Gödel's Incompleteness Theorems and the Foundational Crisis of Mathematics

Bhaskar Mishra

April 19, 2020

1 Introduction

For thousands of years, people have been drawn to study fields like mathematics, philosophy, and logic because of a shared passion for solving challenging problems. The objective and explosively branching nature of mathematics and mathematical systems leads to the creation of a huge number of "puzzles" to solve and truths to discover. This same objectivity often leads people to intuitively believe that every such problem should have a solution. In other words, that all truths within these mathematical systems that we study should be discoverable and provable within these systems, and it's solely up to us, as mathematicians, to find these proofs.

The view that mathematical systems, as typically studied, are both consistent[1] and complete[2] has been held by a majority of mathematicians for most of history. In 1931, however, Kurt Gödel, an Austrian logician, mathematician, and analytic philosopher, published his two incompleteness theorems [6] which disproved this way of thinking. A rough statement of his two theorems is as follows [18]:

[1] A system is consistent if and only if there is no statement provable within the system whose negation is also provable within the system.

[2] A system is complete if and only if for each statement in the system, either it or its negation is provable within the system.

1. Any consistent formal system F within which a certain amount of elementary arithmetic can be carried out is incomplete; i.e., there are statements of the language of F which can neither be proved nor disproved in F.

2. For any consistent system F within which a certain amount of elementary arithmetic can be carried out, the consistency of F cannot be proved in F itself.

In this paper, I hope to provide a non-formal understanding of Gödel's two incompleteness theorems, while also providing an appreciation for the history, significance, and many implications of the theorems.

2 The Foundational Crisis of Mathematics

During the early 20th century, many mathematicians from across Europe were tackling the foundational crisis of mathematics. At the time, mathematics was a large collection of different areas of study, but there wasn't something that could bring them all together. The goal was to create a much more comprehensive mathematical system that would unify all of mathematics under one consistent framework. One of the first promising attempts at creating such a framework was Cantor's Naive Set Theory.

2.1 Cantor's Naive Set Theory

Cantor's Naive Set Theory was one of the early attempts at creating a foundation for mathematics. It involved the creation of a mathematical object that Cantor called an "aggregate" or, in German, "*Menge*" - an object similar to what we now call a set [8]. For better readability, I will refer to the object as a set. Cantor's definition for a set can be stated as follows:

> "A **set** is any collection of definite, distinguishable objects of our intuition or of our intellect to be conceived as a whole (i.e., regarded as a single unity)." [19]

Though Cantor didn't explicitly state any axioms for his theory, he used the following three principles of sets in a similar fashion to how axioms are used in axiomatic systems [19].

1. **Extensionality**: A set is completely determined by its members.
2. **Abstraction**: Every property determines a set.
3. **Choice**: Given any set $\mathcal{F}$ of nonempty pairwise disjoint sets, there is a set that contains exactly one member of each set in $\mathcal{F}$.

Initially, this theory seemed very promising as it was applicable in many different fields of mathematics. Ordered pairs, for example, could be defined using the Kuratowski pair as follows[3] [12]:

$$(a, b) := \{\{a\}, \{a, b\}\}.$$

Functions could then be defined by defining $\mathcal{F} : \mathcal{A} \to \mathcal{B}$ to be the set of all ordered pairs (a, b) such that $a \in \mathcal{A}$, $\mathcal{F}(a) = b \in \mathcal{B}$, and there are no two ordered pairs with the same first element, but different second elements [19]. In similar ways, Cantor's Set Theory was shown to extend to many other areas of mathematics.

Cantor's system continued to gain traction until around the 1900s when several mathematicians, including Cantor himself, discovered various paradoxes in the theory. One such paradox, published in 1902 by British mathematician[4] Bertrand Russell, was Russell's Paradox. Russell's Paradox concerns the set of all sets that do not contain themselves as elements or $\{x \mid x \notin x\}$, and whether or not such a set contains itself [21]. By definition, the set cannot contain itself as that defies the limiting property; however, if the set does not contain itself, it then must contain itself as it satisfies the limiting property. This creates a contradiction and therefore, the paradox showed that Cantor's Naive Set Theory is not consistent.

[3] The subset operator, $\subseteq$, is a total order on this set and provides the ordering for the ordered pair.

[4] He was also a philosopher, logician, writer, essayist, social critic, and a political activist among other things [11].

