Limits and Randomness of Reason

Talk delivered on August 16 and 17, 2011

There are a few tools for thinking about this talk and they are isomorphism, paradoxes, formal systems, recursion, self-reference and infinity.

An isomorphism is essentially a correspondence or similarity between two ideas. Two rooms are isomorphic to each other. A car is isomorphic to a skateboard. Et cetera. The rest of the words are self-explanatory.

My talk will exclusively focus on the structure of logic, using a few tools from mathematics to tell you how there are inherent limitations on mathematics. People who inspired me to be here are Douglas Hofstader when I read his book Godel, Escher Bach: Eternal Golden Braid, Gregory Chaitin's lectures and the life of Godel and Alan Turing. How many people are familiar with the names? Gregory Chaitin's book "The Quest for Omega" is another good read. You might also be interested in "A World Without Time: Forgotten legacy of Einstein and Godel" by Palle Yourgrau. Einstein and Godel were very good friends.

My motivation won't be to discuss any philosophical issues associated with this topic. That is going to be open for the room to decide because there are lots of ways of looking at this talk. I'll be focusing on the theoretical approaches to this talk, even though the main motivation that drove these persons was intensely philosophical. For example, Hofstader's book essentially aims to discuss how and why meaningless little symbols take complete meaning and then basically to think why atoms give us a conscience, how they make the "I". I won't be discussing any of that. That is going to be your corner.

The messages to take home from my talks are two: thinking and metathinking. Thinking is like thinking in a system and metathinking is thinking outside the system. I think this is very important because these concepts have given rise to Montessori education and because it's by thinking out of the system that Karl Marx gained his reputation.

There are three types of paradoxes. Falsidical, Veridical, Antimony. Falsidical paradoxes are the result of a false reasoning and seem very counter-intuitive like the Birthday Paradox. Veridical's are based on a weak form of reasoning because of the absence of tools. Examples of such paradoxes are Zeno's paradox or the Twin Paradox. I'll be focusing on the third type. How many over here have taken a course on Discrete Mathematics? You might need a little grip on logic. I'll revisit that area.

Before I indulge in telling you about this tool of paradoxes, let me just skim over the theory of formalism.

Physics: Theories ➔ Experiments and Calculations ➔ Predictions for observation
Computing: Programme ➔ Execution on a computer ➔ Output
Mathematics: Axioms ➔ Reasoning ➔ Theorems

We'll be visiting this chain again and again.

When I first read Einstein's paper on general relativity, it was in German. Then, I got a translated version and then thought that this might as well have been in German all the same! Reason: Mathematics uses a formal system of rules which only a person trained in the mechanism of it can decipher, which is what we as mathematicians have learned.

Let's say we have a system that is only indicated by x p y q z where x, y, z are variables and p and q are fixed symbols. Let's say that a theorem in this system holds is true if x + y = z. If x, y, z represent a string of hyphens only, then, for example, -- p --- q ----- is a theorem while - p - q - is not. Notice that this entails our usual concept of addition if we believe p represents plus and q represents equal. This system may not be the most useful of all systems but it provides a good insight as to why mathematicians like formalising or mechanising their language. It has to do with a representation of ideas. The same system, however, would be very useless if x y and z were other objects. For instance, if x = happy, y= horse and z= apple, then, in our sytem, a theorem would read happy p horse q apple apple or happy happy p horse horse horse q apple apple apple apple apple.

Moral: All formalism has a motivation.

Logic follows the same principles. We all would like to thank Aristotle, the Stoics, Kurt Godel, John Venn, David Hilbert, de Morgan and all the rest who have helped formalise our thinking processes. Question is: was it empirical? Or was it a product of hard thinking? That is where people divide and I will not delve into the topic and leave the question open.

Now, let me just introduce a few principles. p ➔ q. Equivalently, ~q ➔ ~p. For example, *If something is a bat, then it is a mammal*." The contrapositive is, "*If something is not a mammal, then it is not a bat*." One can essentially switch between both. The symbol ⇔ means if and only if, compactly written iff which means both p ➔ q and q ➔ p, so essentially they're the same thing. The if and only if is a stronger version. It's like saying that p and q are equivalent.

