Gödel's Incompleteness Theorem

Patlin, Gryphon
College of Computing
Georgia Institute of Technology
Atlanta, Georgia
gpatlin3

Kuperman, Hunter
College of Computing
Georgia Institute of Technology
Atlanta, Georgia
kup

Karthikeyan, Mukilan
College of Computing
Georgia Institute of Technology
Atlanta, Georgia
mkarthikeyan26

Abstract—**The expansion of mathematics has led throughout its history to discussions of its limits. Gödel's Incompleteness Theorem demonstrates a key boundary to mathematical understanding by using self-reference to disrupt the concept of completeness. In what amounts to a mathematically rigorous version of "This function is not provable", Gödel demonstrates that mathematics cannot uncover the truth values of all propositions. This finding has been foundational to a myriad of discoveries and is at the core of the modern perspective on mathematics.**

I. Introduction

A. The Piña Colada Situation

Imagine, for a moment, that Professor Brito and his wife, Lianet, have made too many of their famous non-alcoholic piña coladas.

In an effort to free up fridge space, Brito has come up with a plan to drink more piña coladas! Each day, Brito will choose a random person, let's call them P, and watch them throughout the day. As he watches P, Brito will adhere to the following rules:

- Anytime P is not drinking a piña colada, Brito will be drinking a piña colada
- But to lower his sugar intake, anytime P is drinking a piña colada, Brito will not be drinking a piña colada

One day, Brito wakes up in a silly mood and decides that he will choose to watch himself for the day. Wait, but how does he follow his rules?

If he isn't drinking a piña colada then P isn't drinking piña colada (as P := Brito), which means that he must be drinking a piña colada to follow the rules. But if he is drinking a piña colada then P is drinking piña colada (as P := Brito), which means that he must not be drinking a piña colada to follow the rules.

Brito thinks long and hard about what to do, but ultimately decides that there's no way to follow his rules in this situation; his system of rules is incomplete.

The next day, he throws away the remaining piña coladas.

B. From Drinks to Mathematics

Brito's dilemma, while amusing, hints at a much greater philosophical problem: the paradox of self-reference. When logical systems create statements that refer to themselves, paradoxes often emerge. In the case of Brito's piña colada situation, self-reference was introduced when Brito used his own state to make decisions his own state.

If we can formalize another logical system (mathematics, for example) in a way that allows for self-reference, we can create a similar paradox. Gödel achieves this by using Gödel numbers to represent mathematical statements. From there, he formally constructs the self-referential statement, "this statement cannot be proven," which is often labeled G.

If G is false, then it can be proven that G cannot be proven; this cannot be the case, because it yields a contradiction. That means that G must be true, and there exist mathematical statements, namely G, that cannot be proven (i.e. mathematics is incomplete).

II. Background

Naturally, the previous few sections provide some nice intuition towards Gödel's Incompleteness Theorem, but do not provide a rigorous proof. We offer such a proof in Section III, to formalize the notion of self-reference. However, before we can begin, we must cover some necessary background information on completeness and Gödel Numbering.

A. Completeness

In order to understand Gödel's Incompleteness Theorem, must first to define what it means for a formal system to be complete (and what a formal system is, in the first place).

A formal system is a system of axioms combined with rules of inference which enables the creation of new theorems. A consistent system, the only kind of system relevant here (as inconsistent systems are trivially complete), is one in which a theorem and its negation can't both be proven (i.e. the Law of Non Contradiction is maintained). A formal system is complete if every statement or its negation can be derived in the system.

B. Gödel Numbers

In order to formalize self-referential statements in mathematics, we utilize Gödel Numbering. Let k mathematical symbols correlate each to a natural number from 1 to k, and x variables correlate each to a prime number greater than k as such:

We seek to assign a unique numerical representation to each mathematical formula (aka. number them) so that we can analyze the properties of the set of all formulas.

We can assign each mathematical formula of length l a number by creating the prime factorization:

Symbol	Number	Symbol	Number
'0'	1	'¬'	9
'′'	2*	'∀'	10
'+'	3	'x'	11
'x'	4	'y'	13
'='	5	'z'	17
'('	6	'a'	19
')'	7	'b'	23
'→'	8	'c'	29

*Note: This symbol denotes successor such that '0′' = '1'

$$2^{i_1} * 3^{i_2} * 5^{i_3} * ... * p_l^{i_l}$$

Where each i represents the number correlating to each symbol in the mathematical formula. For example, the formula 0=0 can be represented with $2^1 * 3^5 * 5^1$ which is equivalent to the number 2340. Because of the Fundamental Theorem of Arithmetic, each formula will be given a different number as each will have a unique prime factorization. We will denote the Gödel Number of a Formula A with $\lceil A \rceil$

The next key ingredient to Gödel's proof is the meta-analysis of symbols made possible through Gödel numbering. For example, if we wish to discuss the set of formulas that begin with the ¬ symbol, we could build the set of relevant Gödel numbers with:

$$\{x \in \mathbb{N} \ : \ 2^9 \mid x \ \wedge \ 2^{10} \nmid x\}$$

This logic carries forward into functions that transform formulas into their negations ($f(\lceil A \rceil) = \lceil \neg A \rceil$), send two formulas A and B to their implication ($impl(\lceil A \rceil, \lceil B \rceil) = \lceil A \rightarrow B \rceil$), and so forth for all syntactic operations and analyses. There are even functions that identify whether given formulas are truly given axioms and rules of inference. These functions can be chained and combined into a function that identifies whether a series of formulas exists such that a given formula can be proven. We will call this formula $Prov_F$.

