	=
.	=	=	=

Example

$$\begin{array}{c}
 \frac{\text{most}^\uparrow : pr \rightarrow np^+ \quad \text{dogs}^\downarrow : pr}{\text{most dogs}^\uparrow : pr \rightarrow t} > \quad \frac{\frac{\text{chased}^\uparrow : e \xrightarrow{+} pr}{\text{chased}^\uparrow : np^+ \xrightarrow{+} pr} \text{ K} \quad \frac{\text{most}^\uparrow : pr \rightarrow np^+ \quad \text{cats}^\downarrow : pr}{\text{most cats}^\uparrow : np^+} > \\
 \frac{\text{most dogs}^\uparrow : pr \rightarrow t \quad \text{chased most cats}^\uparrow : pr}{\text{most dogs chased most cats}^\uparrow : t} >
 \end{array}$$

Example

$$\begin{array}{c}
 \frac{\text{chase}^\downarrow : e \xrightarrow{+} pr}{\text{chase}^\downarrow : np^+ \xrightarrow{+} pr} \text{ K} \quad \frac{\text{every}^\downarrow : pr \rightarrow np^+ \quad \text{cat}^\uparrow : pr}{\text{every cat}^\downarrow : np^+} > \\
 \frac{\text{didn't}^\uparrow : pr \rightarrow pr \quad \text{chase every cat}^\downarrow : pr}{\text{didn't chase every cat}^\uparrow : pr} > \\
 \frac{\text{Fido}^\uparrow : e \quad \text{didn't chase every cat}^\uparrow : pr}{\text{Fido didn't chase every cat}^\uparrow : t} <
 \end{array}$$

Introducing the armadillo



Example

$$\begin{array}{c}
 \frac{\frac{\text{ch}^\uparrow : e \xrightarrow{+} pr}{\text{ch}^\uparrow : np^+ \xrightarrow{+} pr} \quad \text{some cat}^\uparrow : np^+}{\text{ch}^\uparrow : np^+ \xrightarrow{+} pr} \quad \text{some cat and no armadillo}^\uparrow : np < \\
 \frac{\text{and} : np^+ \xrightarrow{+} (np^+ \xrightarrow{+} np^+) \quad \text{some arm}^\uparrow : np^+}{\text{and some arm} : np \xrightarrow{+} np} > \\
 \frac{F^\uparrow : e \quad \text{chased some cat and some armadillo}^\uparrow : pr}{\text{Fido chased some cat and some armadillo}^\uparrow : t} <
 \end{array}$$

Example

$$\begin{array}{c}
 \frac{\text{ch}^= : e \xrightarrow{+} pr}{\text{ch}^= : np \xrightarrow{+} pr} \text{K} \quad \frac{\text{some cat}^\uparrow : np^+}{\text{some cat}^\uparrow : np} \text{M} \quad \frac{\text{and} : np \xrightarrow{+} (np \xrightarrow{+} np) \quad \frac{\text{no arm}^\downarrow : np^-}{\text{no arm}^\downarrow : np} \text{M}}{\text{and no arm} : np \xrightarrow{+} np} > \\
 \frac{\text{ch}^\uparrow : np \xrightarrow{+} pr}{\text{ch}^\uparrow : np \xrightarrow{+} pr} \text{W} \quad \frac{\text{some cat}^\uparrow : np}{\text{some cat}^\uparrow : np} \text{M} \quad \frac{\text{and no arm} : np \xrightarrow{+} np}{\text{some cat and no armadillo}^\uparrow : np} < \\
 \frac{F^\uparrow : e \quad \text{chased some cat and no armadillo}^\uparrow : pr}{\text{Fido chased some cat and no armadillo}^\uparrow : t} >
 \end{array}$$

We use (M) twice
in order to conjoin **some cat** and **no armadillo**.

In effect, we are weakening the type of **some cat** from np^+ to np ,
and also weakening the type of **no cat** from np^- to np

We do this in order to conjoin them,
since the type of **and** is $x \xrightarrow{+} (x \xrightarrow{+} x)$.

Note also that we used (W) in order to change the polarity of **chased** from $=$ to \uparrow .

Example

$$\begin{array}{c}
 \text{fewer than}^{\uparrow} : \text{num} \xrightarrow{+} (pr \xrightarrow{-} np^{-}) \quad \text{three}^{\uparrow} : \text{num} \\
 \hline
 \text{fewer than three}^{\uparrow} : pr \xrightarrow{-} np^{-} \quad \text{cats}^{\downarrow} : pr \\
 \hline
 \text{fewer than three cats}^{\uparrow} : np^{-} \quad \text{swam}^{\downarrow} : pr \\
 \hline
 \text{fewer than three cats swam}^{\uparrow} : t
 \end{array}$$

We are going to move from sets to preorders in semantics

In order to talk about **inference by replacement**, we need sensible notions of “bigger” and “smaller”.

It's clear enough how to do this for “simple” semantic objects like intransitive verbs, but others (complex NPs), it's not that easy.

It's better to change the way we do formal semantics by moving from **sets** to **preorders** as much as possible.

Definition

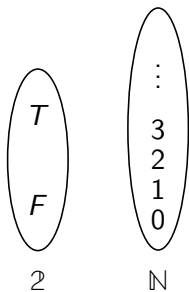
A **preorder** is a pair $\mathbb{P} = (P, \leq)$ consisting of a set P together with a relation \leq which is **reflexive** and **transitive**.

This means that the following hold:

- ▶ $p \leq p$ for all $p \in P$.
- ▶ If $p \leq q$ and $q \leq r$, then $p \leq r$.

The preorder $\mathbb{2}$ of truth values: $\{T, F\}$, with $F \leq T$.

The preorder \mathbb{N} of natural numbers: $0 \leq 1 \leq 2 \leq \dots$



Going up in these pictures means getting “bigger” in the preorder.

