Proof Theory of Martin-Löf Type Theory

Anton Setzer

Swansea University, Swansea UK

Oxford, 15 October 2012

From Hilbert to Gentzen

- Hilbert's program.
  - Proof consistency of mathematical theories by finitary methods.
- Doesn't work because of Gödel's Incompleteness theorem.
- Gentzen: Reduction of consistency of PA to well-foundedness of ordinal notation systems.

Gentzen's Argument (Modern Version)

- Essentially
  - interpret proofs of PA as infinitary proofs in a semi formal system, e.g. induction interpreted as

$$\dfrac{A(0) \qquad \dfrac{A(0) \qquad A(0)\rightarrow A(1)}{A(1)} \qquad \dfrac{\dfrac{A(0) \qquad A(0)\rightarrow A(1)}{A(1)} \qquad A(1)\rightarrow A(2)}{A(2)} \qquad \cdots}{\forall x.A(x)}$$

Gentzen's Argument (Modern Version)

- If you omit the cut rule

$$\frac{\Gamma, A \qquad \Gamma, \neg A}{\Gamma}\ (\mathrm{Cut})$$

  the calculus is just a truth definition, which can only prove true formulas.
- Proof that cuts can be eliminated using induciton over the height of trees.
- Height can be measured by ordinals, therefore induction over trees can be replaced by transfinite induction over an ordinal notation system of strength ϵ_0.
- So

$$\mathrm{PRA} + \mathrm{TI}(\epsilon_0) \vdash \mathrm{Cons}(PA)$$

Wellfoundedness of Ordinal Notation System

- In case of ϵ_0 some insight into the well-foundedness of the ordinal notation system can be achieved.
- Works as well for theories up to strength $(\Pi^1_1 - \mathrm{CA})_0$.
  - Sufficient by reverse mathematics for proving most mathematical theorems.
  - Argument carried out in articles on "Ordinal systems" by the author.
- Beyond $(\Pi^1_1 - \mathrm{CA})_0$ it becomes difficult to get a direct insight into the well-foundedness of the ordinal notation system.

Wellfoundedness of the Ordinal Notation System

- Need for a second theory in which we can prove the well-foundedness of the ordinal notation system.
  - That theory will then (if it contains PRA) prove the consistency of the theory in question.
    - In fact usually more follows, at least validity of Π^0_2-sentences.
  - Requires a theory for which we have an insight that everything proved in it is valid.
  - Most successful (but not necessarily only) approach: constructive theories.
  - Candidates could be
    - Frege structures,
    - Feferman's systems of explicit mathematics
    - Martin-Löf Type Theory.

Wellfoundedness of the Ordinal Notation System

- Most effort has been taken to develop Martin-Löf Type Theory for that purpose.
  - Argument for validity of its judgements: meaning explanations.
    - If argument formalised mathematically we prove the consistency of the theory in question.
    - Requires by Gödel's Second Incompleteness theorem more strength than the theory itself.
    - Therefore consistency argument needs to be necessarily philosophical in nature.
    - Each consistency argument neeeds to rely on a philosophical argument (even if it is not reflected properly).
  - According to Martin-Löf his type theory is the most serious attempt to create a theory where we have an insight into the validity of its judgements.

Two Step Approach

So we arrive at a two step approach:

- First step:
  - Prove the consistency of a theory T in PRA extended by transfinite induction up to over an ordinal notation system OT.
- Second Step:
  - Proof well-foundedness (transfinite induction) of OT in an extension of Martin-Löf's Type Theory ML^+.
- Since PRA can be embedded into ML^+, we obtain that

$$\mathrm{ML}^+ \vdash \mathrm{Cons}(T)$$

- We obtain even usually more namely that all Π^0_2-statements of T are provable in ML^+.

Need for Proof Theoretically Extensions of Martin-Löf Type Theory

- Needed: development of strong extensions of Martin-Löf Type Theory with an insight into the validity of what can be shown into it.
- Applications:
  - Discovery of advanced data structures for use in programming.

Some Type Theoretic Notations

- We have judgements:

$$a : A \qquad\qquad A : \mathrm{Set}$$

- The latter expresses that A is a small type (= Set)
- We have the dependent function type:

$$(x : A) \to B$$

  - Elements are functions f mapping $a : A$ to $f\ a : B[x := a]$.
  - Example Matrix multiplication:

$$\mathrm{matmult} : (n, m, k : \mathbb{N}) \to \mathbb{R}^{n,m} \to \mathbb{R}^{m,k} \to \mathbb{R}^{n,k}$$

Version of MLTT of Strength PA

- MLTT (which includes $\mathbb{N}$) and a small universe we obtain the strength of PA (ϵ_0).
  - We call this microscopic universe Atom and the theory $\mathrm{MLTT} + \mathrm{Atom}$.
  - Only problematic principle: $\mathbb{N}$.
  - Trust in the validity requires an insight into the understanding the concept of a least set introduced by finitary introduction rules.
  - Insight closely related to the concept of time.

Principles for Adding Strength

- Two principles added in order to increase the strength of type theory:
  - The W-type.
  - Universes.

