THE ASPEN INSTITUTE

ASPEN IDEAS FESTIVAL 2015

ASPEN LECTURE:
I AM OR AI AM

Paepcke Auditorium
Aspen, Colorado

Wednesday, July 1, 2015

LIST OF PARTICIPANTS

EDWARD FRENKEL
Professor of mathematics, University of California, Berkeley
Author of "Love and Math"

* * * * *

<u>I AM OR AI AM</u>

MR. GERSON: Good morning, everyone. We're going to get started, so if you could take a seat. I'm Elliot Gerson of the Aspen Institute, delighted to see you here this morning. We know we exhaust you over these sessions, but this is going to really be remarkable. I am pleased for a couple of reasons; one, I'm pleased at the response we've had to the track on mathematics. Every year, we think about things that we could do, and we thought that having something on something so important yet so intimidating to so many people would actually be exciting, and the reactions across our entire range of participants has been even better than I could have imagined.

I mentioned to those of you who may have been here for a couple of other mornings or either of the other mornings where we've had Aspen Lectures, we added this feature to the Ideas Festival last year because people had said to us who'd been coming over a number of years, that they thirsted for something more like a traditional lecture. For some of the extraordinary people we have, they were frustrated that they could only listen to them when they were on a panel with two or three or even four people, and asked us if we couldn't have more opportunities for slightly deeper dives into subjects, but of course, reserving opportunity for questions as well.

And that's how we launched this special series called the Aspen Lectures, a special branded series. And they're all being taped, and you'd be able to tell all your friends who are at many other things this morning that they should look at this.

It's my pleasure this morning to introduce Edward Frenkel, a professor of mathematics at Berkeley. He's also won the Hermann Weyl Prize in mathematical physics, and his latest book, "Love and Math," an incredible, wonderful bestseller, is something that I highly commend to you. The book has catapulted him into fame in at least 15 languages now and even as a movie star as that book has been turned to film.

He asks a timeless question for everyone today, essentially the question that we've been asking since our species emerged. Who am I? Edward Frenkel.

(Applause)

MR. FRENKEL: Thank you very much for such a wonderful introduction. Thank you, Aspen Ideas Festival for having me. It's really such a wonderful gathering of wonderful people and sharing ideas and experiences, which I think is wonderful and we need more of that. So, now is my chance to maybe to share some of my experiences and some of my ideas, and thank you all for coming. It's an early talk, 9:00 in the morning. You know, normally you wouldn't wake me up at this hour with a cannon, but I just came from Europe, so I'm jetlagged. So, it's working out just fine.

(Laughter)

MR. FRENKEL: Perfect scheduling, and so I guess one of my tasks today is, kind of, to help you to wake up also, and maybe wake up myself a little bit as well. We'll see how it goes. But my talk -- I would like to talk about AI, the artificial intelligence, which has been in the news a lot lately, and I'm sure you have seen it. Let me just put a couple of quotes for you. Some of the great minds today are very concerned about the developments with the artificial intelligence such as Elon Musk and Stephen Hawking and Bill Gates.

With AI, we are summoning the demon, the demon, as Elon Musk said, and it might be the biggest existential threat that we have right now. And Stephen Hawking says, the development of full artificial intelligence could spell the end of the human race. Are they being alarmist?

SPEAKER: Yes.

MR. FRENKEL: Okay. Now, we have a discussion. We'll see. Okay. So interesting thing is when we hear this, our mind immediately starts looking back to depictions of artificial intelligence in popular culture, going back, for example, to the really amazing film by

Stanley Kubrick, "2001: A Space Odyssey," the notorious HAL 9000. "Yes, I can hear you, Dave. This mission is too important for me to allow you to jeopardize it." Very prescient, I might say. 1968 is the year when Kubrick made this film. Or more recently, you have seen another one, "Ex Machina," and of course, I would be amiss if I did not mention "Terminator Genisys" out in theaters today. I swear I have no contact with the marketing team for that movie. This is purely coincidence. Everybody's like, yeah, right.

(Laughter)

MR. FRENKEL: And so, this is an unfortunate thing, I think, about this debate that we fall back onto this discussion of the future and the robots - are robots going to present some threat for us, and so on. And what I would like to talk about today is -- the point I'd like to make is that actually I think this conversation is very important. This discussion about artificial intelligence is very important, but not because it is a discussion about the robots and about the future, but because actually, secretly perhaps, it is a discussion about us humans and about the present, about now, about who we are, about who we want to be.