Definition

For any preorder \mathbb{P} , there is an **opposite preorder** \mathbb{P}^{op} .
Its domain is the same domain as that of \mathbb{P} .

$$p \leq q \text{ in } \mathbb{P}^{op} \quad \text{iff} \quad q \leq p \text{ in } \mathbb{P}$$

Definition

For any preorder \mathbb{P} , there is an **flattened version** \mathbb{P}^b .
Its domain is the same domain as that of \mathbb{P} .

$$p \leq q \text{ in } \mathbb{P}^b \quad \text{iff} \quad p = q$$

Definition

For any preorders \mathbb{P} and \mathbb{Q} , there is a **product preorder** $\mathbb{P} \times \mathbb{Q}$.
Its domain is the cartesian product the given domains.

$$(p, q) \leq (p', q') \text{ in } \mathbb{P} \times \mathbb{Q} \quad \text{iff} \quad p \leq p' \text{ in } \mathbb{P}, \text{ and } q \leq q' \text{ in } \mathbb{Q}$$

Going up in the pictures means \geq

Example



2



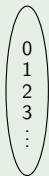
2^{op}



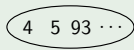
2^b



\mathbb{N}



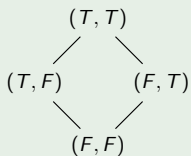
\mathbb{N}^{op}



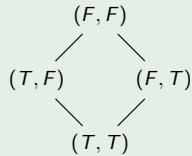
\mathbb{N}^b

Going up in the pictures means \geq

Example

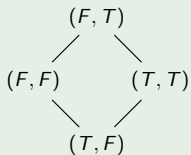


2×2

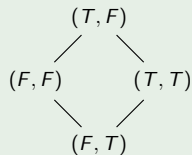


$(2 \times 2)^{op}$

Example



$2^{op} \times 2$



2×2^{op}

Monotone $f : \mathbb{P} \rightarrow \mathbb{Q}$

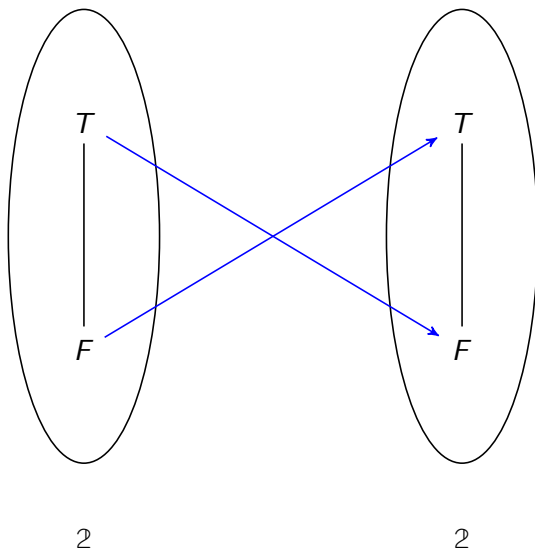
If $p \leq q$ in \mathbb{P} , then $f(p) \leq f(q)$ in \mathbb{Q} .

We write $f : \mathbb{P} \xrightarrow{+} \mathbb{Q}$.

Antitone $f : \mathbb{P} \rightarrow \mathbb{Q}$

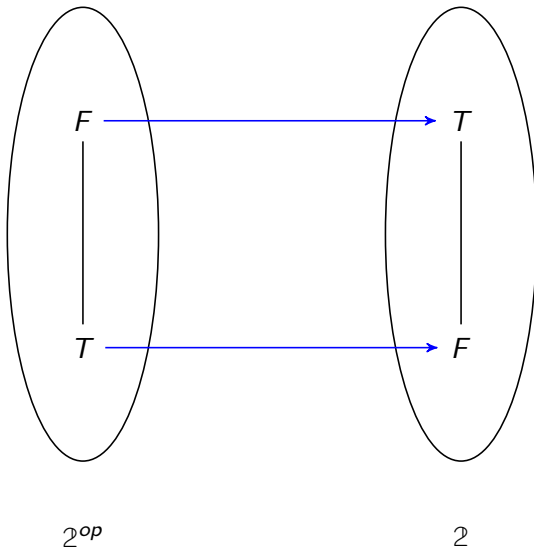
If $p \leq q$ in \mathbb{P} , then $f(q) \leq f(p)$ in \mathbb{Q} .

We write $f : \mathbb{P} \xrightarrow{-} \mathbb{Q}$.



