

The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high quality educational resources for free. To make a donation or view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at [ocw.mit.edu](http://ocw.mit.edu).

**PROFESSOR:** All guys. I'm going to go ahead and get started here. Sorry about missing last time, I had to go home and visit the family. It's been a couple of months.

Today's going to be kind of a review session of a bunch of things. And we're also going to go through a dialogue, probably my favorite dialogue, maybe two, a little harmonic labyrinth. But I want to start out with kind of just entertaining any questions that people might have, burning to ask me right away, confusion over the past two lectures, or what else? What have you? Anything?

All right. I'm sure questions will develop. All right.

So chapter four, which I asked you guys to have read for the previous lecture, even though an entire lecture on recursion, was about three things-- consistency, completeness, and geometry. I'm just going to quickly define these three terms as much as they can be defined, and then lead into what the whole point of chapter four was about, and what the [INAUDIBLE] was trying to introduce you to. And that's really kind of getting these two to go to Godel's theorem.

So can anyone tell me what consistency means? Anyone? Sure.

**AUDIENCE:** The set is like no theorems contradict each other.

**PROFESSOR:** Yeah, exactly. So what Felix said was no theorems contradict each other. So basically, if we were to put this in terms of a formal system, if we're deriving things, if we're playing with  $\mu$  or whatever formal system we had, and we happen to derive a proposition,  $p$ , we couldn't somehow simultaneously derive from our set of axioms  $p$  and-- this means and-- not  $p$ , that being not  $p$ .

So basically, if we have a set of operating assumptions and we are trying to somehow formally predict today's weather, and somehow the computer spit out, well, today it's going to rain and not rain at the same time, that would be an example of an inconsistent system. So you have

things where you derived a contradiction directly.

And the thing which is really interesting about this-- and I'm not going to go into all the details of it-- but if your formal system produces anywhere in it a contradiction, you can derive anything. And this is really kind of a bad thing.

And a lot of philosophers have grappled with this question. Why is it that if we derive a statement, say, this table is red and not red, how can we deduce from that that the universe is infinite, right? But somehow when you have a contradiction, everything goes haywire. You can derive anything and problems abound. And this is going to be one of the things which Godel's two incompleteness theorems will tell us about.

But of course, I use the word "incompleteness." What does "completeness" mean? And this is kind of a harder concept to get across. And it's really, really counter-intuitive at first. And that's something I want to talk about today. Does anyone have a good definition of completeness? Any takers? Sandra? You have an idea?

OK. That's all right. So completeness is, I think, probably one of the hardest concepts. And it has to go back to a picture which I drew.

Hello. Go ahead and come on in. I've got a handout for you.

So completeness goes back to a picture I drew on the first day of lecture. And I don't know if you guys remember it. But it actually appears in a chapter which I didn't assign you all to read. And it's that idea that if we had a truth box and this truth box was somehow a graphical display of all of our theorems and the things we could prove, but things we also knew to be true-- so let's take this to be the true box. And let's take this to be the not true box.

And as we talked in the past couple days, if we had some axioms, principle points to start building truths from, we can derive all sorts of ideas from these axioms.

And this is just kind of a weird graphical tree of deductions, right? Just like when we were playing with the mu system, we started with  $\mu_1$ , and then that was our axiom. And then we applied all of our rules of inference to get all the possible theorems here. And these are things which are provable. Sorry if this is incomprehensibly small. But it says "provable."

But then we have all of this space here, which we already said were true things but provable things. And this is really kind of a counter-intuitive idea. And some of you might go, yeah, that's

exactly captures what I feel. And that's the idea that there are truths, things we know to be true, which aren't provable. And this is really kind of hard to wrap your head around. Suddenly we have things which we know are true, but how do we know they are true? It's not that we have proof of them, we just know that they are true.

And this really is going to go into Godel's theorem big time, Godel's two incompleteness theorems. And so completeness, if you want to give it a short definition, is that every true system is derive-able from the axioms, or from the system. There is no incompleteness. Here, based on this graphical drawing here, you've got obvious incompleteness. We have all of the space. We have all of these true statements which aren't reachable from our axioms.

Can anyone think of something which might be true but not provable, or they have an idea?

**AUDIENCE:** Like if there's anything outside the universe?

**PROFESSOR:** If there's anything outside the universe. Exactly. That's one idea. I was actually just reading the other day Seth Lloyd's *Programming of the Universe*. That the universe, if we kind look at it, expanded like so from a big bang. But the rate at which it expands is four times the speed of light.

