Some Notes on Gödel, Escher, Bach: an Eternal Golden Braid

Ibraheem Khan

February 2019

Chapter -1

Preface

"In a word, *GEB* is a very personal attempt to say how it is that animate beings can come out of inanimate matter." In many ways this book could have been "Gödel's Theorem and the Human Brain" as opposed to *GEB* as it is not concerned with these three men, nor particularly in connections between math, art, and music, but rather in the investigation of "strange loops" or "tangled hierarchies" which arise from meaningful patterns of meaningless symbols in particular systems that ultimately beget inanimate matter and later animate beings. The book is a study of consciousness (although I posit this idea could be extended for existence and the birth of the world in general). Hofstadter has one main point to make here: the connection between Godel's Theorem and Consciousness:

> The Gödelian strange loop that arises in formal systems in mathematics (*i.e.*, collections of rules for churning out an endless series of mathematical truths solely by mechanical symbol-shunting without any regard to meanings or ideas hidden in the shapes being manipulated) is a loop that allows such a system to "perceive itself", to talk about itself, to become "self-aware", and in a sense it would not be going too far to say that by virtue of having such a loop, a formal system *acquires a self.*

Another central thesis of the book is that the aforementioned "meaningless symbols" are not completely meaningless. In his own words, "meaning cannot be kept out of formal systems when sufficiently complex isomorphisms arise." For example, suppose we create some symbol, "a", this symbol may undergo some isomorphism f (which implies $\exists\ f^{-1}$ as f is an isomorphism on a), we could imagine this isomorphism being sufficiently complex in that it "tracks" or "mirrors" some phenomena in the real world, this complexity thus has already applied meaning to it. It is then a question of whether or not all such sufficiently complex isomorphisms on symbols must always model real world phenomena, if the real world is or isn't restricted to isomorphic actions on symbols, or if it is necessary for the functions themselves to be completely isomorphic (are morphisms enough?). In any case, meaning arises from a sufficiently complicated description- the less complicated description, the less meaningful. It is another question if description is enough to ascribe meaning.

Hofstadter continues to elaborate that consciousness is not unique to humans nor does it require some particular and undiscovered biological prerequisite- instead consciousness arises from certain patterns, the aforementioned "strange loops" rather than some keen biological or neuro-electro-chemical process. Such a view, creates hope for real artificial, electromechanical consciousness as there is no anthropocentrism involved in the creation of these strange loops- just that human evolution has done a good job of fashioning it. Consciousness only requires a media capable of supporting such strange, self-referential, loops.

Delving deeper here, we receive a soft-introduction to Russel and Whitehead's famous *Principia Mathematica* which aimed to solve the conundrum of the self-referentiality of the foundations of mathematics. We find later that

Kurt Gödel not only completely and provably rejects this idea but also shatters Hilbert's dreams of converting mathematics into mere "symbol shunting." Although, this idea seems universally accepted now, that Mathematics is semantically meaningful even outside of its application, this was not always the case and many mathematicians (the giant Hilbert himself included) were of the opinion that math was a mere plaything for humans- a toy exercise in logic and ideas or how Hilbert himself described "poetry made up of ideas, not words."

Gödel achieves this by employing a particular mapping scheme from metamathematical statements to those of mathematical number theory via his Gödel Numbering- that is all statements about mathematics are in fact, mathematical statements in themselves. Hofstadter notes that this is in fact very similar to the leap between inanimate to animate and in general is recursive- mathematical statements peer into themselves by their very nature. An important thing to note here is the central incompleteness of mathematics as demonstrated by the sentence G: "G is not provable inside *PM*" which is a *true* statement about mathematics but is still unable to be proven. That is, truth and provability are two totally different concepts. This also asserts that there is some causal power to meaning. This is a nonintuitive statement. We would assume that we can arbitrarily apply any meaning to any rule-bound string of symbols and it would not cause any thing. That is, a rule-bound string of symbols is internally complete [1] or that given our set of rules and symbols we can form all sentences, yet clearly adding meaning to a sentence (consider sentence G) causally implies the incompleteness of our rule-bound string of symbols.[2]

The preface further clarifies the nature of consciousness in that it is not a binary phenomenon, in fact in can be said that everything is conscious- it is a but a matter of how conscious. The "pattern called "I" can shove around inanimate particles in the brain no less than inanimate particles can shove around patterns." It remains to be said what such a view of identity and consciousness says of freedom of will, freedom of action, and freedom of thought. This is later explored in Hofstadter and Daniel Dennet's book *The Mind's I.*

[1] In mathematical logic and metalogic, a formal system is called complete with respect to a particular property if every formula having the property can be derived using that system, i.e. is one of its theorems; otherwise the system is said to be incomplete.

[2] Examples of rule-bound string of symbols include mathematics or really any formal language- better discussion is to be seen in Formal Languages and Automata

Chapter 0

Introduction: A Musico-Logical Offering

0.1 Bach

The book starts off amusingly by introducing the reader to perhaps the most magnanimous and great king of all time: King Frederick the Great of Prussia- a patron of the arts and science employing Euler, Voltaire, and Bach in his court. The book then jumps into a discussion of Bach's *Musical Offering*. In this seemingly unrelated discussion of music, Hofstadter formally defines what he means by "isomorphism"-in his case, an information preserving transformation. He also gives, in a rather roundabout but interesting way nonetheless, our first example of a strange loop: an endlessly rising canon composed by Bach simply named "Canon per Tonos" which modulates its music six times only to arrive back at the original key of the canon, C Minor. Hofstadter then formally states that

> The "Strange Loop" Phenomenon occurs whenever, by moving upwards (or downwards) through the levels of some hierarchical system, we unexpectedly find ourselves right back where we started.