Negation is **antitone**

This is the same as a monotone function from \mathcal{P}^{op} to \mathcal{P}

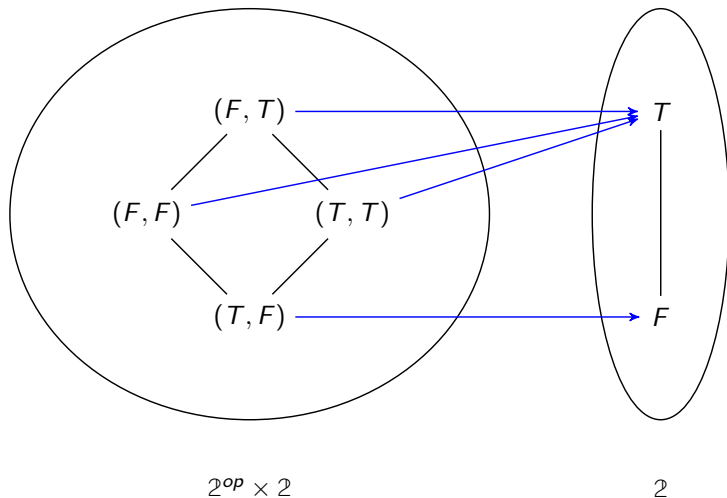


We write these as

$$\neg: \mathbb{2} \xrightarrow{-} \mathbb{2}$$

$$\neg: \mathbb{2}^{op} \xrightarrow{+} \mathbb{2}$$

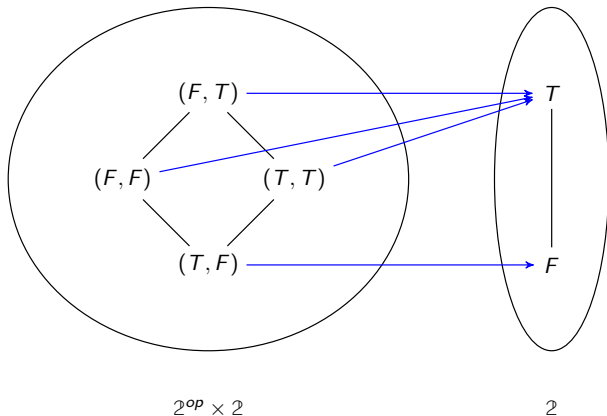
Standard implication \rightarrow is a monotone function,
but how?



Standard implication \rightarrow is a monotone function, but how?

We write this as

$$\rightarrow: (2^{op} \times 2) \xrightarrow{+} 2$$



The monotonicity summary of logical implication \rightarrow

\rightarrow is antitone in its first argument,
and monotone in its second argument.

This should be familiar.

In a mathematical result like

Theorem

(If \downarrow , then \uparrow) \uparrow .

We get a better result by **weakening the hypothesis**,
or by **strengthening the conclusion**.

The monotonicity summary of logical implication \rightarrow

\rightarrow is antitone in its first argument,
and monotone in its second argument.

$$\rightarrow: (2^{op} \times 2) \overset{+}{\rightarrow} 2$$

curried forms

$$\rightarrow: 2^{op} \overset{+}{\rightarrow} (2 \overset{+}{\rightarrow} 2)$$

$$\rightarrow: 2 \overset{-}{\rightarrow} (2 \overset{+}{\rightarrow} 2)$$

Definition

For any preorder \mathbb{P} and any set X ,
we have a new preorder called $X \rightarrow \mathbb{P}$.

The domain of this preorder is the **function set**

$$X \rightarrow P$$

The order on $X \rightarrow \mathbb{P}$ is the **pointwise order**:

$$f \leq g \text{ iff for all } x \in X, f_x \leq_{\mathbb{P}} g_x.$$

Monotone, Antitone, and “no information”

Monotone function f

If $p \leq q$ in \mathbb{P} , then $f(p) \leq f(q)$ in \mathbb{Q} .

We write $f : \mathbb{P} \xrightarrow{+} \mathbb{Q}$.

Antitone function f

If $p \leq q$ in \mathbb{P} , then $f(q) \leq f(p)$ in \mathbb{Q} .

We write $f : \mathbb{P} \xrightarrow{-} \mathbb{Q}$, or equivalently $f : \mathbb{P}^{op} \xrightarrow{+} \mathbb{Q}$.

Arbitrary function f with no special properties

We write $f : \mathbb{P} \xrightarrow{\cdot} \mathbb{Q}$.

So this means “in general, neither monotone nor antitone.”

It is equivalent to $f : \mathbb{P}^b \xrightarrow{+} \mathbb{Q}$.

Every function from a flat preorder is automatically monotone.

Monotone and Antitone function sets

For sets X , Y , there is a set $X \rightarrow Y$ of all functions from X to Y .

Moving to preorders, this idea becomes more involved.

Given preorders \mathbb{P} and \mathbb{Q} , we have

- ▶ A preorder $\mathbb{P} \xrightarrow{+} \mathbb{Q}$ of all monotone functions
- ▶ A preorder $\mathbb{P} \xrightarrow{-} \mathbb{Q}$ of all antitone functions
- ▶ A preorder $\mathbb{P} \xrightarrow{\cdot} \mathbb{Q}$ of all functions

The order on each of these is the pointwise order.

Preorder enrichment of categorial grammar

We need to change the entire architecture of CG.

Standard CG

Define

$k : \text{Syntactic categories} \rightarrow \text{Semantic types}$

by

$$\begin{aligned}k(S) &= t \\k(N) &= e \rightarrow t \\k(NP) &= (e \rightarrow t) \rightarrow t \\k(X/Y) &= k(Y) \rightarrow k(X) \\k(X \setminus Y) &= k(Y) \rightarrow k(X)\end{aligned}$$

Then in a model we choose a set S_e for e , we always take $S_t = \{T, F\}$ and $S_{X/Y}$ is

$(S_Y \rightarrow S_X) = \text{the set of all functions from } S_{k(X)} \text{ to } S_{k(X)}$

We'll see soon how this works in the preorder enrichment.

More semantic types brings more power and more responsibility

It used to be that the “ e, t ” type for NP was $(e \rightarrow t) \rightarrow t$.

Now we have 9 semantic types:

$(e \xrightarrow{+} t) \xrightarrow{+} t$	$(e \xrightarrow{-} t) \xrightarrow{-} t$	$(e \xrightarrow{\cdot} t) \xrightarrow{+} t$
$(e \xrightarrow{+} t) \xrightarrow{-} t$	$(e \xrightarrow{+} t) \xrightarrow{\cdot} t$	$(e \xrightarrow{\cdot} t) \xrightarrow{-} t$
$(e \xrightarrow{-} t) \xrightarrow{+} t$	$(e \xrightarrow{-} t) \xrightarrow{\cdot} t$	$(e \xrightarrow{\cdot} t) \xrightarrow{\cdot} t$

Which do we actually want for NP now??

More semantic types brings more power and more responsibility

We want to interpret semantic types by preorders, not sets.

For t , we want $\mathbb{2}$.

For e , we want an arbitrary **flat** preorder $\mathbb{P} = \mathbb{P}^b$.

By flatness $e \overset{+}{\rightarrow} t$, $e \overset{-}{\rightarrow} t$ and $e \overset{\cdot}{\rightarrow} t$ give the same preorder.