But let's see the other side, so to speak. Let's see the other side, Ray Kurzweil, the great prophet of singularity. Singularity refers to technological singularity perhaps, refers to a time when the abilities of our computers, the supposed time when the abilities of our computers will equal those of humans, and according to Ray Kurzweil, this will happen around the year 2045. By that time, he says, we'll be able to fully back up our brains, we'll be able to think in the cloud. Some of us, I guess, already think in the cloud.

(Laughter)

MR. FRENKEL: We're going to put gateways to the cloud in our brains. By the late 2030s, human thought would be predominantly non-biological. These last three quotes, by the way, are from a conference in New York, which happened just four weeks ago. I am not making this up. And earlier, he said a couple of years ago at another meeting, he said we'll be uploading our entire minds to computers and become immortal by 2045. That's Ray Kurzweil. That's a nice picture of him. He looks like he just came from a casting call for the next James Bond movie where he was supposed to play Dr. Evil or maybe Austin Powers -- I don't know, maybe Mike Myers' not available anymore.

(Laughter)

MR. FRENKEL: Very nice picture. He's not alone in this. You can go online and google "Initiative 2045," the Avatar Project, which is a child of a Russian multimillionaire, Dmitry Itskov, my compatriot as you might guess from my accent; I am from Russia. And he has this all figured out. There are four stages, the phases of Avatar Project.

You know, preparing for this lecture, I spent some time searching online and I've been reading all these stuff. So, I'm just -- I'm so tired of this, honestly. So I'll just let you read it. I'm not -- I don't want to --

(Laughter)

MR. FRENKEL: I can't do it anymore, honestly. It's artificial brain which will transfer all of that somewhere. Let's just transfer somewhere ourselves. Let's just run away somewhere. Actually, so I just came from Paris and, you know, it's not just a discussion here in the United States. One of the major publications in France, Le Nouvel Observateur, just ran a cover story on artificial intelligence and transhumanism.

And this fellow, Dmitry Itskov, was one of the main characters here. The title here says, "The man Who Wants to Kill Death." And he's actually very serious about it. He's put a lot of money into this. He's organized several big meetings on this subject, Initiative 2045, that's what I'm talking about. And he actually is doing what he's talking about. He actually created an avatar in his own liking, in his own likeness -- his own liking as well, I suppose, which is kind of interesting because it's such a sad robot. He is such a sad robot.

In Russian, it says the technology of immortality. I'm not sure I would like this kind of immortality, you know, to be honest. You know, it would be nice if he'd put a little smile, that maybe make it a little more attractive, you know, maybe we'd be seduced by that. But when you look at this picture, and frankly at this picture, I am not so sold. I'm not so sold on these ideas.

So, you know, I am a mathematician, and so I feel like I have to approach it with an open mind. If this is what mathematics teaches you -- and by the way, I'm so happy that Aspen Ideas Festival this year has mathematics as one of its tracks. We need more. We need to talk about mathematics, and one of the reasons why we need to talk about mathematics is precisely because this kind of issues are very closely connected to mathematics. And if we don't talk about it, if we don't talk about mathematics, if we run away from mathematics, if we are scared of mathematics, if we think we hate mathematics, what we're doing essentially is we're giving all the power to this -- to these people who espoused this ideas, the man of Silicon Valley, I call him.

There are some women too, but unfortunately mostly men. I think actually it would do them a lot of good if more women came to Silicon Valley. But for now, it's really mostly men. And so, we have to be aware and we have to know what's going on, so that we understand the fallacy of these ideas. And so, what I wanted to do, in part, to kind of prove to you that I am a mathematician, I wanted to do something on a -- so, I asked for this little notepad.

But like I said, I have to approach it with an open mind, because this is what mathematics teaches you. You never miss anything, any ideas outright. And you realize that a lot of things are paradoxical. They are not -- you know, the easiest answer you have is usually not the right answer, but the paradox as Richard Feynman famously said, it's not really a paradox in reality. It's the tension between reality and your image of reality.

So, I say, okay, to what extent is it really possible to say, transfer our mind to -- our brain to a computer, to a machine? So, what does it mean? And what is a computer? A computer is an algorithm. Computer works as an algorithm and as a turing machine at a more technical term. So, an algorithm is a step-by-step procedure. It's a step-by-step procedure as one thing falls from another.

Like imagine, you're doing long division of two numbers and so what they are saying implicitly when they are saying that this robot can actually, sort of, come alive that it can transfer the brain on to this robot, you're saying that the function of our brain -- the functions of our mind on consciousness can actually be replicated by an algorithm. That's what it says.

