

A unified view towards diagonal arguments

Qiangru Kuang

February 25, 2019

Abstract

Cantor's diagonal argument is a simple yet deep theorem. Perhaps less known is that the same idea in the proof powers many famous results, including but not limited to: Russell's paradox, halting theorem, undefinability theorem and Gödel incompleteness theorem. This talk aims to state and prove an abstract form of diagonal argument and derive all of the above as corollaries if time permits. The talk will be self-contained. IA Numbers and Sets or knowing Cantor's diagonal argument will be helpful but not necessary.

1 Acknowledgement

I would like to thank Archimedean for organising this talk. My gratitude goes especially to Jeremy, who invited me to give a talk at an early stage of the event, and whose continual checking of my progress eventuates in this article, instead of another draft in the limbo of unfinished blog posts.

I am also grateful to Rob, Linshu and Jenny for listening to an earlier version of the talk and for kindly providing advice which substantially improved the talk.

2 Exponential

We start by staring at the notation Y^Z where Y and Z are sets, which is by definition the set of all functions $Z \rightarrow Y$. There are two things worth remarking: firstly Y^Z is a set itself. This statement does not hold in a general category. For example given two manifolds M and N , the set of all smooth maps $M \rightarrow N$ is not obviously a manifold. Given two groups G and H , it is not clear how we can endow the space of all homomorphisms $G \rightarrow H$ a group structure (however if G and H are abelian groups then $\text{Hom}(G, H)$ is obviously an abelian group, although pursuing in this direction will lead us astray from today's topic).

The second fact is a consequence of our definition. Consider another set X . There is a canonical bijection

$$\{X \times Z \rightarrow Y\} \leftrightarrow \{X \rightarrow Y^Z\}$$

given by what computer scientists call "currying" and "uncurrying": given a map $f : X \times Z \rightarrow Y$, we may plug in some $x \in X$ to get

$$g(-) := f(x, -) : Z \rightarrow Y.$$

In this case we say g is *represented* by x . Conversely, “uncurrying” is the process of, given $X \rightarrow Y^Z$, i.e. a family of maps $Z \rightarrow Y$ parameterised by X , we turn it into a function $X \times Z \rightarrow Y$.

We may ask the question: when does X fully represent Y^Z ? Obviously we can just take X to be Y^Z . A more meaningful question would be when X fully represents Y^X . Cantor’s theorem says that if Y has more than one element then there can be no surjection $X \rightarrow Y^X$, so there is always an element that is not represented.

3 Lawvere fixed point theorem

We need to develop a little bit category theory to state Lawvere fixed point theorem in full. If you are not familiar with category theory, you can safely take \mathbf{C} to be **Set**, every instance of X, Y, Z to be a set, a point $x : 1 \rightarrow X$ to be an element of X and operations to be the obvious set theoretical operations.

A category \mathbf{C} is *cartesian closed* if there are finite products and for every object Z there is a right adjoint to $(-) \times Z : \mathbf{C} \rightarrow \mathbf{C}$. The adjunction is usually written as $(-) \times Z \dashv (-)^Z$, ergo the name exponential. Let the unit and counit of this adjunction be

$$\begin{aligned} \eta_Y^Z &: Y \rightarrow (Y \times Z)^Z \\ \varepsilon_Y^Z &: Y^Z \times Z \rightarrow Y \end{aligned}$$

respectively. The counit is usually called ev , which is the internal language version of taking an ordered pair $(f, x) \in Y^Z \times Z$ and sending it to $f(x)$.

There is an bijection between $f : X \times Z \rightarrow Y$ and $\tilde{f} : X \rightarrow Y^Z$. More specifically, given $f : X \times Z \rightarrow Y$, \tilde{f} is the composition

$$X \xrightarrow{\eta_Y^Z} (X \times Z)^Z \xrightarrow{f^Z} Y^Z$$

and given $\tilde{f} : X \rightarrow Y^Z$, f is the composition

$$X \times Z \xrightarrow{\tilde{f} \times 1_Z} Y^Z \times Z \xrightarrow{\varepsilon_Y^Z} Y.$$

By abuse of notation we take the liberty to use f to denote both morphisms. In particular, we identify a morphism $f : 1 \times Z \cong Z \rightarrow Y$ with $f : 1 \rightarrow Y^Z$, i.e. an element of Y^Z .

