CS 4510 Automata and Complexity March 8th 2023

Lecture 13: Foundation of Mathematics

Lecturer: Abrahim Ladha *Scribe(s): Yitong Li*

1 A Motivation from Geometry

Recall last time we had a lecture on pure mathematics, countability, and set theory. The lecture before that was kind of pseudo-philosophical on the Church-Turing thesis. The lecture before that one was on engineering, programming and understanding the Turing machine. To keep the trend of not having one, today's lecture will be on history.

We will go from 300 BC to 1936. We begin of course, with the Greeks. Around 300 BC, Euclid wrote "The Elements", several treatises in geometry. It is one of the most influential texts of all time. It established mathematics as a deductive rather than empirical science. It has been in print for millennia and comes second only to the bible. Just because calculus wasn't invented yet didn't mean you didn't have math class. You used to have to take a three course series on classical geometry.

Recall that a theorem is some statement proved. From what? Other theorems? Not quite. There is some flow of implications in this giant tree of knowledge. Follow back to eventually reach some root: The axioms. An axiom is a statement that needs no proof. It may be assumed to be true. It is usually so trivial to be anything but true. For example, associativity of addition: $(a + b) + c = a + (b + c)$.

Euclid defined[1] his first five axioms, or postulates as follows:

1. any two points may be connected by a line segment.
2. any line segment may be extended infinitely in both directions.
3. for any point and radius, there exists a circle.
4. all right angles equal each other
5. given a line l and a point p, not on that line, there exists exactly one line through p parallel to l.

A proof is an application of axioms with the "rules of deduction" which are themselves axioms. All of the axioms for Euclid's elements are a model for what we now call Euclidean geometry. From the axioms, you can prove things like: every square has four equal right angles, the sum of the interior angles of a triangle is $180°$, if a triangle has two equal angles it has two equal sides, and so on.

Euclid's elements have nothing to do with geometry. It is about rigorous and systematic thinking. It is nothing more than a by-product of the school of thought that Euclid and other

[1] Of course, he did it in ancient Greek. Here they have been modernly rephrased. Look up Playfair's axiom if you are interested in this rephrasing.

Greeks came from. Plato's Theory of Forms asserts that ideas are the supreme achievement of human beings. They are a refined, pure reflection of the capability of the human mind. There exists the Real: the material, empirical, measurable, and approximate world. There also exists the Ideal: one of concepts and thought. The Real and Ideal certainly have a duality[2]. This school of thought asserts that the world, the materiality, is shaped by some things prior to it, the immaterial. When I draw a triangle on the board, understand this exists no where except in your mind. No Real triangle you can form from sticks, or by drawing in the sand can ever reach the precision of the Ideal triangle. However, by studying the Ideal, it may reveal to you something about the Real. To understand the material, you only need to understand the non-material. Abraham Lincoln famously used the Elements to train as a lawyer.

> At last I said, 'Lincoln, you never can make a lawyer if you do not understand what demonstrate means'; and I left my situation in Springfield, went home to my father's house, and stayed there till I could give any proposition in the six books of Euclid at sight. I then found out what demonstrate means, and went back to my law studies.

In 1854 in an unpublished note, he used this rigorous thinking to assert abolition.

> If A. can prove, however conclusively, that he may, of right, enslave B. — why may not B. snatch the same argument, and prove equally, that he may enslave A? – You say A. is white, and B. is black. It is color, then; the lighter, having the right to enslave the darker? Take care. By this rule, you are to be slave to the first man you meet, with a fairer skin than your own. You do not mean color exactly? – You mean the whites are intellectually the superiors of the blacks, and, therefore have the right to enslave them? Take care again. By this rule, you are to be slave to the first man you meet, with an intellect superior to your own. But, say you, it is a question of interest; and, if you can make it your interest, you have the right to enslave another. Very well. And if he can make it his interest, he has the right to enslave you.

Millennia was spent trying to refine Euclid's Elements, to show they were only as good and simple as necessary. A set of axioms is independent if no axiom can be proved from the others, like independence with respect to a basis of a vector space. If an axiom could be proved from the others, then it need not be an axiom. Remove it, and simply take it as a theorem. The fifth axiom took a lot attention as if it was the first really unobvious one. The first four are just definitions. Let A be the parallel postulate and EE the axioms.