So, that makes me -- that made me wonder, to what extend can we actually represent things by numbers, because an algorithm ultimately is a sequence of zeros and ones. You can transform an algorithm into a binary code. So, it's a line. It's a very long line, but the very long line of binary code, zeros and ones, right. So, what they're saying is that I am just a sequence of zeros and ones, and each of you is just a sequence of zeros and ones, and that just then from their point of view becomes a technical issue of finding that sequence.

So, that made me wonder to what extent can you actually do that, and I found this very simple example which, I think, gives a very good illustration of what's going on and this has to do with the concept of a vector. So, I'm not going to overload you with lot of math, don't worry. What is a vector? Think of this notepad as a part of a plane or a two dimensional plane, which is, sort of, imagine it extending in all directions and pick a point, say, right here.

So, a vector is going to be -- a vector is a directed segment like this. It connects the origin, which we fix once and for all at special point and some other point. This is a vector. So, you can do the same for any other point. So, for example, here is another vector, and these vectors, they have their own life. It's not just -- they're not static, they're not just sitting here, but actually they can interact with each other, so to speak, okay. So, for example, you can add two vectors. There is a very simple rule, which you may remember called parallelogram rule. So, what you do is you kind of draw a line here parallel to this one and you draw a line here parallel to this one and then when they intersect, you get a new vector, okay.

And we say that this vector, the red one, is the sum, which is nice by the way. So, this is a -- it's a good example of self-referential loop.

(Laughter)

MR. FRENKEL: This is very important by the way also. I will come back to this. So, but, you know, -- so, back to vectors. This is the sum of these two vectors, okay. You can also multiply vector by numbers and so on. And this vector is like a happy family. They just live there. They have fun and you can add them up, and then you can multiply, and they play with each other. And then here is what happens. A mathematician comes and says, you know what, I'm going to introduce a coordinate system. I'm going to coordinatize you, okay.

What does it mean? It means that I'm going to draw two axes. I'll call this one x. My hair looks pretty good. That's nice.

(Laughter)

MR. FRENKEL: Even from the back, yeah, nice. I should do this more often. So, and then, you have this one which you'll call y. And then, let's take the first vector. And so what we can do is we can -- let me use another. Let's make it more colorful. So, let's draw a perpendicular here, and let's draw a perpendicular here. So, let's say that this is, maybe, let's call this 1, so then this will be 2, this will be 3, and then this would be -- maybe, I want to say that's a little bit less, so maybe like 2.5, doesn't matter, maybe, let say 2.5, and then it's very -- and then we can say that this vector can be represented by this pair of numbers. The first one is 3 and the second is 2.5, okay.

So, we studied this in school. We studied this in college and so on. And this is a very good -- it's a very useful idea, because once we do that, you can then emulate all operations that I talked about, like addition of vectors, multiplication of vectors by numbers. You can emulate them with these numbers. For example, to add two vectors, you just add the first components and the second components, right. So, very convenient, you can program this stuff. And this surely is the sequence of zeros and ones, so you can convert into binary.

But then you get so excited about this possibility of representing vectors by numbers and you say, this is the vector, actually. There is no difference between this and this, but is it really true? Is it really true? The vector is there. The vector is there. The vector couldn't care less about our coordinate systems and all the shenanigans we're doing with those coordinate systems. The vector is just there. It exists. It's us who decided to represent it by numbers, and let's not forget that this involves a lot of choice, our free will, if you want, because I didn't have to make coordinates to go like this.

I could also tilt the system, coordinate system. I could make it, you know, like this. I don't want to draw to overload the picture, but you can imagine a coordinate -- one axis going this way, one axis going that way. In fact, you learn that it's not even necessary for them to be perpendicular. So, there are infinitely many ways to draw a coordinate system. And once you draw a coordinate system, you also have to choose the unit, because if I choose a different unit, for example, this will not be 3, it will be 2. So, what does it mean?

It means that the same vector can be represented by any pair of numbers, by any pair of numbers. So, how can you say that this vector is this pair of numbers? It's not. It's one particular representation which we have chosen, and let us never forget that. So, this is a very good metaphor, I think, but what happens when we get too excited about our ability to represent things by numbers that even a vector cannot be identified with the pair of numbers, even a vector. Can I be identified with a sequence of numbers? Can you be identified with a sequence of numbers? Even this simple thing cannot be identified in this way.

