Quantificational paradoxes and methodology

(joint w Zach Weber)

Ellie Ripley (they/she)

Monash University
https://negation.rocks

The plan

Fjellstad's threat

Bad news?

Can the robots be our friends?

Fjellstad’s threat

Good old Curry paradox begins from a sentence κ that is in some way equivalent to $T\langle\kappa\rangle \Rrightarrow \bot$

Here we assume a noncontractive solution to this,
one where $T\langle\kappa\rangle \Rrightarrow T\langle\kappa\rangle \Rrightarrow \bot$ is provable
but $T\langle\kappa\rangle \Rrightarrow \bot$ (and so κ itself) is not.

- $(B \Rrightarrow C) \Rrightarrow (A \Rrightarrow B) \Rrightarrow A \Rrightarrow C$
- $(A \Rrightarrow B \Rrightarrow C) \Rrightarrow B \Rrightarrow A \Rrightarrow C$
- $A \Rrightarrow B \Rrightarrow A$
- $A \Rrightarrow B \Rrightarrow A \otimes B$
- $(A \otimes B \Rrightarrow C) \Rrightarrow A \Rrightarrow B \Rrightarrow C$
- $A \Rrightarrow A \vee B \qquad B \Rrightarrow A \vee B$
- $(A \Rrightarrow C) \Rrightarrow (B \Rrightarrow C) \Rrightarrow (A \vee B \Rrightarrow C)$
- $\bot \Rrightarrow A$
- $A \Rrightarrow B, A \therefore B$

- $\forall x A \Rrightarrow A$
- $\forall x(Z \Rrightarrow B) \Rrightarrow Z \Rrightarrow \forall x B$ (x not free in Z)
- $\forall x(Z \vee B) \Rrightarrow Z \vee \forall x B$ (x not free in Z)
- $\forall x(A \otimes B) \Rrightarrow \forall x A \otimes \forall x B$
- $x = x$
- $x = y \Rrightarrow A(x) \Rrightarrow A(y)$

- $A \therefore \forall x A$

- $(B \Rrightarrow C) \Rrightarrow (A \Rrightarrow B) \Rrightarrow A \Rrightarrow C$
- $(A \Rrightarrow B \Rrightarrow C) \Rrightarrow B \Rrightarrow A \Rrightarrow C$
- $A \Rrightarrow B \Rrightarrow A$
- $A \Rrightarrow B \Rrightarrow A \otimes B$
- $(A \otimes B \Rrightarrow C) \Rrightarrow A \Rrightarrow B \Rrightarrow C$
- $A \Rrightarrow A \vee B \qquad B \Rrightarrow A \vee B$
- $(A \Rrightarrow C) \Rrightarrow (B \Rrightarrow C) \Rrightarrow (A \vee B \Rrightarrow C)$
- $\bot \Rrightarrow A$
- $A \Rrightarrow B, A \therefore B$

- $\forall x A \Rrightarrow A$
- $\forall x(Z \Rrightarrow B) \Rrightarrow Z \Rrightarrow \forall x B$ (x not free in Z)
- $\forall x(Z \vee B) \Rrightarrow Z \vee \forall x B$ (x not free in Z)
- $\forall x(A \otimes B) \Rrightarrow \forall x A \otimes \forall x B$
- $x = x$
- $x = y \Rrightarrow A(x) \Rrightarrow A(y)$

- $A \therefore \forall x A$

In "A multiplicative ingredient for ω-inconsistency" (2025), Fjellstad works a paradox against certain noncontractive theories of arithmetic and truth.

Fjellstad's threat works by iterating conjunction.

This is multiplicative/intensional/fusion:
A is very much not the same as $A \otimes A$,
so iterating matters!

Let A^1 be A, and A^{n+1} be $A \otimes A^n$.

I won't go through the threat here in detail, but:
it uses a sentence C equivalent to $(\forall n C^n) \Rrightarrow \bot$,
and it requires the use of both $\forall x(A \otimes B) \Rrightarrow \forall x A \otimes \forall x B$
and arithmetic induction, in the form
$A(0) \otimes \forall x(A(x) \Rrightarrow A(Sx)) \Rrightarrow \forall x A(x)$

Now, we like this logic,
we would like to do arithmetic in it (with induction),
and we would like to have a naive truth predicate.

So something seemingly has to give.

Here's an idea: the precise form of induction in play is at fault.

Induction is a principle about numbers,
but the formulation Fjellstad uses is about everything.

We should instead be able to express numberhood
and restrict induction to numbers.

Let $\mathbb{N}$ be the 'is a number' predicate,
and let $\mathbb{\forall} x A(x)$ be $\forall x(\mathbb{N}x \Rrightarrow A(x))$.

Then the induction we expect is not

$$A(0) \otimes \forall x(A(x) \Rrightarrow A(Sx)) \Rrightarrow \forall x A(x)$$

but instead

$$A(0) \otimes \mathbb{\forall} x(A(x) \Rrightarrow A(Sx)) \Rrightarrow \mathbb{\forall} x A(x)$$

(Other arithmetic axioms should also be restricted to numbers.)

Good news! (pt 1)

Restricted induction is enough to prove that numberhood contracts:
$\mathbb{N}x \Rrightarrow \mathbb{N}x \otimes \mathbb{N}x$ is a theorem.

This makes basic arithmetic still work,
even with the explicit restriction to numbers.

Good news! (pt 2)

Even though $\forall$ distributes over $\otimes$,
there is no reason to expect $\mathbb{И}$ to do the same.