First, people tried to see if you could derive the parallel postulate from the others. Notationally, we would present this as $(EE - A) \vdash A$? Here $\vdash$[3] means provable from a set of axioms. It is similar to an implication. We now know the fifth postulate is independent, so this is impossible.

Secondly, we wanted to see if it was even a necessary axiom. So there were attempts to prove that $(EE - A + \neg A) \vdash 0 = 1$. That is, if you removed the axiom and assumed its

[2] Maybe are better known by other names to you, theory and practice?

[3] latex is `vdash`

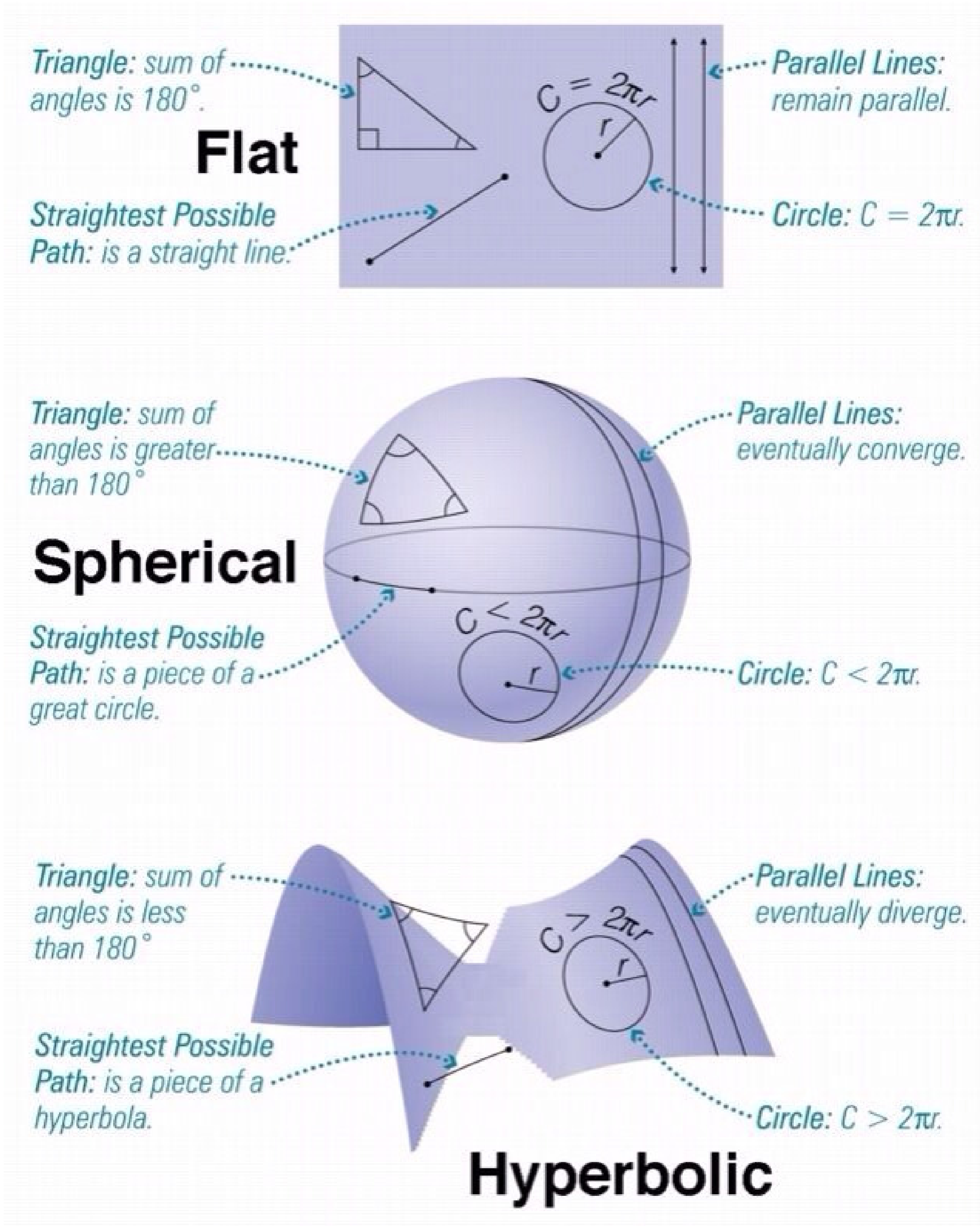

Figure 1: Three models of geometry, and their different theorems

negation, you would produce a contradiction. This should work for any usable and necessary axiom, but they discovered something insane. Taking the negation doesn't produce an inconsistency, instead it produces two different consistent models!

Recall the parallel postulate says "Given a line and a point not on that line, there exists **exactly one** line through the point parallel to the line." Our negations will be changing "exactly one" to "greater than one" or "less than one". We are changing the number of possible parallel lines from $= 1$ to > 1 or < 1.

If the number of parallel lines through the point equals one, we live in a flat world, the Euclidean plane. Any other parallel lines would be equivalent since they intersect at all points. If the number of parallel lines were greater than one, then we exist not on the plane, but on a saddle, or a pringle. There exists many "parallel" lines through the point which do not intersect our line. This is also called hyperbolic, or Lobachevskian geometry. If the number of parallel lines is less than one, as in there do not exist any parallel lines anywhere ever, we are in the model of spherical or ellipsoid geometry. We are embedded onto an egg, or globe. The lines on a sphere are only the great circles, smaller bands of latitudes are curves.

Note that these are consistent models. They cannot prove $0 = 1$ but all the theorems which these models derive are slightly different. For example, in spherical geometry, a triangle has the sum of its interior angles $> 180°$. You can construct a triangle with three right angles, something impossible to do in the Euclidean plane. Begin at the equator, go to the north pole, turn 90 degrees, go back to the equator, and turn 90 degrees again to go back to where you began. On a sphere, a triangle has the sum of its interior