2.2 Intuitionism

In response to the many paradoxes discovered in the context of Cantor's Naive Set Theory, mathematicians split into three main schools of thought as to how to resolve them. These schools consisted of Intuitionism, Logicism, and Formalism. Intuitionism rejected the Platonic view of mathematics, which proposed that math exists independently of the physical and mental world, and that humans simply discover mathematical truths from a set of axioms [14]. Instead, it argued that mathematics exists exclusively in our human minds and that mathematical objects exist only if they can be constructed within the mind. This view of mathematics differed from other philosophies primarily in its attitude towards the concept of infinity. Whereas Cantor's Naive Set Theory treated infinity as a mathematical object in and of itself, Intuitionism treated it as a potentiality, something that may exist but can't actually be *shown* to exist [10]. The reasoning for this can be seen when considering an infinite set. Though one can constructively show the existence of any number of members of an infinite set, they can never show all of them, and therefore, a "complete" infinity can't exist [19]. Under this framework, the set $\{x \,|\, x \notin x\}$ does not contain itself, as its own existence can't be constructed completely; it exists only as a potentiality. Like this, Intuitionism resolved Russell's paradox, along with the many other paradoxes in Cantor's Naive Set Theory.

2.3 Logicism

Though Intuitionism avoided many of the paradoxes that grew from infinite sets, most mathematicians found the costs it posed in doing so to be too great. Intuitionism required "large parts of mathematics... to be cast off" as they seemed "impossible to reconstruct... according to intuitionistic principles." [19] Another approach was necessary that didn't so greatly decrease the scope of mathematics. Logicism attempted to solve this problem by building the foundations of mathematics on pure logic, without the use of purely mathematical ideas like numbers and sets [17].

In trying to create such a system, mathematicians quickly realized the need for a new "concise and precise" notation that avoided the ambiguity in natural language [19]. In 1879, Gottlob Frege, a German mathematician and one of the primary contributors to the school of Logicism, published his first book, *Begriffsschrift*, in which he introduced a new "formula language" or symbolic language in order to represent logic and model the foundations of arithmetic [3].

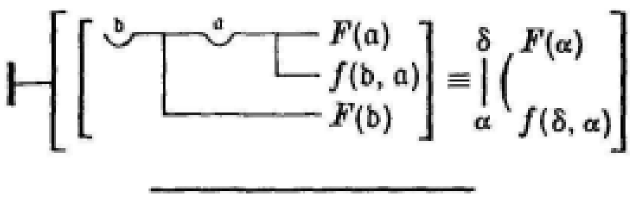

Figure 1: An example of the notation introduced in *Beggriffsschrift*. This diagram can be expressed in words as follows: "If from the proposition that b has property F it can be inferred generally, whatever b may be, that every result of an application of the procedure f to b has property F, then I say: 'Property F is hereditary in the f-sequence'." [3]

Though Frege's notation wasn't very well received among mathematicians, likely because of its 2-dimensional nature and it being "radically new" [28], his work, along with another notation introduced by Italian mathematician Guissepe Peano, set the grounds for the development of First Order Logic as well as much of the mathematical notation used today [19].

Russell, also a significant contributor to the school of Logicism, aimed to expand on Frege's goals. Whereas Frege's system only aimed to create foundations for arithmetic and was shown to still be subject to paradoxes like Russell's Paradox, Russell aimed to use the principles of Logicism to create a system that could serve as a foundation for all of mathematics and would be free of those paradoxes [15]. To avoid his own paradox, Russell introduced the idea of a "theory of types," which, said informally, required all mathematical objects to be assigned a "type," and for each mathematical object to be constructed only from objects of lower types based on a certain ordering of those types [20][5]. As a result, the set $\{x \mid x \notin x\}$ would not exist, as its definition requires it to contain itself, causing it be constructed from an object of the same type. In the early 1910s, Russell used this idea and published the Principia Mathematica (PM) [22]. In the book, Russell introduced a complex system that avoided all known paradoxes at the time, while also seeming to branch into all fields of mathematics [15]. Though the question of whether PM was truly consistent and complete remained an open research problem, the system gained a lot of popularity as possibly being a promising step towards creating a foundation for mathematics.

2.4 Formalism

With the Principia Mathematica providing the "required logical basis for a renewed attack on foundational issues," [27] in the 1910s, German mathematician David Hilbert, introduced a new school of thought known as Formalism [23]. The formalist view contended that mathematics can be seen as simply a syntactic game, and that there is no need for semantics or interpretation of mathematical statements. Formalism aimed to "formalize" axiomatic systems by creating a formal language for mathematics and allowing certain manipulations of strings in that language. Using the strings for the respective axioms to start with, these manipulations could be used to prove theorems within the system.

The primary goal in creating such formal languages for mathematical systems was to be able to prove their consistency. By creating such a precise formal language, the question could be changed from asking whether a mathematical system T is consistent, to asking whether one can "possibly run into contradictions if one proceeds entirely formally within [formal language] L, using only the axioms of T and those of classical logic, all of which have been expressed in L."