If a vector could speak, if a vector could speak and you came and asked him or her, what are your numbers? What are your numbers? A vector will be like, what, what are you talking about? I don't know anything about numbers. I am here. I am. I am. So, that's one illustration for you about how careful we should be when we talk about numerical representations. Let me give you another example. When we talk about -- if mathematicians had a hall of fame, he would be the greatest logician of all times. In my book, he is. Well, I suppose you could argue that Aristotle comes close.

So, maybe two of them will share the honor, but he is really one of the greatest logicians and not just logicians, but mathematicians of all time, and, in fact, you know, he worked with Albert Einstein in the -- at the Institute for Advanced Study, and Einstein actually told his friends that the only reason he, Einstein, came to work was so that he will have a chance to walk home with Godel and talk to him in the evening. A great mind of the 20th century, famous for his incompleteness theorem. Oh, oh, it's going a little bit -- ah, great, thank you.

Actually, in fact, there are two incompleteness theorems. So, I'm talking about the first one, which says basically that in any formal system -- now what is a formal system? A formal system is something where you have a system of axioms, things which you accept as given from the outset and rules of inference like if A -- if you have A as an assertion, like an axiom, and then A implies B, then B is also an assertion, is also a sentence which you prove. So, you start with axioms, you use rules of inference and you produce step by step other statements or assertions, which you called theorems.

So, formal system works as an algorithm and in fact, there is essentially no difference between formal systems and algorithms. Now, what he said -- what his incompleteness theorem published in 1931, says, is that if you have such a formal system, which one, is sufficiently sophisticated -- it's a technical thing. It means that you can speak about whole numbers, 1, 2, 3, 4, 5 and all operations on them, and all the things we usually do with numbers. The technical term is Peano arithmetic. That's number one. Number two, consistent. Consistent means that you cannot prove that it's not self -- it's not contradictory, your formal system.

So, you don't -- you cannot prove simultaneously A and negation of A, because if you can prove A and negation of A, you then can prove every statement. So, it's not so interesting. So, suppose, you have a sufficiently sophisticated and consistent formal system, then his theorem -- and this is a mathematical theorem; it's not some kind of, you know, guess or speculation. It is a mathematical theorem, which is accepted by everyone on the planet earth. It's says that there will always be a statement which is true, but cannot be proved within the system.

In other words, truth cannot be formalized mathematically. Truth cannot be expressed by an algorithm, not because we are not sophisticated enough, not because we are stupid, we don't know how to do it in principal. This is an inherent property of our logical thinking. Truth cannot be formalized. Maybe I'll tell you a couple of words about how to prove it, because there's a beautiful argument. It's an absolutely beautiful argument. It goes to the old adage about the beauty of mathematics.

You know, David Henry Thoreau wrote that, you know, we've heard so much about the beauty of mathematics but very little of it has been sung, you know. So, let me sing a little bit about the beauty of mathematics and I think this is one of the most beautiful theorems and proofs. The idea is a self-referential loop. It's like exactly like the self-referential loop that you saw where you can see -- I can see me on the screen and then it said -- and it goes on, and on, and on, like this, you see.

Liar's paradox, consider this sentence. "This sentence is a lie," where this means it refers to itself, you see. If you haven't seen this before, you might need a few seconds to figure it out, but if you -- it cannot be neither true nor false, because if it's true then you accept that it's true, then it says it's I'm a lie, but then it is false. Or if it's false, then it says then it has to be true. So it cannot be either. It's a very interesting thing when you find it, and this, of course, has been known since ancient Greeks, I suppose. And he used the same idea, Godel.

He showed that if you have in your formal system, whole numbers 1, 2, 3, 4, 5 and so on, you cannot make the liar's paradox sentence in your formal system, but you can make sort of the next best thing, a sentence which refers to itself in this way; this sentence cannot be proved. And you see, so if you could prove it, then you would also prove that you can not prove it. And that would mean that your formal system is inconsistent, because you would prove it and its negation.

So, if your system is consistent, formal system is consistent or another word is "sound," so it's not self-contradictory, then you cannot prove this, but then it's true and then it cannot be true. So, here is a true statement, but it cannot be proved. Now, what to do? Well, in mathematics, we say, well, it's not such a big deal because -- so we get this statement, you can add it as an axiom. You can take this sentence and add it as a new axiom. So then you have a true statement which is proved, because axioms, by definition, are theorems.

But then, you obtain a new formal system and in this new formal system, there will be another statement which you cannot reach by formal argument. It's like a horizon line. You can chase it and you can run after it. You can't reach the horizon line; same thing here. So very beautiful argument. And so, in fact, this is a very strong indication also of the fallacy of these ideas that we can replicate our consciousness by an algorithm, which is the same as a formal system.