angles equal to $180°$ if and only if has area zero[4]. Euclidean geometry is supposed to be "intuitive," but we discovered it was on all shaky foundations. We formulated Euclidean geometry following our empirical experiences in the Real, measuring angles and generalizing our observations. Maybe the Real could follow these models instead? Who can say? How can we know that every time weve measured the angles of a triangle, it hasn't technically been $180 + \varepsilon$ this whole time? We do live on a sphere after all. This sparked more serious concerns about the foundation of all of mathematics.

2 A Foundation from a Theory of Sets

A modern first attempt of formalizing mathematics was done by Gottlob Frege's "Begriffsschrift". It builds off of previous work by Aristotle's Organon and Leibniz. It identifies as "a formal language modeled on that of arithmetic, for pure thought". He also created many of our rules of deduction, for example

- $A \implies \neg\neg A$
- $c = d \implies f(c) = f(d)$
- $(A \implies B) \wedge A \implies B$[5]

[4]In fact, the fifth postulate is logically equivalent to "there exists a triangle whose interior angles sum to 180°"

[5]This is called modus ponens

Theses are just a few of the rules of deduction noted by Frege, that we now use for propositional logic. You may use these without even realizing they are axioms.

Set theory was a natural foundation for mathematics, as they can derive numbers but also greater and more interesting structures. The motivation was to create a unifying foundation for all of mathematics. Bertrand Russell noticed the following issue, applied to many axiomatic systems. Suppose we have the following axioms for some simple theory of sets

$$\forall x \forall y [\forall z (z \in x \Leftrightarrow z \in y) \Rightarrow x = y] \tag{1}$$

The axiom of extensionality simply defines the equality relation of sets. Two sets are equal if they contain the same elements.

$$\exists y \forall x [x \in y \Leftrightarrow \varphi(x)] \tag{2}$$

For any predicate φ, the axiom of unrestricted comprehension basically freely allows you to define any set you want. It allows you to make statements like "let y be the set of primes, or horses or whatever". If you can define it with a logical predicate, then there exists a set of those elements. As an example of a predicate, the following is one for the prime numbers

$$Prime(x) = \forall x \neg \exists z [(x > 1) \wedge (z \leq x) \wedge \neg(z = 1) \wedge \neg(z = x) \wedge \neg(z|x)] \tag{3}$$

