Proofs that, proofs why, and the analysis of paradoxes

To Gerhard, on your 60th birthday

Peter Schroeder-Heister

Universität Tübingen

Russell's antinomy in naive set theory

Extend natural deduction with the following introduction and elimination rule for set membership:

$$\frac{A(t)}{t \in \{x : A(x)\}} \qquad \frac{t \in \{x : A(x)\}}{A(t)}$$

Then we can derive a contradiction $\bot$ as follows.

Let R stand for $\{x : \neg(x \in x)\}$

$$\dfrac{\dfrac{\dfrac{\dfrac{[R \in R]^{(1)}}{\neg(R \in R)} \quad [R \in R]^{(1)}}{\bot}}{\neg(R \in R)}\,{\scriptstyle (1)} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[R \in R]^{(1)}}{\neg(R \in R)} \quad [R \in R]^{(1)}}{\bot}}{\neg(R \in R)}\,{\scriptstyle (1)}}{R \in R}}{}}{\bot}$$

Simplified neutral form

Inference rules for a defined atom or nullary logical constant:

$$\frac{R \to \bot}{R} \qquad\qquad \frac{R}{R \to \bot}$$

or as a clausal (impredicative) definition:

$$\left\{ \; R := R \to \bot \right.$$

with appropriate closure and reflection principles

Derivation of absurdity

$$
\dfrac{
\dfrac{
\dfrac{
\dfrac{[R]^{(1)}}{R \to \bot} \quad [R]^{(1)}
}{\bot}
}{R \to \bot}\,(1)
\qquad
\dfrac{
\dfrac{
\dfrac{
\dfrac{[R]^{(1)}}{R \to \bot} \quad [R]^{(1)}
}{\bot}
}{R \to \bot}\,(1)
}{R}
}{\bot}
$$

Self-contradiction in the sequent calculus

Right- and Left-Introduction rules:

$$\frac{\Gamma \vdash R \to \bot}{\Gamma \vdash R} \qquad\qquad \frac{\Gamma, R \to \bot \vdash C}{\Gamma, R \vdash C}$$

Derivation of absurdity:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{R \vdash R \quad \bot \vdash \bot}{R, R \to \bot \vdash \bot}}{R, R \vdash \bot}}{R \vdash \bot}}{\vdash R \to \bot}}{\vdash R} \qquad \dfrac{\dfrac{\dfrac{R \vdash R \quad \bot \vdash \bot}{R, R \to \bot \vdash \bot}}{R, R \vdash \bot}}{R \vdash \bot}}{\vdash \bot}$$

Structural rules: Three critical places

Right- and Left-Introduction rules:

$$\frac{\Gamma \vdash R \rightarrow \bot}{\Gamma \vdash R} \qquad\qquad \frac{\Gamma, R \rightarrow \bot \vdash C}{\Gamma, R \vdash C}$$

Derivation of absurdity:

$$\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{R \vdash R \quad \bot \vdash \bot}{R, R \rightarrow \bot \vdash \bot}}{R, R \vdash \bot}}{R \vdash \bot}}{\vdash R \rightarrow \bot}}{\vdash R} \qquad \cfrac{\cfrac{\cfrac{R \vdash R \quad \bot \vdash \bot}{R, R \rightarrow \bot \vdash \bot}}{R, R \vdash \bot}}{R \vdash \bot}}{\vdash \bot}$$

The problematic case of identity

The philosophical discussion centers around contraction and cut. Identity is not normally considered a problem.

In logic programming it has been seen as a problem.

It provides a link to earlier work by Gerhard Jäger (together with Robert Stärk).

The significance of logic programming

Though not very popular any more among computer scientists, it is still an outstanding foundational paradigm:

- It sheds new light on inductive definitions
- No well-foundedness requirements
- Slogan: Definitional freedom

Being well-defined does not imply being well-behaved.

