Toward a Set Theoretic Foundation Closer to Mathematical Text

Ciarán M. Dunne: http://www.macs.hw.ac.uk/~cmd1/

Supervised by J. B. Wells and Fairouz Kamareddine

June 25, 2020

Foundations of Mathematics

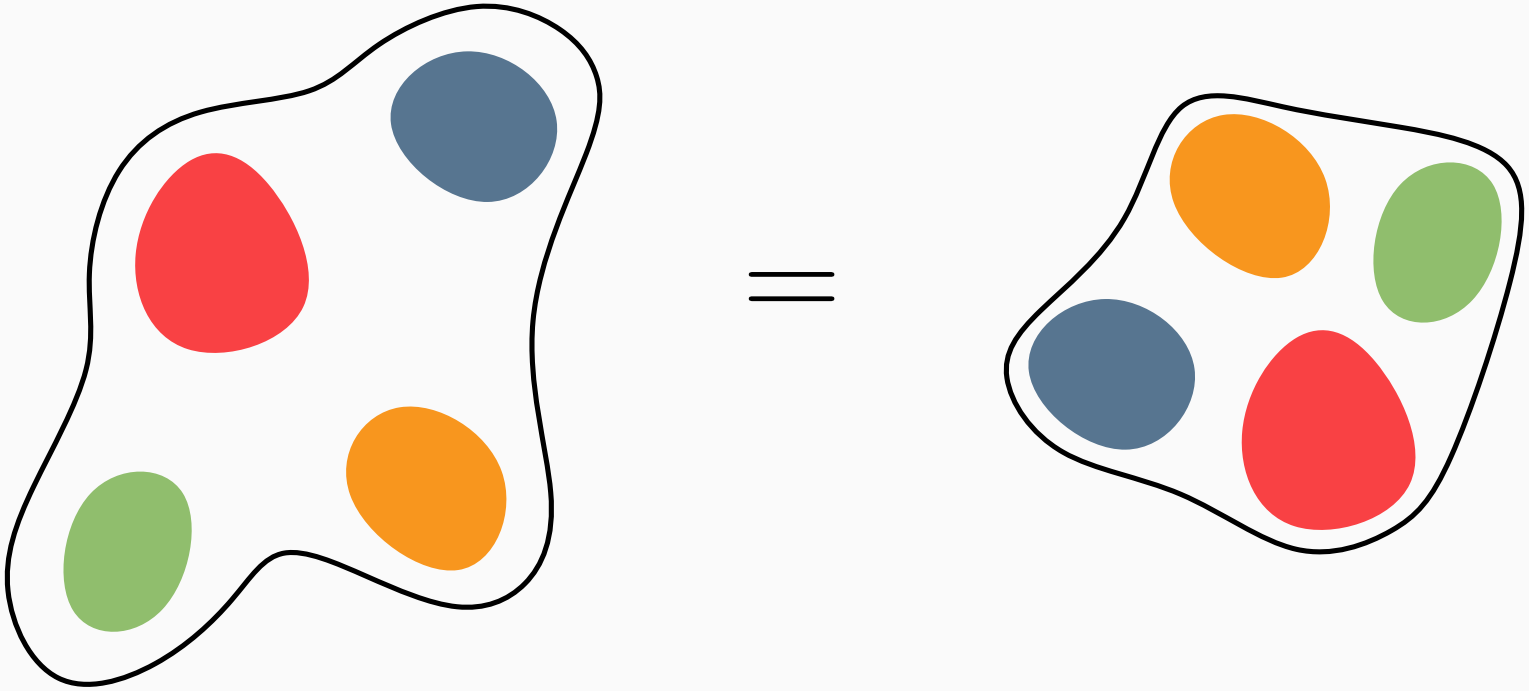
- Mathematical foundations are systems that model the practice of mathematicians using a small number of fundamental concepts.
- They account for mathematical objects such as numbers, vectors, functions, groups and their properties.
- Axiomatic set theory — particularly Zermelo-Frankel (ZF) — with first-order logic (FOL) are broadly accepted by mathematicians.
- Type theories are foundations of mathematics popular in machine assisted theorem proving, and also in programming languages.
- Translation of mathematical text into such systems is called formalisation.

