A game of two symbols: An introduction to the arithmetic hierarchy

Maths essay competition 2025

1 Arithmetic and logic

Arithmetic is maths using the natural numbers. These are counting numbers $1, 2, 3, \ldots$ that can be used to answer the question "how many?", for example; 6 apples, 2 miniature dachshunds, 10 neutron stars or $20,000,000$ watermelons. (Opinions on whether 0 is to be included are divided but will make very little difference here.)

Logical statements about these numbers can be true or false depending on the values variables take on. For example,

$$a = b - 3$$

is true for $a = 1, b = 4$ but false for $a = 1, b = 5$: it can never be both. More abstractly, moving away from specific values, we may want to ask if any inputs can make a statement true or if it maybe is true whatever the input. There is a pair of inputs that works already found to be $a = 1, b = 4$ (in fact there are infinitely many). We also know that the statement is not always true whatever the inputs are as $a = 1, b = 5$ does not work. However, different questions could be asked about the different variables. Whatever the value of a, is there at least one value of b that makes the statement true? This is possible as we can always choose $b = a + 3$. How about the other way round: whatever b is, can a valid a be chosen? This is where the domain of numbers we have access to is important as negative numbers are not included in arithmetic so there is no a that make the equation true for $b = 1$, for example.

These questions can become very wordy, especially as equations become more complex so we can introduce symbols: the two symbols which are the stars of this essay are called quantifiers.

$\exists a$ asks "is there at least on value of a that works", or more commonly read as "there exists an a".

$\forall b$ says "whatever the value of b", or more commonly read as "for all b".

The questions above can therefore be written as $\exists a \exists b(a = b - 3)$ which is true, $\forall a \forall b(a = b - 3)$ which is false, $\forall a \exists b(a = b - 3)$ which is true and $\forall b \exists a(a = b - 3)$ which is false, respectively.

We can define a set of numbers by including a variable not involved with a quantifier. Lets call this variable n for any natural number. Each value of n is added to the set only if the statement is true when that specific natural number is in the place of n. Consider a different statement with two variables and only one quantifier (of which there can be any number):

$$\exists a(a^2 = n).$$

Every possible value of n is cycled through and shown to be in the set only if the logic statement is true:

n	possible a so that $a^2 = n$	$\exists a(a^2 = n)$?
1	1	True
2	none	False
3	none	False
4	2	True
5	none	False
$\vdots$	$\vdots$	$\vdots$

Therefore 1 and 4 are in our set which continues to include all the perfect square numbers. Other sets of natural numbers can be described, for examples $\forall a(n \neq 2 \times a)$ defines the odd numbers and $\forall b \exists a(a^n = b)$ is only satisfied when $n = 1$ so the set contains just that element.

The order of the quantifiers matters. The statement

$$\forall a \exists b(a^n = b)$$

for a given value of n means that whatever a is, $b = a^n$ is a natural number which is indeed true for any n. Therefore, the set it defines is the whole natural numbers. However, flipping the order of quantifiers gives

$$\exists b \forall a(a^n = b)$$

which means that a single b can be chosen so that for any a and given n, $b = a^n$. That would suggest all a^n are equal for a specific n which can never be true always, when a varies, for any value of n. So, the set defined is empty. By reading the symbols in the typical way, the meaning between these two examples changes from "for all a, there exists a b" to "there exists a b, for all a". These sound nearly identical in meaning but actually one describes every number in the domain whereas the other describes none.

2 The arithmetical hierarchy

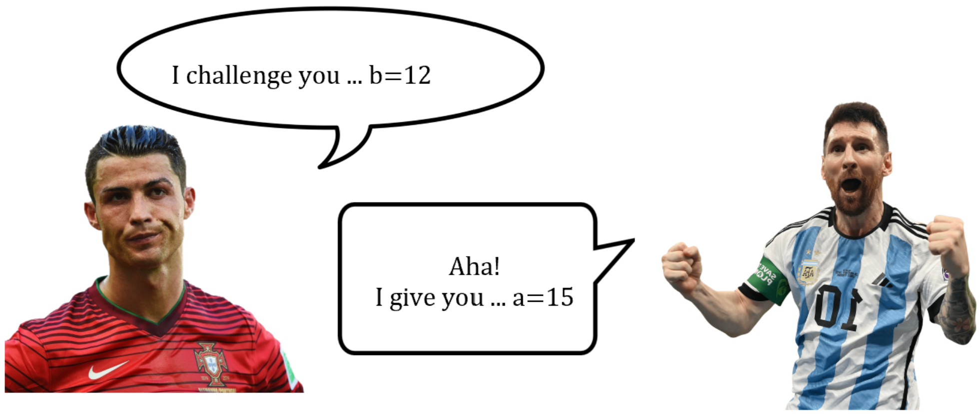
The previous exercise shows that much care must be taken when interpreting the quantifiers. Instead, they can be worked through sequentially like a game against an opponent. This would involve $\exists$ symbols allowing you to pick a number and $\forall$ symbols being numbers chosen that you have no control over, from the other player. Consider a question consisting of quantifiers $\exists x \forall y \exists z$ followed by some statement involving x,y,z and n. The first "move" of the game is choosing an x then your opponent can respond with any y and based on that you try to pick a z to satisfy the statement (for a given n). If it is true, despite the best efforts of the other player to make it false, then n is in the set. A winning strategy (for a valid n) would exist where no matter the choices of your opponent, you can guarantee winning by making it true.