Of course, the only things we can perceive travel as fast as the speed of light. So we've got this light cone. So everything inside this shaded region are things which we can perceive. But ultimately, the universe is expanding faster than that. So there's all sorts of these things about the universe which we'll never know.

And that's a nice physical example. But it's not exactly what I mean in terms of formal systems and completeness, and things have been true but not provable. And this is the heart of Godel's theorem.

So before we get there, I just want to talk briefly about geometry. We're going to meet it in a variety of two settings. In the chapter, you met Euclidean, non-Euclidean. And I want to elaborate on that and show you some of the cool things.

So I'm going to have Euclidean and not Euclidean. But we'll get back to that. Right now, I kind of want to harp on what Godel's theorem is, who this guy, Kurt Godel, was, and why it is that we have one of our three title names named after him.

So Kurt Godel was born in Vienna. He was a mathematician. And he grew up in a time where

mathematics was being directly influenced by, really, a variety of paradoxes which were popping up. And a man named David Hilbert, who really wanted to clear up all these ideas of paradoxes which arrived in mathematics-- David Hilbert felt that mathematics was our most sure and certain source of knowledge, that if there was any flaws in mathematics, we were doomed as far as human beings in terms of knowing true things.

And the paradoxes which I speak of really refer to two main things. One is this issue of Euclidean versus not Euclidean. This was a revolution which started happening in the 1800s, and people slowly started dealing with over those 100 years entering 1900.

But then there was, right towards the turn of the century, the two set theory paradoxes. And these were paradoxes which I talked about earlier. And that was the idea of the barber paradox. And for those of you who weren't here for the first lecture, the barber paradox says, suppose we have a town where a barber shaves all people, and only those people, who don't shave themselves. Well, then does the barber shave himself or does he not?

And by the definition, by the way we set up the town, it appears to be a contradiction. Because if the barber does shave himself, then according to who the barber shaves, he doesn't shave himself. And if he doesn't shave himself, then he should. So it's a contradiction.

And we thought that this was a paradox deriving from set theory. And we felt that set theory was going to be our sure and certain foundation of knowledge. We thought we could deal with these paradoxes.

David Hilbert was huge. He said, guys, look. There is no unknown. Mathematics has to have a sure and certain foundation. And we should be able to establish as a model system consistency and completeness of number theory. And he felt that if anything is true, it's got to be number theory.

And I always like doing this poll. But I want to ask you guys to vote. And I want you to decide the truth of the following. The sky is blue. And one plus one equals two.

So imagine, of all the possible worlds, which do you feel like is more true, the fact that the sky is blue or that one plus one equals two?

**AUDIENCE:** The second one.

**PROFESSOR:** So you feel like the second one should be true. Do you feel like it's true in all possible worlds?

**AUDIENCE:** It has to be.

**PROFESSOR:** So you don't think there's any universe out there where one plus one could not equal two? OK. What about anybody else? Does anyone feel like, no, come on, this isn't even perceptual.

What is this statement about? It's about these abstract entities which I just kind of created and wrote down on a piece of paper. It has absolutely no perceptual foundation. The sky is blue. I look outside and what do I see? I see the sky is blue. So surely, what I see has to be more true than this. Is anyone willing to defend "sky is blue" over "one plus one equals two?"

**AUDIENCE:** [INAUDIBLE]

**PROFESSOR:** True. So Rene Descartes was huge on this. He was like, what if this is all a dream? You know, what if this is the matrix, Neo? Right, like, this is exactly what he said.

But what if we were to base our arithmetic on something else? Suppose we based our arithmetic on rain drops. So suppose we have one raindrop and another one. And we define addition as when they meet. So one raindrop plus another raindrop just gives me another raindrop. So in this system, one plus one equals one.

And what's wrong? I mean, this is perceptually validated. When I'm driving in my car at 60 miles an hour and I have rain hitting my windshield, and I see raindrops merging together--

**AUDIENCE:** Maybe you should measure the size of yours.

**PROFESSOR:** OK, exactly. So there's all sorts of problems with identity. There's problems, maybe, with how the system was formed. But even in mathematics, we have to be very clear with what we're stating, what kind of field we're working over here.

Because suppose I'm actually working over the integers, mod two, modular arithmetic. And that just says if it's divisible by something, I say it's zero. So when I do modular arithmetic, when I do mod two arithmetic, two mod two is zero. So one plus one over the integers mod two is not two, but it's actually zero.

So I mean, there's all sorts of things we have to deal with. And there's some uncertainties. But still, these are all rigorously defined. I could say, well, I'm just working over the integers. So I know this is fine. One plus one is equal to two in all possible worlds.