We call this ***pr***, for **property**.

More semantic types brings more power and more responsibility

We want to interpret semantic types by preorders, not sets.

For t , we want $\mathbb{2}$.

For e , we want an arbitrary **flat** preorder $\mathbb{P} = \mathbb{P}^b$.

By flatness $e \xrightarrow{+} t$, $e \xrightarrow{-} t$ and $e \xrightarrow{\cdot} t$ give the same preorder.

We call this **pr**, for **property**.

Notation

$$np^+ = pr \xrightarrow{+} t$$

$$np^- = pr \xrightarrow{-} t$$

$$np = pr \xrightarrow{\cdot} t$$

Preorder enrichment of categorial grammar

Preordered CG

Define

$k : \text{Syntactic categories} \rightarrow \text{Semantic types}$

by

$$k(S) = t$$

$$k(N) = e \rightarrow t$$

$$k(NP) = (e \rightarrow t) \overset{+}{\rightarrow} t$$

For the types σ corresponding to categories like transitive or intransitive verbs, we have **choices** for \mathbb{P}_σ .

Anyways, when we form function types, we always have three choices:

$$\mathbb{P}_\sigma \overset{+}{\rightarrow} \mathbb{P}_\sigma$$

$$\mathbb{P}_\sigma \overset{-}{\rightarrow} \mathbb{P}_\sigma$$

$$\mathbb{P}_\sigma \overset{\cdot}{\rightarrow} \mathbb{P}_\sigma$$

If $f : \mathbb{P} \xrightarrow{+} \mathbb{Q}$ and $g : \mathbb{Q} \xrightarrow{+} \mathbb{R}$, then $g \circ f : \mathbb{P} \xrightarrow{+} \mathbb{R}$.

If $f : \mathbb{P} \xrightarrow{+} \mathbb{Q}$ and $g : \mathbb{Q} \xrightarrow{-} \mathbb{R}$, then $g \circ f : \mathbb{P} \xrightarrow{-} \mathbb{R}$.

If $f : \mathbb{P} \xrightarrow{-} \mathbb{Q}$ and $g : \mathbb{Q} \xrightarrow{+} \mathbb{R}$, then $g \circ f : \mathbb{P} \xrightarrow{-} \mathbb{R}$.

If $f : \mathbb{P} \xrightarrow{-} \mathbb{Q}$ and $g : \mathbb{Q} \xrightarrow{-} \mathbb{R}$, then $g \circ f : \mathbb{P} \xrightarrow{+} \mathbb{R}$.

If either was $\dot{\rightarrow}$, the composition would also be $\dot{\rightarrow}$.

Not to miss a generalization

$$\frac{f : \mathbb{P} \xrightarrow{m} \mathbb{Q} \quad g : \mathbb{Q} \xrightarrow{n} \mathbb{R}}{g \circ f : \mathbb{P} \xrightarrow{mn} \mathbb{R}}$$

We use a “multiplication” operation on the markings

$$m, n \mapsto mn$$

given in the chart:

mn	+	-	·
+	+	-	·
-	-	+	·
·	·	·	·

For all preorders \mathbb{P} and \mathbb{Q} , and $f_1, f_2 : P \rightarrow Q$, and all $p_1, p_2 \in P$:

- ① If $f_1, f_2 : \mathbb{P} \xrightarrow{+} \mathbb{Q}$, and $f_1 \leq f_2$, and $p_1 \leq p_2$, then $f_1(p_1) \leq f_2(p_2)$.
- ② If $f_1, f_2 : \mathbb{P} \xrightarrow{-} \mathbb{Q}$, and $f_1 \leq f_2$, and $p_2 \leq p_1$, then $f_1(p_1) \leq f_2(p_2)$.
- ③ If $f_1, f_2 : \mathbb{P} \xrightarrow{\cdot} \mathbb{Q}$, and $f_1 \leq f_2$, and $p_2 = p_1$, then $f_1(p_1) \leq f_2(p_2)$.
- ④ If $f_1, f_2 : \mathbb{P} \xrightarrow{+} \mathbb{Q}$, and $f_2 \leq f_1$, and $p_2 \leq p_1$, then $f_2(p_2) \leq f_1(p_1)$.
- ⑤ If $f_1, f_2 : \mathbb{P} \xrightarrow{-} \mathbb{Q}$, and $f_2 \leq f_1$, and $p_1 \leq p_2$, then $f_2(p_2) \leq f_1(p_1)$.
- ⑥ If $f_1, f_2 : \mathbb{P} \xrightarrow{\cdot} \mathbb{Q}$, and $f_2 \leq f_1$, and $p_2 = p_1$, then $f_2(p_2) \leq f_1(p_1)$.
- 7-9. If $f_1, f_2 : \mathbb{P} \xrightarrow{\cdot} \mathbb{Q}$, and $f_1 = f_2$, and $p_1 = p_2$, then $f_1(p_1) = f_2(p_2)$.

We have seen the summarizing notation

Pattern

$$\frac{f^d : \mathbb{P} \xrightarrow{m} \mathbb{Q} \quad p : \mathbb{P}^{md}}{(f(p))^d : \mathbb{Q}} >$$

We combine markings and polarities as in the table below:

md	$+$	$-$	\cdot
\uparrow	\uparrow	\downarrow	$=$
\downarrow	\downarrow	\uparrow	$=$
$=$	$=$	$=$	$=$

Rules to build trees (explanation below)

Rules: x, y are semantic types, b is a boolean type

$$\begin{array}{c}
 \frac{u^d : x \xrightarrow{m} y \quad v^{md} : x}{(uv)^d : y} > \quad \frac{u^d : x \xrightarrow{m} y \quad v^{md} : y \xrightarrow{n} z}{(uv)^d : (x \xrightarrow{mn} z)} B \quad \frac{u^{md} : x}{u^d : (x \xrightarrow{m} y) \xrightarrow{+} y} T \\
 \\
 \frac{u^{md} : e \rightarrow b}{u^d : np^m \xrightarrow{+} b} K \quad \frac{u^d : x \xrightarrow{m} y}{u^d : x \rightarrow y} M \quad \frac{u^= : x \xrightarrow{m} y}{u^d : x \xrightarrow{m} y} W
 \end{array}$$

The $>$ rule is function application.