Formalisation of Mathematics

- Foundations are idealisations of mathematical practice, so formalisation requires significant translation.
- Awkward compromises must constantly be made to fit the mathematics to the foundation.
- Current foundations fail to accurately capture the practice of mathematicians.
- We consider the consequences founding a piece of mathematical text in ZF set theory to demonstrate one of these issues, and propose a solution.

Zermelo-Fraenkel Set Theory

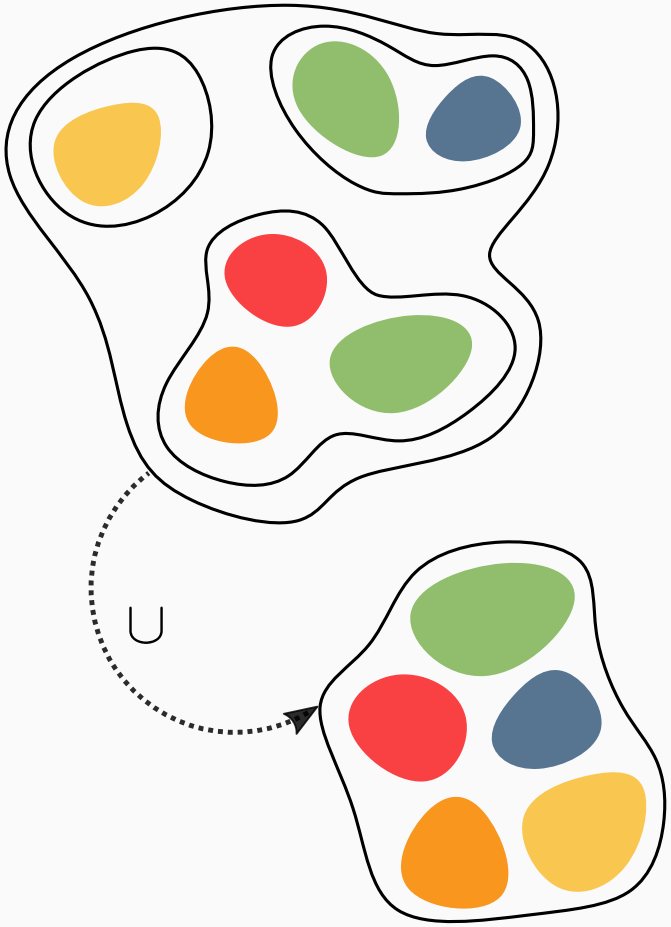
- ZF is a collection of axioms stated in first-order logic.
- The outcome of many decades of work started by Ernst Zermelo (above), and Abraham Fraenkel (below) in the early 20th century.
- Describes a domain of objects called sets, and the behaviour of the set membership relation $\in$.
- Logical deduction from these axioms provides us with a rich theory of sets.
- Sets are used to represent mathematical objects, so that for any mathematical theorem, there is a corresponding theorem about sets provable from ZF.



Axiom of Extensionality

$$\forall\, x, y : (\forall\, a : a \in x \leftrightarrow a \in y) \rightarrow x = y$$

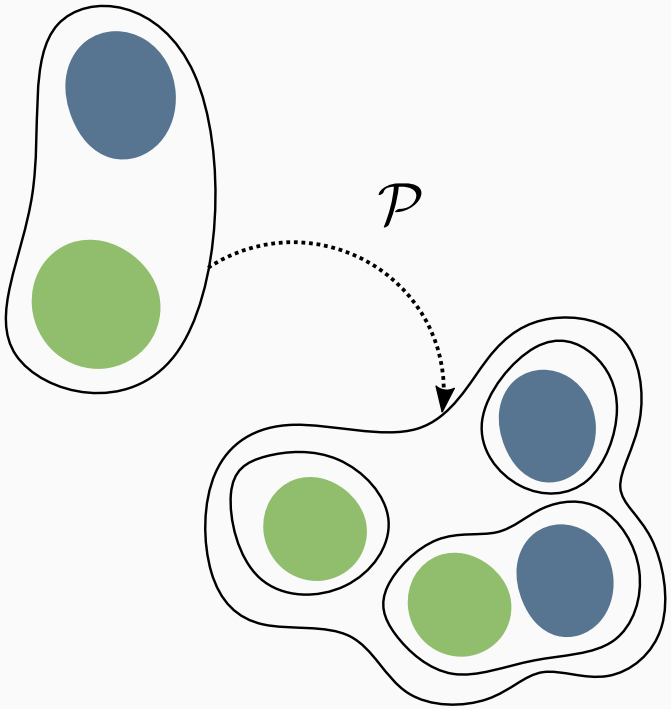
"If two sets contain exactly the same members, then they are equal."