Recall that $\mathbb{И}xA$ is $\forall x(\mathbb{N}x \Rrightarrow A)$

But we do not have $(A \Rrightarrow B \otimes C) \Rrightarrow (A \Rrightarrow B) \otimes (A \Rrightarrow C)$;
the $\forall$ distributes but the $\Rrightarrow$ does not.

So Fjellstad's argument is blocked.

Bad news: the robots are coming for us

Key methodological point

We are working without a net: no nontriviality proofs.
They're hard to find and fragile once you have them;
focusing on nontriviality stifles exploration!

Rather, we work to show what is provable,
both desired and undesired.

Fall into triviality? Whoops! Tweak and try again.

Simonelli fed all this to an LLM,
and found an issue with even this restricted arithmetic.

They showed that $A \Rrightarrow A \otimes A$ is provable;
adding naive truth or sets would then bring back Curry paradox.

The proof works with the predicate $F(x)$ defined as:

$$x = 0 \vee (x = 1 \otimes A) \vee (A \otimes A)$$

It shows by restricted induction

that $A \Rrightarrow \mathbb{?}xF(x)$

This uses $\vee$L, but does not distribute $\forall$ over $\otimes$

Instantiating at 2, we get $A \Rrightarrow A \otimes A$

A tentative diagnosis:
it's induction's fault again.

Even restricted induction conceals contraction:

when we conclude $\forall x A(x)$ by induction,
and then instantiate it at a number n,
we are implicitly using the inductive step n times,
although induction only registers one use.

But now the choices cut deeper.

One possibility is to move to rule form induction,
which would block this issue
and still support at least some mathematical reasoning:

$$A(0), \mathbb{N}x(A(x) \Rrightarrow A(Sx)) \therefore \mathbb{N}xA(x)$$

But this would break the deduction theorem,
messes with at least some of the spirit of axiomatic theorising,
and we're not sure how much reasoning it would really support.

But then while we're trying to figure out how to handle that, in comes another problem from Hanson and an LLM.

This one doesn't use induction at all!

In the full setting, there is a negation $\neg$,
not definable by $\Rrightarrow \bot$.

Their proof uses just LEM from the logic: $A \vee \neg A$,
together with a predicate B such that both
$(\forall x B(x)) \Rrightarrow \bot$ and $(\forall x \neg B(x)) \Rrightarrow \bot$
are provable.

The logic alone doesn't contain such predicates,
but the naive set theory does: let $B(x)$ be $x \in x$

1	$A \otimes B(x) \Rrightarrow (A \vee B(x)) \otimes (A \vee \neg B(x))$	provable
2	$A \otimes \neg B(x) \Rrightarrow (A \vee B(x)) \otimes (A \vee \neg B(x))$	provable
3	$A \otimes (B(x) \vee \neg B(x)) \Rrightarrow (A \vee B(x)) \otimes (A \vee \neg B(x))$	1, 2
4	$A \Rrightarrow (A \vee B(x)) \otimes (A \vee \neg B(x))$	3, LEM
5	$A \Rrightarrow \forall x((A \vee B(x)) \otimes (A \vee \neg B(x)))$	4, quantifiers
6	$A \Rrightarrow \forall x(A \vee B(x)) \otimes \forall x(A \vee \neg B(x))$	5, $\forall/\otimes$
7	$A \Rrightarrow (A \vee \forall x B(x)) \otimes (A \vee \forall x \neg B(x))$	5, $\forall/\vee$
8	$A \Rrightarrow (A \vee \bot) \otimes (A \vee \bot)$	7, assumption
9	$A \Rrightarrow A \otimes A$	8, $\vee$L

This uses more logical resources:
all of $\vee$L and $\forall/\otimes$ and $\forall/\vee$.

But it avoids induction entirely:
answering Simonelli with rule-form induction won't touch this.

	Fjellstad	Simonelli	Hanson
$\forall/\otimes$	Y	N	Y
$\forall/\vee$	N	N	Y
$\forall/\Rightarrow$	N	Y	Y
$\vee$L	N	Y	Y
induction	Y	$Y^{\mathbb{N}}$	N
LEM	N	N	Y

A heterogenous mess?

Can the robots be our friends?

Right now, LLMs seem to be able to find
novel proofs of triviality in naive systems.

They do not, however, do much at all to shed any light
on what they find.

This leaves us still scrambling to understand the last problem when the next one arrives.

As long as these LLMs remain widely available, this is unlikely to stop.

Key analogy: exploits in software systems.

Also, the LLMs output tons of crap, too!

And it is not fun to dig through their outputs trying to figure out if there's anything there.

I’m optimistic here about the role of proof assistants.

LLMs are clever and untrustworthy;
proof assistants are hard to work with but trustworthy.

Make the LLMs pilot the proof assistants.

Then we don't have to dig through LLM slop
looking for verifiable gems;
the LLM gets itself verified.

This is already afoot in maths,
and it can spread to nonclassical maths too.

One possible future, then, is one where
LLMs pilot proof assistants.

We could ask such a system “what if I tried this axiom”?
and get verifiable answers like “that would trivialise your system”.

This is kind of already what we do with each other via email!

(A shift in research power from social capital to capital capital?)

Conclusion

- The robots are coming for us.
- This can help us detect that trivial theories are trivial! (yaaay)
- But it does not help us understand or fix anything yet. (uh-oh)
- Time to get serious about proof assistants!