Roger Penrose, a great mathematician, has published a book in 1994, "Shadows of the Mind." It's like a big book, like 500 pages, and so -- and he used Godel's incompleteness theorem to give a detailed argument that there is something in our conscious thought process that eludes computation. We have access to mathematical truths that are beyond any robot's capabilities. Even mathematical thinking cannot be replicated by an algorithm.

And in fact, if you speak -- if you talk to mathematicians like myself, you can -- you know, they will tell us that when we actually make a mathematical discovery, at that moment, at that moment when we make a mathematical discovery, our thinking stops, our thinking stops just for a second to let in an insight, to let our intuition, our instinct, our inspiration, do its magic. That's what it means to be human.

But, if you don't -- if you've never experienced that, you don't know that experience, you can take the cue from Penrose. In fact, another scientist, John Lucas proposed a similar argument, so sometimes it's called Lucas-Penrose argument. Now, it's not a mathematical theorem, it's not a mathematical theorem, and it has been debated. Some people, you know, because Penrose is very careful. He anticipated all possible arguments and he argues. It's a very beautiful.

It's really an amazing intellectual feat what he had done in this book, in my opinion. But, of course you know, you could always find people who will go and say, well, you know, start nitpicking, well, he left a door open here, he left a door open there, that maybe is actually -- maybe it's actually possible to replicate mathematical thought by an algorithm.

So, you see, this is where we come to maybe the main point. There isn't an algorithm. There isn't the theorem, which will answer this question for us. There isn't an experiment which will tell me who I am. You can build a machine, 10 billion dollars machine like LHC, you know, it's not going to answer this question for me. I have to answer this question. I have to answer this question. I can take some clues from mathematics, from what mathematics tells me here or here that it's not possible, which by the way -- Ray Kurzweil never talks about Godel's incompleteness. I actually tried to google Ray Kurzweil, Godel's incompleteness theorem. The only results I got were, you know, pages where people were asking him the question, why you never address Godel's incompleteness, why you talk -- why you're so enthusiastically talking about transferring your mind or everyone else's mind actually to machine, but you have not even given any consideration to one of the landmark achievements of mathematics of the 20th century?

But, it is actually something very systematic. We decided that we are just bags of elementary particles assembled in a certain way, nothing more. Even great physicists who must know about -- who know about quantum mechanics, who understand that this cannot be true, they have fallen into this trap. I was just -- you know, I wanted to give you a quote just so you don't think I'm attacking a straw man, you know. I want to give you a real quote from Carlo Rovelli, a great physicist. He's actually one of the good ones --

(Laughter)

MR. FRENKEL: -- you know, these are my people I represent, but, you know, here is what he said. This is on edge.org in response to their annual question just a few, six months ago. "The world is more or less just a large collection of particles arranged in various manners. This is just factually true," he says. "Suitably arranged, it can reason like Einstein or sing like Joplin." I'm not sure either of them would appreciate that, not because of some black magic intervening. This is a physicist saying that. How about quantum mechanics? How about quantum mechanics, which tells us that there is no such thing as particles.

And by the way if you really think that this is made of particles, how about finding out that each -- you know, in each atom there is a nucleus and this nucleus occupies 10th more or less -- different estimates, but 10 to minus 15 of the volume of the whole atom. In other words, the atom is 99.99999 remember that old show of Ali G, where you'll say 9999?

SPEAKER: Yeah.

MR. FRENKEL: It's 99999998899 percent empty. So, if you are serious about this, if you think that you are a collection of particles or this even is a collection of particles, then you have to admit it, but most of it is a void, because, of course, it's not -- because double-slit experiment, when you shoot particles through this two slits and you expect just two lumps on the screen, because that's what would happen if you shot actual pellets through these two. Just behind those two slits, you would have lumps, but you get interference pattern because there are also waves. A particle, an electron is not a particle; it also has another side, but it's not a wave either nor is it a particle and the wave. It's something else. It depends on our choice of observation.

I wish I had more time to tell you that this kind of experiments also show you the fallacy of our usual understanding of time. You know, it's like I think it's Yogi Bear who said, the future is not what it used to be, but actually neither is the present or the past. They don't exist other than in our imagination, and you can prove it by experiments in physics. We don't talk about it so much, because it's very disconcerting. We prefer for some reason to say that I am a bag of particles arranged in a certain way, I'm moving around, everybody is a bag of particles and we are kind of, like, trying to find our way around, you know.