Axiom of Union

$\forall x : \exists y : \forall a : a \in y \leftrightarrow (\exists z \in x : a \in z)$

"The union of a set exists."



Axiom of Power Set

$\forall x : \exists y : \forall z : z \in y \leftrightarrow z \subseteq x$

"The power set of a set exists."

Axiom Schema of Replacement

$$\forall\, c_1, c_2, x : (\forall\, a \in x : \exists!\, b : \varphi) \to (\exists\, y : \forall\, b : b \in y \leftrightarrow \exists\, a \in x : \varphi)$$

"For any formula φ with at most c_1, c_2, a, b free, if φ acts like a function f over x then the image of f over x is a set."

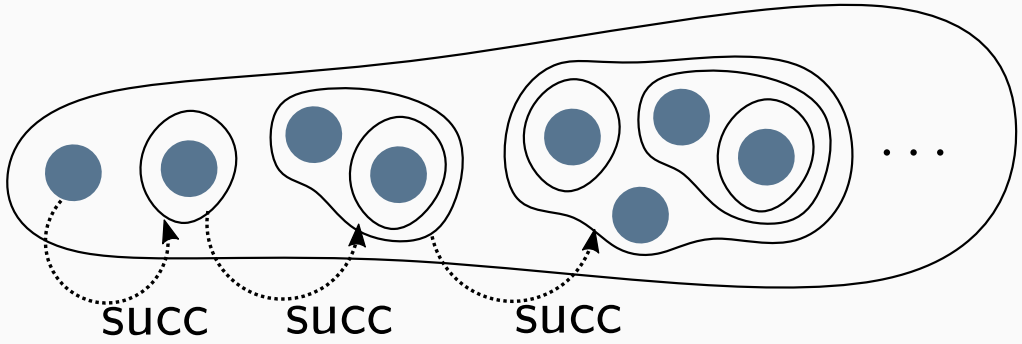
- Allows us to prove the existence of $\emptyset$, $\{A, B\}$, $\{\, x \in A \mid \varphi \,\}$.
- Also needed for results in set theory, measure theory.

Axiom of Foundation

$$\forall\, x : x = \emptyset \vee (\exists\, y \in x : y \cap x = \emptyset)$$

"Every set is well-founded."

- Rules out infinitely descending chains of sets $X_0 \ni X_1 \ni \cdots$.
- No self-containing sets $X \in X$.



Axiom of Infinity

$$\exists y : \emptyset \in y \wedge (\forall x \in y : \mathrm{succ}(x) \in y)$$

"There exists a set containing $\emptyset$, closed under the successor operation."

$$\mathrm{succ}(x) := x \cup \{x\}$$

- The von Neumann natural numbers can be defined as:

$$0 := \emptyset,\ 1 := \{0\},\ 2 := \{0, 1\},\ 3 := \{0, 1, 2\}\ \ldots$$

- Infinity allows us the prove the existence of $\mathbb{N}$ containing all such sets.
- The set membership relation $\in$ acts like the $<$ ordering on $\mathbb{N}$.

Ordered Pairs in Set Theory

- An ordered pair (a, b) is a container that holds two objects.
- Any definition of ordered pairs must satisfy:

Characteristic Property of Ordered Pairs

$$(a, b) = (c, d) \text{ if and only if } a = c \text{ and } b = d.$$

- Projection relations or functions – π_1, π_2 – are used to reason about the contents.

Definition

Widely used set-theoretic definition given by Kuratowski in 1921:

$$\langle a, b \rangle = \{\{a\}, \{a, b\}\}$$

Definition

Tuples can be defined recursively: $\langle a_1, \ldots, a_n \rangle := \langle a_1, \langle a_2, \ldots, a_n \rangle \rangle$

Functions in ZF

- Relations and functions are defined as sets of ordered pairs.
- For relations, we define $xRy := (x, y) \in R$.
- For functions, $f(x)$ is the unique y such that $(x, y) \in f$.
- So there must be only one value y for which $(x, y) \in f$.