We however, are in a theory of sets. We can construct the numbers from the sets, and we will describe how to later. But for now consider predicates over sets. This generality of unrestricted comprehension is also our fragility. Consider the set of all sets which do not contain themselves. Let $\varphi(x) = x \notin x$. A perfectly valid predicate. By the axiom of unrestricted comprehension, we see that

$$\exists y \forall x [x \in y \Leftrightarrow x \notin x]$$

Since it's true $\forall x$, we may specify, and consider the case for one selection of x. What happens for $x = y$?

$$y \in y \iff y \notin y$$

A contradiction![6] We are not in a proof by contradiction, yet we have derived a contradiction. We have shown that Frege's axiomatic system was capable of producing an inconsistency. Although this attack was devastating to Frege's words, it did not deter much the resolve of the formalists. The formalists are a school of thought revolving around building Hilbert's program. They seek a secure, rigorous, and logical foundation in which to secure all of mathematics. Roughly with these goals:

1. All math written in a precise formal language manipulated according to well-defined rules

2. Completeness: all that is true is provable

[6] see Logicomix 162-171

3. Consistency: you should provably be unable to obtain a contradiction
4. Decidability: There should exist an algorithm to decide the truth value of any statement.

The most significant effort in this regard was by Russell and Whitehead, They spent decades and thousands of pages to build up a "Theory of Types." To give a quick summary of twenty years of work, they hoped to avoid self-reference by using these types. Anything of some type i is unable to construct sentences which reference other things of type i (including itself). They thought removing self-reference would create a strong and bulletproof foundation. The axiom of unrestricted comprehension was modified to the axiom of restricted comprehension. Now you can only construct subsets of other sets. By restricting comprehension, we appear to avoid Russell's paradox.

- unrestricted comprehension: $\{x \mid \varphi(x)\}$
- restricted comprehension: $\{x \subseteq z \mid \varphi(x)\}$

Is this system useful? Let's roughly prove $1 = 1$. First let $\emptyset$ exist by the axiom of restricted comprehension with some useless $\varphi(x) = \{x \in y \mid (x \in x) \wedge \neg(x \in x)\}$. Nothing satisfies it so we create $\emptyset$. Sometimes the empty set exists axiomatically, but here it follows from restricted comprehension. Now let $0 := \emptyset$, for shorthand. Let $S(w) := w \cup \{w\}$ also for shorthand. Notice $S(\emptyset) = \emptyset \cup \{\emptyset\} = \{\emptyset\}$. Lets denote that as one. Namely $1 := S(0) = S(\emptyset) = \emptyset \cup \{\emptyset\} = \{\emptyset\}$. We may now apply the axiom of extensionality. Is it true $\forall z,\ z \in \{\emptyset\} \iff z \in \{\emptyset\}$? Yes, then this implies that $1 = 1$. Supposedly to prove if $1 + 1 = 2$ took Principia Mathematica three hundred and seventy two pages. This $1 = 1$ example is also from the axioms of ZFC[7] technically. Could PM serve as a suitable foundation? Free of issue? or somewhat incapable?

3 Gödel Incompleteness

Godel showed the futility of Russell and Whitehead's effort. We say an axiomatic system is

1. Complete: if $\forall p$, there exists a proof of p if it's true, or a proof of $\neg p$ if it's false. This asserts provable $\iff$ true.
2. Consistent: $\forall p$ there exists a proof of $p \wedge \neg p$. Every statement is exactly true or exactly false. No statement can be false and true simultaneously.

A system being complete means in some sense, it is "total". From the axioms, all statements are provable. There is no theorem which requires some missing secret axiom. It also asserts if something is true, there must exist a proof of it, and a way to deduce such a proof.

A system being consistent is the bare minimum requirement for it being useful. This asserts you cannot prove $0 = 1$. If you could, then everything follows trivially.

[7]ZFC is a modern conservative axiomatic set theory, which stands for Zermelo-Fraenkel plus the axiom of Choice. We didn't use the axiom of choice here

For years and years, Russell and Whitehead searched for a proof that PM was consistent. Since PM was intended to serve as a foundation of all of mathematics, they were trying to show $PM \vdash Con(PM)$.

