Speaker Name: Ryan Carpenter
Date: 4/16/2023

$1 + 1 = 2$: I Read *Principia Mathematica* So You Don't Have To
Graduate Student Conference in Logic XXIII

1 Overview

Overview of this talk:

- A bit of history
- The proof that $1 + 1 = 2$
- The legacy of Russell's logicism

2 The History

With the rise of formal precision in the nineteenth century, especially in analysis and in the study of infinity, a philosophical outlook about the content of mathematics called logicism was able to take form. In brief, to be a "logicist" means to believe that all of mathematics is reducible to logical foundations; that is to say, every theorem of mathematics is a theorem of strict logic. This philosophical school of thought was essentially pioneered by Gottlob Frege, a mathematician and philosopher during the late nineteenth century. Unfortunately for Frege, the logical power he formalized was largely ignored until Bertrand Russell brought it to light in his *Principles of Mathematics* (1903) and in his joint project with A. N. Whitehead, *Principia Mathematica* (Vol. I originally published in 1910, followed by Vols. II in 1912 and III in 1913). The goal of Whitehead and Russell's *Principia* was to lay the logical foundations upon which all of mathematics could be shown reducible to. Included in the theorems they proved in their three-volume work is the basic arithmetic truth that $1 + 1 = 2$, and it is this proof which concerns us in this (brief) talk. (Fun fact: after the conclusion of the proof of the theorem, it is remarked by the authors that "[t]he above proposition is occasionally useful.")

3 The Proof

First a bit of important definitions:

- "Class":
  A collection consisting of all terms satisfying some propositional function; hence, every propositional function defines a class. Further, two classes are identical exactly when their defining propositional functions are formally equivalent (note the similarity here to the ZFC axiom of foundation).
  For the purposes of this talk, we may think of the classes we're considering as acting as sets, although they are defined to be proper classes. Lowercase Greek letters will typically stand in for classes and the notation $\hat{\alpha}\{\varphi\}$ means the class of all classes α such that φ.

- "Λ":

$$\Lambda := \hat{x}(x \neq x)$$

  The symbol Λ is to denote the null class. Note that in a certain interpretation of the range of the variable x, this would define the empty set in ZFC under the axiom scheme of comprehension.

- "1":

$$1 := \hat{\alpha}\{(\exists x)(\alpha = i'x)\}$$

  Also, $*52.1$:

$$\vdash ((\alpha \in 1) \Leftrightarrow ((\exists x)(\alpha = i'x)))$$

- "2":

$$2 := \hat{\alpha}\{(\exists x, y)(x \neq y \wedge \alpha = i'x \cup i'y)\}$$

Proposition $*54.43$**:**

$$\vdash (\alpha, \beta \in 1) \Rightarrow ((\alpha \cap \beta = \Lambda) \Leftrightarrow (\alpha \cup \beta \in 2))$$

Proof:

$$\begin{array}{llll}
(\vdash *54.26) \Rightarrow (\vdash (\alpha = i'x \wedge \beta = i'y) \Rightarrow (\alpha \cup \beta \in 2 & \Leftrightarrow & x \neq y)) & \\
[*51.231] & \Leftrightarrow & i'x \cap i'y = \Lambda)) & \\
[*13.12] & \Leftrightarrow & \alpha \cap \beta = \Lambda)) & \textbf{(1)} \\
(\vdash \textbf{(1)} \wedge *11.11 \wedge *11.35) & & & \\
\Rightarrow (\vdash (\exists x, y)(\alpha = i'x \wedge \beta = i'y) \Rightarrow ((\alpha \cup \beta \in 2) \Leftrightarrow (\alpha \cap \beta = \Lambda)) & & & \textbf{(2)} \\
(\vdash \textbf{(2)} \wedge *11.54 \wedge *52.1) \Rightarrow (\vdash \text{Prop}) & & &
\end{array}$$

The proof itself directly contains reference to seven other previously proven propositions; most of these, in turn, reference other previously proven propositions. In fact, a grand total of (more than I cared to count on two separate occasions) references are made to propositions ultimately leading back to either the deductive rules assumed or definitions. It would be a grand task in these twenty minutes to convince ourselves of the legitimacy of each of these referenced propositions, so instead, we'll focus on the seven directly referenced in the demonstration:

- $*54.26$

$$\vdash ((i'x \cup i'y \in 2) \Leftrightarrow (x \neq y))$$

  Intuitively, this proposition is trying to claim that the sum of two unit classes is a couple if and only if the unit classes contain different elements. Crucially, the proof of this proposition relies on a basic distributivity law in symbolic logic that:

$$P \wedge (Q \vee R) \Leftrightarrow (P \wedge Q) \vee (P \wedge R)$$

  Using this, we see the entailment in the first line of the proof (we are simply naming $\alpha = i'x$ and $\beta = i'y$).