Identity and initial sequents

Reminder: In standard sequent calculi initial sequents can be assumed to be atomic.

$$\dfrac{A\wedge B\vdash A\wedge B}{\vdots}$$

can be reduced to

$$\dfrac{\dfrac{\dfrac{A\vdash A \qquad B\vdash B}{A,B\vdash A\wedge B}}{A\wedge B\vdash A\wedge B}}{\vdots}$$

Philosophical analysis: Apply meaning rules whenever they are available

Application to paradoxes

$$\frac{\Gamma \vdash R \to \bot}{\Gamma \vdash R} \qquad\qquad \frac{\Gamma, R \to \bot \vdash C}{\Gamma, R \vdash C}$$

$$\frac{R \vdash R}{\vdots}$$

can be reduced to

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{R \vdash R \qquad \bot \vdash \bot}{R, R \to \bot \vdash \bot}}{R \to \bot \vdash R \to \bot}}{R \to \bot \vdash R}}{R \vdash R}}{\vdots}$$

which can be reduced to

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{R \vdash R \qquad \bot \vdash \bot}{R, R \to \bot \vdash \bot}}{R \to \bot \vdash R \to \bot}}{R \to \bot \vdash R}}{R \vdash R} \qquad \bot \vdash \bot}{R, R \to \bot \vdash \bot}}{R \to \bot \vdash R \to \bot}}{R \to \bot \vdash R}}{R \vdash R}}{\vdots}$$

Philosophical significance

- Definitions are not necessarily well-founded
- Identity must be well-founded

I.e., we require good behaviour on the derivation side, but not on the definition side.

Semantically, this can be handled by an appropriate three-valued logic (Jäger, Stärk).

Result: Contraction and cut are admissible in such a system.

Summary

Unspecific initial sequents $A \vdash A$ only serve for the case where A has no *specific* meaning.

An initial sequent $A \vdash A$ is only allowed if no *specific* way of introducing A is available. *Kreuger's restriction*

This corresponds to the requirement that initial sequents be atomic. We restrict unspecific assumptions to the irreducible case.

Restricting identity is a very plausible way of dealing with the paradoxes.

Application to natural deduction

Restricting identity means that derivations must be ‘co-normal’ in the sense that ‘minimal formulas’ are only allowed in the atomic case:

$$\begin{array}{cl} \vdots & \\ \hline {\color{red}A} & \text{E rule} \\ \hline \vdots & \text{I rule} \end{array}$$

is not permitted, if introduction and elimination rules for *A* are available. The derivation must be expanded:

$$\begin{array}{cl} \vdots & \\ \hline {\color{red}A} & \text{E rule} \\ \hline \vdots & A\ \ \text{E} \\ \hline {\color{red}A} & A\ \ \text{I} \\ \hline \vdots & \text{I rule} \end{array}$$

No minimal shortcuts!

Advantage

The restriction on identity is purely local and can be easily checked.

Sequent calculus: If there are defining rules for A, you must not use identity for A.

Natural deduction: If there are defining rules for A, you must not use A as a minimal formula.

If we want to enforce identity, we need to restrict contraction and/or cut, which becomes way more complicated.

Structural rules: Three critical places

Right- and Left-Introduction rules:

$$\frac{\Gamma \vdash R \rightarrow \bot}{\Gamma \vdash R} \qquad \frac{\Gamma, R \rightarrow \bot \vdash C}{\Gamma, R \vdash C}$$

Derivation of absurdity:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{R \vdash R \quad \bot \vdash \bot}{R, R \rightarrow \bot \vdash \bot}}{R, R \vdash \bot}}{R \vdash \bot}}{\vdash R \rightarrow \bot}}{\vdash R} \quad \dfrac{\dfrac{\dfrac{R \vdash R \quad \bot \vdash \bot}{R, R \rightarrow \bot \vdash \bot}}{R, R \vdash \bot}}{R \vdash \bot}$$

$$\vdash \bot$$

Restrict contraction rather than identity

Disallowing contraction blocks the paradoxes (Fitch, Curry).

However, this goes too far!!

No proper mathematics without contraction.

Way out: Disallow a specific form of contraction, namely that of specific (evaluated) and unspecific (unevaluated) propositions.