Example

$$f := \{ \langle x, x^2 \rangle \mid x \in \mathbb{N} \} = \{\langle 0, 0\rangle, \langle 1, 1\rangle, \langle 2, 4\rangle, \langle 3, 9\rangle, \ldots\}$$

Satisfies the equation $\forall x \in \mathbb{N} : f(x) = x^2$

Functions in ZF

- A user may want to define a function acting on a domain containing sets and ordered pairs. Let $D = (\mathbb{N} \times \mathbb{N}) \cup \mathcal{P}_{\text{fin}}(\mathbb{N})$ and let $g : D \to \mathbb{N}$ be a function such that:

$$g(x) = \begin{cases} n + m & \text{if } x = \langle n, m \rangle \text{ for some } n, m \\ \sum_{i=1}^{k} n_i & \text{if } x = \{n_1, \ldots, n_k\} \end{cases}$$

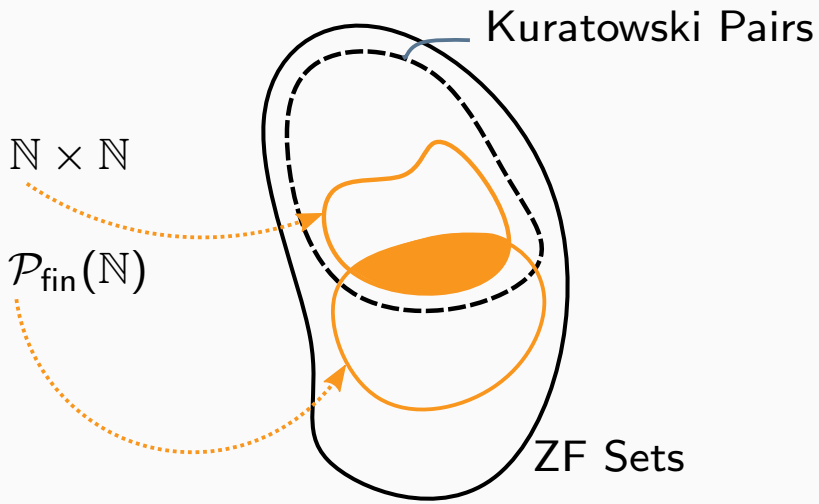
- Does this specification succeed in defining a function?
- What is $g(\langle 0, 1 \rangle)$? $g(\{1, 2\})$?

Functions in ZF

- From the specification, we might expect that $g(\langle 0, 1\rangle) = 1$, and $g(\{1, 2\}) = 3$.
- However, by definition of Kuratowski ordered pairs: $\langle 0, 1\rangle = \{\{0\}, \{0, 1\}\}$.
- By definition of von Neumann natural numbers: $\{1, 2\} = \{\{0\}, \{0, 1\}\}$.
- So by definition of function, there must be only one y such that $\langle \{\{0\}, \{0, 1\}\}, y\rangle \in g$, and hence only one result for $g(\langle 0, 1\rangle)$ and $g(\{1, 2\})$.
- What went wrong?

- The behaviour of the function given by the specification is based on cases of the 'type' of the object.
- We assumed that these cases were mutually exclusive, since $\mathbb{N} \times \mathbb{N}$ and $\mathcal{P}(\mathbb{N})$ are considered to be disjoint, since ordered pairs are not usually thought of as sets.

- If a naive translation is used where the sets of the text are mapped to sets in ZF, it is unavoidable that there is some overlap between sets and representations.
- Reasoning on a mixture of sets and objects thus requires us to ensure the domain of sets is disjoint from the other domains of representations.



Types?

- In type-theoretic foundations, every typeable term x has at least one type τ, written $x : \tau$.

Example

$$42 : \mathsf{Nat} \qquad (2, 5) : \mathsf{Nat} * \mathsf{Nat} \qquad (\cdot)^2 : \mathsf{Nat} \to \mathsf{Nat}$$

- Types are used to organise the operations that may be performed on objects, and characterise the objects that can exist.
- Overloading of operations can be safely performed since type information can often be inferred by an algorithm.