- $*51.231$

$$\vdash ((i'x \cap i'y = \Lambda) \Leftrightarrow (x \neq y))$$

  Here, we see the connection to the previous reference: the product of two unit classes is the empty class if and only if the elements of the classes are different. Using this, we may rewrite the right side of the iff on the first line.

- $*13.12$

$$\vdash ((x = y) \Leftrightarrow (\psi x = \psi y))$$

  Here, the proposition is meant to express that if two things are identical, then any property holds of one if and only if it holds of the other. This is one of the basic properties of identity that Russell and Whitehead define. This proposition is used to replace $i'x$ and $i'y$ with α and β again.

- $*11.11$

  If $\varphi(w, z)$ is true whatever possible arguments w and z may be, then $(\forall x, y)\varphi(x, y)$ is true.

  This proposition is essentially definitional for the universal quantifier, but it is actually a reformulation of an earlier definition.

- $*11.35$
$$\vdash ((\forall x, y)(\varphi(x, y) \Rightarrow P) \Leftrightarrow ((\exists x, y)(\varphi(x, y)) \Rightarrow P))$$
  Combining this proposition with the previous yields the justification for the existential introduction. So far, it wasn't established that there must be some individuals x and y which comprise the unit classes α and β, and we now instantiate them. This proposition is less obvious until one notices it's only making formal the movement of quantifiers over implication.

- $*11.54$
$$\vdash (((\exists x, y)(\varphi x \wedge \psi y)) \Leftrightarrow ((\exists x)\varphi x \wedge (\exists y)\psi y))$$
  In other words, the symbol $(\exists x, y)$ is shorthand notation.

- $*52.1$ This is again the characterization of unit classes that we saw before. Combining these last two propositions with **(2)** just yields the exact formulation of the desired proposition. We may notice that the real work of the proof is done here by the combination of $*54.26$ and $*51.231$; together, they make precise the connection between a sum of classes being a couple and the product of classes being empty.

The attentive audience member will note that we haven't actually proven that the sum of one and one is two; in fact, this is proposition $*110.643$ in Vol. II. However, the proposition $*54.43$ is the key ingredient in that proof after the authors define cardinal numbers and the sum operation on cardinals (NB: the "1" and "2" we used in the proof above are *not* cardinals as they appear in the proof; they happen to satisfy R+W's definition of cardinal numbers after they're defined, though).

4 The Legacy

The logicist project of showing that the whole of mathematics is reducible to (or is a subset of) formal logic was unofficially ended (though this is a bit unfair to say) with the proofs of Godel's Incompleteness Theorems. A large part of the motivation behind formalizing mathematics the way Whitehead and Russell desired was to guard against the paradoxes that cropped up in the mathematics of their time; once Godel showed that any sufficiently powerful system of mathematics (such as PM) could never be shown to be consistent within the system, the paradoxes that the logicists believed could be avoided were proven to be unavoidable; moreover, the fact that there are formally undecidable sentences in any such formal system was thought to give evidence that the power of formal systems such as PM was far less than what logicists expected. There are today "neologicists" who tend to maintain the philosophical outlook that mathematics is, in some important sense, reducible to logic as Frege and Russell wanted to show. Although, the neologicist standpoint only arrives at this outlook by first introducing so-called "abstraction principles," axioms of a more empirical or analytic flavor than strictly logical axioms, such as Hume's Principle:
$$\#F = \#G \Leftrightarrow F \cong G$$
The neologicists also tend to argue for the logical status of much of these abstraction principles and sometimes for non-logical abstraction principles such as existence claims (e.g., the existence of any set). In the end, though the PM was by many regards a failed project, it did indeed breathe a life that had never before been enjoyed into mathematical logic, and it set the stage for many meta-logical results of, for example, Godel, Church, and Turing.