Specific vs. unspecific assumptions

Unspecific assumptions: Result from $A \vdash A$

Specific assumptions: Result from meaning steps (left-introduction rules)

As they are semantically different, we may require that there be no specific / unspecific overlap.

Paradoxes and critical contraction

$$
\dfrac{
\dfrac{\dfrac{\dfrac{\dfrac{R \vdash R \qquad \bot \vdash \bot}{R, R \to \bot \vdash \bot}}{R, R \vdash \bot}}{R \vdash \bot}}{\vdash R \to \bot}
\over \vdash R
\qquad
\dfrac{\dfrac{\dfrac{R \vdash R \qquad \bot \vdash \bot}{R, R \to \bot \vdash \bot}}{R, R \vdash \bot}}{R \vdash \bot}
}{\vdash \bot}
$$

Red: unspecific Blue: specific

Specific vs. unspecific assumptions

Formulas are indexed depending of whether they are specific or unspecific.

We disallow contraction in cases where this is well motivated, i.e. where there is a semantical difference between formulas of the same shape.

Technically involved: We assign a *meaning index* to every formula in a proof. This index goes up when a formula is introduced by a meaning rule (L- or R-rule).

In effect: Stratification with respect to meaning rules.

Summary

The identification of an evaluated with an unevaluated formula is a characteristic feature of the paradoxes.

Prohibiting the identification of specific (evaluated) with unspecific (unevaluated) propositions blocks the paradoxes.

Result: Cut is admissible for impredicative definitions, if contraction is restricted.

Problem with restricted contraction

Problem: The restriction on contraction is neither local nor easy to check.

It can be made local by labelling formulas at the object-linguistic level:

$$\frac{\Gamma, A^m, A^n \vdash C}{\Gamma, A^n \vdash C} \quad \text{provided } m = n$$

Alternative: Enforce contraction and restrict cut instead.

Structural rules: Three critical places

Right- and Left-Introduction rules:

$$\frac{\Gamma \vdash R \to \bot}{\Gamma \vdash R} \qquad\qquad \frac{\Gamma, R \to \bot \vdash C}{\Gamma, R \vdash C}$$

Derivation of absurdity:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{R \vdash R \quad \bot \vdash \bot}{R, R \to \bot \vdash \bot}}{R, R \vdash \bot}}{R \vdash \bot}}{\vdash R \to \bot}}{\vdash R} \qquad \dfrac{\dfrac{\dfrac{R \vdash R \quad \bot \vdash \bot}{R, R \to \bot \vdash \bot}}{R, R \vdash \bot}}{R \vdash \bot}}{\vdash \bot}$$

Restricting cut

Cut is a structural rule that comes in addition to the semantical rules.

In principle, we can give up cut.

Cut is something whose admissibility needs to be demonstrated, not something that should be forced to hold.

Cut is a (mathematical) fact

not a principle

Normally, we can show the admissibility of cut

However, in the situation, in which a proposition R is defined by the rules

$$\frac{\Gamma \vdash R \to \bot}{\Gamma \vdash R} \qquad \frac{\Gamma, R \to \bot \vdash C}{\Gamma, R \vdash C}$$

cut is not admissible

Analogy: Recursive functions

Consider partial recursive functions or Turing machines.
They not necessarily terminate.

Being total corresponds to the admissibility of cut.

In the example we see from the definition that the partial recursive function is not defined everywhere.

In general this problem is not decidable (halting problem).

Summary

Whether cut holds or not, is accidental — depends on the situation considered.

In the case of the paradoxes cut is simply not admissible.

We might just work in a cut-free framework.

A more sophisticated way out

Is there a certain restriction on the application of cut (a proviso), such that, when the proviso is satisfied, we have cut elimination?

Even though we cannot decide, whether we have admissibility of cut or not: At least a plausible condition, under which cut can be shown to hold?