Maybe Not

- Overloading allows us to define $g_1 : (\mathsf{Nat} * \mathsf{Nat}) \to \mathsf{Nat}$ and $g_2 : \mathcal{P}(\mathsf{Nat}) \to \mathsf{Nat}$, and then:

$$g(x) = \begin{cases} g_1(x) & \text{if the type of } x \text{ is Nat} \\ g_2(x) & \text{if the type of } x \text{ is } \mathcal{P}(\mathsf{Nat}) \end{cases}$$

- However, the typing rules are often very complicated, and require machine assistance to carry them out.
- Also, formalisation of mathematics into a type theory requires removal of set-theoretic dependencies.

Tagging: Definition and Diagram

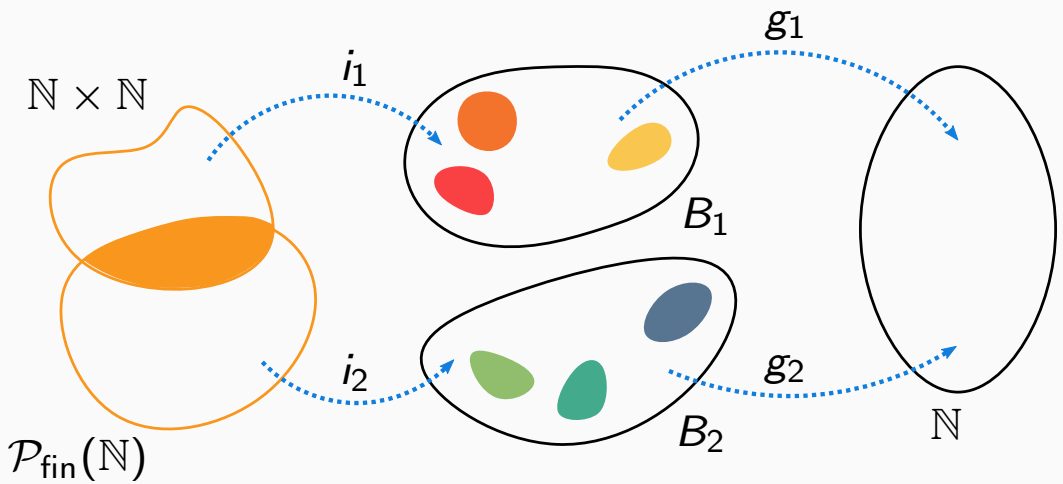
- What if we want to stay within set theory?

Definition

$$i_1 : \mathbb{N} \times \mathbb{N} \to B_1,\ i_2 : \mathcal{P}(\mathbb{N}) \to B_2$$
such that i_1, i_2 are injective and $B_1 \cap B_2 = \emptyset$.

Definition

$$g_1 : B_1 \to \mathbb{N},\ g_2 : B_2 \to \mathbb{N}, \text{ such that}$$
$$g_1(i_1(\langle n, m \rangle)) = n + m, \quad g_2(i_2(\{n_1, \ldots, n_k\})) = \sum_{i=1}^{k} a_i.$$



Tagging: Results

- The function g can then be taken as $g_1 \cup g_2$, and i_1 and i_2 are used to tell g whether to use g_1 or g_2.
- The mappings $x \mapsto \langle 0, x \rangle$, $x \mapsto \langle 1, x \rangle$ are often used for i_1, i_2.
- Tagging is ugly and annoying.
- Trying to reason about the tagged version of g in a theorem prover would be frustrating.

What else?

- If you don't want to do tagging, there is not much choice for resolving this issue whilst staying within set-theoretic foundations.
- Some set theories have non-set objects called atoms or urelements, though they usually have no internal structure.
- We propose the use of an information hiding mechanism in which the representation of mathematical objects is hidden away by using genuine non-sets rather than sets.
- Special non-sets will act like a set representation, but without the properties the user considers irrelevant.
- As a first instance of this aim, we arrive at ZFP (ZF with Ordered Pairs), which hides the representation of ordered pairs.

