Ray Kurzweil Interview

Interviewer:
Dag Spicer

Recorded: July 13, 2009
Mountain View, California

CHM Reference number: X5461.2010

Dag Spicer: Okay. Good afternoon, we're here with Dr. Raymond Kurzweil. It's July 13th 2009, and Dr. Kurzweil, thanks so much for being with us today.

Ray Kurzweil: It's my pleasure.

Spicer: I'd like to start by asking you how you would explain your vision of the future to an intelligent 12-year old.

Kurzweil: Same way I would explain it to anyone else. In fact, the 12-year old would probably get it very quickly. Even a 12-year old has been around long enough to see the exponential growth of information technology, the accelerating pace of change. I mean, only a few years ago, maybe when this 12 year old started using computers, we really didn't use Wikis or blogs or social networks. We didn't use "tweets"[Twitter messages] a few hours ago; that's a small exaggeration. Think back a decade, maybe that's before the 12-year old's time, most people didn't use search engines. That sounds like ancient history, and that's less than a decade ago. The pace of change is getting faster and faster. The first changes, paradigm shifts, stone tools, fire, the wheel, took tens of thousands of years; the printing press took two centuries; the telephone only took half a century to reach a quarter of the population; and now we have major paradigm shifts in just a few years' time. And that's the nature of an evolutionary process, and technology is an evolutionary process. It's survival of the fittest, just like biological evolution. And biological evolution or technological evolution will evolve a capability and then adopt that capability and use it to evolve the next stage, and that's why the next stage goes more quickly, it builds on the shoulders of the previous generation and there's exponential growth in the capability of these technologies. And if you measure the power of these technologies, let's say per dollar, the amount of computation you get, the number of MIPS [Millions of Instructions Per Second]per dollar, or the number of bits that you can buy per dollar, or the number of bits you move around on the Internet, or the number of bits of brain data we're getting, or the number of base pairs of DNA that we're sequencing in a year, the cost of sequencing a base pair of DNA; these fundamental measures of information technology grow in an exponential manner. Now, maybe a 12-year old hasn't become familiar with exponential numbers but if I take 30 steps linearly, I go one, two, three, four, five; 30 steps later I'm at 30. If I count exponentially, I double each time. Two, four, eight, 16; 30 steps later I'm at a billion. It ultimately explodes. We start out doubling little numbers but finally we're doubling big numbers, and there's a huge difference between the linear perspective and the exponential perspective. It's the nature of human intelligence, or basically if you ask, "What is intelligence?" Intelligence makes predictions about the future. So we are constantly anticipating what will happen next, but our intuition, what's hard-wired in our brains, is a linear prediction about the future. When we walked through the fields 1,000 years ago, we saw something coming at us through the corner of our eye; we made a linear prediction where that animal would be and what to do about it, and that's hard-wired. And that works quite well, but it doesn't work well in anticipating the nature of information technology. So even sophisticated scientists will use their intuition, which is *linear*, to think about where technology will be in five, 10, 20 years, but the true nature of it is *exponential* and that's why people's imagination fails them when they think about the future. When I was a student at MIT, we all shared a computer that took up half a building; it's an IBM 7094. Maybe it's here [at the CHM], it'd cost tens of millions of dollars. The computer in your cell phone today is a million times cheaper and a thousand times more powerful. That's a billion fold increase in the amount of computation you get per dollar since I was a student, and we'll do it again in the next 25 years. So this is a very revolutionary characteristic of information technology, and it's not just Moore's law; Moore's law talks about shrinking the size of components on an integrated circuit. Moore's law is just one of these paradigms. It was not the first paradigm to bring exponential growth to computing. We had four paradigms before Moore's law.