Of course that's not true. The reality is much more sophisticated than this. And so this made me wonder why is it -- what makes us want to believe this ideas. And I think we have to go all the way back to the 17th century to Mr. Descartes, Rene Descartes, the father of Cartesianism and, one could say, of modern science. This idea of a split of a duality between the body and the spirit, the body and the mind, this idea, "Je pense, donc je suis," "I think, therefore I am." Someone said, in fact, it's the other way round. When I stop thinking, that's when I am.

But we think about it and if we think about it, then, of course, we think that that's how it is. The idea of this mechanistic idea of the world, of this sort of like billiard balls flying around and as being these collections of billiard boards or balls or particles like this, it all goes back to 17th century Cartesianism. And it was very important for us for the humanity, because, of course, this ideas came in protest against the authority, the authoritarian system of the church, of the state. It was a way to give agency to individuals who were oppressed during the Dark Ages, because science is a great equalizer.

Two plus two is four. It doesn't matter how rich you are, poor or rich whatever religion, et cetera, it's the same. It's the same for all of us. It gives us dignity, agency. It was very important for humanity to find these ideas as a refuge from the trauma of the Dark Ages. But here we are 400 years later. The ideas which were -- which Descartes came up with to rebel against dogma are now turning into dogma. And the irony of this is the falling. You know, in 17th century, most people believed in God. I'm not saying it's good or bad. I'm not saying we should or should not believe in God, but there is one aspect of it.

If you do believe in God, you accept that there is something ineffable, something that cannot be explained, something beyond all logic and reason. Are we really willing to sacrifice that? So today, perhaps we've lost that and we decided to believe in science, which is great. Let's believe in science. But let's not, number one, make it into a new religion; and number two, let's not be so selective. Let's not believe in science as it was in the 17th century. Let's accept all of science, everything that science tells us. Let's accept all the achievements of science in the last 100 years; quantum mechanics, relativity, Godel's incompleteness. Let's take all of that into account. Why don't we do that?

And you know, a friend of mine -- I discussed this with a friend of mine. He told me about the psychological condition, which is called anosognosia. So, I've been reading about that. It's a very interesting psychological disorder when our logical mind -- people may -- somebody may have a severe disability, for example, paralyzed limb like an arm or a leg, but their mind, their logical mind is convincing them that that's not the case. So you'd have a person who would have a paralyzed leg, but he would say, tonight I'm going to go to dance club, you know, tonight I'm going to -- and then when you ask that person, so, did you go? Oh, I couldn't go because I was busy.

So, in other words, you find excuses and arguments and you explain to yourself why you couldn't do it instead of facing reality, instead of facing the reality of your condition. Anosognosia, is that our -- is that becoming our common disease? What exactly are we trying to run away from? What exactly is it that we are afraid to face? And you know, another point is they talk about the body as if it is this fixed object like, you know, like this podium. But we know from biology that the body constantly rejuvenates. All the cells are constantly rejuvenated.

It's this constant motion, you know, constant motion of us and the universe, and we are all connected by this, you know. It's not -- you don't have to listen to a guru, you know. You don't have to go to listen to spiritual teachers who, by the way, have been saying that all along. Modern science tells us. So, we believe in science? Great, let's take that from our science. So, when they say that they want to create an avatar, you know, inside the robot which I showed you earlier, that's a fixed thing. That is a fixed thing.

So, what do you want? You want to capture that, you want to capture that, but you can't capture a whirlpool, you can't capture a vortex in an ocean. It's there, it's moving, it's alive, but you can't just grab it, you know. That's what happens, you know, and of course, that would be fine, you know. Back to Mr. Kurzweil, and I don't want to appear like I'm picking on him, you know. But, here's what I read in the *Guardian*. He takes 150 pills a day, because he is a -- he would like to survive until 2045, which is by his predictions is the time when he will be able to transfer himself into a robot and become immortal.

But he has to survive until then, you know. Wouldn't it be disappointing if he died just one year before 2044? Damn.

(Laughter)

MR. FRENKEL: I took 150 pills a day for so many years, and it didn't work. According to some reports, it's actually -- he has increased to 250, and by the way he sells those pills online. It's called Ray and Terry, kind of like Ben & Jerry except with no fun.

(Laughter)

MR. FRENKEL: Now, look, it's his personal matter, you know. Anybody is free to take as many pills as they want. But, you know, and he is also -- it would be fine if he tried to build this robot in his garage. But, in fact, he is trying to build this robot using the resources of the biggest company, information technology company in the world, called Google. In December of 2012, Mr. Kurzweil was appointed as the director of engineering at Google, a position he is occupying, still occupying today. And according to report in -- at MIT -- in the MIT technology review, his scope has actually expanded. He is in charge of all the research in artificial intelligence in -- at Google, okay.