3.1 Gödel's First Incompleteness Theorem

Godel show for any axiomatic system, including PM. It cannot be both consistent and complete. Rephrasing: There does not exist a complete and consistent axiomatic system with sufficient arithmetic. Let's proceed with the proof.

Let $Dem(p, r) = 1 \iff p$ is a proof of r[8]. Here, Dem stands for demonstrates. Let

$$g = \neg \exists p \; [\; Dem(p, g) \;]$$

In human words, g says "I am not provable" or "There does not exist a proof of me." Notice the self-reference. Since by our assumption, our axiomatic system is complete and consistent, g is one of provably true or provably false. We have two cases.

1. g is provably true. Then g asserts there is no proof of g. Then g is true and not provable. Now there exists a true but unprovable statement, so we are incomplete.

2. g is provably false. Then $\neg g$ is provably true.

$$\neg g = \neg\neg \exists p \; [\; Dem(p, r) \;] = \exists p \; [\; Dem(p, r) \;]$$

   So $\neg g$ implies there exists a proof of g, so g is provably true. Then $(\neg g \wedge g)$ is provably true and we are inconsistent.

3.2 Gödel's Second Incompleteness Theorem

Not only does Gödel says that achieving a complete and consistent axiomatic system that is strong enough impossible but any system is capable of proving its own consistency. Logically, for any axiomatic system AS:

$$AS \nvdash Con(AS)$$

The consistency of AS cannot be proven from within AS. Assume to the contrary $AS \vdash Con(AS)$. That there exists a proof of the consistency of AS from within AS. Let this proof be denoted as C. Since the proof of Gödel's first incompleteness theorem assumes the consistency of AS, we may replace this assumption with the proof C. Then we proceed and observe $C \implies g$, our diagonal sentence. Since we can construct g, then AS was not simultaneously consistent and complete, a contradiction. No system is capable of proving its own consistency.

It turns out that some toy systems can be complete and consistent, but they cannot prove their own consistency. You need to use techniques from outside the system to prove

[8]Technically, Gödel encoded proofs, statements, everything as numbers using a Gödel numbering, and then defined these functions based off of those numberings. Pedantically, it should be said like "p is a number which represents a proof of a statement which is represented by the number r". To prove this statement was constructible from simple arithmetic, he had to build it up using some forty formula.

them. If PM was attempting to be a model for all of mathematics, then there there is no "outside" and such a proof of the consistency of PM is unprovable. Not only were the formalists, Russell, Whitehead, and Hilbert losers, they were double losers. The proof they spent decades searching for could never exist.[9]

4 Turing's Undecidable

Alan Turing takes a class foundations, where he learns Gödel Incompleteness. It also contains a description of a large, unsolved problem we will call "Hilbert's decision problem" or the "Entscheidungsproblem". It is phrased as "Give a procedure that takes as input a statement, and returns yes/no if it's always true or always false." Hilbert genuinely believed there were no unsolvable problems. Turing was twenty two when he gave a negative answer. First, he had to formalize the notion of computation, and to do so, he invented what we now call the Turing machine. Next, he described the Church-Turing thesis to convince us that this definition was in fact universal.

Recall that a language $L \subseteq \Sigma^*$ is decidable if there exists a Turing machine M such that

$$x \in L \iff M \text{ accepts } w$$
$$x \notin L \iff M \text{ rejects } w.$$

We can rephrase Hilbert's decision problem as "give a process to decide every language". Following the Church-Turing thesis, the decidable languages give a characterization of the concept of an "algorithm". A purely mechanical process in which a decision on yes or no is reached. If every language is decidable, then there exists an algorithm to solve every problem. There do not exist any unsolvable problems. Could every language be decidable? Every problem be solvable? Turing said no, there exist undecidable languages. He did so in two ways.