Finally, a morphism $f : A \rightarrow B$ is *point-surjective* if for every point $b : 1 \rightarrow B$ there exists $a : 1 \rightarrow A$ such that $fa = b$. The corresponding notion in **Set** is surjectivity.

Theorem 1 (Lawvere fixed point theorem). *In a cartesian closed category \mathbf{C} , if $f : X \rightarrow Y$ is point-surjective then every $\alpha : Y \rightarrow Y$ has a fixed point $1 \rightarrow Y$.*

This is proven by William Lawvere in 1970s, an influential category theorist who made contributions to category theory and foundation of maths, as well as philosophy. He is credited with introducing, for example, elementary topos and ETCS.

Proof. Define $q : X \rightarrow Y$ to be the following composition

$$X \xrightarrow{\Delta} X \times X \xrightarrow{f \times 1_X} Y^X \times X \xrightarrow{\text{ev}} Y \xrightarrow{\alpha} Y$$

By point-surjectivity of f exists $x : 1 \rightarrow X$ that represents q , i.e. $f(x_0, -) = q$. Now if we plug x_0 back in we get

$$\begin{aligned} q(x_0) &= \alpha(f(x_0, x_0)) \quad \text{by definition of } q \\ &= \alpha(q(x_0)) \quad \text{by definition of } f(x_0, -) \end{aligned}$$

so $q(x_0) \in Y$ is a fixed point of α . □

It might be instructive to consider the contrapositive:

Theorem 2. *If there exists $\alpha : Y \rightarrow Y$ with no fixed point then $f : X \rightarrow Y^X$ is not point-surjective.*

By the way this is named by Lawvere “Cantor’s theorem”, which we’ll justify in a second.

Proof. The same argument as above will exhibit a function $q : X \rightarrow Y$ that is not represented by X : suppose for contradiction it is represented by $x_0 \in X$, i.e. $f(x_0, -) = q$. But

$$q(x_0) = f(x_0, x_0) = \alpha(q(x_0)) \neq q(x_0),$$

contradiction. □

The idea is that if there exists a point-surjection $X \rightarrow Y^X$, i.e. every map $X \rightarrow Y$ can be represented by an element of X then Y has a particularly simple “internal structure” in the sense that every endomorphism has a fixed point. When we are in the setting of sets, the only such Y is singletons.

4 Cantor’s diagonal argument

By the trivial observation that $2 = \{0, 1\}$ admits a self map with no fixed point, it follows from Lawvere fixed point theorem that a nonempty set X cannot surject its power set 2^X , which is Cantor’s diagonal argument. As the proof is constructive, we may well apply it and see what nonrepresentable element we get.

We will use as example the classical Cantor’s diagonal argument which proves the nonexistence of surjection $\mathbb{N} \rightarrow 2^{\mathbb{N}}$. Any attempted enumeration of $2^{\mathbb{N}}$ can be seen as a map

$$\begin{aligned} f : \mathbb{N} \times \mathbb{N} &\rightarrow 2 \\ (n, k) &\mapsto k\text{th digit of } n\text{th binary sequence in the enumeration} \end{aligned}$$

so suppose f looks like

n, k	1	2	3	...
1	1	0	1	...
2	0	0	1	...
3	1	1	1	...
⋮				

then $q \in 2^{\mathbb{N}}$ is obtained by negating elements along the diagonals.

5 Russell’s paradox

Recall Russell’s paradox: we cannot have a set of all sets as otherwise consider the set $\{A : A \notin A\}$.

Suppose we have the set of all sets **Set**. We have a membership function

$$f = \in : \mathbf{Set} \times \mathbf{Set} \rightarrow 2$$

$$(A, B) \mapsto 1_{A \in B}$$

so the nonrepresentable element $q \in 2^{\mathbf{Set}}$ we get is the indicator function $1_{A \neq A}$. By Yoneda lemma this is the same as the set $\{A : A \notin A\}$, which is what we get in classical proof of Russell’s paradox.

6 Entertainment maths

Many self-referential paradoxes find their way into entertainment maths. You might have heard of this one. Some adjectives in English describe themselves. For example “English” is an English word, “polysyllabic” is itself polysyllabic, but “monosyllabic” is not¹. Call such adjectives describing themselves “autological” and “heterological” otherwise. But then you will find that heterological is heterological if and only if it is not.