So Hilbert felt like, really, this has to be it. Mathematics has to be it. Number theory has to be so sure and certain that there's got to be no problem. But what Kurt Gödel did is he established two things. And these are going to be his two incompleteness theorems.

One we're not going to really talk about so much. And I'm not going to work through all the proofs of these. But we're going to try to get an idea of what each of them mean.

The first one is that any system as powerful as number theory-- and I'm just going to let NT be number theory-- which can prove its own consistency, that system is necessarily inconsistent.

**AUDIENCE:** [INAUDIBLE]?

**PROFESSOR:** Exactly. Somehow, the the second a system as powerful as number theory is able to talk about itself, things go haywire. If it can actually prove that-- here I am, number theory, saying, look, I promise you guys, you can actually prove from me that I'm internally consistent. Well, any system which can talk about itself in that way is necessarily inconsistent. And you know, this is really kind of a nitty-gritty proof, and I can't even give you all the details.

So the second one-- I'm just going to go ahead and start it over here-- is any system as powerful as number theory-- so I'm just kind of going to go "ditto"-- is necessarily incomplete.

So this means that any system which is as powerful as number theory automatically looks like this. There are true statements which we can formulate which are not provable. And I'm going to go ahead and give you guys an English language example of a statement which is-- yeah, go ahead, Rishi.

**AUDIENCE:** What does it mean to be as powerful as number theory?

**PROFESSOR:** Good question. And I'm going to explain this a little bit. But the idea is that in order to prove Gödel's incompleteness theorems, he had to use a very interesting trick. And that's called Gödel numbering. But first, I want to give you this sentence. And then I'll tell you where that comes into play.

So here's a statement. So I'm going you as star. And if we forget what the star means, it's Rishi's question. So let's just try and remember that.

So let's consider the following statement. I talked about the liar paradox, right? I said, what happens with this sentence? So this statement is false.

**AUDIENCE:** [INAUDIBLE] in your language, you can't say things like that?

**PROFESSOR:** OK, very good. And why wouldn't we be able to say things like this?

**AUDIENCE:** Because if you have a logical and perfect language you can say, you shouldn't be able to say things like that. By the end of the problem, it would be like, how can you tell that this is consistent?

**PROFESSOR:** So what is the problem, exactly, with it? If you were to design your perfectly logical language, what would you rule out? What would you prevent sentences from doing that would prevent something like that?

**AUDIENCE:** Self-reference.

**PROFESSOR:** There you go. Exactly. Self-reference is key. And there's actually two very, very, very smart guys, Bertrand Russell and Alfred North Whitehead, who derived a book. They wrote a book. And it's not a fun book to read. I've heard it's about as interesting as reading the treads of a tire. So it's not interesting at all. But that book was called *Principia Mathematica*. And Douglas Hofstadter will talk about this book a lot.

And in this book, they develop a system, they develop exactly that perfect language that you speak of. And the basic idea, if we want to formulate what you're thinking, is that we create a level language, L1, and we only allow certain terms. So we create a kind of bag of terms. In certain sentences in L1, it would be like "the sky is blue," et cetera, and "snow is white."

These are perfectly upstanding citizen sentences, right? They never break the law. But then what we do is we prevent certain terms, we prevent these sentences from talking about themselves. But whenever we're in this class and we're talking about those sentences, we're actually speaking in another language-- L2. And L2 contains L1 as a subset. And then we can start saying things like "the sentence 'snow is white' is white." Let me do that, yeah.

So suddenly I've got the sentence "snow is white," which belongs in L1. And I'm talking about it. But that's only something I can talk about in L2. Because in order to talk about something, you can't talk about yourself. You have to leap outside of it and then, in order refer to it, you have to stand from somewhere else. It's just like you can't really see yourself until you use something else to look at yourself.

So they developed this system. But even this had flaws. And what Godel did was he actually

used *Principia Mathematica*. And he took a statement similar to this, but in fact one much more clever, in order to prove his incompleteness theorem. And it's this.

This statement is not provable. And we can specify in what system. And one of the things we'll talk about is in PM, which means *Principia Mathematica*. Well, we could say this statement is not provable in number theory.

But the bottom line is, what does this statement say? It's not like this statement because what happens if this statement's false? Well, if it's false, then whatever it says about itself is not true. So that means it's provable. So if we say it's false, then that means it is provable.

But if there's one thing which we are certain of-- yes, go ahead.

**AUDIENCE:** So the truth is, in this case, like, [INAUDIBLE].

**PROFESSOR:** Careful-- that only goes in one direction. So it is certainly true that a sentence-- OK, sorry. What's your name, again?