There was this man called Goerg Cantor who single-handedly invented set theory in the 20^{th} century. Before that, 19^{th} century mathematics was just hard analysis and formulae. Like Taylor series, polynomials and number theory but then Cantor introduced Set Theory and revolutionised mathematics and made everything wishy-washy. What are sets? You just take an objects, throw them into a collection and you have a set. You just need to define your "box", so things are in simple language, too.

Why did Cantor do that? Well, he was mostly interested in theology. He wanted to know what infinity was. In fact, he ended up introducing smaller infinities and bigger infinities and called God as the absolute infinity. Let me show you how and why.

$1, 2, 3, \ldots, \omega, \omega + 1, \omega + 2, \ldots, \omega\cdot 2, \omega\cdot 2 + 1, \ldots, \omega^2, \ldots, \omega^3, \ldots, \omega^\omega, \ldots, \omega^{\omega\omega}, \ldots, \varepsilon_{0, ..}$

We can already see why this was mostly theological. This is where things start to blow up.

This had some serious issues. In 1897, the Burali-Forti paradox was stumbled upon. It essentially asks us to imagine a set of all ordinal numbers Ω. Ω itself is an ordinal number because it carries all the properties of ordinal numbers and would be in the set. Also, Ω + 1 would be in the set. Thus we end up with $\Omega < \Omega + 1 <= \Omega$. It essentially says that there isn't a biggest ordinal number and that is theologically very troubling. I won't go into the technical details of the matter because I'll be discussing an equivalent paradox. The heart of the problem lies in the word "all". How? We'll see shortly.

Even though this paradox existed, mathematicians loved Set Theory. David Hilbert called it a paradise that no one could expel mathematicians from. A few mathematicians didn't. Like Henri Poincaire, the famous French mathematician called Set Theory a disease from which he expected future generations to recover but not all mathematicians took the same approach. Topology was completely formulated in terms of sets. Everything began to take shape in the form of sets. Here are two beautiful concepts as a result of set theory: We can have sets and subsets. Each subset has the cardinality 2^n, where n is the size of the original set. Since the real numbers are infinite (their size is actually denoted by the Hebrew letter aleph with a nought in the subscript), then all the subsets of the real numbers have a size of 2^aleph-nought. This is a bigger infinity. It's called aleph-one. There an infinite amount of them even between two decimals. Also, we can say that the size of the natural numbers is the same as that of the integers. It introduced the concept of things that we can count even if we can count them forever and that of the things that we couldn't possibly count, no matter what. Set theory also gave mathematicians the idea of space and dimension. A space is basically a set is with a given structure. That's how mathematicians define space. Lastly, well, look. A line, if doubled in one dimension, creates two. A square, if doubled in two dimension, gives us four squares. A cube, if doubled in three dimensions, gives us eight cubes and so on. For a triangular fractal, the Sierpinski triangle, this is $2^d = 3$ because you halve the width and height and you get three triangles. This gives us d=1.55 approx. I think this is pretty cool.

Set theory was pretty interesting. It solved many problems and opened new dimensions. Mathematics since then has been grounded on set theory, logic and number theory.

I am going to start with a brief introduction on paradoxes. The antimony paradox is a statement "p iff not p". For example, let's start with the Barber's paradox. Suppose there is a town where there is just one barber. The barber shaves all those men and only those men that do not shave themselves. The question is, does the barber shave himself? Each possibility implies its negation "This sentence is false" is a paradox. "I am a liar" is another. If meaningless is meaningless, then meaningless is meaningful but if meaningless is meaningful, then meaningless is meaningless. If a paradox is a paradox, then a paradox is not a paradox. If a paradox is not a paradox, then a paradox is a paradox.

It is reasonable to question if our system of logic is consistent i.e. if a proposition and its negation carry the same value or not. A paradox p is both true and false. If our system is consistent, then a paradox has no place in our system of logic. You wouldn't want two contradictory statements stemming from the same axioms. If it does, then you're in trouble.

Logicism believes that all mathematics is reducible to logic. If the foundations of set theory are under attack, mathematicians have a good reason to panic. Let's attack mathematics, then. Let's just say that there is a set of all sets S which are not a member of themselves. Does S belong to S? The answer is yes and no.