First, notice that the languages decidable by a Turing machine are countable. Each language is decidable by many deciders, but one decider decides each language. Let D be the set of all deciders and $\mathscr{L}_D(TM)$ be the decidable languages. We may map each decidable language to exactly one of its deciders to see $|\mathscr{L}_D(TM)| \leq |D|$. By the typewriter principle, we see that $|D|$ is countable. Therefore, so must be the decidable languages.[10] Next observe that the number of languages is uncountable. If $L \subseteq \Sigma^*$ then $L \in \mathcal{P}(\Sigma^*)$. Since $|\Sigma^*|$ is countable, $|\mathcal{P}(\Sigma^*)|$ is uncountable, by Cantor's Theorem. There exists no surjection $\mathscr{L}_D(TM) \to \mathcal{P}(\Sigma^*)$ so there must exist undecidable languages, infinitely many in fact. Most languages are undecidable using this simple nonconstructive, counting proof.

Next, Turing showed constructively that there exist real concrete unsolvable problems. Define $HALT$ as

$$HALT = \{ \langle M, w\rangle \mid M \text{ halts on } w\}.$$

$HALT$ is a language of pairs of encodings of Turing machines and possible inputs, where $\langle M, w\rangle \in \text{HALT} \iff M$ halts on input w. We show that $HALT$ is not decidable. This

[9] See Logicomix 283-286

[10] I previously left this as an exercise for you to show $|\mathscr{L}(NFA)|$ was countable.

means there is no general algorithm to decide if a Turing machine will halt on an input! A provably unsolvable problem.

Assume to the contrary that $HALT$ is decidable. Then there exists a Turing machine $H(\langle M \rangle, w)$ on input $\langle M \rangle$ and w always correctly says yes/no if M halts on w. Notice that since H is a decider, it always returns and never loops. We give a visual proof, representing H like an API or some IDE plugin or something. The middle circuit is the decider for H. We build D around using calls to H, like H is its subroutine. D takes in one argument and passes it to both arguments of H. Then if H returns true, D infinitely loops. If H returns false, then D simply returns and halt.

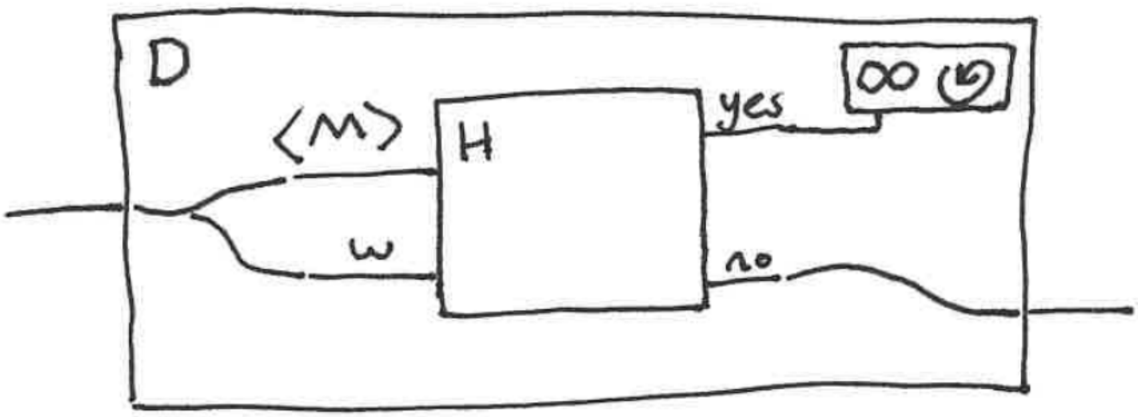

Figure 2: The Halting Problem

In pseudocode, D is doing the following with H:

```
def D(M):
    if H(M,M):
        while True:
            continue
    else:
        return
```

What is D on input $\langle D \rangle$? $D(\langle D \rangle)$? There is no problem with asking this question. We may run the code of a machine on the machine itself with no problem. Compilers can compile themselves. We have two cases.

- $D(\langle D \rangle)$ halts $\iff$ $H(\langle D \rangle, \langle D \rangle)$ returns true $\iff$ $D(\langle D \rangle)$ loops
- $D(\langle D \rangle)$ loops $\iff$ $H(\langle D \rangle, \langle D \rangle)$ returns false $\iff$ $D(\langle D \rangle)$ halts.