**AUDIENCE:** Lativ.

**PROFESSOR:** How do you say it?

**AUDIENCE:** Lativ.

**PROFESSOR:** Lativ?

**AUDIENCE:** Yeah.

**PROFESSOR:** OK. So one of the things Lativ said was that in this case, are we saying that when we're doing a derivation, when we have something which is provable we know it's true, and also vice versa. But what I'm cautioning against is that's only true in one direction.

So we certainly know that things which are provable aren't true. If we can prove it, if I can say, I can prove to you that this is so, then it automatically is so. And this is why people put so much trust in mathematics, is that the second we have a proof of something, we know it's true.

But what we just were asking about was, what about the other way? Does true always imply provable?

Well, let's ask. Let's ask this statement. What if this statement is true? Well, if it's true, what it

says about itself must be true, and that's that it's not provable. So the only way that this statement is true is if it's not provable. So suddenly, we know that we can't go the other way.

And the trick that Godel used-- and this is why we get to the star question-- is that this is not a statement in mathematics. But what Godel did is he essentially took a statement like this and he said, well, we're going to let every letter and logical symbol stand for a number. So we're going to let  $p$  be 101010. And we're going to give a unique number to every symbol including spaces.

And then once we have this, we have a Godel number for this statement. And then what we can start doing is we can start giving certain operations. Remember what we are doing  $\mu$ ? And we were saying, well, what we can always do is if we have three "I"s we can cancel it. Or if we have a string of hyphens or whatever, or a string of letters after  $m$ , we can double it.

So what Godel did is he turned each of these rules of inference into rules of arithmetic. So I'm going to go ahead and hop over here.

So what he did is he made rules of inference. And he made them equivalent to our special isomorphism symbol, to rules of arithmetic.

So this is kind of counter-intuitive. And I can't go into all the details. But we will meet them in chapter nine. And that's the idea that-- suppose we have a logical thing like the statement  $P$ . And then we also have the statement that  $P$  implies  $Q$ . So we know that the statement if  $P$  is true. So if it is cloudy, then it's going to rain.

And then if we have, well, I'm looking outside and it's cloudy, then this is equal to  $Q$ , right? So if we know that "if it's cloudy, then it will rain" is true, and we have that it's cloudy, then we can immediately deduce that it's going to rain.

And what Godel did is he said, well, these statements I can actually make into numbers. And I can make the logical symbol "and" into an operation, like addition. And I can make implies-- well, this would also be a symbol. And I can have this total operation of detachment, of pulling out  $Q$ , into a statement almost like one plus one equals two.

And then what he did is he captured the idea of provability into a property of numbers, like something being prime. So then, really, what this statement comes down to is such and such number, that number which codes for this statement, does not have a property. And that's why you need something as strong as number theory in order to do this numbering trick.

But that's just kind of a first glance at Godel's theorem. And I don't want to go into too much detail about it. However, I do want to go back and talk a little bit more about the things which we mentioned, and we kind of glanced upon chapter four. And that was the ideas of geometry. And this is really cool. And it has something to do with interpretation.

Now, I want you to remember what we mean by interpretation. And I think it's a term I briefly defined on the first day of lecture. For those of you who were here, can anyone tell me what interpretation's about? Go ahead.

**AUDIENCE:** Sort of like choosing the real world [INAUDIBLE].

**PROFESSOR:** Exactly, exactly. So-- I'm sorry, I forgot your name. Is it Lativ?

**AUDIENCE:** Yeah.

**PROFESSOR:** Lativ. What Lativ said was that it's basically like giving an example of an interpretation, to be giving a real world model for what you're doing. And we saw an example of an interpretation which was not true, right? When we assigned to our symbols "one plus one equals two" to "raindrop meld with raindrop," it didn't give us two raindrops. Instead, it gave us just one.

So that's an example of an interpretation which doesn't hold and doesn't work. But we gave some other interpretations. And when we were playing with the P-Q system, we didn't give it a real world interpretation, but instead we gave it a mathematical interpretation of addition. I always say that hyphen P, hyphen Q, hyphen, hyphen is "one plus one equals two."

And so that's not so interesting. But what is interesting is that several thousand or so years ago-- not several, but at least two and 1/2-ish-- a guy named Euclid said OK, you know what? Geometry is for us as true and certain as anything is going to get. But what I want to do is go ahead and write down the rules, write down everything we know, so that we can proceed and deduce directly from these statements and know that everything we say is true.

But in order to get his feet off the ground-- I mean, he couldn't lift himself up by his own bootstraps-- he made a series of requests. And these are known as the postulates of Euclid.