The Barber Paradox and the set paradoxe are a result of "faulty" axioms. It says that we can't have a universal set. Formally, there's a function of choice that takes care of these matters. It just doesn't allow the construction of such sets. Okay, that's one way of sweeping the problem under the carpet. There have been ways to overcome these horrendous problems by correcting a few places of the system, such as the Zermelo–Fraenkel set theory with the axiom of choice but how far can we go in producing such a system? When will it be "complete"? i.e. when will we have derived all the possible truths from a few axioms? This question was posed by David Hilbert as a second problem in his famous 23 problems in 1900 that asks to have arithmetic with axioms that would derive all theorems. David Hilbert came to rescue mathematics because in the early 20th century, right after the Burali-Forti paradox, many paradoxes were emerging and that wasn't a healthy sign for mathematics. There was a need for an artificial language that could live free of paradoxes – a language that could be objective and mechanical and black and white and unlike the real world that is pretty messed up.

Another goal behind Hilbert's programme was to have a mechanical proof checker. So, basically, if one has a proof-checker, that essentially means it can recognise all the arguments of a proof. That's also how mathematics would get complete by being able to differentiate between an axiom and a theorem.

Godel answered Hilbert's question in negative because of Godel's incompleteness theorem which states that "a system cannot be both consistent and complete". That's the first incompleteness theorem. Yes, they were two! Consistent means that in the same system, a statement and its negation can never have the same truth values. A system is complete if it contains all the theorems it can contain. A theorem is always a true statement in that system. Godel proposed two theorems. Any system cannot be both consistent and complete. Secondly, if a system is self-verifying, then there are statements that cannot be proven in the system. This second theorem is a stronger version of the first theorem. Why? Because he showed that there exists a theorem that is not provable viz. "This theorem G is not provable" for if G was provable, then the system would be inconsistent but if it was otherwise, then lo and behold, we're limited!

There is, however, this proof called Gentzen's consistency proof in Proof Theory that uses various arguments to prove that says that essentially it's not about proving whether a system is consistent but it's about whether the logical principles are firmly grounded. This is one reason mathematicians still have hope. It is argued, however, that what Godel showed implied that a system which satisfies the necessary hypothesis Godel chose cannot prove its own consistency and that a different system is needed to prove the consistency of that system. Also, ironically, Godel had a completeness theorem. It said that any system that uses the principles of logic is bound to be consistent and hence complete. Should mathematicians hope? Can they have a higher mathematics? You and I need to find out.

What now?

There cannot exist a theory of everything. For, suppose there exists a theory of everything that explains all those theories that cannot explain themselves. Question is, can the theory of everything explain itself? If the theory of everything can explain itself, then the theory of everything does not need the theory of everything to explain itself or that the theory of everything cannot explain itself. If the theory of everything cannot explain itself, then by the definition of the theory of everything, it can explain itself. The existence of a story of everything leads to a contradiction. In 2002, Stephen Hawking gave a lecture titled "The end of physics" arguing why he gave up hope of ever having a theory of everything. Much like the mathematical community, the physics community just shrugged their shoulders and continued. This may partly be because people are willing to give up some answers at the cost of self-verification and consistency.

This had implications for Turing. Can we have a computer programme that decides which programme will run for an infinite time and which programme will halt after a finite time? In other words, can we have a programme checker? Suppose there exists a computer programme X that takes all other programmes as input I. X prints "halts" if a programme I will not run indefinitely and X prints "loops forever" if I does not halt. What happens if I=X? If X loops forever, then X will halt but if X halts, then X will loop forever. This is related to the Incompleteness theorem by the following: there cannot be any method to check whether any theorem is true or not.

This paper was published in 1936 titled "On Computable numbers and applications to the *Entscheidungsproblem*". This was the *Entscheidungsproblem* which Hilbert had proposed. The second part of this talk will now focus on what it has to do with computable numbers.

