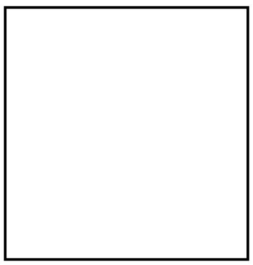
Gödel's Proof "On Formally Undecidable Propositions of *Principia Mathematica* and Related Systems"*

Introductory Exercise:



Assuming this square has a side length of 2, what is the area?

What would be the side length of a square with double the area?

More importantly, what is the *geometrical relationship* between the side of the double area square to the side of the original square?

What bearings does this have on human problem-solving as opposed to a formal-deductive method of problem-solving?

Part I: Sets

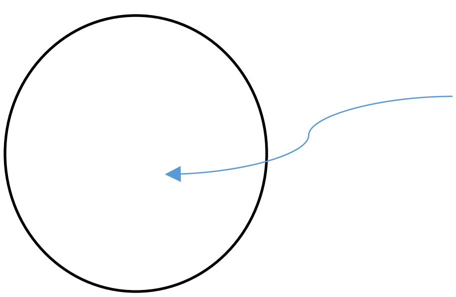
Question: can sets fully account for the categorization of knowledge?

Example:

Consider two sets, "normals and non-normals"

NORMALS are all of those things which are not a member of themselves
NON-NORMALS are all of those things which are members of themselves

N=the set of all normals



N?

Do we include N in the set of all normals, that is, **is N a member of itself**?

If so, then N, is normal, which means it is not a member of itself.

If not, then N **is** a member of itself, and is therefore **non-normal**.

But if N is non-normal, then it must be a member of itself, and so on, *ad infinitum*.

---

* Compiled from ideas, and some direct statements, from the book, *Gödel's Proof* by Nagel and Newman.

Part II: Defeating the Sets Paradox: Enter Peano, Hilbert, Russell, and Whitehead

Solution: meta-mathematical statements cannot be smuggled into a formal system; they are of a different order. Furthermore, there *can* exist a **formal system** (a language, or *sentential calculus*), which, beginning from a few axioms, can express all the truths of number theory, i.e., a *complete* system.

Peano: tried to develop a system of arbitrary symbols, and a few axioms, from which all of number theory could be expressed

Hilbert: improved upon Peano's system

Russell and Whitehead: wrote *Principia Mathematica* (henceforward PM), to create, once and for all, a system of pure sentential calculus (merely an arbitrary symbolic language) which could produce all of mathematics

---

Part III: How PM Works

Note: though the symbols are considered arbitrary (meaningless), it can be shown that, by strictly following the rules of derivation, all derived statements correspond to mathematical truths (this is called the **correspondence lemma**).

Symbol	**Meaning**
~	negation
∨	or
⊃	If…then
·	and
p, q, r, . . .	variables
(,),[,], etc.	groupings

Rules:
Substitution: any variable can be replaced with a number or formula
Detachment: for any formula S_1 and $S_1 \supset S_2$, S_2 can be derived

Axioms: the given formulas from which all others can be derived by following the rules

1. (p∨p)⊃p Example: "If either it is raining or it is raining, then it is raining."
2. p⊃(p∨q) Example: "If it is raining, then it is either raining or not raining."
3. (p∨q)⊃ (q∨p) Example: "If it is either raining or not raining, then it is either not raining or raining."

4. (p⊃q)⊃((r∨p)⊃(r∨q)) Example: "IF 'if it is raining (p) then there are clouds (q)' THEN 'if it is either sunny (r) or it is raining (p), then either it is sunny (r) or there are clouds (q).'"