So we've got Euclid's postulates. And I'm not going to go through them all. They're actually listed, I believe, in chapter four. But there was one postulate which really got on Euclid's nerves. And that was known as the fifth postulate. And he tried to derive it from the previous

four, but he couldn't.

And that's the idea that if you have a line and a point not on the line, I can give you a line, there's a unique line that goes to that point but never intersects this line. And of course, in geometry we say that lines extend on forever, and line segments kind of terminate. But these are good old infinite lines.

But he could never prove it. And there were efforts for well over 1,500 years to try to prove the fifth postulate. But it's such an intuitive and obvious statement, right? I mean, does anyone feel like this isn't right, that there's any reason why we can't assume this? Good.

**AUDIENCE:** What if you like [INAUDIBLE]?

**PROFESSOR:** Exactly, exactly. So this gives us the idea of non-Euclidean geometry. And just to give you a quick example, suppose we're on the surface of a sphere. And we define our lines to be the great circles. So the way you make a great circle is you take your sphere with your center,  $O$ , and you cut through it. And you make a plane slice.

So that goes through the origin. And you define a line to be this great circle which is formed. So any line on here that you can form, any line that you necessarily have-- and these are kind of like your lines for longitude or latitude, except not necessarily because we couldn't do something like this. But if we had any line that goes through that, which has the same radius as our sphere, they necessarily intersect in at least two spots.

So in spherical geometry, things obviously don't behave the same way. And one of the things you could derive using Euclid's axioms was that the sum of the internal angles of a triangle is always 180 degrees. But on a sphere, if you draw a triangle-- let's say we go from somewhere on the equator and we travel up to the North Pole. And we pivot 90 degrees and head back down to the equator. So we've got a right angle here, and a right angle here. And we also have a right angle here.

So for a spherical triangle, we can actually get up to 90 plus 90 plus 90-- 270 degrees. So obviously, this doesn't hold in the spherical geometry. And similarly, we have hyperbolic geometry. And this is something which is a very beautiful subject.

And you have several models of how hyperbolic geometry works. You can think of them as projections. And one is kind of the upper-half plane model. And then one's just kind of your unit disk here. And what we define our lines to be is these segments which end with right

angles on the outside of your circle.

So what we can do, then, is actually construct two lines. We can actually construct, infinitely, many lines that don't intersect each other.

So in here, you had two intersection points. And here, if you had a line like that, you would only have one intersection point. But here, you could have a whole family of lines with no intersection points.

But the weird thing is that we can give our same terms, our same statements, like point and line. And we can do a lot of the same geometry which Euclid did, except if we give them different interpretations, like, we'll define the line to be like this, or we'll define a line to be like this, then different things happen. Yes?

**AUDIENCE:** So the [INAUDIBLE] necessary, but the way you interpret it is [INAUDIBLE].

**PROFESSOR:** So here is a fact. It's true that with the four original postulates of Euclid-- sorry, what Lativ said was that, necessarily, what's true in your formal system is the interpretation you give them.

That is true in this example, right? Here, the truth of your statement directly depended on how you interpret your terms like point, and line, and things like that. And the problem was is that in the assumption, the axiom, which was Euclid's fifth postulate, was that what Euclid's fifth postulate said only it could be interpreted consistently with the other four postulates, and when he did it in simple, plane geometry like we're working on the top of this table.

But the second you interpreted all five of his postulates in this setting, the fifth one was inconsistent with the previous four. And what it said was inherently wrong. And similarly, things with this-- you had to be very specific about your interpretation and what you assumed. Otherwise, you could get an internally inconsistent interpretation.

So this is all part of a family of things called hyperbolic geometry. And inherently, what this has to deal with is the beauty of complex numbers. And you can do things in hyperbolic geometry which just completely boggle the mind. Like suppose you had a circle, a line-- this is a line, remember, because it ends with perpendicular points on the real axis.

And you can find a mapping which takes it up to here and preserves the distance between these two. And there's all sorts of different things you can do. And it's just completely a gorgeous subject. I encourage you all to learn more about it.

But this was one of the examples. Because what we thought for well over 1,500 years was that what Euclid said was as sure and certain as any knowledge that we could have. And people would often try to base their arguments and try to derive them to geometry.

It's kind of a funny anecdote but Karl Marx and Frederick Engels, when they were writing their texts, they actually tried to reduce what they were saying to mathematics. Because they felt like if they could prove the system they were advocating in terms of mathematics, then people would have to accept it.

But what happens in mathematics itself is inconsistent. If you get paradoxes like the set theory paradoxes, or you get these possible interpretations where things are sometimes true or not true-- what happens then if mathematics is not a sure footing?