CODE: For more exploration with converting expressions to Gödel numbers refer to the github page: https://github.com/MukilanKarthikeyan/GodelNumbers

C. Introducing Self-Reference

The final ingredient to Gödel's Theorem is a self-reference. Let us analyze exclusively formulas with a single variable (Without Loss of Generality, we will use x). Gödel numbering will allow us to identify a unique representation number for each formula, and each formula will contain at least one prime raised to 11, the number representing x:

$$2^{i_1} * 3^{i_2} * 5^{i_3} * ... * p_x^{11} * ... * p_l^{i_l}$$

For example, the formula x=x′ can be represented as $2^{11} * 3^5 * 5^{11} * 7^2 = 1.1907 * 10^{15}$. As we can substitute any number for x and still have a valid formula, in any equation with x, we can substitute all powers representing x (All powers of exactly 11)

with a different number. To achieve self-reference, we simply substitute in a formula's Gödel number into itself:

$$x = x' \rightarrow 1.1907 * 10^{15} = 1.1907 * 10^{15} + 1$$

$$B(x) \rightarrow B(\lceil B \rceil)$$

While in the above operation this self-reference is arbitrary, as the equation x=x′ has no concept of the meaning of Gödel Numbering, this substitution becomes critically important when the formula being altered can make reference to the rules of Gödel Numbering.

III. PROOF

The proof of the Incompleteness theorem uses a self-referencing Gödel function which in Layman's terms translates to "This statement is not provable" to demonstrate that mathematics is incomplete. We will begin with the construction of this statement.

Let A(x) be an arbitrary function and S(x, y) be the function that determines whether y is the Gödel Number created by substituting x into itself (ie. $S(\lceil H(x) \rceil, \lceil H(\lceil H(x) \rceil) \rceil)$ is true where H(x) is an arbitrary function). B(x) is defined:

$$B(x) = \exists y(A(y) \wedge S(x, y))$$

Note that this essentially translates, given the rules of Gödel numbering, to there is a y which is the number of a self referencing Gödel formula (ie. $y = \lceil H(\lceil H(x) \rceil) \rceil$) and A(y) is true. We will now substitute $\lceil B(x) \rceil$ (Represented here as b) for x to yield:

$$B(b) = \exists y(A(y) \wedge S(b, y))$$

Because Gödel numbers are unique, we can identify that the only valid y for this equation is $\lceil B(\lceil B(x) \rceil) \rceil$. For sake of simplicity, we will denote the self-referential $B(\lceil B(x) \rceil)$ as D, such that $y = \lceil D \rceil$.

By substituting D under the laws of F, we get:

$$F \vdash D \leftrightarrow A(\lceil D \rceil)$$

Note that $\vdash$ represents "proves", and $\leftrightarrow$ replaces the equals sign from the prior equation. Since A is an arbitrary function, we can substitute $\neg Prov_F$ for A. We will call this new statement G:

$$F \vdash G \leftrightarrow \neg Prov_F(\lceil G \rceil)$$

In layman's terms, this translates to: F proves that "This Statement is not provable under F" where G is the statement in quotes. From this point, it is not difficult to identify the contradiction through which Gödel identifies Incompleteness. If G is true, than it is not provable. If G is false, than it is provable and, thus, true. This is a contradiction; therefore, F is incomplete.

IV. DISCUSSION

A. Implications for Mathematics

On a broad scale, Gödel's results have tremendous ramifications for mathematics as a whole. Over the course of this class, we've discussed, referenced, or touched upon dozens of open problems. The Twin Prime Conjecture, the Riemann Hypothesis, the Collatz Conjecture, and NP=P are just a few of the thousands of currently-unproven mathematical theorems. Solving just one of these (consider NP=P, for example) would have a drastic effect on mathematics and the sciences as a whole, likely leading to a surge in technological progress and development.

Yet, Gödel's special statement indicates that any consistent system of mathematics contains true statements that cannot be proven! As such, not only might we be currently lacking the appropriate mathematical intuition or lemmas to prove any of the above statements; it might be physically impossible to prove those statements under any system of mathematical reasoning at all.

No matter how many careers are spent analyzing the Twin Prime Conjecture or the Riemann Hypothesis, we may never be able to confidently convince ourselves that these statements are true. We may never be able to use them to logically intuit other facts about mathematics. We may never be able to build off of them to generate new knowledge. We may never be able to utilize them to model the natural world. Our system of mathematics—any system of mathematics—cannot be complete.