The W-Type

- **Formation rule:**

$$\frac{A : \mathrm{Set} \qquad B : A \to \mathrm{Set}}{\mathrm{W}(A, B) : \mathrm{Set}}$$

- **Introduction rule:**

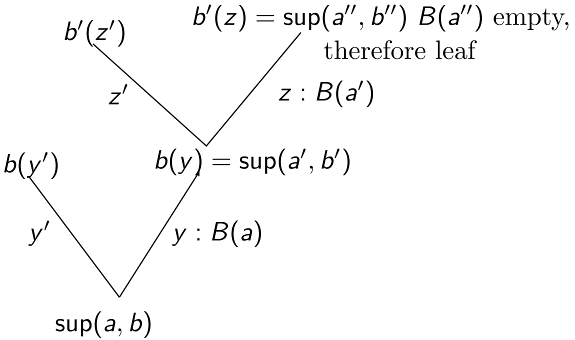
$$\frac{a : A \qquad b : B(a) \to \mathrm{W}(A, B)}{\mathsf{sup}(a, b) : \mathrm{W}(A, B)}$$

- **Elimination/equality rule:**

  Induction over trees.

W-Type

$\mathrm{W}(A, B)$ is the type of well-founded recursive trees with branching degrees $(B(a))_{a:A}$.



Understanding of the W-Type

- Understanding the meaning of the W-type requires understanding the existence of a type containing exactly those elements introduced by the introduction rules.
- Leastness more difficult than elements of the W-type not introduced in finitely many steps.
  - We can only draw an analogy from the concept of time.

Strength of the W-Type

- $\mathrm{MLTT} + \mathrm{Atom} + \mathrm{W}$ has strength of finitely iterated inductive definitions or $(\Pi^1_1 - \mathrm{CA})_0$.
  - In proof theory considered as truely impredicative theory.
  - Requires substantially more complex techniques and ordinal notation systems in order to understand them.
    - As mentioned before some direct insight into the well-foundedness can still be obtained.
- By the above mendioned results of reverse mathematics sufficient for proving most mathematical theorems.
- So most mathematics actually used is secured, if we trust in validity of judgements in $\mathrm{MLTT} + \mathrm{Atom} + \mathrm{W}$.

Universes

- A universe is a family of sets
- Given by
  - a set $\mathrm{U} : \mathrm{Set}$ of **codes** for sets,
  - a **decoding function** $\mathrm{T} : \mathrm{U} \to \mathrm{Set}$.

Universes

- **Formation rules:**

$$\mathrm{U} : \mathrm{Set} \qquad \mathrm{T} : \mathrm{U} \to \mathrm{Set}$$

- **Introduction and Equality rules:**

$$\widehat{\mathbb{N}} : \mathrm{U} \qquad \mathrm{T}(\widehat{\mathbb{N}}) = \mathbb{N}$$

$$\frac{a : \mathrm{U} \qquad b : \mathrm{T}(a) \to \mathrm{U}}{\widehat{\Sigma}(a, b) : \mathrm{U}}$$

$$\mathrm{T}(\widehat{\Sigma}(a, b)) = \Sigma(\mathrm{T}(a), \mathrm{T} \circ b)$$

  Similarly for other type formers (except for U).

- **Elimination/equality rules:** Induction over U.
  (Not needed in order to obtain their strength).

Understanding of Universes

- Insight into validity of rules for universes is not much more complex than insight into validity of rules for W-type.
- $\mathrm{MLTT} + \mathrm{W} + \mathrm{U}$ has strength KPI^{+} (AS) which is a slight extension of KPI which is Kripke Platek set theory plus one recursively inaccessible ordinal.
  - KPI corresponds roughly speaking to an inductive definition which refers to closure under another inductive definition.
- Proof theoretically there are some complications but not dramatic ones.

Understanding of Universes

- In the ordinal notation system, something happens when making the step from $(\Pi^1_1 - \mathrm{CA})_0$ to $(\Pi^1_1 - \mathrm{CA})$.
  - Strength of KPI is $\Delta^1_2 - \mathrm{CA} + \mathrm{BI}$ which is much stronger than $(\Pi^1_1 - \mathrm{CA})$.

Generalisation to Inductive-Recursive Definitions

- Inductive-Recursive Definitions originally defined by Dybjer, closed formalisation by Dybjer + AS.
- Definition of a type theory containing all standard (inductive) definitions, universes, and many generalisations.
- Generalise the principles.

Induction-Recursion

- We have one set $\mathrm{U} : \mathrm{Set}$ with constructors:

$$\mathrm{C} : \underbrace{(a : A)}_{\text{non-inductive argument}}$$

$$\rightarrow \quad \underbrace{(b : B\ a \rightarrow \mathrm{U})}_{\text{inductive argument depending on } a}$$

$$\rightarrow \quad \underbrace{(c : (x : D\ a) \times \mathrm{T}\ (b\ (f\ x)))}_{\text{non-inductive arguments depending on } a \text{ and } \mathrm{T} \circ b}$$