The exponential growth of computing goes back decades before Gordon Moore was even born. I put lots of computers, going back to the 1890 American census, the first census to be automated with electromechanical equipment on a logarithmic graph, and there's a smooth progression for 110 years. We've made a trillion-fold increase in the amount of computation you can buy per dollar over the last century, and it was not affected by any of the things that happened. You don't see any evidence of impact of the Great Depression, World War I or II or the Cold War; it's just this inexorable progression in the power of computers. At the same time, we've shrunk them to be smaller at a rate of 100 3-D volume per decade[? D.S.], and it's true not just of computers but anything where you can measure the information content of a technology, and we can talk about some of the other aspects of information technology, like our biology. So this exponential growth is quite revolutionary. As powerful as these computers are today and other forms of information technology, 25 years from now, they'll be a billion times more powerful again and 100,000 times smaller, so you get some idea of what will be feasible.

<crew talk>

Spicer: Okay, let's start. How has science fiction influenced your thinking, if at all?

Kurzweil: Well, I read the Tom Swift Jr. series when I was eight, nine and 10. That impressed upon me the idea that you can find ideas and inventions to overcome any problem; that was kind of the philosophy of my family. So the plot of every one of those novels was the same; Tom Swift Jr. would get into some problem and the fate of the human race hung in the balance. He would disappear into his basement and the sort of tension of each novel was, "What idea would he come up with to save the day?" And invariably, he'd come up with some very clever idea that you wouldn't have thought of that overcame the problem. And it did represent my own philosophy which I got from my family, that no matter what kind of problem you encounter, there's a solution out there and you can find it. And it was personalized; "You, Ray, can find this solution and you should look for it. And when you find it, you should implement it," and that's been kind of an imperative in my life.

Spicer: Great answer. Now, a lot of your life's activities have combined humans and technology. We can start with the OCR, optical character recognition stuff, the synthesizers, now the singularity. Is that a leitmotiv in your life, this combining and/or enhancing the human condition?

Kurzweil: It actually goes back to a feeling I had when I was five. My parents gave me all these Erector Sets and construction toys, and I had this idea if you put things together in just the right way, you could create transcendent effects. I did not have that vocabulary, but I do remember the feeling as a five year old that, "Wow, I can put these things together and do something magical that would overcome human problems." And I decided I would be an *inventor*, and I remember having that idea and I did have that much vocabulary. And other kids were wondering what they would be, and I always had this conceit, "I know what I'm going to be," and my invention started to get some traction when I was seven or eight. I created a puppet theater that was kind of a virtual reality world, and I had a command station and I could move the sun and the stars and characters on and off the stage with these mechanical linkages and I could control this small universe from my command station. I discovered the computer when I was 12, and I had the feeling then that you could recreate reality in a computer. You could create virtual realities, you could recreate aspects of our thinking and that sort of animated me, starting when I was 12. I had access to some early [IBM] 1401s and IBM 1620s back starting in 1960, and I also built some of my own computerized devices. And I had this idea that really, the heart of human intelligence was our ability to recognize patterns, and very much the history now of neuroscience confirms that. That is what human beings do very well. We're very good at looking at a chess board and seeing a pattern. We're not very good at doing the logical Minimax algorithm in our head. Kasparov was asked, "How many moves ahead do you look per second when Deep Blue could analyze 300 million board positions in this recursive expanding tree of move/counter-move possibilities? Deep Blue could do 300 million; how many do you do, Mr Kasparov?" and he said, "Less than one." So how is it that he can hold a candle to a computer? Well, he's got deep powers of pattern recognition. That is the heart of human intelligence. By the way, computers are much better now at the pattern recognition aspect, so even with one percent of the computation of Deep Blue, they can do even better than Deep Blue did because we've mastered pattern recognition more than we have in the past.