The B -rule is function composition, backwards.

The T -rule is polarized form of type raising a la Montague.

In the K -rule, b has to be a boolean type.

The soundness of the rule is connected to Keenan and Faltz' Justification Theorem.

The M - and W -rules are forms of weakening.

Keenan and Faltz' "Justification Theorem"

Equivalent to the characterization of free complete atomic boolean algebras

Definition

Caba is the category of **complete atomic boolean algebras**, with morphisms the maps $f : A \rightarrow B$ such that

$$f(\bigwedge S) = \bigwedge_{a \in S} f(a)$$

for all $S \subseteq A$, and also such that $f(-a) = -f(a)$ for all $a \in A$.

We have an adjunction

$$\begin{array}{ccc} \text{Set} & \xrightarrow{2^{2^X}} & \text{Caba} \\ & \xleftarrow{U} & \end{array}$$

Keenan and Faltz' "Justification Theorem"

Equivalent to the characterization of free complete atomic boolean algebras

Definition

Caba is the category of **complete atomic boolean algebras**, with morphisms the maps $f : A \rightarrow B$ such that

$$f(\bigwedge S) = \bigwedge_{a \in S} f(a)$$

for all $S \subseteq A$, and also such that $f(-a) = -f(a)$ for all $a \in A$.

Theorem (Justification Theorem)

Let \mathbb{B} be any **Caba**, and let

$$f : E \rightarrow U\mathbb{B}$$

be any **Set** function.

Then there is a unique **Caba**-morphism

$$f_* : 2^{2^E} \rightarrow \mathbb{B}$$

such that

$$Uf_* \circ \eta_E = f$$

where $\eta_E : E \rightarrow 2^{2^E}$ is the Montague lift.

Notation

2^{2^E} is $E \rightarrow 2 \rightarrow 2$.

$(2^{2^E})^+$ is $(E \rightarrow 2) \xrightarrow{+} 2$.

$(2^{2^E})^-$ is $(E \rightarrow 2) \xrightarrow{-} 2$.

For $m = +$ and $m = -$, we have a **restriction map**

$$r_m : (2^{2^E} \xrightarrow{+} \mathbb{B}) \xrightarrow{+} ((2^{2^E})^m \xrightarrow{+} \mathbb{B}). \quad (1)$$

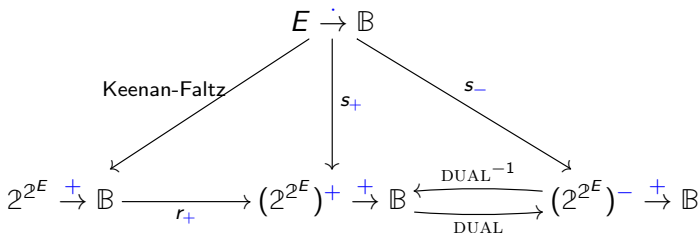
Let

$$s_m : (E \xrightarrow{\cdot} \mathbb{B}) \rightarrow ((2^{2^E})^m \xrightarrow{+} \mathbb{B}) \quad (2)$$

be given by

$$s_m f = f_* r_m$$

That is, $s_m f Q = f_* Q$, for all $f : E \xrightarrow{\cdot} \mathbb{B}$ and all $Q \in (2^{2^E})^m$.



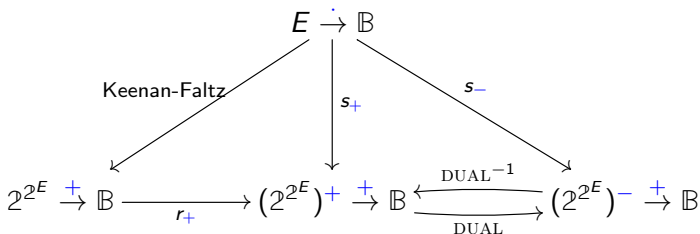
Lemma

s_+ is monotone, and s_- is antitone:

$$s_+ : (E \dot{\rightarrow} B) \xrightarrow{+} ((2^{2^E})^+ \xrightarrow{+} B)$$

$$s_- : (E \dot{\rightarrow} B) \xrightarrow{-} ((2^{2^E})^- \xrightarrow{+} B)$$

Moreover, $s_- = \text{DUAL} \circ s_+$, and $s_+ = \text{DUAL}^{-1} \circ s_-$.



This justifies the rule

$$\frac{f^{md} : e \rightarrow b}{(s_m f)^d : np^m \xrightarrow{+} b} \text{ K}$$

Fact that also comes up

$$(\mathbb{P}^{op} \xrightarrow{+} \mathbb{Q}^{op}) = (\mathbb{P} \xrightarrow{+} \mathbb{Q})^{op}$$

Polarized parse trees, and what they mean

Example (a polarized syntax tree)

$$\begin{array}{c}
 \frac{\text{some}^\uparrow : pr \xrightarrow{+} np^+ \quad \text{dog}^\uparrow : pr}{\text{some dog}^\uparrow : pr \xrightarrow{+} t} > \frac{\frac{\text{chased}^\downarrow : e \xrightarrow{+} pr}{\text{chased}^\uparrow : np^- \xrightarrow{+} pr} \quad \frac{\text{no}^\uparrow : pr \xrightarrow{-} np^- \quad \text{cat}^\downarrow : pr}{\text{no cat}^\uparrow : np^-} >}{\text{chased no cat}^\uparrow : pr} > \\
 \text{some dog chased no cat}^\uparrow : t
 \end{array}$$