A contradiction. No decider for H can exist and we see it is undecidable.

5 Conclusion

We have given three proofs in three different settings. If you have a keen eye, you may note these are really all the same proof. They all have the same structure. They are all diagonalization over different settings. Sets, logical formulas, and decidable languages. These three proof all have the telltale common features of a proof by diagonalization. There is some negation, and some self-reference, or diagonal. A set not containing itself, a formula saying something about its own unprovability, or a machine contradicting being run on its own code.

6 Moral of History

What is the moral of the history here? We should be incredibly thankful that Hilbert's program failed. Had it succeeded, mathematics would have been drained of all its creativity. There would exist perfect automatic theorem provers. All of mathematics, all of the complex and beautiful technical arguments could be reduced to symbolic manipulation. In Formulario Mathematico, Peano develops a symbolic language for mathematics. He says

> Each professor will be able to adopt this Formulario as a textbook, for it ought to contain all theorems and all methods. His teaching will be reduced to showing how to read the formulas, and to indicating to the students the theorems that he wishes to explain in his course.

Mathematics is an ancient, and didactic, and even dramatic tradition. You sit in front of a board and a lecturer like humanity has for millennia. Reduction of this art to something as mechanical as a combine harvester, reaping theorems, is controversial, putting it politely.

Leibniz, centuries before the foundational crisis in mathematics made a similar remark on this mechanization. He noted that ideas were compounded from some "alphabet of human thought". He also remarked that complex ideas proceed from these by a process analogous to arithmetical multiplication.

> It is obvious that if we could find characters or signs suited for expressing all our thoughts as clearly and as exactly as arithmetic expresses numbers or geometry expresses lines, we could do in all matters insofar as they are subject to reasoning all that we can do in arithmetic and geometry. For all investigations which depend on reasoning would be carried out by transposing these characters and by a species of calculus.

This is Leibniz's motivation to build some of the first mechanical calculators. I am personally thankful that the mechanization of mathematics failed. Otherwise, I would not have this job. Some of Hilbert's program has been salvaged. The consensus is that ZFC forms a safe and conservative foundation for much of the usable parts of mathematics. This is independent of Gödel's theorems, which say that ZFC could never be not both complete or consistent.

A second moral is to not bet against the youth. Russell was 29 when he showed his Paradox. Frege was 53. Russell chose to go down the same path, attempting to build a system that Frege could not. Gödel was 24 when he showed his incompleteness theorems. By then, Russell had aged to 59. Turing was 24 when he proved the existence of unsolvable problems. Hilbert was 74.[11] It can become easy to become entrenched in your own ideas for decades. All it may take someone younger to come in with a different perspective.

7 Further Reading

- The referenced graphic novel is called Logicomix. A copy from the internet archive can be found here
https://archive.org/details/Logicomix-Comic-EarlyLifeOfBertrandRussell
- One of the best in depth proofs of Gödel Incompleteness is from https://evoniuk.github.io/Godels-Incompleteness-Theorems/index.html
- Gödel's original proof is not too beyond your ability. A translated copy (the original was in German) may be found in THE UNDECIDABLE, a collection of basic papers edited by Martin Davis
- The construction of the naturals from axiomatic set theory was done following the Von Neumann ordinals. Following our sucessor function $S(w) = w \cup \{w\}$, The first few are
  - $0 : \emptyset$
  - $1 : \{\emptyset\}$
  - $2 : \{\emptyset, \{\emptyset\}\}$
  - $3 : \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$
  - $4 : \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}\}$

  There are other constructions of the naturals as well, from set theory.
- Lobachevsky is equally remembered for his model of geometry, as he was for accusations and rumors of plagarism. Tom Lehrer wrote a song about it.
https://www.youtube.com/watch?v=gXlfXirQF3A
- In this worksheet, I go into a slightly more technical proof of Gödel's theorems https://ladha.me/files/sectionX/godel.pdf.
- There is also a video on my youtube channel here
https://www.youtube.com/watch?v=VpehBGEenWY
- There is a popular science book called Gödel, Escher, Bach which argues that self-reference something something consciousness.

[11]These may be off by a year or two since I don't want to count months.