Real numbers are infinite. No machine can compute the complete list of the numbers. It has to terminate at some point. A computable real number is the one that can be computed with an arbitrary degree of accuracy. However, these are very few. Pi, for example, can be computed by $\sum_{k}^{\infty} \frac{(-1)^k}{2k+1}$ or $e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$. This gave the first idea of having a computer. Turing originally gave the idea of a general purpose computer with his 1936 paper. The original development was an engineering problem, as far as I'm concerned. Turing used the idea of a machine which was later called a Turing Machine.

So the computer is actually the result of a failure. We're all benefitting from the result of a failure. This failure gave us Google. The computer business, the digitalisation of society, has been flourishing ever since and it is outrageous to note that the computer is the result of a failure.

Now, mathematicians worry a lot when they find a statement that has no proof let alone the fact that it can never have a proof. They'd rather die that give up on it. So they tried to change axioms but like I've said, that would've been useless. Many mathematicians were in a state of shock but then shrugged their shoulders and continued with mathematics. In the 1920s, there was a parallel crisis in physics of Quantum Theory that had a lot to do with giving up locality and determinism in the favour of randomness – "God does not play dice" and all. Then the World Wars came up and this whole idea seemed to be buried in dust until Gregory Chaitin revived them. He had been observing all these as a child. He used the idea of randomness and brought it to mathematics. Maybe that instance wasn't the only instance. Maybe there was more to it – tip of the iceberg. Maybe sometimes you can't prove something because it has no solution or because we're not smart enough but because it's not there. May be God does play dice. May be the answer has no pattern, no structure. This idea is not foreign. There is work in mathematics that suggests a probabilistic distribution of primes numbers. Some numbers are also considered to have normally-distributed decimal figures. This is also done with any base so you can see that it's not just numbers but something greater than that.

So like Quantum Mechanics was incorporated into Physics, randomness was incorporated into logic by Gregory Chaitin. The OED defines randomness as "having no definite aim or purpose; not sent or guided in a particular direction; made, done, occurring, etc., without method or conscious choice; haphazard" but what is randomness to a mathematician? This had to be defined.

An idea or statement is random if it can't be compressed any further. This is like theorems. This can be translated in terms of a programming language. The chain in the beginning says all. Also, it's about physics and the world, too!

So it really is the lack of structure or order, order or pattern. So there is no precise theory that explains it. How is this relevant? Because then no axioms that can be used to derive it.

By the way, Occam's Razor? Any one?

It's not exactly unpredictability but something like it.

A computer programme is random if the execution is same as the algorithm. Examples are printing a few irrational numbers.

The subject called Algorithmic Information Theory was developed. It had ideas for complexity. This is just the amount of bits of a programme. Leibniz had clearly mentioned this idea. Complexity is essential because we are going be discussing the size of the algorithm or the theory.

Now we can also take theory as something of a programme or a theorem. So, essentially, this translates into mathematics. And physics for that matter.

What does this imply? Godel's incompleteness theorem every where you turn

- You can't tell if a statement is random or not. Impossible
- You can't tell the complexity of a programme. This is stated directly from the Halting Problem.
- Instead, we can calculate only upper bounds but not lower bounds. This is bad because not all the smart algorithms can be known, then.
- Then we have the Halting probability.

It's donated by the number omega. It's not a constant but varies so it's called Chaitin's Construction rather than Chaitin's Constant. How does it work? Suppose that the computer we are dealing with has only three programs that halt, and they are the bit strings 110, 11100 and 11110. These programs are, respectively, 3, 5 and 5 bits in size. If we are choosing programs at random by flipping a coin for each bit, the probability of getting each of them by chance is precisely $1/2^3$, $1/2^5$ and $1/2^5$, because each particular bit has probability 1/2. So the value of omega (the halting probability) for this particular computer is given by the equation: omega = $1/2^3 + 1/2^5 + 1/2^5$ = .001 + .00001 + .00001 = .00110
This binary number is the probability of getting one of the three halting programs by chance. Thus, it is the probability that our computer will halt. Note that because program 110 halts we do not consider any programs that start with 110 and are larger than three bits—for example, we do not consider 1100 or 1101. That is, we do not add terms of .0001 to the sum for each of those programs. We regard all the longer programs, 1100 and so on, as being included in the halting of 110. Another way of saying this is that the programs are self-delimiting; when they halt, they stop asking for more bits.