And this is a problem which I want you guys to think about as we move along through this book. And what does it mean to provide an interpretation and things like that?

So what I want us to do is take a quick break, because we're going to go into one of my favorite dialogues. And you'll see the purpose of this [INAUDIBLE] later. But because it's a long dialogue I want everyone to kind of take a break and get some food and drink. And we'll then read the dialogue. But I'll need some volunteers for reading.

So let's go ahead and take a five minute break before you start reading, OK?

So a little harmonic labyrinth. Which did you guys think?

**AUDIENCE:** Confusing.

**PROFESSOR:** Confusing? Why do you say it's confusing?

**AUDIENCE:** It's like it switched roles in between [INAUDIBLE].

**PROFESSOR:** So in what way do you mean "roles?"

**AUDIENCE:** It's like the [INAUDIBLE].

**PROFESSOR:** So there's role flipping in terms of the way they treat each other, or--

**AUDIENCE:** It's like the characters [INAUDIBLE].

**PROFESSOR:** Ah, OK. Yeah, and more interestingly, though, they had an opportunity to do that by constantly going down to nested roles. And they could essentially be new people in some ways. So that's good. Does anyone else have-- oh, yes, Sandra?

**AUDIENCE:** They talk about themselves in some of the dialogue.

**PROFESSOR:** Right, so we had this weird playing around with levels. And in some ways, that was kind of hard to capture perfectly in terms of audio. But I'm sure most of you saw as we were reading along that there was indentations in the text. And that was a visual reminder as you're reading what level of the story you were at.

And Douglas Hofstadter used all sorts of really nice tricks where you would have characters like, oh, I think he means tonic. And they would be talking up here on level one. And down on level two, they would say, oh, thank you, yes, tonic is exactly what I needed. Even though in this situation, these guys don't really know about their higher levels of reality.

It's just like the same question of what happened to the weasel when here he was sitting in our everyday normal life, and he took some popping tonic? And he popped up to a higher level of reality. And it kind of makes you wonder.

I know for me, I always get the visual image of playing the universe as a fractal. And we spend all our time living down in this corner of the Sierpinski gasket. And then one day, somebody takes a popping tonic and someone's like, holy mackerel, there's actually all of these levels.

And then just the fact that you've had that one experience of playing around with two levels of reality, it makes you speculate. Well, why can't there be more? Why can't there be an infinite set of realities? And how do I know what mine is?

And there's something I want us to just bring out into the open here. And it's the idea that when I first started this class, on the first lecture-- I know not all of you were here-- I said the fundamental thing we want to answer at the end of this course is, what is an "I?" What makes something conscious from unconscious things? How do we get particles and atoms to start talking about themselves like the way we do?

So fundamentally, in this class we're going to be talking about a lot of really deep and profound questions. And I want everybody to not feel afraid that their opinion might be persecuted. Because even in this story, we managed to meet God during the middle of this dialogue. And if we can't talk about God, it's going to really narrow what we're allowed to talk

about.

And similarly, when I'm saying, what is the mind, how do we get a physical brain to then start operating with mental and conscious thoughts, that's very fundamentally asking questions about the soul. And what your guys opinions are on that become important. And I don't want to feel like anybody is learning in a hostile environment. So I encourage you guys to speak actively.

It's interesting because Douglas Hofstadter presents a very unique picture of God, right? He picks this recursive idea of a stack of infinities.

And just as an anecdote-- and this is actually a little historical fact for you-- the Jesuits, right around the time of the development of calculus, we had Isaac Newton, and Leibniz, and these guys doing their stuff in the 1670s, maybe a little later. But it really took everyone else into like the 1730s. Newton, Leibniz-- these guys that developed the calculus, studying the infinitesimally small.

Remember, when we're playing around with calculus, we're asking what happens when we approximate a function originally in a finite way? And then what happens if we take the limit to something which is infinitely small? And what does that mean?

And when you start playing around with calculus, you can find things like the area under the curve. And the way you approximate this is these blocks and taking infinitely small limits.

The Greeks, they were really close to developing the calculus. Archimedes was, in many ways, conceptually just a few stone throws away from it. But the Greeks were also much smarter than Newton and Leibniz. Because they said, well, dumbos, when you take a bunch of infinitely small things, you can't get something finite, right? Like, how is it that I can take a bunch of two dimensional circles which are infinitely thin and then stacked them on top of each other in order to get you know a cylinder? It doesn't make sense. It does not make sense.