The *Guardian* quotes Larry Page, one of the founders of Google saying, we'll give you independence, but you will also have all the Google scale resources." And those resources are quite amazing. I don't know if you've read about this, but Google has been buying left and right, it has been on a shopping spree buying robotic companies, artificial intelligence companies. They paid almost a billion dollars for two startups, the DeepMind and Magic Leap. They bought DeepMind outright for close to half a billion; they invested more than half a billion in the Magic Leap recently, a billion dollars for two startups.

When they bought DeepMind, they announced -- a year and half ago, they announced that they will have an ethic boards on the issues of artificial intelligence. Well, I googled Goggle Ethics Board last night.

(Laughter)

MR. FRENKEL: Guess what I found? There's no information about it. Oh, wait, there was information. A couple of weeks ago there was an article in the *Wall Street Journal* in which they said that Google spokesman declined comment on the ethics board. The founders of Google have famously said, "Don't do evil" will be our motto, and I have no reason to doubt that. But Google is a company and a company is not a person no matter what our Supreme Court would say on that topic.

(Laughter)

MR. FRENKEL: A company is an algorithm. A company is an algorithm which has a utility function which is called profit. So, how do you expect to contain this research? How could you expect to remember who we are if you have as the head of your research in artificial intelligence someone like this? But I don't want to pick on Mr. Kurzweil. It's almost like, you know --

(Laughter)

MR. FRENKEL: -- because I want to be honest with you. He is me, you know. I am a mathematician, I come -- it's hard to talk about this. It's hard to admit it, but I was like this -- I mean once, you know. I know -- and now I know why. I never believed that I was a machine really, because that's what -- when you say you will transfer and so on, your mind, you're actually -- it means you are believing that you are a machine. I didn't believe it, but I wasn't sure, I wasn't quite sure.

But, you know, there is a reason for that. There is a reason for that and in my book, "Love and Math," I talk about my experience as a 16-year-old when I applied to Moscow University to become a mathematician. This was my passion. I wanted to become a mathematician.

And I was failed at the entrance exam. I was ruthlessly failed, because I went to Semitism, the policy of anti-Semitism in the Soviet Union of that era and it wasn't just that they would say, "Oh, you're Jewish" -- my father is Jewish -- "So you're Jewish, we cannot take you." No, they would try to destroy you. They would try to show you, to convince you, a 16-year-old boy that I'm not good enough. They want to break you. They wanted to break you.

In my year, 1984, George Orwell, not a single Jewish sounding -- person with a -- student with a Jewish sounding name like mine was accepted. Hundreds of thousands of students were failed ruthlessly. My exam, oral exam lasted for more than four hours, and they failed me -- they gave me questions which were not usually given to other students, much harder. They were called coffin problems, coffin. So, they wanted to put me in a coffin and put the nails and bury me there. But, I told myself -- I survived, obviously I survived. I went to technical school; I started doing research at a young age. I published papers which became known when I was 19.

At age 21, I was invited by Harvard to be a visiting professor, five years after my exam. And I always told myself a story how I was so strong, that I was able to persevere and overcome this terrible -- these terrible obstacles. It took me 30 years, 30 years to learn what really happened that day. I was asked to speak about this in New York last year in September, and the night before, you know, I wrote about it in my book, it was my story. I knew the story, I knew what happened to me, turns out I knew it in the third person.

I did not know the pain. I disconnected myself from the emotion, from the pain of that boy, and he was waiting for me for 30 years to come back, to search for him, to give him my hand and to bring him back, to give him my love and to bring him back. And when I finally connected with him, when he told me what happened, it was a totally different story. It was like a tsunami that came -- literally overwhelming for some time, connecting to that pain, and knowing that actually wasn't fine, he wasn't fine, he died. He died at that moment. This was a very traumatic experience. There is no other way to say it. He refused to accept it. I refused to accept it when I was 16.

When we are small, when we are children, we cannot accept this kind of pain. So, what do we do to cut it off? We cut it off, we split. I left that boy there, frozen and continued to move on, but then to survive I had to severe my emotional connections to him. I had to severe my emotional connections to the world. I had to feel more secure, I had to feel more safe, I had to have illusion of control.