$$\rightarrow \quad \cdots$$

$$\rightarrow \quad \mathrm{U}$$

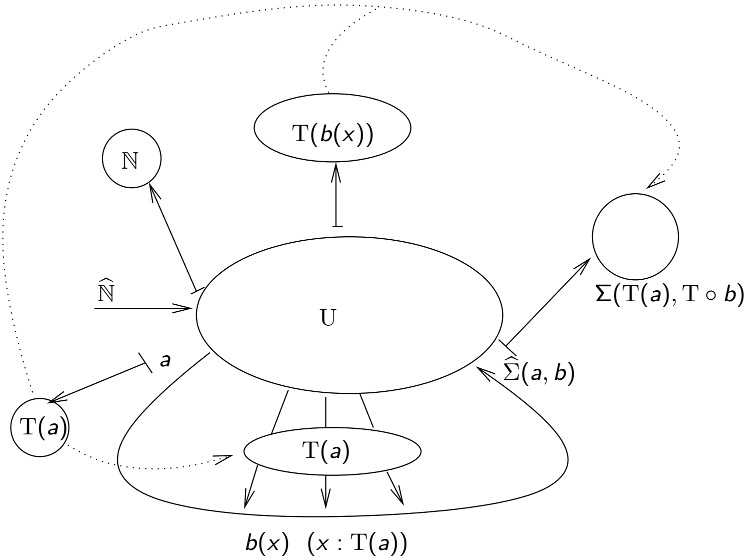
Induction-Recursion

- We have $\mathrm{T} : \mathrm{U} \to \mathrm{Set}$ with recursive equations for each constructor:

$$\mathrm{T}\ (\mathrm{C}\ a\ b\ c\ \cdots) = t[a, \mathrm{T} \circ b, c, \ldots] : \mathrm{Set}$$

- Generalisation to $\mathrm{T}\ u : D$ for some type D.
- If $D = \{*\}$, we obtain the special case of inductive definitions.

Universe


T(b(x))
ℕ
$\widehat{\mathbb{N}}$
U
$\Sigma(\mathrm{T}(a), \mathrm{T} \circ b)$
a
$\widehat{\Sigma}(a, b)$
T(a)
T(a)
$b(x) \quad (x : \mathrm{T}(a))$

Strength of Induction-Recursion

- Proof theoretic strength in $[\mathrm{KPM}, \mathrm{KPM}^+]$.
- $\mathrm{KPM} =$ Kripke-Platek set theory plus one recursively Mahlo ordinal.
- $\mathrm{KPM}^+ =$ slight extension of KPM.
- (Upper bound not formally proved yet).

Steps Towards Mahlo

- First step beyond standard universe
  - The super universe (Palmgren).
  - He introduced a universe V,
  - together with a universe operator $\mathrm{U} : \mathrm{Fam}(\mathrm{V}) \to \mathrm{V}$,
    - $\mathrm{Fam}(\mathrm{V})$ is the set of families of sets in V indexed over elements of V, roughly speaking

$$\{(B_x)_{x:B} | B : \mathrm{V}, \quad x : B \Rightarrow B_x : \mathrm{V}\}$$

  - s.t. for any family of sets A in V, $\mathrm{U}(A)$ is a universe containing all elements of A.

Steps Towards Mahlo

- A Universe is a family of sets closed under constructions for forming sets.
- We can now form a universe, closed under the formation of the next universe above a family of sets.
- (The next slide doesn't exhaust the strength, it shows only universes containing one set, not universes containing family of sets)