The order of the moves is critical, you cannot know what the following one is going to be but you can chose based on previous choices. Flipping the quantifiers gives another game of the form $\forall x \exists y \forall z$ where your opponent goes first, choosing any x they want to which you can respond with a y that, if you have a strategy, will guarantee that, whatever they pick for z, the statement will still be true. Note that having multiple of the same quantifier next to each other makes no difference to the format of the game as it just means a player can make multiple choices for that move.

There are two pieces of information that we can use to classify a game:

1. Who goes first - which type of quantifier is first in the statement
2. How many moves are made - how many times the type of quantifier alternates

Inevitable win playing $\forall b \exists a (a = b + 3)$



Back to focusing only on the logic, this information can be used to classify sets based only on the quantifiers in the formula defining it. The details of the statement after the quantifiers are not relevant as all the information needed to sort based on complexity has already been described by then. Any formulas defining a set of natural numbers can be written in this form with a list of quantifiers followed by a quantifier-free statement. Therefore, just analysing these can be used to classify all arithmetical sets. The notation is assigned as follows:

1. If the first quantifier is a $\exists$, the Greek letter Σ (upper case sigma) is used and if it is a $\forall$ then a Π (upper case pi) is used.
2. If the quantifiers alternate k times, the notation is of the form Σ^0_{k+1} or Π^0_{k+1} (the 0 superscript refers to that fact we are dealing with the natural numbers)

Π statements are looking to show universality, that a statement works no matter what the value of the first input. It is defensive, in the game context, showing truth despite the actions of the opponent. Σ, on the other hand, looks for existence. It is searching for an instance that resolves a question such as a counter-example. For example, $\forall b \exists a (a^n = b)$ is classed as Π^0_2 and $\exists a (a^2 = n)$ as Σ^0_1.

If a set can be defined by both a Σ^0_k and a Π^0_k formula then it is given the classification Δ^0_k (upper case delta). An example of a Δ^0_1 set is the multiples of threes that can be described as

$$\exists a (n = 3a)$$

using the Σ^0_1 format but can also be as a Π^0_1 using

$$\forall a (n \neq 3a + 1 \text{ and } n \neq 3a + 2).$$

There is one more gap in the hierarchy yet to be discussed which is the base: $\Delta^0_0 = \Sigma^0_0 = \Pi^0_0$. These are formulas where the variables after every quantifier are bounded so have finite possibilities that they can take on. Bounded quantifiers are of the form $\exists x < k$ or $\forall x < k$ for a specific value of k (they are implicitly bounded below $0 < x$ as x is a natural number). For a game, this would mean that all the outcomes could just be checked in finite time so there is no need for a strategy. Π^0_1 and Σ^0_1 formulas add a single type of unbounded quantifier to these and then Σ^0_2 and Π^0_2 add a $\exists$ or a $\forall$ to these respectively. This keeps going, building up the hierarchy adding a different type of quantifier, swapping the letter and incrementing the subscript.

3 Computability

Beyond game analogies, there are tangible consequences of the quantifiers used at each level of the hierarchy. We have seen that it loosely allows sets to be classified by the complexity of the logical formula that define them. Specifically, the classification refers to the difficulty in computing a formula. This means to what extent an algorithm, like a computer program, can be followed to verify if a number is a member of a given set. The "computer" we will consider is called a "Turing machine" which is a theoretical computer that can follow algorithms. We will primary look at sets on the first layer of the arithmetic hierarchy (Δ_1^0, Π_1^0 and Σ_1^0) which have the most clearly definable features.

The machine receives an input and determines whether it belongs to a specific set. It uses this to follow the steps of an algorithm. It will then either stop and output the answer or loop forever and never give an answer. Δ_1^0 sets are "decidable" which means that there is an algorithm that a Turing machine can follow that will stop and decide whether the input is a member or not. Σ_1^0 is called "recognisable" so there is a algorithm that will stop and confirm membership if the input is in the set but if it is not, it may output that but could also continue to loop forever. Π_1^0 is the opposite, "co-recognizable", in that the algorithm will definitely stop if the input is not in the set but if it is, the machine may stop or continue to loop. This makes sense why Δ_1^0 can always give answers either way as they are the intersection of recognisable and co-recognisable sets.

All bounded sets are Δ_0^0 as they are decidable - in a fixed amount of time with a simple algorithm using no loops or unbounded searches. Higher levels than 1 follow the same decidable/recognisable/co-recognisable categorisation but algorithms are more complex, requiring additional information from other levels of the hierarchy.

In conclusion, it is astonishing that we can compute all arithmetic sets and that two symbols entirely determine how this is done. In fact, this classification can be extended to others domains of numbers such as negatives and fractions using different superscripts and bold letters. These categories can therefore encompass all logical formulas underpinning the whole of mathematics, or at least its general number systems. This exemplifies one of the most attractive and unique qualities about maths as a field and a subject - that it is perfectly simple. Famous mathematician John Von Neumann suggested that *"if people do not believe that mathematics is simple it is only because they do not realize how complicated life is"*.