Note: the situation is worse for undecidable problems!

This is something very simple. It uses principles which we are thought in our senior years at school. It appeals physicist because it has maximum entropy – each bit is determined independently but not mathematicians because mathematicians don't want logic to be randomised. For now, logic has just been reduced to philosophy and its role in mathematics concerning the development of logic stands outdated. It's like Physicist sneer at them saying, "Aha! You're no better than us!"

This number is maximally unknowable. Since this is the case, there cannot be any algorithm that describes this number. This number is not computable because each time the computer has to perform a coin toss. So this cannot be compressed into anything further. It's random.

Equivalently, if we have a statement that our mechanical proof checker creates that way, there is no way we can decide if it is correct or not because the statement is determined randomly. It cannot be obtained from any "axiom".

So it turns out we can really never have a complete set of axioms. Hilbert wasn't just wrong. He was very wrong.. for now. The stance is that no set of axioms can complete mathematics because even if we could formulate a complete mathematics, we'd need an infinite amount of patience.

There have been many crises in mathematics. Pythagoreans had to face one. The philosophy of Pythagoreans can be contrasted to neo-Pythagoreans (supporters of Digital Physics or Digital Philosphy include Steven Wolfram, Edward Fredkin on www.digitalphilosophy.org, Seth Lloyd, Wheeler with "It for bit" and Tom Toffoli). Infinitesimal calculus was under attack when Bishop George Berkeley published his The Analyst subtitled "A discourse addressed to an infidel mathematician" using his famous line "ghost of departed quantities". Euclidean Geometry with the parallel postulate (states that a line can only have only one another line which does not touch a point that is around the line). Mathematics has come out of those crises. The Church-Turing hypothesis for a Turing Machine that states that "Any algorithmic process can be simulated efficiently using a Turing machine" that was later modified to Any algorithmic process can be simulated efficiently using a *probabilistic* Turing machine" because David Deutsch later realised that the universe obeys the laws of Quantum physics. The Church-Turing hypothesis was also later renamed to Church-Turing-Deutsch hypothesis. Some facts have also been changed. For example, how many of you believe 1+1 = 2? This equals zero in a specific (modular) arithmetic.

Maybe one day we can discover a system or a reason to emerge free from paradoxes or form a mathematics that uses a modified version of the implication.

Just for your interest, there is article in *Scientific American* by Gregory Chaitin titled "Limits of Reason". There's also an article in *New Scientist*, the British version of *Scientific American* called "Random Reality" that uses Gregory's ideas to say that space-time really is random. It turns out that physicist are inspired by the limits of logic. First they weren't but now they are. They can relate this to Maxwell's Demon.

If a system is consistent, then it is incomplete. This statement is isomorphic to our nature. If we believe things make sense and one sense only, we struggle to acquire a complete meaning because it is (currently) incomplete. This theorem tells us that our search is fruitless and that things will never end. It is only that it satisfies our aesthetic considerations. I think this is why mathematicians have had more fervour than ever even after Godel.

If a system is consistent, then it is incomplete. The contrapositive of this statement is "If a system is complete, then it is inconsistent." The above understanding to this equivalent statement applies in a reversed direction. If we know everything, they will stop making sense. Is this where we are headed from Quantum Mechanics, or maybe deeper? Are we stuck in a dichotomy?

Note: We have used logic to understand logic. We have appealed to the system to inform us about the system. Our thoughts think. We write "Urdu" in urdu. We write "English" in english. Is this a free domain? Where things go fuzzy? At a certain point, our desires parallel our ego. Limitless parallels infinity (these are distinct concepts). Making sense means not making sense. This is where I think the head and the tail of the Ouroboros are in contact. We stay on the rest of the body, we're happy. This is a vicious circle. (Our?) tamed -- or is it frenzied? -- logic is frought with ambitions. Or is it?

It is inherent that a system will contain grey areas, no matter how strong their footing will be. Having said that, is it fruitless to appeal to a system? This appeals to religion, science, society and even humanity at large. Even the question "Who am I?" is largely self-referential and thus not free of a paradox.

Are these systems our representations? That opens the door to another debate, which Neitzsche would have gladly been a part of, had he been alive.