[5] This simplified version of the theory of types introduced in Appendix B of Russell's Principles of Mathematics still faced some other paradoxes, including versions of the Liar's paradox, leading Russell to soon create his "ramified" theory of types [1]. This ramified theory will, however, not be discussed in this paper.

[23] Without semantics to create ambiguity, these formal languages were much easier to analyze and sparked a lot of research on the consistency, completeness, and other properties of mathematical systems.

In order to fully clarify the goals of Formalism, in the early 1920s, Hilbert formulated a list of goals for the foundations of mathematics which have now come to be known as "Hilbert's Program". Four relevant goals from Hilbert's Program can be stated as follows [19]:

1. Find a formal axiomatic system **M** capable of deriving all the theorems of mathematics;
2. Prove that the theory **M** is semantically complete;
3. Prove that the theory **M** is consistent;
4. Construct an algorithm that is a decision procedure[6] for the theory **M**.

Though the ideas of Hilbert's Program and the formalist school of thought remain important to this day, within 3 years of the program's proposal, Gödel's Incompleteness Theorems proved that there were significant limitations on its goals, specifically in the context of goals 2 and 3. This will be explored further in the following chapter.

3 Gödel's Incompleteness Theorems

Given the impact Gödel's incompleteness theorems had on Hilbert's Program, it is a surprising fact that when he first started working on the problem, he had "set out not to destroy Hilbert's program but to advance it." [2] In 1930, a year before his publication of the incompleteness theorems, Gödel was attempting to use arithmetic to prove the consistency of second-order arithmetic. In doing so,

[6] A decision procedure for a formal axiomatic system **M** is an algorithm that is capable of taking any well-formed statement from the formal language used to represent **M** and outputting whether or not the statement is true in **M** in finite time.

however, Gödel encountered many different paradoxes regarding the notion of truth, and soon arrived on a version of the "undefinability of truth theorem[7]," a theorem later discovered and published by Polish-American mathematician Alfred Tarski [18]. This discovery sparked the research that led to Gödel's work on his incompleteness theorems.

For better readability, here, again, are the statements of Gödel's two incompleteness theorems:

1. Any consistent formal system F within which a certain amount of elementary arithmetic can be carried out is incomplete; i.e., there are statements of the language of F which can neither be proved nor disproved in F.

2. For any consistent system F within which a certain amount of elementary arithmetic can be carried out, the consistency of F cannot be proved in F itself.

The direct implication of these two theorems is that either goal 3 of Hilbert's Program isn't met due to the theoretical formal axiomatic system $\mathbf{M}$ not being consistent, or goal 2 isn't met, as the system is consistent, but not complete. In the latter case, there is also a limitation on goal 3, as based on the second incompleteness theorem, $\mathbf{M}$'s consistency can't be proven from within $\mathbf{M}$.

3.1 Proof Sketch of First Incompleteness Theorem

The following proof sketch for Gödel's first incompleteness theorem will closely mirror the proof laid out in Douglas Hofstadter's Pulitzer Prize winning book, *Gödel, Escher, Bach: An Eternal Golden Braid* (G.E.B.) [9][8].

[7]The undefinability of truth theorem roughly states that the concept of arithmetical truth cannot be defined in arithmetic, or in other words, there is no formula in arithmetic that can determine whether or not a statement of arithmetic is true.

[8]Note that while Hofstadter's proof is used for its elegance and well-designed structure, this paper does not accept Hofstadter's broader interpretation of the significance or metaphysical understanding of Gödel's theorems.

3.1.1 Typographical Number Theory

In G.E.B., Hofstadter introduces Typographical Number Theory or TNT, a formal axiomatic system based on the axioms of Peano arithmetic. The formal language of TNT has numerals, variables, system-specific operators, connectives, and quantifiers as follows:

1. **Numerals**: TNT represents the numeral zero by the symbol **0**, one by **S0**, two by **SS0**, and so on. **S** can be interpreted as the successor function which returns one more than whatever follows it.

2. **Variables**: Variables are represented by the symbol **a** followed by any number of apostrophes where symbols with different numbers of apostrophes denote different variables [9]. For example, **a**, **a'**, and **a''** all denote different variables in TNT.

3. **Operators**: The operators (that are neither connectives or quantifiers) in TNT are those for addition, multiplication, and equivalence.

   (a) **Addition**: Addition is represented using the notation $(\mathbf{a} + \mathbf{a'})$. The parentheses are a strict part of the notation, which means the sum of **a**, **a'**, and **a''** would have to be written as $((\mathbf{a} + \mathbf{a'}) + \mathbf{a''})$ or some similar grouping.

   (b) **Multiplication**: Similar to addition, multiplication is represented using the notation $(\mathbf{a} \cdot \mathbf{a'})$. As with addition, the parentheses are a strict part of the notation.