Illustration of the Super Universe

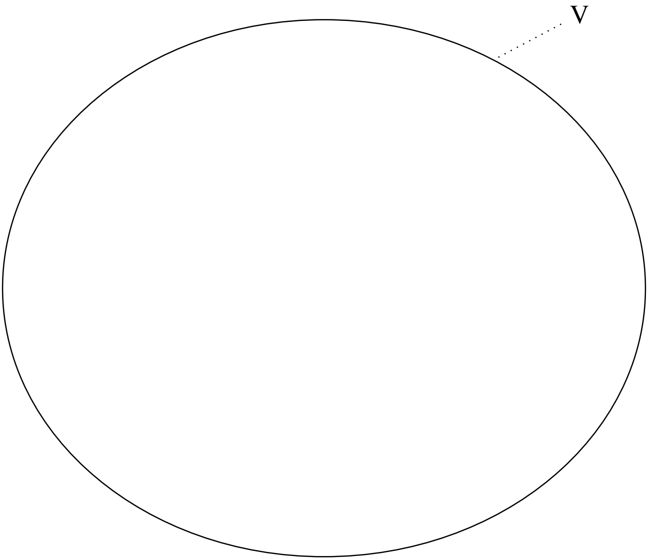


Illustration of the Super Universe

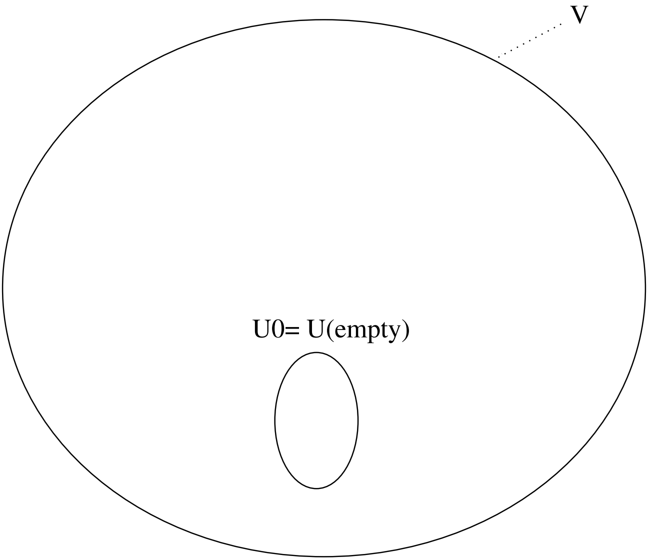


Illustration of the Super Universe

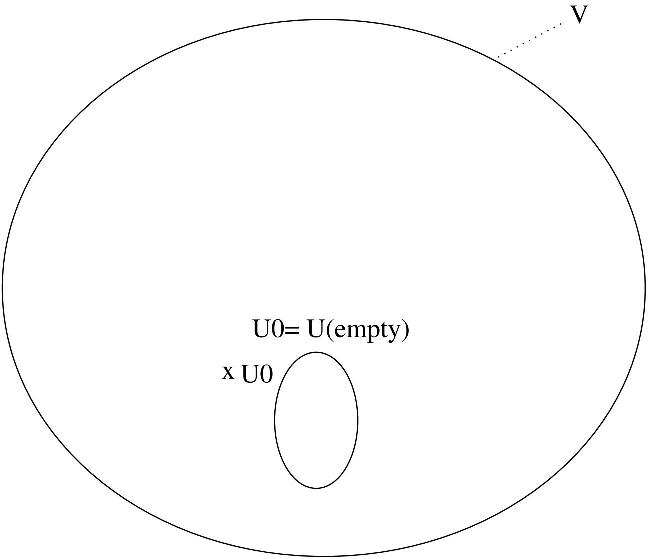


Illustration of the Super Universe

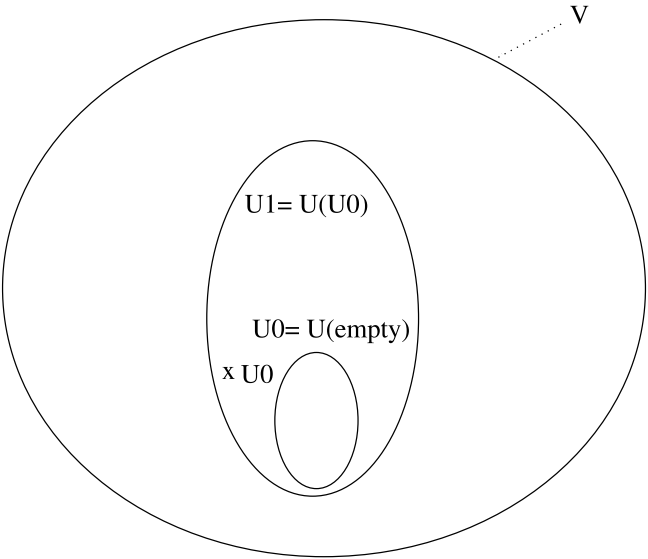


Illustration of the Super Universe

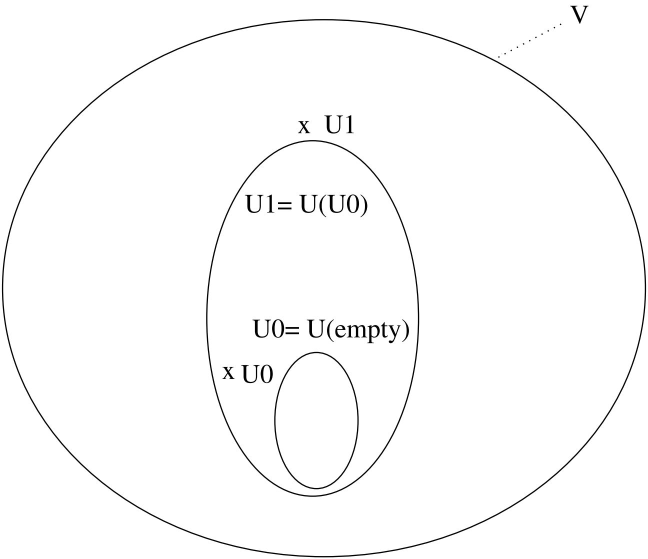


Illustration of the Super Universe

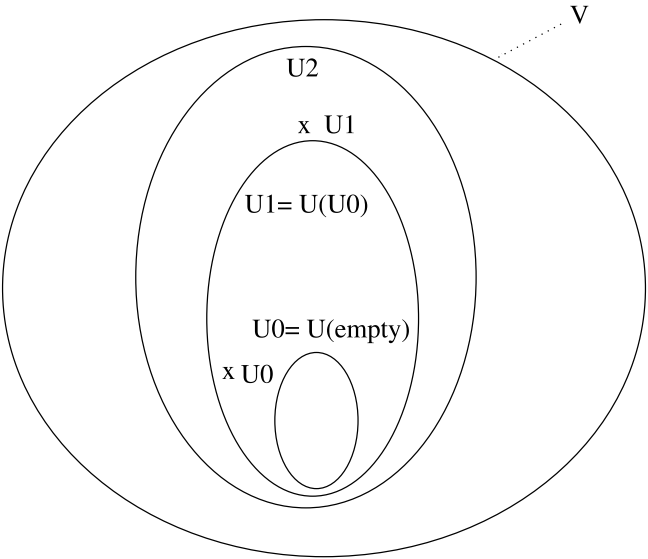
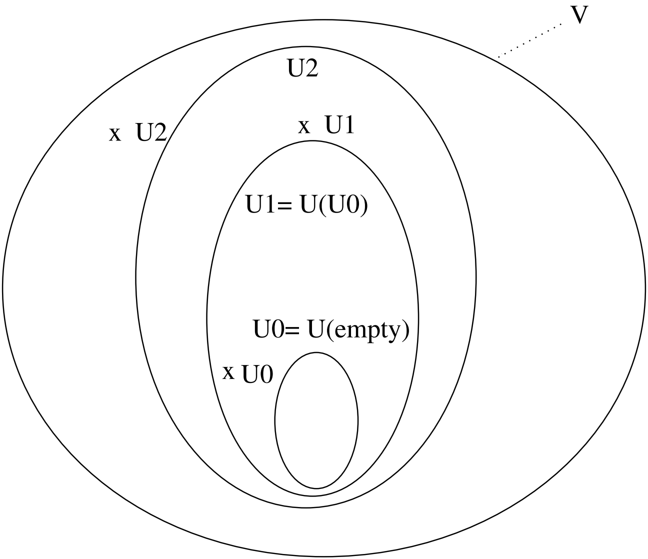