Kurzweil: Computer pattern recognition is still not as good as humans, but it's getting better and better, and that is the heart of what human beings do. And I had this sense, really starting when I was 12. When I was 14, I did a project to analyze melodies, to look for the patterns that a composer would use and I would feed in Chopin and Mozart and I had ways of finding the patterns, and then composing original music using the same patterns. And indeed, it would sound like a student, maybe a third-rate student, of Chopin or Mozart, depending on who I had analyzed. That was my first pattern recognition project. It won some prizes like the Westinghouse Science Talent Search; I was one of the finalists that got to meet President Johnson, and that was my first pattern recognition project. My first major commercial project in pattern recognition was omni-font optical character recognition, and that really dealt with the key issues of finding invariant features in patterns, because that's what pattern recognition does, particularly at the human level. We're able to look at a face and we're able to recognize the invariant features, no matter how the face is pointed, even if it's occluded, even if it's represented through a distorted lens, we can find the invariant features. We can find the invariant features of speech no matter who's speaking, even if the accent is different and even if it's a differently-shaped vocal tract and so on. So we're very good at recognizing patterns, even if there are changes in the information. And so that's what we did with character recognition. We were able to recognize that a shape, say, that we would recognize as a capital A, had a concave region facing south and it had a triangular looped portion in the northern region and it had a north-central to south-east and south-west connection and a crossbar between them; these are the invariant features of a capital A, no matter what typestyle you use, and we were able to successfully do that. That became kind of a solution in search of a problem. I sat next to a blind guy on a plane and he was telling me how he had difficulty accessing ordinary printed material, but otherwise blindness was not a handicap, he said, but he did have this one issue. And so we devoted the omni-font character recognition to the reading problem of the blind, and we had to develop a couple of other technologies: text-to-speech synthesis and a CCD flatbed scanner. We developed those and created the first print-to-speech reading machine for the blind. Later on, I got involved with speech recognition, which is an even more complex pattern recognition problem because if you look at a spectrogram, a picture of speech, of two different people saying the same word, it can look completely different, yet there must be some invariant features there that allows human intelligence to recognize it as the same word. Even the same person saying the same word at different times will look very different. So we developed technology that could handle that, and it got involved with looking for patterns in human language, which is really the heart of human intelligence. Alan Turing based the Turing test on language, recognizing that language, in its hierarchical nature, reflects the hierarchical nature of our intelligence. That's what really unique about humans, is that we have this neocortex that can understand and deal with hierarchies of patterns to reflect the hierarchies of patterns that exist in the natural world.

Spicer: Right. That's great. Could you tell us a bit about your synthesizers and how you came across them? That seems like a bit of an outlier, and yet you were very successful.

Kurzweil: Well, the music synthesizer is an output rather than an input technology, but it's also based on pattern recognition. To really examine the question, what makes a piano sound like a piano? What patterns of sound enable a piano to do that? And so we actually had a model of how a piano produces sound; we combined that with sampling technology, but sampling by itself was not adequate to solve the problem, because you don't have enough memory to accurately record every possible sound and how they interact with each other in a piano. So we would model the signal processing and pattern generation capabilities of a piano to be able to recreate that realistically, and also understand what the human auditory perceptual system does, because it can be fooled by certain things, whereas in other ways, it's extremely sensitive to the slightest variations. But some things, it's not able to recognise very well, so understanding what its strengths and weaknesses are, we could recreate the piano with enough accuracy in the ways in which the human auditory perceptual system was accurate. I got into that really through the reading machine project, because Stevie Wonder was our first customer. He heard me present the reading machine in 1976 on *The Today Show*; he called us up out of the blue, our receptionists didn't really think it was him, but he did come over, and we happened to just finish our first reading machine that we could actually part with, and he left with it after we showed him how to use it. And that was 1976, and that started what is now, you know, a 33-year friendship. A few years later in 1982, he was giving me a tour of his studio, which he called "Wonderland", and he was lamenting the state of the art in musical instruments. On the one hand, there were these 19th century acoustic instruments which were still the instruments of choice in terms of creating very beautiful, rich sounds. From a pattern recognition / signal processing point of view, we would say they're very complex sounds. And on the other hand, there were these electronic sounds that a computer could generate. They were very thin, relatively simple, but you could control them much better. Acoustic instruments, even if you're a virtuoso, you can't play every instrument, and even if you could play every instrument, you can't play them at the same time; most of them, you can only play one note at a time. You can't sit down and play an orchestra. With these electronic instruments at that time, the early 1980s, you could play a line of music, the computer would remember it, you could play it back and then you could play another line over it. You could actually erase notes and replay them, just like you would edit a letter on a word processor. You had much better control methods. You could play a whole orchestra or rock band on your computer, but the sounds that you had to use that were available at that time were very thin. When you selected "piano", it sounded like an organ; when you selected "violin", it sounded like an organ. He said, "Wouldn't it be great if we could really command these very rich, beautiful, complex sounds of acoustic instruments with these very powerful control methods of the computer world?" and I felt that would be feasible using some principles of signal processing and pattern recognition, combined with the sampling world. And we started Kurzweil Music together; he was part of our company as kind of a musical adviser to guide the development. And we worked together and two years later, we introduced the Kurzweil 250 which has been recognized as the first electronic instrument that accurately could recreate the grand piano and other orchestral instruments.