Part IV: An Absolute Proof of Consistency

Goal: prove that the PM system is consistent by showing it cannot produce together the formulas S and ~S

Problem:

1. Given: p⊃(~p⊃q) (this is a formula derivable from the axioms and rules)
2. Substitution: S⊃(~S⊃q)
3. Detachment: ~S⊃q
4. Detachment: q

BUT, according to the rule of **substitution**, *any* formula can replace q, therefore *any formula whatsoever is deducible from the axioms*

Why is this a problem? *Any* formula means the possibility of mathematical untruths

Workaround

Prove that NOT every formula can be a theorem (then PM will be **consistent**)
How can this be done: show that there **is at least one formula** that can be derived from the axioms, yet is not a theorem (therefore showing that not *any* formula can be a theorem)

Considering the two sets, K_1 and K_2, where K_1 represents **tautologies**. It is important to note that ALL FOUR AXIOMS are tautologies (this can be discerned by reading the examples). A principle of a *lawfully derived theorem* is that it must carry with it the hereditary property of the four axioms. *Tautologous* is a hereditary property of the axioms and must therefore be a property of all derived formulas.

If, then, it can be shown that a derived formula is *not* tautologous, then it is not hereditary and is therefore not a theorem of PM.

(p∨q) (detached from axiom 2) satisfies the condition of *not being hereditary* (i.e., not a tautology) and therefore shows that *there is at least one formula that can be lawfully derived from the axioms, yet is not a theorem.*

Here's how . . .
Use the following criteria for sets K_1 and K_2:

K_1 : tautologous (any single element, "S," can be experimentally placed in K_1 OR K_2)	**K_2: non-tautologous** (any single element, "S," can be experimentally placed in K_1 OR K_2)
$S_1 \vee S_2$ only if S_1 and S_2 are NOT BOTH members of K_2	$S_1 \vee S_2$ only if BOTH S_1 and S_2 ARE members of K_2
$S_1 \supset S_2$ only if K_2 criteria are not satisfied	$S_1 \supset S_2$ only if S_1 IS a member K_1 and S_2 is not a member of K_2
$S_1 \cdot S_2$ only if S_1 and S_2 are NOT BOTH members of K_2	$S_1 \cdot S_2$ only if BOTH S_1 and S_2 are NOT members of K_1
~S only if S is NOT a member of K_1	~S only if S IS a member of K_1

For axiom number two, $p \supset (p \vee q)$, there are four possible arrangements:

K_1	K_2	K_1	K_2	K_1	K_2	K_1	K_2
p		p	q		p	q	p
q					q		

The p's and q's have been placed for you. Using the criteria for K_1 and K_2, place $(p \vee q)$ and $p \supset (p \vee q)$ in the appropriate boxes for each of the four arrangements. ($p = S_1$ and $q = S_2$.)

Is there any arrangement for which $(p \vee q)$ is NOT tautologous?

If so, then $(p \vee q)$ is a formula which, though derivable, is NOT a theorem of PM, because it does not possess the hereditary property of being tautologous.

The problem above of the appearance of q (via detachment) has been solved, since the substitution of ANY FORMULA for q has been shown NOT to include EVERY possible derivation from the axioms.

That is, by running every derived formula through a hereditary test, it can be shown whether or not a formula is a theorem of PM (more importantly, showing that NOT all derived formulas are permissible in PM).

Part V: Gödel's Proof

Gödel's sought to prove: "given any consistent formalization of number theory, there are true number theoretical statements that can't be derived in the system." (As an example of such a statement, Goldbach's theorem that "every composite number is the sum of two primes" has never been proven, yet, by all accounts, is true.)

Gödel's used a form of **Richard's Paradox**:

1. Assign a cardinal number to every pure mathematical statement in serial order from the least number of characters (in English) to the greatest. Statements with the same number of characters are ordered alphabetically.
2. Define two classes of numbers
    a. Non-Richardian: those assigned numbers which are described by the statement which accompanies them
    b. Richardian: those assigned numbers which are NOT described by the statement which accompanies them
3. Pose the question: is the assigned number that goes with the statement about being Richardian itself Richardian?
4. Resulting paradox: an assigned number is Richardian, if and only if it is not Richardian (because being Richardian would mean it is NOT described by the description "Richardian")

5. Solution: like with the set of "Normals" above, Richard smuggled in a statement of a different order than those which the system purports to describe. "Richardianism" is not a purely arithmetical property (it's rather a meta-statement about *language*), and therefore does not belong in the system.