Illustration of the Super Universe



Super^n-Universes

- The above can be continued: We can form a
  - super^2-universe V,
  - closed under a super-universe operator, forming a super universe above a family of sets in V.
- And we can iterate the above n-many times, and even go beyond.
- Up to now everything was inductive-recursive

Mahlo Universe

- The Mahlo universe is
  - a universe $\underset{\sim}{\mathrm{V}}$,
  - which has not only subuniverses corresponding to some operators, but subuniverses corresponding to all operators it is closed under:
  - for every universe operator on V,
    - i.e. every $f : \mathrm{Fam}(\mathrm{V}) \to \mathrm{Fam}(\mathrm{V})$,
  - there exists a universe $\underset{\sim}{\mathrm{U}_f}$ closed under f.
  - which is represented in V.

Illustration of the Mahlo Universe

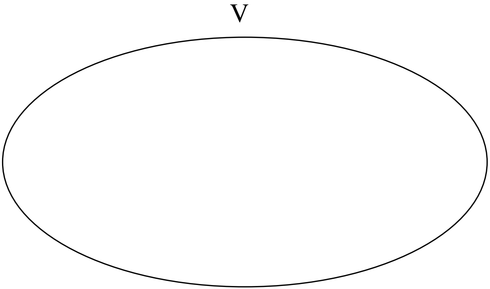


Illustration of the Mahlo Universe

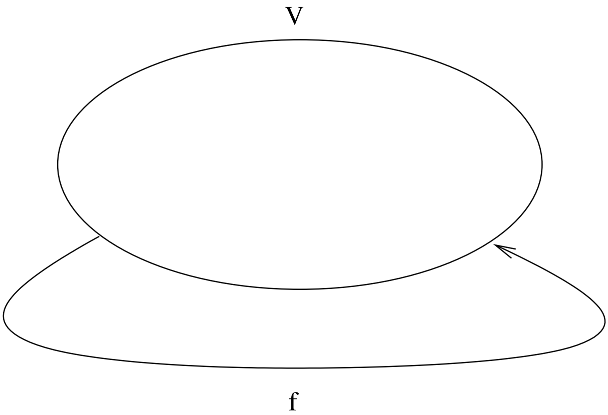


Illustration of the Mahlo Universe

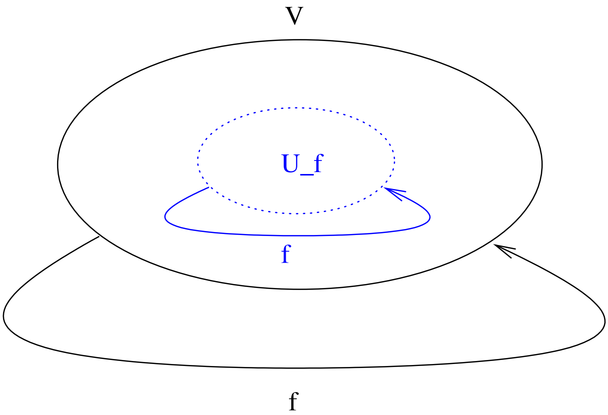
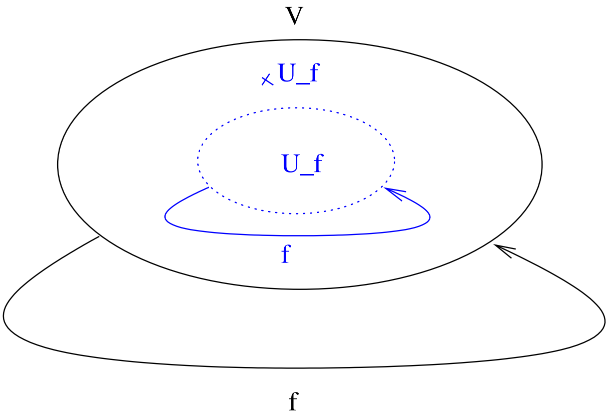


Illustration of the Mahlo Universe



Impredicativity of the Mahlo Universe

- The introduction rule for V introduces for every

$$f : \mathrm{Fam}(\mathrm{V}) \to \mathrm{Fam}(\mathrm{V})$$

  an element

$$\widehat{\mathrm{U}}_f : \mathrm{V}$$

- Depends on the totality of V.

Red Mahlo Universe

- If we only have Set as a Mahlo universe, then we obtain a construction which is an example of inductive-recursive definitions.
- Strenght of IRD primarily based on the fact that we can define a universe U_f closed under a function

$$f : \mathrm{Fam}(\mathrm{Set}) \to \mathrm{Fam}(\mathrm{Set})$$

- So we have

$$f : \mathrm{Fam}(\mathrm{Set}) \to \mathrm{Fam}(\mathrm{Set}) \Rightarrow \mathrm{U}_f : \mathrm{Set}$$

Red Mahlo Universe

- Therefore Set is essentially a Mahlo universe, called "the red Mahlo universe".
- Highly depends on the use of the logical framework.
- Strength of the Red Mahlo universe is expected to be KPM.

