Bertrand Russell (1872-1970)

- 1910-1913: *Principia Mathematica* (with Alfred Whitehead)
- 1918: Imprisoned for pacifism
- 1950: Nobel Prize in Literature
- 1955: Russell-Einstein Manifesto
- 1967: *War Crimes in Vietnam*

“Great spirits have always encountered violent opposition from mediocre minds.”

Albert Einstein (talking about Bertrand Russell)

---

*Principia Mathematica*

Whitehead and Russell (1910–1913)
Three Volumes, 2000 pages

Attempted to axiomatize mathematical reasoning

- Define mathematical entities (like numbers) using logic
- Derive mathematical “truths” by following mechanical rules of inference
- Claimed to be **complete** and **consistent**:
  - All true theorems could be derived
  - No falsehoods could be derived
Perfect Axiomatic System

Derives all true statements, and no false statements starting from a finite number of axioms and following mechanical inference rules.

Incomplete Axiomatic System

Derives some, but not all true statements, and no false statements starting from a finite number of axioms and following mechanical inference rules.

Inconsistent Axiomatic System

Derives all true statements, and some false statements starting from a finite number of axioms and following mechanical inference rules.

Russell’s Paradox

Some sets are not members of themselves e.g., set of all Jeffersonians

Some sets are members of themselves e.g., set of all things that are non-Jeffersonian

\[ S = \text{the set of all sets that are not members of themselves} \]

\[ S \notin S \]

\[ S \in S \]

Russell’s Paradox

\( S = \text{set of all sets that are not members of themselves} \)

Yes?

If \( S \) is an element of \( S \), then \( S \) is a member of itself and should not be in \( S \).

No?

If \( S \) is not an element of \( S \), then \( S \) is not a member of itself, and should be in \( S \).

Ban Self-Reference?

- *Principia Mathematica* attempted to resolve this paragraph by banning self-reference

- Every set has a type
  - The lowest type of set can contain only “objects”, not “sets”
  - The next type of set can contain objects and sets of objects, but not sets of sets
  - Can we define \( S \)?
Epimenides Paradox

Epimenides (a Cretan):
“All Cretans are liars.”

Equivalently:
“This statement is false.”

Russell’s types can help with the set paradox, but not with these.

Kurt Gödel

1931: Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme
(On Formally Undecidable Propositions of Principia Mathematica and Related Systems)

1906 (Brno) – 1978 (Princeton)

1939: flees Vienna Institute for Advanced Study, Princeton (with Einstein, von Neumann, etc.)

Gödel’s Theorem

In the Principia Mathematica system, there are statements that cannot be proven either true or false.
<table>
<thead>
<tr>
<th>Gödel’s Theorem</th>
<th>Gödel’s Proof – General Idea</th>
</tr>
</thead>
<tbody>
<tr>
<td>In any interesting rigid system, there are statements that cannot be proven either true or false.</td>
<td><strong>Theorem:</strong> In the <em>Principia Mathematica</em> system, there are statements that cannot be proven either true or false. <strong>Proof:</strong> Find such a statement.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gödel’s Statement</th>
<th>Gödel’s Proof Idea</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>G:</strong> This statement does not have any proof in the system of <em>Principia Mathematica</em>.</td>
<td><strong>G:</strong> This statement does not have any proof in the system of <em>PM</em>.</td>
</tr>
<tr>
<td>G is unprovable, but true!</td>
<td>If G is provable, PM would be <strong>inconsistent</strong>. If G is unprovable, PM would be <strong>incomplete</strong>.</td>
</tr>
<tr>
<td><strong>Possibilities:</strong></td>
<td><strong>Thus, PM cannot be complete and consistent!</strong></td>
</tr>
<tr>
<td>1. G is <strong>true</strong> ⇒ G has no proof System is <strong>incomplete</strong></td>
<td></td>
</tr>
<tr>
<td>2. G is <strong>false</strong> ⇒ G has a proof System is <strong>inconsistent</strong></td>
<td></td>
</tr>
</tbody>
</table>

**Axiomatic System**

- Derives some, but not all true statements, and no false statements starting from a finite number of axioms and following mechanical inference rules.
- System is **incomplete**

- Derives all true statements, and some false statements starting from a finite number of axioms and following mechanical inference rules.
- System is **inconsistent**

*Pick one:*
**Inconsistent Axiomatic System**

A Derives all true statements, and some false statements starting from a finite number of axioms and following mechanical inference rules.

Once you can prove one false statement, everything can be proven!  false \(\Rightarrow\) anything

**Finishing The Proof**

Turn G into a statement in the *Principia Mathematica* system

Is PM powerful enough to express G?

G: “This statement does not have any proof in the PM system.”

Yes! If you don’t believe me read Gödel’s paper or Hofstadter’s *Gödel, Escher, Bach*

**Gödel’s Proof**

G: This statement does not have any proof in the system of PM.

If G is provable, PM would be inconsistent.
If G is unprovable, PM would be incomplete.
PM can express G.
Thus, PM cannot be complete and consistent!

**Generalization**

All “powerful” logical systems are incomplete: there are statements that are true that cannot be proven within the system.

“powerful” = able to express: “This statement has no proof.”

**Does this solve the Entscheidungsproblem?**

**Decision procedure:** given a well-formed statement, determine whether it can be proven.

Turing, Church (1936): this is impossible!