Whereby he uses the term "Tangled Hierarchy" to describe the kinds of systems in which Strange Loops occur.

0.2 Escher

In this section we are introduced to yet another example of Strange Loops- this time, evident in Escher's mathematically inspired art. As Hofstadter regales Escher's art, he makes a strong but subtle point. Loops, by their very nature, encode a concept of infinity, "for what else is a loop but a way of representing an endless process in a finite way?" Another subtle point he hints at here is that the media of the levels forming tangled hierarchies (counting the number of levels is a current ambiguity that Hofstadter previously notes) is not limited to the key of music or the continuation of a visual theme but also the interpretation of a particular thing lying more as "imagination" or as "reality". He then implies that it is possible for the chain of levels to be nonlinear- rather that they too may form a loop.

0.3 Gödel

So far we have developed an amusing intuition of notions of infinity, loops, finiteness, all intertwined in "a strong sense of paradox." We are now slowly being interested to a key example of a Strange Loop-one brought on by the mathematician Kurt Gödel who mathematically formalized an "ancient intuition" known as *Epimenides paradox* or more commonly known as the *liar paradox.* It is at its sharpest the statement:

This statement is false.

This statement is key as it has a unique property of not being able to be categorized as either true or false. Some may even call it nonsensical or a figment of grammar (as Chomsky or Wittgenstein might try to egg it as) and relates closely to such ideas as constructivism in Mathematics and the validity of the Law of Excluded Middle. Gödel's formal statement made in his 1931 paper *On Formally Undecidable Propositions in Principia Mathematica and Related Systems I.* goes as follows:

> To every ω-consistent recursive class κ of *formulae* there correspond recursive *class-signs* r, such that neither v Gen r nor Neg(v Gen r) belongs to Flg(κ) (where v is the *free variable* of r).

Or in simpler English:

> All consistent axiomatic formulations of number theory include undecidable propositions.

The Strange Loop, however, does not lie here, but rather in the proof of this statement which involves writing a self-referential mathematical statement. The encoding of this statement lies in undecidability or really that there may be true statement unable to be proven by some weak axiomatic system. The problem, or really the theorem, lies in the fact that *any* axiomatic system is too weak to prove all true statements- that at their cores they are incomplete. Provability is inherently a weaker notion than truth.

Soon after a small introduction to the history of mathematical logic that lead up to this, Hofstadter hints at yet another small but subtle point: Strange Loops often do not arise from one particular thing but in combination or in a particular configuration- it is a global property not usually local. For example, while we can consider the Strange Loop "This sentence is false" we can equivalently phrase it as:

The following sentence is false.
The preceding sentence is true.

Now each individual sentence is clearly intelligible and is useful, but together we arrive back at our Strange Loop. And now, something to chew on:

- Are mathematics and logic distinct; does one precede the other?

0.4 Consistency, Completeness, Hilbert's Program

I will now briefly define two important concepts: consistency meaning contradiction-free and completeness meaning that given an axiomatic system for some theory all true statements of that theory can be proven via that axiomatic system. It was these two things that David Hilbert wanted of *Principia Mathematica*. Following Gödel's paper, such a dream was to never be accomplished and never will.

0.5 Babbage, Computers, Artificial Intelligence

This section is somewhat more light and describes the history of computation and artificial intelligence starting from Charles Babbage who dreamt of an "Analytical Engine." However, even here we receive a reiteration of a central thesis of the book: the fine line between human intelligence and artificial intelligence is not really a line at all and ultimately the way computers "think" via rules is just a simpler stripped down way human think yet the two still follow "rules" and at the heart of intelligence lies Strange Loops involving rules that change themselves in some multilevel sense. In the following section there is some interesting historical dialogue concluding that even 200 years later the battle between materialism and on the worthiness of human thought remain. Some questions and thoughts:

- How closely tied is Materialism to *GEB*'s central theses?
- How are ideas of qualia, materialism, and Thomas Nagel's *What is it Like to be a Bat?* connected to the central theses of *GEB*?
- Theories of mind and theories of reality are not all that far apart, why?

Chapter 1

The MU-Puzzle

Chapter 2

Meaning and Form in Mathematics

Chapter 3

Figure and Ground

Chapter 4

Consistency, Completeness, and Geometry

Chapter 5

Recursive Structures and Processes

Chapter 6

The Location of Meaning

Chapter 7

The Propositional Calculus

Chapter 8

Typographical Number Theory

Chapter 9

Mumon and Gödel

Chapter 10

Levels of Description and Computer Systems

Chapter 11

Brains and Thoughts

Chapter 12

Minds and Thoughts

Chapter 13

BlooP and FlooP and GlooP

Chapter 14

On Formally Undecidable Propositions of TNT and Related Systems

Chapter 15

Jumping Out of the System

Chapter 16

Self-Ref and Self-Rep

Chapter 17

Church, Turing, Tarski, and Others

Chapter 18

Artificial Intelligence: Retrospects

Chapter 19

Artificial Intelligence: Prospects

Chapter 20

Strange Loops, or Tangled Hierarchies

Chapter 21

Appendix

21.1 Proof of Gödel's Completeness and Incompleteness Theorems

Here we will go over the central mathematical theorem of the book and will delve into some of the discussion offered by Nagel and Newman's *Gödel's Proof*. The first 5 sections of the book are dedicated to a survey of logic and metamathematics until Gödel's famous 1931 paper.

21.1.1 The Problem of Consistency