So really what happened here, and then getting into Euler and these guys, is they essentially were trying to take our concepts of the infinite and making them rigorous. And the reason why I go on this is that the Jesuits, around this time and getting later, were one of the very first groups in schools to start teaching their students calculus. Because they felt that if they understood mathematically and rigorously, and they could deal with concepts of the infinite, they had a better understanding of God.

So the Jesuits deeply felt that your understanding of calculus was essential to your understanding of God. And I think this very much goes in spirit with the little harmonic labyrinth where you saw this image of God over Jinn. And "Jinn" is actually an Arabic word for genie. And then this going off to God itself and then coming back.

So this really requires wrestling some of the conceptual tools behind dealing with the infinite. And it's not something I can teach you fully in this class. But I encourage you all to pursue it.

Quick question-- you'll notice that each of the genies did it in half the amount of time that it took the previous genie. Can anyone tell me why, or at least why Hofstadter went ahead and paid attention to that detail?

**AUDIENCE:** Because you don't want to do it for everyone?

**PROFESSOR:** OK, so you need something-- and Felix, go ahead.

**AUDIENCE:** [INAUDIBLE]

**PROFESSOR:** Great, one moment.

**AUDIENCE:** [INAUDIBLE]

**PROFESSOR:** Right. OK, so the amount of time it took was one plus  $1/2$  of one genie's time plus half of the previous genie's time plus-- yes, and so on. Yes, sorry. I knew that wasn't right.  $1/16$ , I missed a term. Dot, dot, dot.

And do you know what this equals?

**AUDIENCE:** [INAUDIBLE]

**PROFESSOR:** OK, there you go. We've got some winners.

So this is actually an example of a geometric progression. And this gives you an idea of-- we actually know that this infinite process could converge in a finite amount of time. Really, it took until this time. And if we were to go way back. Let's go ahead and have a logarithmic scale backwards to around, I think it was, 200 or 300 BC.

That Zeno of Elea used, actually, the same argument for saying that motion was inherently impossible, and that we can never go anywhere, and all motion was an illusion because it

would require doing an infinite amount of stuff, and that you could never do an infinite amount of stuff because it would necessarily be infinite. But it took us, as a human collective conscious, well over 1,700, close to 2000 years, to understand and develop the tools necessary to deal with infinities.

So I think that's a really important thing to deal with. And I could go on and give entire courses about infinities, and talk about a lot of the important characters. And there's just one term I want to introduce, and a couple of concepts really quickly. That's the idea that there's never a top infinity. And you can always construct more infinities from smaller ones. And you can actually carry out a formal system for playing around with these infinities.

You have your natural numbers, three, and then dot, dot, dot. And then we can just say, OK, well, let's take all of those guys. And we'll call them omega. Well then, we can take omega. And then we can take omega plus one, and dot, dot. Now I've got two omegas. And then we can actually start carrying out cardinal arithmetic.

And there's kind of a mix of notations and concepts between different fields. But you might also see this first level of infinity as aleph, aleph not, or aleph sub zero. And this refers to the level of infinity which you get from the natural numbers.

Of course, some of you may or may not know this, but we've got all sorts of infinities. We can start constructing tons. But we can even define exponentiation. So one of the big things is that, well, what happens when you have two raised to the aleph not? Well, the claim is that this is aleph one, which is roughly the size of the reals. And we're talking all of your numbers on the real line.

And this is kind of a paradoxical thing because in here we have this many, all right? And what's very strange is-- yeah, go ahead.

**AUDIENCE:** Isn't it like, [INAUDIBLE].

**PROFESSOR:** Yes, exactly. So you just actually stated that a very rigorous form of defining something to be infinite is that you can put it into one-to-one correspondence with itself. What Lativ said is-- he asked, is one of the properties of infinity the fact that you can put it in correspondence with a subset of itself?

And so for example, we can take the normal integers. And fortunately, we have an infinite amount of those guys. And we can put them into one-to-one correspondence with the even

numbers. And we can create a completely bijective map just by division or multiplication by two.

But the weird thing is that we intuitively feel like all of these guys are inside of here so there should be less of them than these. But with infinity, you can do all sorts of things. And that's what's magical.

And what's nice is you can also take the interval zero to one. And you can create a map which sends this to the entire infinite real line. So it's amazing because we can capture an infinite thing in a very finite way.

And we also know, because of a good guy named Cantor, that these are really different quantities, that there's a level of infinity different between the natural numbers and the reals. And that deals with the diagonal argument, which I can maybe show you one day.

So you guys read chapter five, which was recursive structures and processes, for today's lecture, although we've been talking a lot about other things. And this is really motivated because you got an excellent lecture last time from Curran Kelleher who showed you all of the different varieties in which infinity and recursion come together.