Polarized parse trees, and what they mean

Example (a polarized syntax tree)

$$\begin{array}{c}
 \frac{\text{some}^\uparrow : pr \xrightarrow{+} np^+ \quad \text{dog}^\uparrow : pr}{\text{some dog}^\uparrow : pr \xrightarrow{+} t} > \quad \frac{\frac{\text{chased}^\downarrow : e \xrightarrow{+} pr}{\text{chased}^\uparrow : np^- \xrightarrow{+} pr} \text{ K} \quad \frac{\text{no}^\uparrow : pr \xrightarrow{-} np^- \quad \text{cat}^\downarrow : pr}{\text{no cat}^\uparrow : np^-} >}{\text{chased no cat}^\uparrow : pr} > \\
 \text{some dog chased no cat}^\uparrow : t
 \end{array}$$

The rules actually have **combinator constants** which I elided before

Rules

$$\begin{array}{ccc}
 \frac{u^d : x \xrightarrow{m} y \quad v^{md} : x}{(uv)^d : y} > & \frac{u^{md} : x}{(Tu)^d : (x \xrightarrow{m} y) \xrightarrow{+} y} \text{ T} & \frac{u^{md} : u \xrightarrow{n} y \quad v^{md} : y \xrightarrow{n} z}{(Buv)^d : (x \xrightarrow{mn} z)} \text{ B} \\
 \\
 \frac{u^{md} : e \rightarrow b}{(r_m u)^d : np^m \xrightarrow{+} b} \text{ K} & \frac{u^d : x \xrightarrow{m} y}{u^d : x \xrightarrow{\cdot} y} \text{ M} & \frac{u^- : x \xrightarrow{m} y}{u^d : x \xrightarrow{m} y} \text{ W}
 \end{array}$$

Incidentally, I also polarized a few other combinators in addition to the CCG ones.
There almost certainly is a general algorithm to be had.

Polarized parse trees, and what they mean

Example (What it means)

$$\begin{array}{c}
 \frac{v^{\uparrow} : pr \xrightarrow{+} np^{+} \quad w^{\uparrow} : pr}{vw^{\uparrow} : pr \xrightarrow{+} t} > \quad \frac{\frac{x^{\downarrow} : e \xrightarrow{+} pr}{r_{-}x^{\uparrow} : np^{-} \xrightarrow{+} pr} \text{ K} \quad \frac{y^{\uparrow} : pr \xrightarrow{-} np^{-} \quad z^{\downarrow} : pr}{yz^{\uparrow} : np^{-}} >}{(r_{-}x)(yz)^{\uparrow} : pr} > \\
 \tau = (vw)((r_{-}x)(yz))^{\uparrow} : t >
 \end{array}$$

The semantic term τ on the bottom is a combinator.

Soundness

The polarity arrows on the leaves mean that in every model,

$$\llbracket \tau \rrbracket (v, w, x, y, z) : \mathbb{P}_{pr \xrightarrow{+} np^{+}} \times \mathbb{P}_{pr} \times (\mathbb{P}_{e \xrightarrow{+} pr})^{op} \times \mathbb{P}_{pr \xrightarrow{-} np^{-}} \times (\mathbb{P}_{pr})^{op} \xrightarrow{+} \mathbb{P}_t$$

At this point, my talk shifts from theory to practice.

My former PhD student **Hai Hu** designed and implemented an arrow-tagging tool called **ccg2Mono** based on the rules which we have seen, and then he built an inference engine for NLI based on **monotonicity** and **natural logic** called **MonaLog**.

<https://huhailinguist.github.io/>

I am cheating a little, so you can ask me on this

$$\begin{array}{c}
 \frac{}{\text{every} : \text{NP}/\text{N}} & \frac{\text{cat} : \text{N} \quad \frac{\text{that Fido chased} : \text{N} \setminus \text{N}}{\text{cat that Fido chased} : \text{N}}}{} & & \frac{\frac{\text{F} : \text{NP}}{\text{F} : \text{S}/(\text{S} \setminus \text{NP})} \text{ T} \quad \text{ch} : (\text{S} \setminus \text{NP})/\text{NP}}{\text{Fido chased} : \text{S}/\text{NP}} \text{ B} \\
 & \frac{}{\text{every cat that Fido chased} : \text{NP}} & & \frac{}{\text{ran} : \text{S} \setminus \text{NP}} \\
 & \frac{}{\text{every cat that Fido chased ran} : \text{S}} & &
 \end{array}$$
$$\frac{\frac{\frac{\text{every}^\uparrow : pr \rightarrow np^+}{\text{every cat that Fido chased}^\uparrow : np^+} > \frac{\frac{\frac{\text{cat}^\downarrow : pr}{\text{cat that Fido chased}^\downarrow : pr} < \frac{\frac{\frac{\text{that}^\downarrow : (e \rightarrow t) \rightarrow (pr \rightarrow pr)}{\text{Fido chased}^\downarrow : e \rightarrow t} > \frac{F^\downarrow : e \quad \text{ch}^\downarrow : e \rightarrow (e \rightarrow t)}{<}}{\text{ran}^\uparrow : e \rightarrow t} \text{K}}{\text{ran}^\uparrow : NP^+ \rightarrow t} <}}{\text{every cat that Fido chased ran}^\uparrow : t} <$$

49/67

I am cheating a little, so you can ask me on this

$$\frac{\frac{\frac{\text{every}^\uparrow : pr \multimap np^+}{\text{every cat that Fido chased}^\uparrow : np^+} > \frac{\frac{\frac{\text{cat}^\downarrow : pr}{\text{cat that Fido chased}^\downarrow : pr} < \frac{\frac{\frac{\text{that}^\downarrow : (e \rightarrow t) \multimap (pr \multimap pr)}{\text{Fido chased}^\downarrow : e \rightarrow t} > \frac{F^\downarrow : e \quad \text{ch}^\downarrow : e \rightarrow (e \rightarrow t)}{} <}{\text{ran}^\uparrow : NP^+ \multimap t} \text{K} <}{\text{ran}^\uparrow : e \multimap t} <}{\text{every cat that Fido chased ran}^\uparrow : t} <$$

To me, we only have clear order intuitions about “low” types, and so the type-raising treatment is appealing.