Red Mahlo Universe

- The well-ordering proof requires as well that we have large elimination for W-type and $\mathbb{N}$ (elimination into an arbitrary type rather than a set).
  - Can be omitted by forming suitable indexed higher type universes which are sufficiently closed.
- The red Mahlo universe is usually accepted, whereas the black Mahlo universe is often rejected.
- However, the red Mahlo universe is essentially as impredicative as the black Mahlo universe.
- It is accepted because one considers Set as an "open concept".
- However introduction rule for Set depends on the the set of total functions

$$f : \mathrm{Fam}(\mathrm{Set}) \to \mathrm{Fam}(\mathrm{Set})$$

Problems of Mahlo Universe

- This section is joint work with R. Kahle.
- Elements of V are constructed, depending on total functions

$$f : \mathrm{Fam}(\mathrm{V}) \to \mathrm{Fam}(\mathrm{V})$$

- Therefore V has an impredicative definition.
- However, for defining U_f, only the restriction of f to $\mathrm{Fam}(\mathrm{U}_f)$ is required to be total.
- In order to define functions for which this restrictions is total, we need to define candidates of U_f for arbitrary (not necessarily total) f.
- Problem: In Martin-Löf Type Theory all functions are total.
- In Feferman's explicit mathematics reference to arbitrary terms possible.

Extended Predicative Mahlo (in Explicit. Mathematics)

- We will use syntax borrowed from type theory in Explicit Mathematics.
  - but $a \in B$ instead of $a : B$.
- Explicit mathematics more Russell-style, therefore we can have $\mathrm{V} \in \mathrm{Set}$, $\mathrm{V} \subset \mathrm{Set}$.
- We can encode $\mathrm{Fam}(\mathrm{V})$ into V, therefore need only to consider functions $f : \mathrm{V} \to \mathrm{V}$.

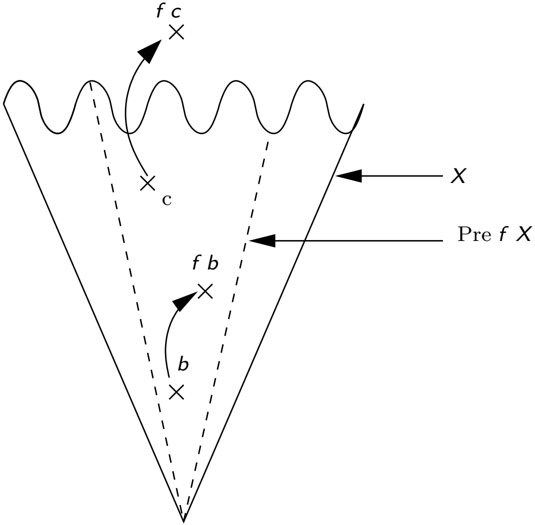
Step 1: Define $\mathrm{Pre}\ f\ X$

- Define V to be closed under universe constructions for explicit mathematics.
- Define for $f, X \in \mathrm{Set}$, $X \subseteq \mathrm{Set}$

$$\mathrm{Pre}\ f\ X \in \mathrm{Set} \qquad \mathrm{Pre}\ f\ X \subseteq X$$

- $\mathrm{Pre}\ f\ X$ is the least subset of X closed under universe constructions and f relative to X.
- So, if
  - result of applying a universe operator to $\mathrm{Pre}\ f\ X$ is in X, then add it to $\mathrm{Pre}\ f\ X$.
  - result of applying f to an elemnet of $\mathrm{Pre}\ f\ X$ is in X, then add it to X.

Pre f X

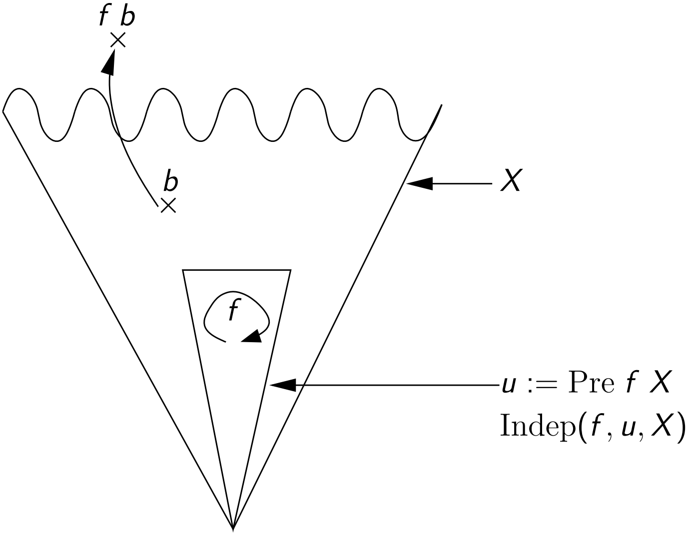


Step 2: Independence of Pre f X

- If, whenever a universe construction or f is applied to elements of $\mathrm{Pre}\ f\ X$ we get elements in X, then $\mathrm{Pre}\ f\ X$ is independent of future extensions of X.