It is mildly interesting on an entertainment level that we can produce an example witnessing the paradox. However, it has the disadvantage of being overly constructive and more importantly, depending on the semantics of the adjective “heterological” in English language. A more convincing presentation of the paradox can be given by Lawvere fixed point theorem: consider the function

$$f : \mathbf{Adj} \times \mathbf{Adj} \rightarrow 2$$

$$(A, B) \mapsto 1_{A \text{ describes } B}$$

Then the nonrepresentable q is the indicator function of the set of adjectives whose element has the property that it describes itself if and only if it doesn’t. “Heterological” is an element of this set under ordinary semantics.

Exercise: exhibit Liar’s paradox as an instance of Lawvere fixed point theorem.

7 Gödel incompleteness theorem

There is actually a moral in the paradox. The function $f : \mathbf{Adj} \times \mathbf{Adj} \rightarrow 2$ comes for free with the category **Adj**. By contrast, in Cantor’s diagonal argument we rely on extra data to define f (membership function in Russell’s paradox can also be seen as “free” if we allow unrestricted comprehension, which is granted since we assume for contradiction there is a set of all sets). In **Adj**, an object

¹Let’s just say for the moment that this notion is well-defined so we don’t have to fight to decide if “short” is indeed short.

A serves an additional syntactical role, namely the ability to form the predicate “ B is described by A ”.

More generally, consider X , a collection of syntactical entities such as nouns, formulae, proofs, programs etc. We can define a function

$$f : X \times X \rightarrow \Omega$$

where Ω should be thought as values predicate can take, that describe all describable properties $X \rightarrow \Omega$.

The innovation of Gödel to introduce Gödel number is an example of this, which is roughly speaking a way to name formulae in PA. This allows us to talk about predicates of the form “the natural number 3 satisfies the formula with Gödel number 5, which is $t^3 = 3t$ ”. We denote by $\ulcorner A \urcorner$ the Gödel number of formula A .

We claim that for $E(x) \in L^1$, there exists a formula C such that

$$\vdash (C \iff E(\ulcorner C \urcorner)).$$

Thus if we take E to be

$$E(x) := (\forall y)\neg\text{Prov}(y, x)$$

where $\text{Prov}(y, x)$ is the predicate “ y is the Gödel number of the proof of a formula whose Gödel number is x ”, then by the assertion we can find closed C such that

$$\vdash (C \iff E(C)),$$

i.e. C is equivalent to its own unprovability. This is the famous Gödel incompleteness theorem.

Proof of claim. We define an auxiliary function $D : \mathbb{N} \rightarrow \mathbb{N}$: for every formula $A(x) \in L^1$,² where x is its only free variable, we have

$$\begin{aligned} D : \mathbb{N} &\rightarrow \mathbb{N} \\ \ulcorner A(x) \urcorner &\mapsto \ulcorner A(\ulcorner A(x) \urcorner) \urcorner \end{aligned}$$

We let

$$\begin{aligned} f : L^1 \times L^1 &\rightarrow L^0 \\ (A(x), B(x)) &\mapsto \ulcorner A(\ulcorner B(x) \urcorner) \urcorner \end{aligned}$$

and

$$\begin{aligned} \alpha : L^0 &\rightarrow L^0 \\ A &\mapsto E(A) \end{aligned}$$

Then we consider the usual composite q which is given by

$$\begin{aligned} q : L^1 &\rightarrow L^0 \\ A(x) &\mapsto E(\ulcorner A(\ulcorner A(x) \urcorner) \urcorner) \end{aligned}$$

² L^1 is the Lindebaum algebra of 1 free variable, defined as the set of all formula with 1 free variable with provably equivalent formulas identified.

We claim that q is represented by $G(x) = E(D(x))$: indeed

$$f(G(x), A(x)) = G(\ulcorner A(x) \urcorner) = E(\ulcorner A(\ulcorner A(x) \urcorner) \urcorner)$$

which is the same as $q(A(x))$. It follows from general Lawvere fixed point argument that $C = q(G(x)) = G(\ulcorner G(x) \urcorner)$ is a fixed point of E , i.e. $C = E(C)$. \square

References

- [1] F. W. Lawvere. *Diagonal arguments and Cartesian closed categories*. Reprints in Theory and Applications of Categories, No. 15, 2006
- [2] N. S. Yanofsky. *A Universal Approach to Self-Referential Paradoxes, Incompleteness and Fixed Points*. arXiv:math/0305282 [math.LO]
- [3] F. W. Lawvere, S. H. Schanuel. *Conceptual Mathematics: A first introduction to categories*. Cambridge University Press, 2011