But fundamentally, we still have a question which we're pursuing. And this derives back to, I think, our two most important tools for thinking which we'll meet in this first part of the book. And that's recursion and isomorphism.

And remember, isomorphisms come about when you're trying to put equivalence relationships between one thing and the other. And you can do it in a well-defined way.

And I want you to think about, what's the relationship between these two? And what really connects these concepts and the way in which we deal with things?

But I want to highlight just at least one section or two from chapter five which I don't know if you guys found interesting as well. It's that idea of, well, we have recursion here. And we can do all sorts of different things, right? Curran showed you how we can construct fractals, and trees, and mountains, and all sorts of beautiful things. And with the recursive transition networks, we can get recursive programs to create sentences and language.

So then it appears that, well, if recursion is at the heart of intelligence, why are humans so bad at it? And two, is it really going to be what leads us to creativity and creating a computer which

we can't distinguish from a man or from a person?

And it's funny, just to show the date on this book, Douglas Hofstadter actually says people said it would only take 10 years before we could create a computer program that would be the world champion in chess. And of course, from then they said it would take another 10 years, and then another 10 years. So he's kind of alluding to the idea that it would never happen.

But sure enough, in the early 90's, IBM had developed Deep Blue who beat Kasparov, then-world champion at chess. And it was kind of a triumph showing that the way in which a human thinks and the way that a human plays chess is very much more intuitive than an analytical problem solving.

And he was talking about replying to a recursive algorithm. You basically analyze all possible moves. And then you choose whichever one would be worst for your opponent. And then in deciding what move they're going to make in the next step, you apply the same method, except you take on your opponent's role and you say, well, analyze all of the possible moves given that one. And what would be the worst move for them? And then you continue this.

And depending on how far out you can search this tree-- we have this conceptual tree of chess moves-- it really helps give you an advantage. But Kasparov doesn't think like that. The way Kasparov thinks is he intuits something. And that's really a magical element of human intelligence which we haven't yet been able to capture with our computer programs. Yet somehow, Deep Blue was able to beat him. And I think this is an interesting example of recursion and what its role in intelligence is, and whether it's going to be the final answer.

Just to give you a quick show of things to do-- I'm not talking much about chapter five just because we've read that, we've done that. We've had an entire lecture on recursion and its possible roles. I want you guys to pay careful attention to chapter six, which is your reading assignment for next time, because they're going to fundamentally ask the question of, how do we get meaning? How do we know that our words mean anything?

So the idea of developing a theory of meaning of language goes back to the idea that, suppose I were to plop you down on an island in the middle of nowhere. And all of these people are going, blah blah, blah. And they're speaking their own kind of language. And you have no idea what it means, right?

Yet, you're stuck there. You depend on them for survival. And you figure it might be a good

idea to go ahead and start learning their language. How would you go about doing that?

So suppose you go out on your hunting missions with this tribe. And every time there's a rabbit which gallops through, one of the guys says, "gavagai." "Gavagai." And you're like, OK, I'm going to create an internal dictionary here. I'm going to write a dictionary for this language. And I know that "gavagai"-- I don't know if that's how you spell it. That's equivalent to rabbit.

But now let's switch roles. Let's suppose you're one of those native people. And you part of your culture is that you never view something as greater than the sum of its parts. So when you say "gavagai," you don't actually mean the whole rabbit. But you just mean un-detached rabbit part. Because the second you start splitting up the rabbit and you're cooking it, and you have its leg over the open fire, it's no longer "gavagai." Just like when we refer to a cow, we usually talk about beef. And the same with a pig, we usually talk about pork once we start eating it because we have this idea of connecting "gavagai" to the full rabbit.

But they actually want that to be un-detached rabbit part, which has a completely different conceptual network for them than it does for us. So this equivalence isn't true.

And this can actually be a very rigorous problem which presents itself in the theory of meaning and language. And we're going begin focusing on that and the role of isomorphisms in the next lecture.

Go ahead, Lativ, you have a question?

**AUDIENCE:** So then how are you going to learn the language?

**PROFESSOR:** How are you going to learn the language? Well, we can obviously develop some sort of functional apparatus, like I can get close to a rabbit. But how do we actually decipher meaning, right? We can obviously learn a language in some ways. And in some ways, it's the same. How do you know what I say is exactly what you want me to say? And--

**AUDIENCE:** [INAUDIBLE]

**PROFESSOR:** And that's exactly it. So then how do we develop a theory of meaning?

But plenty of good questions. Good questions for next lecture. And turn in the surveys. And I look forward to seeing you guys next time.