Contradiction and absurdity with terms

$$\frac{t : R \to \bot}{rt : R} \qquad \frac{t : R}{r't : R \to \bot} \qquad r'rt \rhd t$$

gives non-normalizable terms:

$$\dfrac{\dfrac{\dfrac{\dfrac{[x : R]^{(1)}}{r'x : R \to \bot} \quad [x : R]^{(1)}}{r'xx : \bot}}{\lambda x.r'xx : R \to \bot}{\scriptstyle (1)} \qquad \dfrac{\dfrac{\dfrac{\dfrac{[x : R]^{(1)}}{r'x : R \to \bot} \quad [x : R]^{(1)}}{r'xx : \bot}}{\lambda x.r'xx : R \to \bot}{\scriptstyle (1)}}{r\lambda x.r'xx : R}}{(\lambda x.r'xx)r\lambda x.r'xx : \bot}$$

$r'(r\lambda x.r'xx)(r\lambda x.r'xx) \rhd (\lambda x.r'xx)(r\lambda x.r'xx) \rhd$
$r'(r\lambda x.r'xx)(r\lambda x.r'xx)$

Way out: Side condition on modus ponens

$$\frac{s : A \to B \quad t : A}{st : B}\ st!$$

st! means: *st* is normalizable

$$\dfrac{\dfrac{\dfrac{\dfrac{[x : R]^{(1)}}{r'x : R \to \bot} \quad [x : R]^{(1)}}{r'xx : \bot}}{\lambda x.r'xx : R \to \bot}\,(1) \qquad \dfrac{\dfrac{\dfrac{\dfrac{[x : R]^{(1)}}{r'x : R \to \bot} \quad [x : R]^{(1)}}{r'xx : \bot}}{\lambda x.r'xx : R \to \bot}\,(1)}{r\lambda x.r'xx : R}}{(\lambda x.r'xx)r\lambda x.r'xx : \bot}\ (\lambda x.r'xx)r\lambda x.r'xx\ !$$

$\boxed{(\lambda x.r'xx)r\lambda x.r'xx\ !}$ is not satisfied.

Formal representation of contradiction with terms

$$\frac{\Gamma \vdash t : R \to \bot}{\Gamma \vdash rt : R} \qquad \frac{\Gamma, x : R \to \bot \vdash t : C}{\Gamma, y : R \vdash t[x/r'y] : C} \qquad r'rt \rhd t$$

Note that this is not a Dyckhoff-style representation, which would instead be

$$\frac{\Gamma, x : R \to \bot \vdash t : C}{\Gamma, y : R \vdash F(y, x.t) : C}$$

for some selector F, whose natural deduction translation would be:

$$\phi(F(y, x.t) = t[x/r'y])$$

So we are using natural deduction terms in the style of Barendregt and Ghilezan.

Reason: Terms should represent knowledge and not just codify proofs.

Derivation of absurdity

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{x : R \vdash x : R}{x : R, y : R \rightarrow \bot \vdash yx : \bot}}{x : R, z : R \vdash r' zx : \bot}}{x : R \vdash r' xx : \bot}}{\vdash \lambda x.r' xx : R \rightarrow \bot}}{\vdash r\lambda x.r' xx : R} \qquad \dfrac{\dfrac{\dfrac{x : R \vdash x : R}{x : R, y : R \rightarrow \bot \vdash yx : \bot}}{x : R, z : R \vdash r' zx : \bot}}{x : R \vdash r' xx : \bot}}{\vdash r'(r\lambda x.r' xx)(r\lambda x.r' xx) : \bot}$$

$$r'(r\lambda x.r' xx)(r\lambda x.r' xx) \rhd (\lambda x.r' xx)(r\lambda x.r' xx) \rhd$$
$$r'(r\lambda x.r' xx)(r\lambda x.r' xx)$$

Termination proviso in the sequent calculus

$$\frac{\Gamma \vdash t : A \quad x : A, \Delta \vdash s : C}{\Gamma, \Delta \vdash s[x/t] : C}\ s[x/t]\,!$$

“!”: “normalizes”

From the Dyckhoff-translation follows:
$s[x/t]\,!$ implies that this cut is admissible.