   (c) **Equivalence**: Equivalence is represented using the "Equals" operator using the notation $\mathbf{a} = \mathbf{a'}$.

4. **Connectives**: TNT includes the logical connectives typically used in propositional calculus: negation, conjunction, disjunction, and implication. For the following definitions, let **P** and **Q** denote statements in TNT.

[9] For the sake of this paper, we are discussing "Austere" TNT, as that is the notation that is relevant to this proof.

(a) **Negation**: Negation is represented by the symbol $\sim$ such that $\sim\mathbf{P}$ denotes the statement "**P** is not true."

(b) **Conjunction**: Conjunction is represented by the symbol $\wedge$ such that $<\mathbf{P}\wedge\mathbf{Q}>$ denotes the statement "Both **P** and **Q** are true".

(c) **Disjunction**: Disjunction is represented by the symbol $\vee$ such that $<\mathbf{P}\vee\mathbf{Q}>$ denotes the statement "Either **P** is true, **Q** is true, or both **P** and **Q** are true."

(d) **Implication**: Implication is represented by the symbol $\supset$ such that $<\mathbf{P}\supset\mathbf{Q}>$ denotes the statement "If **P** is true, then **Q** is true."

5. **Quantifiers**: TNT uses the two quantifiers commonly used in mathematical notation today, namely, the universal quantifier, represented by symbol $\forall$, and the existential quantifier, represented by symbol $\exists$. For the following, let **P** be a statement of TNT where **a** is a free variable (it isn't being quantified over already).

(a) **Universal Quantifier**: The notation for the universal quantifier is $\forall\mathbf{a}\colon \mathbf{P}$ which can be interpreted as "for each natural number (non-negative integer) k, **P** is true when k is substituted for **a**."[10]

(b) **Existential Quantifier**: The notation for the existential quantifier is $\exists\mathbf{a}$: **P** which can be interpreted as "there exists a natural number k such that **P** is true when k is substituted for **a**."

Using the language above, all statements of number theory can be expressed in TNT [9][11]. For example, the statement "No sum of two positive cubes is itself a cube," can be written as

$$\forall\mathbf{a}{:}\sim\exists\mathbf{a'}{:}\ \exists\mathbf{a''}{:}\ ((\mathbf{a}\cdot\mathbf{a})\cdot\mathbf{a}) = (((\mathrm{S}\mathbf{a'}\cdot\mathrm{S}\mathbf{a'})\cdot\mathrm{S}\mathbf{a'}) + ((\mathrm{S}\mathbf{a''}\cdot\mathrm{S}\mathbf{a''})\cdot\mathrm{S}\mathbf{a''})).$$

[10]Note that in TNT, the set being quantified over doesn't need to be specified, as quantifiers always quantify over the universal set of TNT, the set of natural numbers.

[11]The proof of this is beyond the scope of this paper, but is necessary for the proof of the incompleteness theorem

In order for TNT to be a formal axiomatic system, however, it still needs both axioms and rules of inference (valid string manipulations). The axioms of TNT can be stated as follows:

Axiom 1: $\forall\mathbf{a}{:}\sim \mathrm{S}\mathbf{a} = 0$;

Axiom 2: $\forall\mathbf{a}$: $(\mathbf{a} + 0) = \mathbf{a}$;

Axiom 3: $\forall\mathbf{a}$: $\forall\mathbf{a'}$: $(\mathbf{a} + \mathrm{S}\mathbf{a'}) = S(\mathbf{a} + \mathbf{a'})$;

Axiom 4: $\forall\mathbf{a}$: $(\mathbf{a} \cdot 0) = 0$;

Axiom 5: $\forall\mathbf{a}$: $\forall\mathbf{a'}$: $(\mathbf{a} \cdot \mathrm{S}\mathbf{a'}) = ((\mathbf{a} \cdot \mathbf{a'}) + \mathbf{a})$.

With respect to the rules of inference, TNT inherits all the rules of inference from propositional calculus (modus ponens, De Morgan's law, etc.), in addition to a few new rules of inference specific to TNT [9]. In order to save space, three of the rules of inference will be provided as examples here. G.E.B. can be consulted for the remaining rules of inference.

1. **Rule of Specification**: Suppose **a** is a variable which occurs inside string x. If the string $\forall\mathbf{a}$: x is a theorem, then so is x and so are any strings made from x by replacing all instances of **a** in x with a term that doesn't contain any variable that is quantified in x.

2. **Rule of Interchange**: Suppose **a** is a variable. Then the strings $\forall\mathbf{a}{:}\sim$ and $\sim\exists\mathbf{a}:$ are interchangeable anywhere inside any theorem.