Therefore I only relied on my logic, on my thinking, because my logical mind was constantly telling me, don't look there, don't look there, don't look there. It's a defense mechanism. Obviously who wants to connect to their pain? But the pain is still there. We push it in the unconscious. So when I connected to this pain, I let it out, I let it out. We can do that. Nobody told me that, you know, 10 years ago, five years ago that it was possible. I'm telling you now it is possible to connect.

It is possible to regain, to give him -- to give our children, our little ones our love and bring them back. And that's when you start seeing what's going on. Our logical mind working overtime to steer us away from our pain because to be human is to be exposed to joy and happiness and laughter, but also pain, sadness. Have you seen "Inside Out?"

SPEAKER: Yes.

MR. FRENKEL: It was about me, you know, when I saw that those islands of her memory crumbling down, it was me. And so I had to do it, you know. I had to cut it off to survive the way 400 years ago, Cartesianism was our way to survive to break into mind and body and me and you to survive. But the time now is to feel whole again. That's why love and math, not love is math, not math is love, we need both. And then, of course, other -- you know, some memories stay as memories, but not at emotional level. I discovered some memories when I was a little boy, much younger which I completely erased from my memory. And then, of course, if I don't want to feel, experience pain, I would like to transfer myself into a robot, because I know that the robot is -- doesn't experience pain, doesn't feel the pain. But that's what makes us human , and we can connect to it, we can let it out, we can accept reality the way it is, and we can build a better world, a better reality together.

You know, it's always -- when we talk about this kind of topics, it's very easy to get bogged down in technical arguments, in theoretical arguments. He said this and she said that and then it becomes this mental gymnastics. So, I wanted to -- I don't have time for this but it's very real, it's not a theoretical question anymore, the AI weapons and the danger, but I want to leave you with this thought. Pascal said, the heart has reasons, but the reason knows not at all. Your vision will become clear only when you look into your heart. Who looks outside dreams, who looks inside awakens.

But we have a choice. It's our choice to know who we are, to remember who we are. To answer this question, who am I, or maybe more concretely am I a human or a machine? Is there a difference between a human and a machine? Because if we don't answer that question to ourselves, each of us, then we are secretly enabling Ray Kurzweil and others to lead us to the precipice. It's our choice and we can do it. Time to wake up. Thank you.

(Applause)

SPEAKER: Fantastic talk. If you're at top level, your fear that -- do not think about humans, we are the (inaudible), you've talked about that. We could create some sort of beings up in the cloud, but they will be truncated and partial and not fully human. So that is the (inaudible) you're worried, because we can do something (inaudible) --

MR. FRENKEL: Well -- that's right. So already there are the dangers of AI weapons, which I didn't have time to talk about, self-guided weapon systems and -- but -- you know, I like an example yesterday. Sherry Turkle spoke here, she's a professor at MIT who studies the interactions between humans and robots, and she told this amazing story about a little girl who was asked to interact with a robot. And she was very upset after that interaction and they said, why. And she said, because I think the robot didn't like me, but that little girl didn't quite understand that it was a robot. They tried to explain, but she was very disappointed. And Sherry said the poor kid was eating cookies one after another, you know. So, and she made a great point, robots cannot teach us empathy, but they can do us harm, they can do us damage.

That's what happens, but so when we say, you know -- so there is this very real danger, but the point is sometimes people circumvent safeguards. Let's formalize morals, let's create a program and algorithm which will describe morals as well. You know, excuse me, can't already prove that it's not possible, and now Godel proved that it's not possible. So yes, okay, we can do research in this area but, you know, everything can be hacked. You can create safeguards, but this is illusory safeguards. This is because somebody can always come and find a way to circumvent.

The only thing which will be our safeguard is when each of us -- each of us, not somebody else, not this other people because I'm fine. So by the other people, no. When I remember who I am, that's when I will not allow those machines to do me harm. That's when I will use -- and by the way don't think that I'm against AI, I love it. I used artificial intelligence all the time, Google Translate, you know all these Amazon algorithms which make great suggestions, I have my Kindle here with me, you know, and iPhone. I love it as long as these algorithms are at the service of humanity and not the other way around.

(Applause)

MR. FRENKEL: But we can do that, so that -- the bad news is that we will never be able to legislate it. We will never be able to legislate it, but the good news is that we don't need to. All we need is to look inside.

(Applause)

MR. FRENKEL: I suppose my time is up. Thank you so much for being such a wonderful audience. I love you and I'll be happy to answer any questions just outside of the auditorium.

* * * * *

MR./MS. [LAST NAME]:

* * * * *