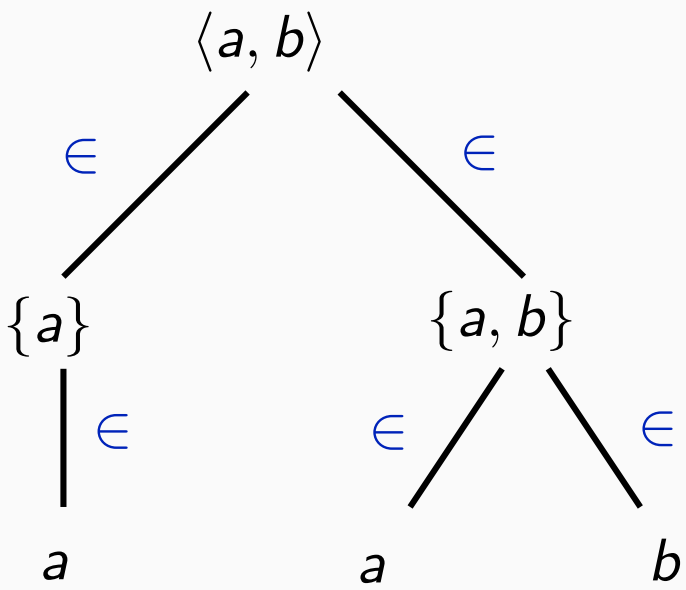
ZFP: Zermelo-Fraenkel Set Theory with (Ordered) Pairs

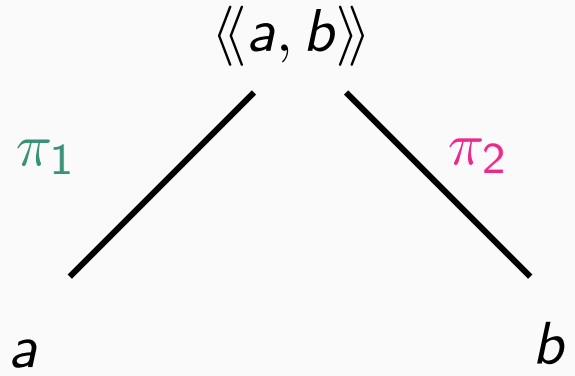
Introduction to ZFP

- ZFP extends ZF with ordered pairs as non-set objects[1].
- The domain of objects is split into pairs and sets.
- For any objects A, B, there exists a 'primitive' ordered pair $\langle\langle A, B \rangle\rangle$ such that $\forall c : c \notin \langle\langle A, B \rangle\rangle$.
- Gained by modifying ZF's axioms to allow non-sets, and axiomatising relation symbols π_1, π_2 with desired properties.
- So that $\langle\langle A, B \rangle\rangle$ is the unique p such that $A\ \pi_1\ p$ and $B\ \pi_2\ p$.
- The domain of ZFP contains all of the 'pure' sets of ZF, including Kuratowski pairs $\langle a, b \rangle$.

[1] Adding an Abstraction Barrier to ZF Set Theory. Ciaran Dunne, J. B. Wells, Fairouz Kamareddine, Conference for Computer Intelligent Mathematics 2020
Pre-print: `https://arxiv.org/abs/2005.13954?fbclid=IwAR1wdCEJbYGVfLJg7hekLK9ePkRHEeE3iaZj9zQbOZF4P_kk3zLNxw--580`

Diagram





Facts?

- Sets can have pairs as their members, and pairs can have sets as their projections.
- Cartesian product can defined using Replacement nested inside Replacement:

$$A \times B = \cup \{ z \mid \exists c \in A : z = \{ p \mid \exists d \in B : p = \langle\langle c, d \rangle\rangle \} \}$$

- If $p = \langle\langle a, b \rangle\rangle$, then $\mathcal{P}(p) = \emptyset$ and $\cup p = \emptyset$, etc.