$$\begin{aligned}\mathrm{Indep}(f, \mathrm{Pre}\ f\ X, X) := {} & (\forall a \in \mathrm{Pre}\ f\ X.\ \forall b \in a \to \mathrm{Pre}\ f\ X.\ \mathrm{j}\ a\ b \in X) \\ & \wedge\ \cdots \\ & \wedge\ \forall a \in \mathrm{Pre}\ f\ X.\ f\ a \in X\end{aligned}$$

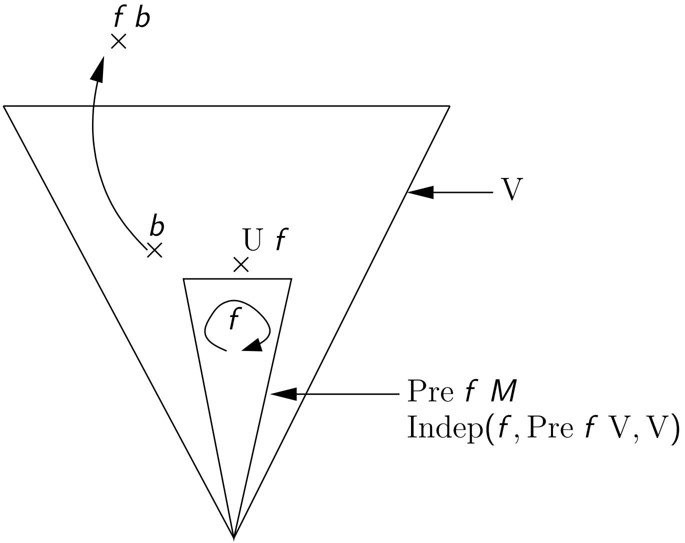
Indpt



Step 3: Introduction Rule for V

- $$\forall f.\,\mathrm{Indep}(f, \mathrm{Pre}\ f\ \mathrm{V}, \mathrm{V}) \to (\mathrm{U}\ f \in \mathrm{Set} \wedge \mathrm{U}\ f =_{\mathrm{ext}} \mathrm{Pre}\ f\ \mathrm{V} \wedge \mathrm{U}\ f \in \mathrm{V}) \quad .$$

Introduction Rule for V



Interpretation of Axiomatic Mahlo

- One can show:

$$\forall f \in \mathrm{V} \to \mathrm{V}.\ \mathrm{Indep}(f, \mathrm{Pre}\ f\ \mathrm{V}, \mathrm{V})$$

therefore

$$\forall f \in \mathrm{V} \to \mathrm{V}.\ \mathrm{U}\ f \in \mathrm{V}\ \wedge\ \mathrm{Univ}(f)\ \wedge\ f \in \mathrm{U}\ f \to \mathrm{U}\ f$$

- So V closed under axiomatic Mahlo constructions.
- Therefore extended predicative Mahlo has at least strength of axiomatic Mahlo.

4. Π_3-Reflection

- First step: Formation of Hyper-Mahlo Universes:
  - A hyper Mahlo universe is a unverse V, T s.t. for every

$$f : \mathrm{Fam}(\mathrm{V}) \to \mathrm{Fam}(\mathrm{V})$$

  there exists a subuniverse of V
    - closed under f,
    - which is Mahlo.

Illustration of the Hyper Mahlo Univ.

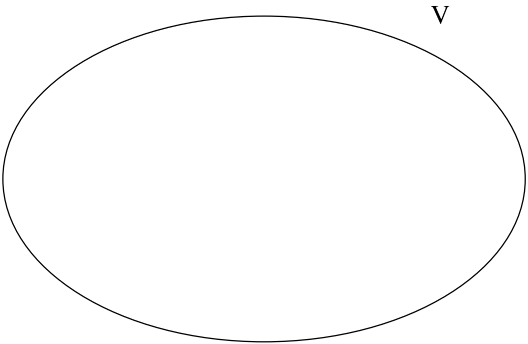


Illustration of the Hyper Mahlo Univ.

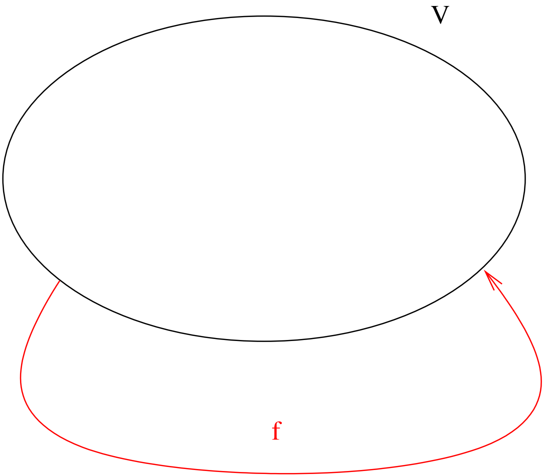


Illustration of the Hyper Mahlo Univ.

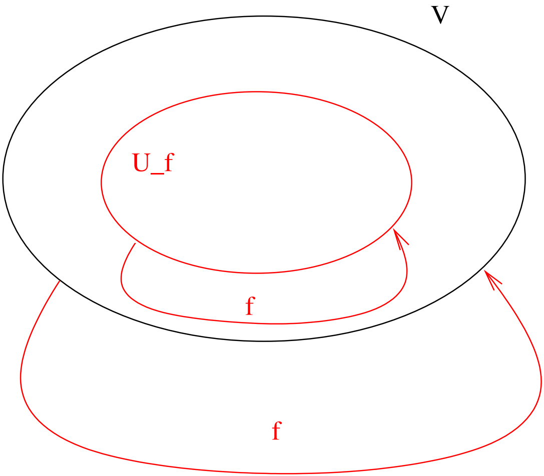


Illustration of the Hyper Mahlo Univ.