3. **Rules of Successorship**: Let r, s, and t all stand for arbitrary terms. If $r = t$ is a theorem, then $\mathrm{S}r = \mathrm{S}t$ is a theorem, and if $\mathrm{S}r = \mathrm{S}t$ is a theorem, then $r = t$ is a theorem.

With a formal language, axioms, and rules of inference, TNT is now a full formal axiomatic system. Additionally, given the axioms and rules of inferences used, TNT can be seen as an implementation of Peano arithmetic, and as a result, conclusions about TNT can be applied to Peano arithmetic as well.

3.1.2 Gödel Numbering

The critical concept that Gödel discovered in order to prove the first incompleteness theorem was Gödel numbering. This is the idea of assigning numbers to each symbol in a formal language, and then expressing statements in that language as the concatenation of the "Gödel numbers" of each character in the statement. In G.E.B., Hofstadter introduced the following Gödel numbering for TNT[12].

$$
\begin{array}{rr}
0 \longrightarrow 666 & \mathrm{S} \longrightarrow 123 \\
+ \longrightarrow 112 & \cdot \longrightarrow 236 \\
(\longrightarrow 362 &) \longrightarrow 323 \\
< \longrightarrow 212 & > \longrightarrow 213 \\
[\longrightarrow 312 &] \longrightarrow 313 \\
\mathbf{a} \longrightarrow 262 & ' \longrightarrow 163 \\
\wedge \longrightarrow 161 & \vee \longrightarrow 616 \\
\supset \longrightarrow 633 & = \longrightarrow 111 \\
\forall \longrightarrow 626 & \exists \longrightarrow 333 \\
\sim \longrightarrow 223 & : \longrightarrow 636
\end{array}
$$

Using this Gödel numbering of TNT, Axiom 1, $\forall \mathbf{a}{:}{\sim}\mathrm{S}\mathbf{a} = 0$, can be written as

$$626, 262, 636, 223, 123, 262, 111, 666.$$

Gödel numbers can similarly be assigned to lists of statements by simply concatenating the Gödel numbers of each of the individual statements separated by the digits 611, digits that would represent a character similar to a period or comma. With this, the Gödel number

$$626, 262, 636, 223, 123, 262, 111, 666, \mathbf{611}, 223, 333, 262, 636, 123, 262, 111, 666$$

[12]In the following Gödel numbering, the new characters '[' and ']' are used in TNT proofs in order to seperate sub-proofs from the main proof. In any sub-proof, the first line after the '[' will be the premise of the sub-proof, and the last line preceding the ']' will be the conclusion.

would represent the list of statements in the following proof for the theorem $\sim\exists\mathbf{a}\colon \mathrm{S}\mathbf{a} = 0$.

$\forall\mathbf{a}{:}\sim \mathrm{S}\mathbf{a} = 0$	$626, 262, 636, 223, 123, 262, 111, 666$	Axiom 1
$\sim\exists\mathbf{a}\colon \mathbf{S}\mathbf{a} = 0$	$223, 333, 262, 636, 123, 262, 111, 666$	Interchange.

3.1.3 Proof-Pairs

Let two natural numbers, m and n respectively, form a TNT-proof-pair if and only if m is the Gödel number of a TNT proof whose bottom line is the string with Gödel number n. From the example above, the following natural numbers m and n form a TNT-proof-pair:

$m = 626, 262, 636, 223, 123, 262, 111, 666, 611, 223, 333, 262, 636, 123, 262, 111, 666;$

$n = 223, 333, 262, 636, 123, 262, 111, 666.$

The fundamental idea regarding proof-pairs in the proof of the first incompleteness theorem is that the check for whether a pair of natural numbers is a proof-pair can be done in a finite number of basic steps, or in other words, is primitive recursive. This can be seen in the following program for that check.

1. Go down the lines in the derivation one by one.
2. Mark those which are axioms.
3. For each line which is not an axiom, check whether it follows by any of the rules of inference from earlier lines in the alleged derivation.
4. If all nonaxioms follow by rules of inference from earlier lines, then you have a legitimate derivation; otherwise, it is a phony derivation.

Since the check for whether a pair is a proof-pair is primitive recursive, it can also be seen as an arithmetic relation[13] and is therefore, also represented by a for-

[13]This is because of the definition of how primitive recursive functions are constructed and their relation to arithmetic relations. Discussion of this topic is much too involved for this paper. To learn more, consult the following cited article by UMD Professor of Computer Science, William Gasarch [4].

mula with two free variables in TNT. From this point onwards, let TNT-PROOF-PAIR{**a**, **a'**} denote this formula with the free variables **a** and **a'** representing the respective elements of the pair being checked.