<u>Gödel's Way around the Richardian Error</u>:

1. Assign a "Gödel number" to every symbol and statement in the following manner:

Number	Sign	Meaning
1	~	Not
2	∨	Or
3	⊃	If…then
4	∃	There is an…
5	=	Equals
6	0	Zero
7	s	The immediate successor of
8	(	Punct. mark
9	)	Punct. Mark
10	,	Punct. Mark
11	+	Plus
12	-	Times

variables	
number	sign
13	x
17	y
19	z
Next prime	…

Example: (∃x)(x=s0) means "there is an x such that x is the successor of 0"

sentential variables (these replace formulas)	
number	sign
13^2	p
17^2	q
19^2	r
Next prime^2	…

predicate variables (replace statements like "is prime" or "is greater than")	
number	sign
13^3	P
17^3	Q
19^3	R
Next prime^3	…

2. Assign Gödel numbers to formulas

E.g.: to find Gödel number for (∃ x) (x = s y)

use the product of $2^8 \times 3^4 \times 5^{13} \times 7^9 \times 11^8 \times 13^{13} \times 17^5 \times 19^7 \times 23^{17} \times 29^9$

(i.e., the product of each successive prime raised to the power of the symbol's Gödel number)

In a similarly recursive fashion, *unique* Gödel numbers can be assigned to each sequence of formulas and even statements about formulas.

In addition, it can be determined IF a number is or is not a Gödel number. For example,

is 243,000,000 a Gödel number?
It is the product of $64 \times 243 \times 15625$
which $= 2^6 \times 3^5 \times 5^6$
which are serially ordered primes raised to a power,
which correspond to the symbols $0 = 0$
so it is a Gödel number

3. Create the formula G (which CAN be created using the axioms, symbols, and derivation rules of PM[†])

Formula G: "The Formula G is not demonstrable using the rules of PM."

Formula G itself has a Gödel number, and can therefore be stated, "The formula that has Gödel number g is not demonstrable." (Since Formula G does have a Gödel number, all of which numbers follow arithmetical rules, Formula G can be seen as an *arithmetical property*, therefore avoiding Richard's mistake.)

4. Seek to prove whether or not Formula G is demonstrable.

Formula G is demonstrable if and only if formula ~G is demonstrable

(because the demonstration of ~G demonstrates
the non-demonstrability of G)

Both Formulas G and ~G ARE demonstrable because they stem from the axioms, follow the rules of derivation, and hold the hereditary properties of the axioms.

5. But now remember Russell and Whitehead's criterion for consistency from part IV above: "Goal: prove that the PM system is consistent by showing it cannot produce together the formulas S and ~S"

*A consistent system cannot have both G and ~G, therefore Formula G is **formally undecidable***

Since it is *formally demonstrable* yet *formally undecidable*, PM must be INCOMPLETE.

6. Suppose, however, that PM were repaired to be able to account for Formula G. Would that make PM complete?
Gödel showed that if that were the case, then Formula G' could be derived which would then encounter the same problem.
What if it were repaired again?
Then G'' could be derived, and so on, *ad infinitum*.

THE CONSISTENCY OF PM CANNOT BE PROVEN WITHIN PM.

[†] Gödel used the symbols of PM out of convenience, and then showed how his prove could apply to any formal symbolic system.

Conclusion: “Arithmetical truth cannot be brought into systematic order by laying down once and for all a fixed set of axioms and rules of inference from which *every* true arithmetical statement can be formally derived.”

Part VI: Implications of Gödel’s Proof

- Can the dream of Peano, Hilbert, Russell, Whitehead, et al., ever be attained?
- Can a machine (a computer) mimic the mind? Consider the introductory problem of the square and its solution.
- Can the processes of the mind be formalized?