B. Gödel's Second Proof

After publishing his first proof on the incompleteness of any mathematical system, Gödel went on to use his established numbering system to prove that any consistent mathematical system that contains Peano Arithmetic (basic arithmetic operations on the natural numbers) can only prove its own consistency if it is, in fact, inconsistent. This serves as an additional blow to mathematics and demonstrates that we have no way of demonstrating that our mathematical systems are free of logical errors.

In order for a set of axioms to prove its own consistency (in other words: prove that the system will never create a contradiction given the constraints of the system) there must exist a sequence of statements that lead to the logical equivalence equivalent of "The system of axioms in which we operate is consistent." Given Gödel's First Incompleteness Theorem, however, we know that a consistent system is incomplete which is the same as the formula for G, where G is logically equivalent to "there exists such a statement that is a true formula and it cannot be proved." Now having established that a set of axioms can never prove G, Gödel created a proof by contradiction. In order for a set of axioms to prove that they were consistent they would have to be able to prove G which we have established is not possible. This would mean that no set of axioms can prove that they themselves are consistent and thus Gödel's Second Incompleteness Theorem marked the end of the search for a mathematical system that was both complete and consistent.

C. Hilbert's Program for Mathematics

During the early 1900s, many mathematicians were re-examining earlier mathematical proofs and were horrified to find multiple logical errors in their reasoning. To solve this problem, David Hilbert (who invented a space-filling curve, along with some other neat stuff) laid out his vision for a formal, unambiguous method of proofs called Hilbert's Program. This system lays out the rules, axioms, and propositional logic that we use for mathematical proofs to this day. Hilbert hoped that this new formal, unambiguous system of mathematics would satisfy three basic properties:
noitemsep
1) *Complete:* all true statements can be proven by the system
2) *Consistent:* no contradictions exist within the system
3) *Decidable:* there exists an algorithm that can decide the truth or falsehood of every mathematical statement

Gödel's first incompleteness theorem disproved the "complete" quality, demonstrating that a consistent mathematical system cannot be a complete one. Gödel's second incompleteness theorem demonstrated that it is impossible for any consistent system of mathematics to prove its own consistency. Turing's work on the halting problem later put the third desired quality, decidability, into question.

D. The Halting Problem

During and after the publications of Gödel's theorems, Alan Turing was studying a similar problem in computer science. Turing envisioned a machine that could read, write, and move between inputs from an infinitely long set of 1s and 0s using instructions from a state-based program; the device he imagined can actually do everything that a modern computer can, a property called Turing-completeness, but proving it is perhaps beyond the scope of this project.

Programs fed to this machine could either run infinitely or stop after some finite number of instructions. Turing was trying to determine if it was possible to create a program that could determine whether any other program would halt or run infinitely given a particular input. Using a similar self-referencing tactic to Gödel's, Turing was eventually able to create a contradiction using the following program outlined in Algorithm 1, thus proving that no such $halts()$ program can exist:

Algorithm 1 Turing's Contradictory Program

---
if halts(self) **then**
    $loop\_infinitely()$
else
    $halt$
end if

Beyond their strategic similarity, Gödel's Incompleteness Theorem and Turing's solution to the Halting Problem are fundamentally intertwined. One may consider Turing's results as a reformulation of Gödel's work in terms of algorithms. Conversely, one may consider that Gödel's Incompleteness Theorem works by looking at mathematics as a Turing-complete system and creating an analogue to the halting problem. Indeed, one may construct a proof for a slightly weaker version of Gödel's Incompleteness Theorem based on the results of the Halting Problem. This allows us to generalize (with some restrictions) the basic intuition behind Gödel's results to any turing-complete system.

E. Quantum Systems (and Magic the Gathering)

More recently, a team of several quantum physicists analyzing the spectral gap (the gap between the lowest-energy and next-lowest-energy states of a material) were trying to generate an algorithm method to determine whether the gap existed for a particular material given its structure.

To start, the physicists began by analyzing a hypothetical infinite 2D crystal lattice of atoms. While working with the substance, they discovered that the lattice formed a Turing-complete system. Thus, it is impossible to know whether the algorithm to determine the spectral gap would ever halt, so the question of whether the gap exists remains undecidable.

The results of this study are fascinating and generalize Gödel's and Turing's results to real-world physical systems. Interestingly, many other physical systems are similarly Turing-complete and undecidable (including *Magic: The Gathering*).

REFERENCES

[1] https://www.quantamagazine.org/how-godels-incompleteness-theorems-work-20200714/
[2] https://plato.stanford.edu/entries/goedel-incompleteness/
[3] https://www.youtube.com/watch?v=HeQX2HjkcNo
[4] https://www.scientificamerican.com/article/what-is-godels-theorem/
[5] https://towardsdatascience.com/the-limits-of-knowledge-b59be67fd50a
[6] https://plato.stanford.edu/entries/hilbert-program/#1
[7] https://mathworld.wolfram.com/GoedelsSecondIncompletenessTheorem
[8] https://www.nature.com/articles/nature.2015.18983
[9] https://www.youtube.com/watch?v=O4ndIDcDSGc
[10] https://www.youtube.com/watch?v=mccoBBf0VDM
[11] https://www.youtube.com/watch?v=7DtzChPqUAw