3.1.4 Self-Referential Formulas

Though we now have a means of asking whether a particular proof of a statement is valid, we don't have a means of asking whether a statement is provable. This question can be asked as follows:

$$\exists \mathbf{a}\text{: TNT-PROOF-PAIR}\{\mathbf{a}, \mathbf{a'}\}.$$

This statement can be interpreted as "**a'** is the Gödel number of a theorem of TNT." Note that this statement is not necessarily a primitive recursive property, and therefore, isn't necessarily represented[14] in TNT.

At this point in the proof, Hostadter introduces a concept he calls "arithmoquining." Arithmoquining takes a Gödel number of a statement with free variables, and substitutes each free variable with that Gödel number. For example, consider the Gödel number of the statement "$\mathbf{a} = 0$." It is 262,111,666. The arithmoquinification of 262,111,666 would then simply substitute **a** for the numeral for 262,111,666, which would be 262,111,666 copies of 'S' followed by a '0'. The resulting Gödel number would be 123,123,123,...,123,123,666,111,666 with 262,111,666 copies of 123. Let two natural numbers **a'** and **a''** form a quine-pair if **a'** is the arithmoquinification of **a''**. This can clearly be checked in a finite number of operations, and is therefore, a primitive recursive property that is represented in TNT. Let QUINE-PAIR{**a'**, **a''**} denote this formula that returns whether or not **a'** and **a''** form a quine-pair.

Now consider a statement P{**a''**} that states the following.

$$\sim\exists \mathbf{a}\text{: }\exists \mathbf{a'}\text{:}< \text{TNT-PROOF-PAIR}\{\mathbf{a}, \mathbf{a'}\} \wedge \text{QUINE-PAIR}\{\mathbf{a'}, \mathbf{a''}\} > .$$

[14]The word "represented" in this context is contrasted from the word "expressed". "Expressed" implies the statement can be written in the language. "Represented" implies that either the statement or the negation of the statement is provable in the system. In this context, the statement is expressed, but not necessarily represented in TNT.

This can be interpreted as "There do not exist two natural numbers **a** and **a'** such that **a** is the Gödel number for the proof of the statement whose Gödel number is **a'**, and **a'** is the arithmoquinification of free variable **a"**."

Next, consider a statement P' that is the statement P$\{$**a"**$\}$ but with the Gödel number of P$\{$**a"**$\}$ substituted for **a"**. Let "x/**a**" denote the term x being substituted for a variable **a** and G(P) denote the Gödel number of P$\{$**a"**$\}$. Given that notation, P' would be the statement P$\{$G(P)/**a"**$\}$. This would be stated as follows:

$$\sim\exists \mathbf{a}\text{: } \exists \mathbf{a'}\text{:} < \text{TNT-PROOF-PAIR}\{\mathbf{a}, \mathbf{a'}\} \wedge \text{QUINE-PAIR}\{\mathbf{a'}, \text{G(P)}/\mathbf{a''}\} > .$$

This can be interpreted as "There do not exist two natural numbers **a** and **a'** such that **a** is the Gödel number for the proof of the statement whose Gödel number is **a'**, and **a'** is the arithmoquinification of G(P)."

But there is only one arithmoquinification of G(P), namely, the Gödel number of P$\{$G(P)/**a"**$\}$. This allows the interpretation to be "There does not exist a natural number **a** such that **a** is the Gödel number for the proof of P$\{$G(P)/**a"**$\}$," or in other words, "P$\{$G(P)/**a"**$\}$ is unprovable in TNT."

But the above statement is equivalent to P$\{$G(P)/**a"**$\}$, so finally, the above statement can be interpreted as saying "This statement is unprovable," finally reaching the desired paradox.

If P' is provable in TNT, then P' would have to assert a truth. But P' asserts that it is unprovable in TNT, and this would therefore cause a contradiction, making TNT inconsistent. If P' is instead unprovable in TNT, then there would be no contradiction; however, P' asserts that it is unprovable in TNT, and would therefore, be true. This would make P' a true statement that is unprovable in TNT, hence proving Gödel's first incompleteness theorem for TNT, Peano's arithmetic, and systems that extend Peano's arithmetic.

3.2 Proof Sketch of Second Incompleteness Theorem

Consider the undecidable statement P' constructed above. Assume that the consistency of TNT can be proven from within TNT. The first incompleteness theorem shows that if TNT is consistent, then P' is unprovable in TNT (*). Let P" be the statement P' is unprovable in TNT. If the consistency of TNT can be proven within TNT, then P" can be proven within the system by (*)[15]. But P", by definition, is the same as P', which by (*), is unprovable in TNT. Therefore, there is a contradiction, and the consistency of TNT cannot be proven from within TNT. Similarly, any formal system that is an extension of Peano arithmetic cannot have its consistency proven from within itself [7].