**“Truth” procedure:** given a well-formed statement, determine mechanically whether it is true.

Gödel (1930): this is impossible!
Defining Computable Numbers

A number is computable if its decimal can be written down by a machine.

I give some arguments showing that the computable numbers include all numbers which could naturally be regarded as computable. In particular, I show that certain large classes of numbers are computable. They include, for instance, the real parts of all algebraic numbers, ... the numbers π, e, etc.

Computable convergence: \[\pi = 4\left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots\right)\]

How many computable numbers are there?

How many strings in \(\Sigma^*\) are there?

\[\Sigma = \{0, 1\}\]

\[\varepsilon, 0, 1, 01, 10, 11, 00, 001, \ldots\]

\[\varepsilon, 0, 1, 2, 3, \ldots\]

How many computable numbers are there?

"Je le vois, mais je ne le crois pas."

Georg Cantor

1. \(0.0000000000000\ldots\)
2. \(0.25000000000000\ldots\)
3. \(0.333333333333\ldots\)
4. \(0.666666666666\ldots\)

\[\ldots \ldots\]

57236.141592653589793\ldots

\[\ldots \ldots\]

How many computable numbers are there?

How many strings in \(\Sigma^*\) are there?

\[\begin{array}{c}
(Q, \Sigma, \delta, \epsilon, q_0, q_{accept}, q_{reject}) \\
\preceq \mathbb{N}_0
\end{array}\]

Countable:

Natural Numbers
Any set that has 1-to-1 mapping to whole numbers

Set of strings in \(\Sigma^*\)

Not Countable:

Real Numbers
How many computable numbers are there?

Each TM produces at most one new number!

**Countable:**
- Natural Numbers
- Any set that has 1-to-1 mapping to whole numbers
- Set of strings in \( \Sigma^* \)
- Set of all Turing Machines
- Computable Numbers

**Not Countable:**
- Real Numbers

---

How many languages in \( \Sigma^* \) are there?

\( \mathbb{N}_0 \)

**Countable:**
- Natural Numbers
- Any set that has 1-to-1 mapping to whole numbers
- Set of strings in \( \Sigma^* \)
- Set of all Turing Machines
- Computable Numbers

**Not Countable:**
- Real Numbers

---

How many languages in \( \Sigma^* \) are there?

\( \mathbb{N}_0 \)

**Countable:**
- Natural Numbers
- Any set that has 1-to-1 mapping to whole numbers
- Set of strings in \( \Sigma^* \)
- Set of all Turing Machines
- Computable Numbers

**Not Countable:**
- Real Numbers
- Set of all languages in \( \Sigma^* \)
Remember Gödel
All “powerful” logical systems are incomplete: there are statements that are true that cannot be proven within the system.

Systems break when they have to answer deep questions about themselves!

Turing Machine Questions
The set of all Turing Machines is countable:

\[ \Sigma = \{0, 1\} \]
\[ < M > \in \Sigma^* \]

Can we describe all TM’s using an encoding in \((0, 1)^*\)?

\[ \Sigma^* = \epsilon, 0, 1, 00, 01, 10, 11, 000, \ldots \]

Are all strings in \(\Sigma^*\) valid descriptions of TMs?

The Language of TMs
\[ \Sigma^* = \epsilon, 0, 1, 00, 01, 10, 11, 000, \ldots \]
\[ M(w) = \text{if } w \text{ is a valid description of a TM, } M(w) \text{ is the TM described by } w. \]
If \(w\) is not a valid description of a TM, \(M(w)\) is this TM:

\[ \text{q_reject} \]

Self-Rejecting Language
\[
\text{SELF-REJECTING} = \{ w \in \Sigma^* : w \notin \mathcal{L}(M(w)) \}
\]

Suppose there is a TM \(M_{SR}\) that recognizes SELF-REJECTING:
\[ M_{SR} = M(w_{SR}) \text{ for some } w_{SR} \in \Sigma^* \]

Is \(w_{SR}\) in SELF-REJECTING?

Is \(w_{SR}\) in SELF-REJECTING?
\[
\text{SELF-REJECTING} = \{ w \in \Sigma^* : w \notin \mathcal{L}(M(w)) \}
\]

Assume TM \(M_{SR} = M(w_{SR})\) recognizes SELF-REJECTING.

Yes?
\[ w_{SR} \in \text{SELF-REJECTING} \]
\[ \Rightarrow w_{SR} \notin \mathcal{L}(M(w_{SR})) \]
\[ \Rightarrow w_{SR} \notin \text{SELF-REJECTING} \]

Contradiction!

No?
\[ w_{SR} \notin \text{SELF-REJECTING} \]
\[ \Rightarrow w_{SR} \in \mathcal{L}(M(w_{SR})) \]
\[ \Rightarrow w_{SR} \in \text{SELF-REJECTING} \]

Contradiction!

The assumption leads to a contradiction: thus, \(M_{SR}\) must not exist!
Languages that can be recognized by *any* mechanical computing machine

- Thursday morning office hours are now 8:45-10am
- Reading this week: Chapter 4
- PS5 will be posted soon (due April 6, postponed 1 week from original due date)
- Next class:
  - Diagonal argument for Self-Rejecting
  - A language that is *Turing-recognizable* but not *Turing-decidable*