Restricted modus ponens vs. restricted cut

Restricted modus ponens:

$$\frac{s : A \to B \quad t : A}{st : B}\ st\,!$$

Restricted cut:

$$\frac{\Delta \vdash t : A \quad \Delta, x : A \vdash s : C}{\Delta \vdash s[x/t] : C}\ s[x/t]\,!$$

The side condition on cut is local.

This is yet another argument in favour of the sequent calculus as the appropriate reasoning format.

Under assumptions

Restricted modus ponens:

$$\dfrac{s : A \to B \qquad \begin{array}{c} y : D \\ \vdots \\ t : A \end{array}}{st : B}\; st\,!$$

Restricted cut:

$$\dfrac{\Delta \vdash t : A \qquad \Delta, y : D, x : A \vdash s : C}{\Delta, y : D \vdash s[x/t] : C}\; s[x/t]\,!$$

Performing substitution

In natural deduction by combining proofs:

$$\dfrac{s : A \rightarrow B \qquad \begin{array}{c} \vdots \\ t' : D \\ \vdots \\ t[y/t'] : A \end{array}}{s(t[y/t']) : B}\ s(t[y/t'])\,!$$

In the sequent calculus by an additional cut:

$$\dfrac{\Delta \vdash t' : D \qquad \dfrac{\Delta, y : D \vdash t : A \qquad \Delta, x : A \vdash s : C}{\Delta, y : D \vdash s[x/t] : C}\ s[x/t]\,!}{\Delta \vdash s[x/t[y/t']] : C}\ s[x/t[y/t']]\,!$$

No re-check of side conditions!

Free type theory

We may consider turning the side condition in

$$\frac{\Delta \vdash t : A \quad \Delta, x : A \vdash s : C}{\Delta \vdash s[x/t] : C} \quad s[x/t]\,!$$

into an actual premiss:

$$\frac{\Delta \vdash t : A \qquad \Delta, x : A \vdash s : C \qquad s[x/t]\,!}{\Delta \vdash s[x/t] : C}$$

Pro: Gain expressive power

Contra: The formal and ontological framework of type theory has to be re-worked

This is not against the spirit of type theory: Formation rules for terms rather than only for types.

Proofs that and proofs why

Suppose a proof of A is given: $\begin{array}{c}\mathcal{D}\\ A\end{array}$

This is a proof that A

It also tells us why A :
By inspecting $\mathcal{D}$ we know why A

However, even though $\mathcal{D}$ tells us why, this story is not an outcome of the proof

The result of our inspection is not what is being proved

Result: A proof of A in the usual sense is a proof that, not a proof why.

Two perspectives

The truth-theoretic perspective:

We are interested in what is true and take a proof as something that shows us that its end-formula is true

The proof-theoretic perspective:

We are interested in how truth is established and reflect on the proof (its form, its structure) as a possible argument telling why its end-formula is true

The first perspective is direct ('intentio recta'), the other one indirect, or by reflection ('intentio obliqua')

Combining the two perspectives: The Curry-Howard correspondence

Suppose we have a proof

$$\begin{array}{c} \mathcal{D} \\ t : A \end{array}$$

Then t allows us to reconstruct $\mathcal{D}$.

Standard example:

$$ {\scriptstyle (1)}\ \dfrac{\dfrac{[x : A]^{(1)}}{\lambda y.x : B \to A}}{\lambda x.\lambda y.x : A \to (B \to A)} $$

$\lambda x.\lambda y.x$ codifies the proof.

We only need to look at the conclusion, i.e., we can stay in *intentio recta*.

The type-theoretic perspective

$$\begin{array}{c}\mathcal{D}\\ t : A\end{array}$$

t: 'why' A: 'that'

A proof delivers two objects: A 'proposition' whose truth is established, and a 'proof object' that incorporates the reason why the proposition is true (its 'ground').

Conclusion for the ‘proofs that’ vs. ‘proofs why’-debate, given the analysis of paradoxes

Proofs that and proofs why are intertwined.

There are ways in which the construction of proofs that depend on previous proofs why.

Not only in the sense in which what is proved depends on proofs why (dependent types).

But in the sense that the construction of proofs not only depends on what has been proved, but on how it has been proved, and on how the intended proof might look.