Examples of polarized sentences from MonaLog

system available at <https://github.com/huhailinguist/monalog>

Every[↑] man[↓] and[↑] some[↑] woman[↑] sleeps[↑]

Every[↑] man[↓] and[↑] no[↑] woman[↓] sleeps⁼

If[↑] some[↓] man[↓] walks[↓], then[↑] no[↑] woman[↓] runs[↓]

Every[↑] man[↓] does[↓] n't[↑] hit[↓] every[↓] dog[↑]

No[↑] man[↓] who[↓] likes[↓] every[↓] dog[↑] sleeps[↓]

Most[↑] men⁼ that⁼ every⁼ woman⁼ hits⁼ cried[↑]

Every[↑] young[↓] man[↓] that[↑] no[↑] young[↓] woman[↓] hits[↑] cried[↑]

A[↑] special[↑] report[↑] found[↓] no[↑] incriminating[↓] evidence[↓]

Challenge

Assume: all skunks are mammals

Which of the following follows?

- ▶ All who love all who hate all skunks
love all who hate all mammals.
- ▶ All who love all who hate some skunks
love all who hate some skunks.

Toy example, but too hard for most people

Solution

Clearly we have

- ▶ All who love all who hate all skunks
love all who hate all skunks.

This is just All who are X are X

Toy example, but too hard for most people

Solution

ccg2Mono tells us

- ▶ All who love all who hate all skunks \downarrow .
love all who hate all skunks \uparrow .

We assumed $\text{skunks} \leq \text{mammals}$.

So we can move the second skunks up to mammals .

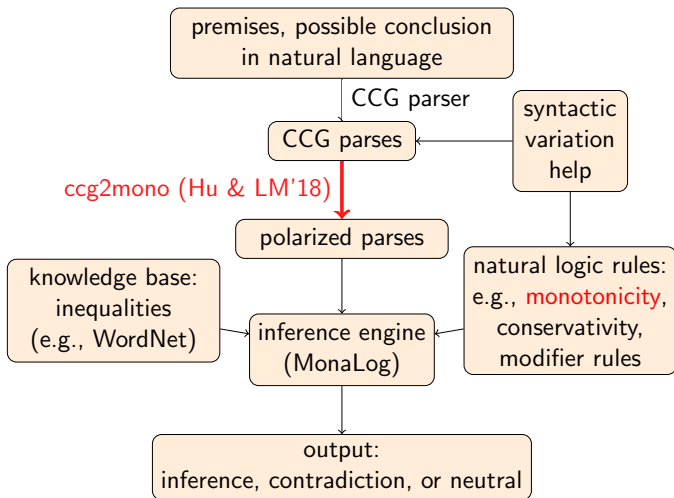
We conclude

- ▶ All who love all who hate all skunks \downarrow .
love all who hate all mammals \uparrow .

-
- ▶ All who love all who hate some skunks \uparrow .
love all who hate some skunks \downarrow .

So here we would make a different inference.

How arrow tagging fits in to our natural logic inference (NLI) system MonaLog



entail, contradict or neutral?

P: A schoolgirl with a black bag is on a crowded train

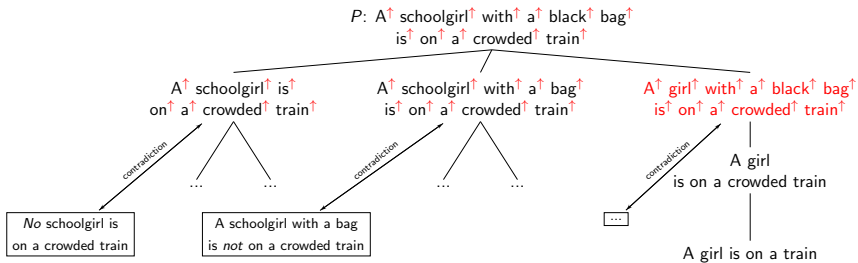
H: No schoolgirl is on a crowded train

entail, contradict or neutral?

P: A schoolgirl with a black bag is on a crowded train

H: A girl is on a train

How the algorithm works, roughly



Results on the SICK dataset

system	P	R	acc.
majority baseline	–	–	56.36
Natural-logic-based: MonaLog [†]			
MonaLog + pass2act	89.42	72.18	80.25 [†]
MonaLog + \exists transformation	89.43	71.53	79.11 [†]
MonaLog + all rewrite help	83.75	70.66	77.19
MonaLog + all rewrite help	89.91	74.23	81.66[†]
Hybrid: MonaLog + BERT	83.09	85.46	85.38
Hybrid: MonaLog + BERT	85.65	87.33	85.95[†]
ML/DL-based systems			
BERT (base, uncased)	86.81	85.37	86.74
BERT (base, uncased)	84.62	84.27	85.00 [†]
Yin & Schütze'17	–	–	87.1
Beltagy et al '16	–	–	85.1
Other logic-based systems			
Bjerva et al '14	93.6	60.6	81.6
Abzianidze'17	97.95	58.11	81.35
Martínez-Gomez et al '17	97.04	63.64	83.13
Yanaka et al '18	84.2	77.3	84.3

[†] = running on a corrected version of the SICK dataset.

[‡] = P / R for MonaLog averaged across three labels.

This is a demo of Lasha Abziniadze's system.

Two talented undergraduate students

Zeming Chen and Qiyue Gao

built a polarity-tagging system using **universal dependency** parses rather than CCG parses.