Spicer: Shifting a little bit towards the future and your vision of a couple of things; one is, I think, in the year 2039, we achieve parity with our-- our computers achieve parity--

Kurzweil: 2029.

Spicer: 2029, sorry. Computers become as intelligent, for some definition of intelligence, as humans. And then later, in 2050, I believe, they actually can be merged through nanotechnology or other technologies with humans ourselves and can augment us. There's a lot in that question, but could you maybe lay it out for us a bit?

Kurzweil: The first thing it's important to recognize is that computers are growing exponentially in power. They're doubling in less than a year for the same cost, and that's actually a trajectory that's been going on for 110 years, and will continue. And it will continue past Moore's law. Moore's law was not the first paradigm to bring exponential growth to computing; it was the fifth paradigm. The exponential growth of computing started decades before Gordon Moore was born. We had electromechanical calculators, ralay-based computers such as Alan Turing's machine which cracked the German Enigma code, we had vacuum tube-based computers in the 1950s. CBS predicted the election of Eisenhower in 1952 with a vacuum-tube based computer. And then they were shrinking vacuum tubes every year, making them smaller and smaller, to keep this exponential growth going. That finally hit a wall. They got to a point where they couldn't shrink the vacuum tubes any more and keep the vacuum, and that was the end of the shrinking of vacuum tubes. It was not the end of the exponential growth of computing. We went to the fourth paradigm; transistors, and finally integrated circuits and Moore's law and the shrinking of component sizes on an integrated circuit. As paradigms go, that's been a great paradigm, but it is not the sum total of this exponential growth of computing. It will come to an end, you know, by around 2020, the key feature sizes then will be on the order of four nanometers, which is about 20 carbon atoms, and we won't be able to shrink them any more, but we'll then go to the sixth paradigm, which is three-dimensional molecular computing, self-organizing circuits. I talked about this in my 1999 book, *The Age of Spiritual Machines*. At that time, it was considered a very controversial notion; it's now a very mainstream notion. If you speak to the scientists at Intel, such as Justin Rattner, their CTO, he will tell you they have these kinds of circuits working in experimental form, they feel the crossover will be in the teen years [2010-2019] well before we run out of steam with flat integrated circuits. They will keep this exponential growth going for a long time. By 2019, $1,000 of computation will be equal to the most conservative estimates of the amount of computation we need to simulate the entire human brain, which I estimated about 10^{16}; 10 million billion calculations per second. We'll achieve that for $1,000 by 2019, even using conventional silicon without even going to the next paradigm. So that's the hardware side of the equation. I think that's pretty non-controversial. The more interesting issue is: *will we have the software*? And we've made a lot of progress in artificial intelligence without looking inside the human brain. In fact, up until recently, we really could not look inside the human brain with enough precision to figure out what was going on. These colorful fMRI [Functinoal Magetnic Resonance Imaging] images would tell you, you know, where things are going on; if you're solving a logic puzzle or looking at visual information, there's something going on here or there, but that doesn't give you enough precision to see what's going on. The spatial resolution of brain scanning has been doubling every year. The amount of data we're getting is doubling every year. We can model individual neurons. We're getting more and more information and we're showing that we can turn this information into working models and simulations of brain regions. There's already about 20 regions of the brain that have been modeled and simulated; regions of the auditory cortex, the visual cortex, the cerebellum, where we do our skill formation like catching a fly-ball, we have figured out how that works. Even slices now of the cerebral cortex, which is the most important region.