Consistency of ZFP

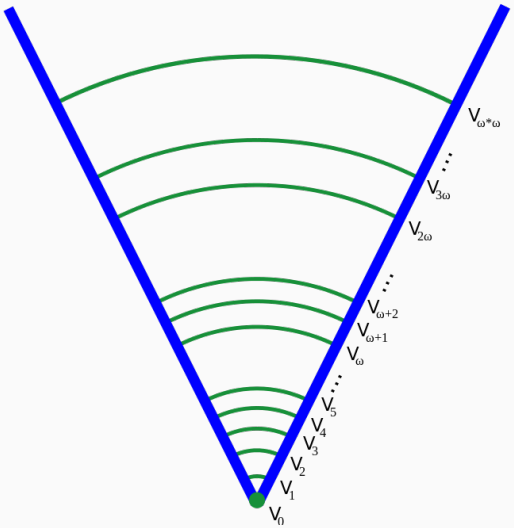
- The axioms of ZF are well studied and believed to be consistent.
- That is, it is not possible to prove a formula φ and its negation $\neg\varphi$.
- Can the same be said about ZFP?
- We can build a model of ZFP using sets in ZF.
- This provides us with a method of interpreting the formulas of ZFP of ZF.
- So if ZF is consistent, so is ZFP.

Models of ZF

- ZF is powerful enough to model its own domain using the Von-Neumann hierarchy.
- An infinite collection of sets obtained via iteration of the power set operator.

Von-Neumann Hierarchy:

$$V_0 = \emptyset,\ V_1 = \{\emptyset\},\ V_2 = \{\emptyset, \{\emptyset\}\},\ V_3 = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}$$



- V_6 contains 2^{65536} elements, greater than the number of atoms in the known universe. . .
- Whilst everything is constructed out of the empty set, it still satisfies the axioms of ZF.

A Model of ZFP

- For ZFP, we use the Kuratowski ordered pairs and sets of ZF.
- Objects need to be tagged at every stage to ensure a disjoint domain of pairs and sets.
- We consider an object a to be a member of the domain **W** if there exists an ordinal α such that $x \in W_\alpha$.
- If $x \in W_\alpha$, then $x = \langle 0, x' \rangle$ or $x = \langle 1, \langle a, b \rangle \rangle$.

Definition

$$
\begin{aligned}
W_0 &= \emptyset \\
W_{\beta^+} &= (\{0\} \times \mathcal{P}(W_\beta)) \cup (\{1\} \times {W_\beta}^2) \\
W_\lambda &= \bigcup_{\beta \in \lambda} W_\beta
\end{aligned}
$$

Definition

$$
\begin{aligned}
a \mathrel{\widehat{\in}} x &:= \exists\, x' : x = \langle 0, x' \rangle \wedge a \in x' \\
a \mathrel{\widehat{\pi}_1} p &:= \exists\, b : p = \langle 1, \langle a, b \rangle \rangle \\
b \mathrel{\widehat{\pi}_2} p &:= \exists\, a : p = \langle 1, \langle a, b \rangle \rangle
\end{aligned}
$$

Interpretation of ZFP in ZF

- The domain **W** together with the relations $\widehat{\in}, \widehat{\pi}_1, \widehat{\pi}_2$ satisfy the axioms of ZFP.
- The proof for each axiom has been machine checked in Isabelle/ZF.[2]
- Any formula of ZFP φ can be interpreted in ZF by restricting quantification to **W** and replacing occurrences of relation symbols with $\widehat{\in}, \widehat{\pi}_1, \widehat{\pi}_2$.
- ZFP can be seen to be an axiomatisation of a fully tagged domain, with all of the nasty details hidden away and dealt with automatically.

[2] See `http://www.macs.hw.ac.uk/~cmd1/cicm2020/ZFP.thy` for the source, and `http://www.macs.hw.ac.uk/~cmd1/cicm2020/ZFPDoc/index.html` for the HTML.

Future work: Information Hiding Mechanism

- What if we want more kinds of non-sets other than ordered pairs?
- The consistency of ZFP depends on the consistency of ZF, and its ability to provide a representation for ordered pairs, and a tagging mechanism.
- The methods used for gaining primitive ordered pairs could be carried out for numbers, functions, and other mathematical objects.
- We propose designing a set theory which provides a mechanism for this, allowing the user to form new non-set objects with any desired properties, provided a valid representation exists.
- Relative consistency can be proved by showing that all of this can be simulated in ZF.

Conclusion

- ZF models mathematical text by representing mathematical objects as sets.
- Reasoning performed based on cases of the 'type' of object requires a messy translation to be considered founded in ZF.
- Types may not be the solution to this, due to the restrictions that come with it.
- Representation of ordered pairs can be hidden using ZFP.
- A generalisation of these methods may allow a theory of mathematical objects in which representations are hidden.