- + More people can use it.
- + It performs better than ccg2Mono, partly because the parses are better, and partly because ccg2Mono misses some arrows, such as on attitude verbs: **John refused to dance**[↓].
- This form of grammar doesn't have a connection to formal semantics, so one can't really prove soundness results the way we can with ccg2Mono.

Two talented undergraduate students

Zeming Chen and Qiyue Gao

built a polarity-tagging system using **universal dependency** parses rather than CCG parses.

It's called **Udep2mono** and you can get it at <https://github.com/eric11eca/Udep2Mono>.

Udep2Mono won a best paper award at IWCS 2021.

Some hard sentences

More dogs than cats sit.

Less than 5 people ran.

A dog who ate two rotten biscuits was sick for three days.

Every dog who likes most cats was chased by at least two of them.

Even if you are addicted to cigarettes you can smoke two a day.

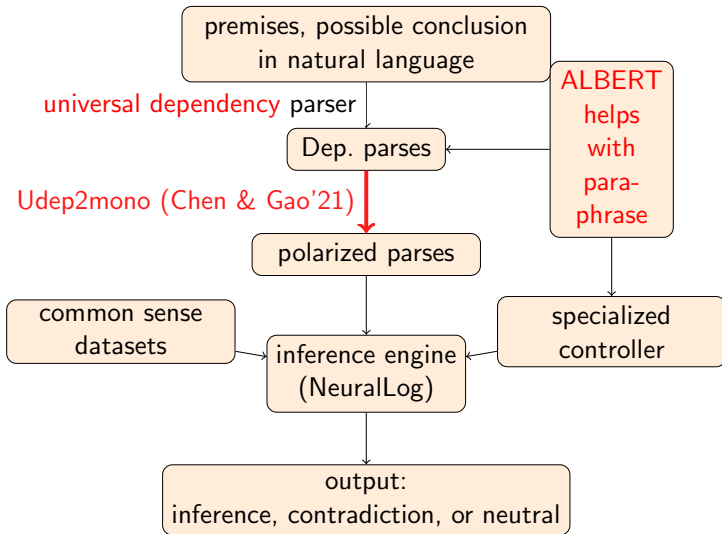
Performance, roughly

NatLog (=MacCartney & Manning 2009) 70%

ccg2Mono 80%

Udep2Mono 97%

A hybrid system: NeuralLog



Model	P	R	acc.
ML/DL-based systems			
BERT (base, uncased)	86.8	85.4	86.7
?	–	–	87.1
?	–	–	85.1
Logic-based systems			
?	98.0	58.1	81.4
?	97.0	63.6	83.1
?	84.2	77.3	84.3
?	83.8	70.7	77.2
?+BERT	83.2	85.5	85.4
?	94.3	67.9	84.4
New System: NeuroLog (Chen, Gao, LM 2021)			
NeuralLog (full system)	88.0	87.6	90.3
– ALBERT-SV	68.9	79.3	71.4
– Monotonicity	74.5	75.1	74.7

What about asking an LLM to explain its reasoning?

entail, contradict or neutral?

Premise: *We compute some simple examples and explore the elementary properties of these invariants.*

Hypothesis: *We explore the elementary properties of these invariants and prove that no other invariants exist.*

Answers

True Label: neutral

Machine's answer: contradiction

Machine answer when you ask for reasons: contradiction

The LLM reasoning in this pair: The premise indicates that simple examples are computed and elementary properties are explored, but it does not make any statement about the exclusivity or completion of these invariants, as in proving no other invariants exist. The hypothesis, by claiming no other invariants exist, introduces information that directly contradicts the premise, which does not mention any such completeness or exclusivity. Therefore, the relationship between the premise and the hypothesis is one of contradiction.

What about asking an LLM to explain its reasoning?

entail, contradict or neutral?

Premise: *We compute some simple examples and explore the elementary properties of these invariants.*

Hypothesis: *We explore the elementary properties of these invariants and prove that no other invariants exist.*

Answers

True Label: neutral

Machine's answer: contradiction

Machine answer when you ask for reasons: contradiction

Give up on strict logicity and on all grammar and semantics, but use data to the exclusion of everything else

This is what contemporary machine-learning-based approaches do.

What I advocate, what I am trying to do

I am trying to put **inference in language**
at the center of the enterprise,
not as a secondary matter.

The study of language involves so many disciplines,
it seems strange to only draw on one!

- ▶ linguistics
- ▶ logic
- ▶ artificial intelligence/cognitive science
- ▶ mathematics
- ▶ philosophy
- ▶ rhetoric

Logic does have a place in the contemporary world of language processing.

It won't be the main event, but it will be there somehow.

We shouldn't be afraid to devise new approaches to many things in logic;
in fact, this is the only way to make progress on a number of hard issues.

Other things we have done and continue to do

- ▶ We extended monotonicity from vanilla CG to CCG.
- ▶ We have an order-enriched version of the typed lambda calculus.
- ▶ We have a running system that can polarize input sentences. Much more should be done on this.
- ▶ We built a MonaLog and now NeuralLog that can solve a large NLI dataset.
Again, we could go back to this.
- ▶ We can generate high-quality sentence pairs, helpful to a ML model.
- ▶ We have hybridized logic and machine learning.
- ▶ We have re-annotated SICK by hand and have a deep study of NLI annotation.

Dissertations on this general topic (inference)

Surely I have left off some important ones

Victor Sánchez Valencia 1991

Rafaella Bernardi 2002

Bill MacCartney 2009

Ka Fat Chow 2012

Lasha Abzianidze 2016

William Tune 2016

Hitomi Yanaka 2018

Katerina Kalouli 2021

Hai Hu 2021

⋮

Tharindu Madusanka 2025

What is happening in this general area?

Check out the upcoming workshop

Natural Logic Meets Machine Learning