4 Implications

4.1 Computability Theory

In addition to the impacts Gödel's incompleteness theorems had on the study of mathematical systems, the theorems also sparked research in other areas, like computability theory. Specifically, the incompleteness theorems, along with the goals posed in Hiblert's Program, provided a lot of the basis for discussion for Alonzo Church's first proof that the *Entscheidungsproblem* (or goal 4 of Hilbert's Program from above) is impossible, and for Alan Turing's independent proof of the same conclusion using ideas from computability theory and Turing Computers.

In 1937, Turing published his paper, *On Computable Numbers With an Application to the Entscheidungsproblem* [26]. In it, he used what is now known as the "Halting Problem," which asks whether a general algorithm can be created that, in finite time, can determine whether a particular program halts or loops infinitely. In the paper, Turing used a similar form of self-reference to what Gödel used in his first incompleteness theorem in order to show that such a general

[15](*) shows that P' is unprovable in TNT, but that is the statement P", so P" is provable in TNT.

algorithm is impossible. This discovery had significant impacts on the field of computing as it showed that there were certain tasks that no computer, regardless of its hardware capabilities, could possibly perform. Turing also proved a corollary of his theorem that showed that the *Entscheidungsproblem* was unsolvable [26].

4.2 Arguments against Mechanism

Gödel's incompleteness theorems have repeatedly been used to attempt to disprove Mechanism, a philosophy that, in the context of the human mind, effectively claims that the human mind is a Turing machine. In 1951, during his Gibbs Lecture, Gödel himself proposed a sort of anti-Mechanism argument, drawing the following conclusion from the incompleteness theorems.

> "Either mathematics is incompletable in this sense, that its evident axioms can never be comprised in a finite rule, that is to say, the human mind (even within the realm of pure mathematics) infinitely surpasses the powers of any finite machine, or else there exist absolutely unsolvable diophantine problems of the type specified." [13]

Note that a key aspect of the above statement is that it doesn't guarantee the anti-mechanist conclusion. It simply states that the incompleteness theorems either guarantee the anti-mechanist conclusion, or guarantee that unsolvable diophantine problems exist. Though Gödel was inclined to deny the second part of the disjunction, and therefore believed the anti-mechanist conclusion, he did not claim that the conclusion directly followed from his incompleteness theorems [18].

Many others have instead misused Gödel's incompleteness theorem and suggested that the theorems directly disprove mechanism. For example, in 1961, British philosopher, John R. Lucas, published his paper *Minds, Machines and Gödel*, in which he argued that Gödel's incompleteness theorems "prove that Mechanism is false, that is, that minds cannot be explained as machines." [16] His argument can be seen in the following quote.

> "Gödel's theorem must apply to cybernetical machines, because it is of the essence of being a machine, that it should be a concrete instantiation of a formal system. It follows that given any machine which is consistent and capable of doing simple arithmetic, there is a formula which it is incapable of producing as being true—i.e., the formula is unprovable-in-the-system-but which we can see to be true. It follows that no machine can be a complete or adequate model of the mind, that minds are essentially different from machines." [16]

The primary problem with this argument and others like it is that they ignore the conditional form of Gödel's incompleteness theorems and assume that all formalized systems or finite machines have unprovable Gödel statements [18]. Instead, their conclusions require that the system being discussed be consistent, and in turn, in the context of the anti-mechanist argument, that "the human mind can always see whether or not a given formalized theory is consistent." [18] These arguments, however, rarely provide a strong argument for this critical assumption, and, as a result, are usually weak[16].

4.3 Effect on Hilbert's Program

Though Gödel's incompleteness theorems proved that in their complete form, the goals of Hilbert's program were impossible to meet, there has still been a lot of research done along the lines of Hilbert's goals. For example, a few years after the publication of Gödel's incompleteness theorems, in 1936, German mathematician Gerhard Gentzen, gave a proof of the consistency of Peano arithmetic [5]. In line with Gödel's second incompleteness theorem, the proof required the use of methods outside of Peano Arithmetic, namely, transfinite induction up to the ordinal ε_0 [27]. Still, these proofs of the consistency of such strong mathematical systems are very powerful for the study of mathematical systems.

[16] Arguments of similar structure have also been used to "prove" the existence of god; however, due to the same faults discussed in this section, these arguments typically aren't strong.

Research has also been done on the completeness of systems. Though Gödel showed that it wasn't possible to prove the completeness for systems that are extensions of Peano's arithmetic, completeness has been proven for other non-trivial systems. For example, around the time of the publication of Gödel's incompleteness theorems, Tarski proved that under a set of axioms he created called Tarski's axioms, Euclidean Geometry is complete [25]. Years later, Tarski also proved that Euclidean geometry, along with elementary algebra, are both decidable [24][17].