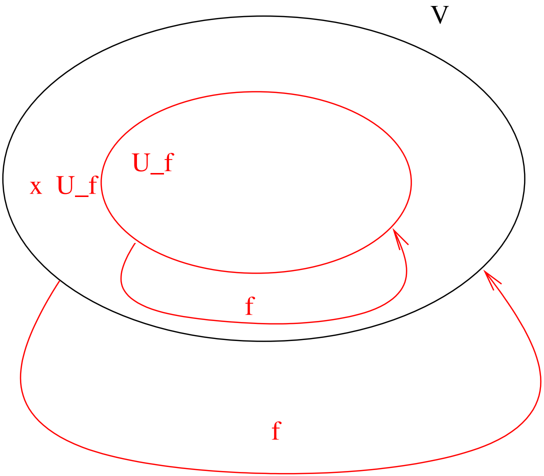


Illustration of the Hyper Mahlo Univ.

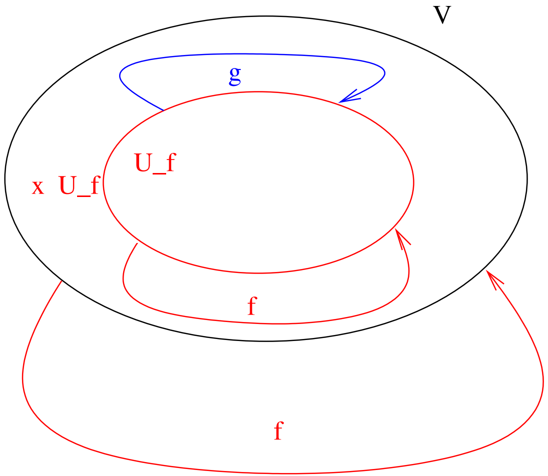


Illustration of the Hyper Mahlo Univ.

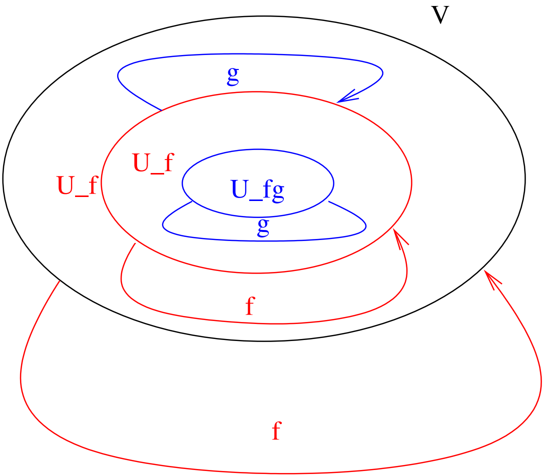
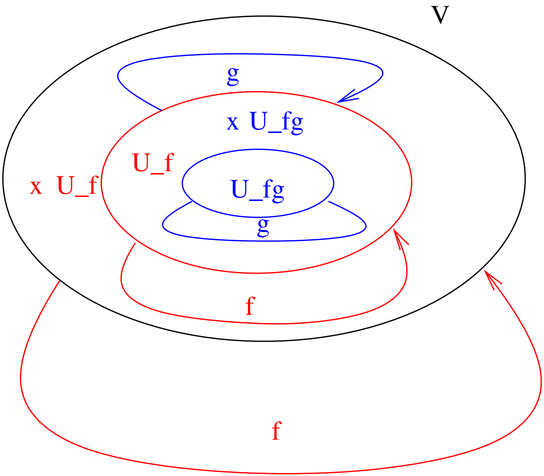


Illustration of the Hyper Mahlo Univ.



Extensions

- Hyper^n-Mahlo Universe: Straightforward Extension.
- Autonomous Mahlo Universe consists of
  - A universe V, S.
  - The set of Mahlo degrees M which is essentially $\mathrm{W}x : \mathrm{V}.\mathrm{S}(x)$.
  - Rules expressing that V is Hyper^w-Mahlo for any $w : \mathrm{M}$.

Π_3-Reflecting Universe

- In case of the Π_3-reflecting universe, we obtain Hyper^d-Mahlo universes for Mahlo degrees d.
  - Mahloness of a hyper^d-Mahlo universe depends locally on the hyper^d-Mahlo universe as well.
- Mahlo degrees are introduced by an introduction rule similar to that of the Mahlo universe:
  - For every

    $$f : \mathrm{Fam}(\mathrm{V}) \to \mathrm{Fam}(\mathrm{V}, \mathrm{MDegree})$$

    we define a new

    $$\mathrm{degree}(f) : \mathrm{MDegree}$$

Π_3-Reflecting Universe

- If Mahlo universe is acceptable, step towards to the Π_3-reflecting universe is essentially a technical difficulty.

Conclusion

- Consistency problem leads naturally to proofs of consistency using transfinite induction.
- Because of lack of direct insight into well-foundedness need to prove well-foundedness of ordinal notation system in constructive theories.
- Up to inductive-reducrsive definition analysis not controversal.
- Mahlo universe is controversal because of impredicative characater.
- Solution by giving an extended predicative Mahlo universe.
- Extension to Π_3-reflecting universe.