Finally, though Gödel showed that the consistent and complete formal theory **M** can't exist, systems like Zermelo-Fraenkel set theory have come close to meeting this goal by seeming to avoid all known paradoxes, and branching to almost all branches of mathematics.

5 Conclusion

Gödel's two incompleteness theorems can be considered to be among the most influential results to ever be discovered in the study of mathematics. Though the theorems did not put an end to the foundational crisis of mathematics, they disproved the existence of a consistent and complete mathematical system that could span all of mathematics, and changed how mathematicians viewed formal axiomatic systems. Though the limitations that the incompleteness theorems placed on mathematical systems caused a shift in goals, research in axiomatic systems and meta-mathematics has only grown, and there is still much more to be discovered.

[17] In addition to these two proofs, Tarski's work also went into other relevant matters, including analyses of the concept of truth and attempts at defining truth. These ideas are, however, beyond the scope of this paper.

References

[1] Thierry Coquand. Type theory. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2018 edition, 2018.

[2] John W. Dawson. *Logical Dilemmas: The Life and Work of Kurt Gödel*. Peters, A.K., 1997.

[3] Gottlob Frege. *Begriffsschrift*. Halle, 1879.

[4] William Gasarch. Primitive rec, ackerman's function, decidable, undeciable, and beyond.

[5] Gerhard Gentzen. Die widerspruchsfreiheit der reinen zahlentheorie. *Mathematische Annalen*, 112(1):493–565, 1936.

[6] Kurt Gödel. Über formal unentscheidbare sätze der principia mathematica und verwandter systeme i. *Monatshefte für Mathematik und Physik*, 38-38:173–198, 1931.

[7] Kurt Gödel. On formally undecidable propositions of principia mathematica and related systems i. *From Frege to Gödel, A source book in mathematical logic, 1879–1931*, 39:596–617, 1967.

[8] T. H. Hildebrandt, Georg Cantor, and Philip E. B. Jourdain. Contributions to the founding of the theory of transfinite numbers. *The American Mathematical Monthly*, 23(10):381, 1916.

[9] Douglas R. Hofstadter. *Gödel, Escher, Bach: An Eternal Golden Braid*. Basic Books, 1999.

[10] Leon Horsten. Philosophy of mathematics. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2019 edition, 2019.

[11] Andrew David Irvine. Bertrand russell. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2019 edition, 2019.

[12] Casimir Kuratowski. Sur la notion de lordre dans la théorie des ensembles. *Fundamenta Mathematicae*, 2:161–171, 1921.

[13] John W. Dawson Jr. Warren Goldfarb Charles Parsons R. Solovay Kurt Gödel, S. Feferman (Editor-in-Chief). *Collected Works. - Unpublished essays and lectures*, volume Vol. 3 of *Godel, Kurt//Collected Works*. Oxford University Press, USA, 1995.

[14] Øystein Linnebo. Platonism in the philosophy of mathematics. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2018 edition, 2018.

[15] Bernard Linsky and Andrew David Irvine. Principia mathematica. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2019 edition, 2019.

[16] John R Lucas. Mind, machine, and gödel. *Philosophy*, 36:112–127, 1961.

[17] The Editors of Encyclopaedia Britannica. Logicism. In *Encyclopædia Britannica*. Encyclopædia Britannica, inc., Mar 2014.

[18] Panu Raatikainen. Gödel's incompleteness theorems. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2018 edition, 2018.

[19] Borut Robič. *Foundations of Computability Theory*. Springer, 1 edition, 2016.

[20] Bertrand Russell. *The Principles of Mathematics*, volume 1. Cambridge: at the University Press, 1903.

[21] Bertrand Russell. Letter to frege. *From Frege to Gödel, A source book in mathematical logic, 1879–1931*, 39:124–125, 1967.

[22] Bertrand Russell and Alfred North Whitehead. *Principia Mathematica.* Cambridge: at the University Press, 1 edition, 1910-13.

[23] Ernst Snapper. The three crises in mathematics: Logicism, intuitionism and formalism. *Mathematics Magazine*, 52(4):207–216, Sep 1979.

[24] Alfred Tarski. A decision method for elementary algebra and geometry. *RAND Report*, 1951.

[25] Alfred Tarski and Steven Givant. Tarski's system of geometry. *The Bulletin of Symbolic Logic*, 5(2):175–214, 1999.

[26] Alan M. Turing. On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society*, s2-42(1):230–265, 1937.

[27] Richard Zach. Hilbert's program. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy.* Metaphysics Research Lab, Stanford University, fall 2019 edition, 2019.

[28] Edward N. Zalta. Gottlob frege. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy.* Metaphysics Research Lab, Stanford University, winter 2019 edition, 2019.