Kurzweil: We've recently simulated substantial slices of the cerebral cortex where we do our logical thinking, our hierarchical thinking, our invariant feature detection for pattern recognition; that's really the heart of human intelligence. We're beginning to understand how that works. I make the case in my book *The Singularity is Near* that we will have all the models and simulations of the human brain within 20 years. So by 2029, we will understand enough about the human brain to recreate its fundamental methods, its algorithms, and the computers at that time will be far more powerful than is necessary to simulate the human brain. And I've been very consistent about the date 2029 when we will have both the hardware and the software to simulate human intelligence. We already have hundreds of examples of software, AI software, in use in our economic infrastructure that does things that used to require human intelligence. Every time you send an e-mail or connect a cell phone call, intelligent algorithms route the information, pick up any product that was designed in part with intelligent computer-assisted design, manufactured at robotic factories, inventory levels controlled by just-in-time inventory systems that are intelligent, automatic detection of credit card fraud, lots of financial decisions made by computers. I get an electrocardiogram, it comes back with an intelligent analysis by computers that rivals that of doctors; the same for blood cell images. Intelligent algorithms fly and land airplanes, guide intelligent weapons systems. So these systems are doing what used to require human intelligence, and very often doing them better than humans consistently, very inexpensively. And these were actually research projects not so long ago. If all the AI in the world stopped tomorrow, our whole civilization would grind to a halt; that was not the case 15 years ago. Now, these are narrow examples of AI. We call them “narrow AI” because they’re doing some particular task; playing chess or analyzing an electrocardiogram and otherwise they don’t have the suppleness and subtlety and flexibility of human intelligence, but the narrowness is gradually getting less narrow. As we’re learning more about how the best example of human intelligence works, which is the human brain itself, it’s not hidden from us, and that’s another exponential progression. And I've been very consistent about this date, 2029. It goes back through several of my books that by that date, we will have machines that operate at a human level in terms of the entire flexibility of human intelligence. And one way to measure that is with the Turing test, which I think has held up very well. Turing described this 50 years ago. Fifty years ago, Alan Turing described a test whereby a human judge would interview a computer, an AI and a human foil, over what he called “teletype lines”; basically instant messaging. So the human judge would not see who he or she is interviewing, and if after a few hours -- he didn’t actually specify the rules vey well -- but if after some period of time, the human judge can’t tell who’s the machine and who’s the human, we say that the machine has passed the Turing test. And no machine has yet passed the Turing test. There are Turing tests run every year. Loebner runs one series of tests. And the machines are getting better every year. According to Loebner, if a computer can fool the human judges 30 percent of the time, we will say that they have passed the test, and the last test they fooled the judges 25 percent of the time. I actually think that those rules are too easy, but I do feel that within 20 years, and that’s probably conservative, computers will pass valid forms of the Turing test. And they will then combine what are now strengths of human intelligence, which is our tremendous ability to recognize patterns, with ways in which machines are already superior. I mean, your cell phone can remember millions or billions of things accurately, machines can find knowledge instantly out of billions of possibilities. They can do logical thinking much better than we can. Few mathematicians, or no mathematicians, can really hold up to Mathematica in terms of manipulating equations and solving theorems and this kind of thing. Machines are better than we are at logical thinking, remembering things, transferring knowledge. If a machine has a skill, it can transfer that skill and that knowledge and that information in that database at electronic speeds, which is a million times faster than the electrochemical transmission that goes on in our brains or the transmission from one person to another using language. So these are superiorities today of machines; when we combine that with a machine that can actually match human intelligence in terms of our ability to recognize patterns, use language and understand it, that will be a very powerful combination.

But it’s not going to be a matter of us competing with these machines, because this is not some alien invasion of intelligent machines from Mars; it’s part of our human machine civilization. We’ve always built our machines to extend our own reach. I mean, ever since we picked up a stick to reach a higher branch, we’ve extended our physical reach with our machines, our tools; now we already extend our mental reach. You know, I can take a device out of my pocket and access all of human knowledge with a few keystrokes; that makes me smarter. That extends my mental reach. Very few people can do their jobs today without these mental extenders represented by our computers and all the knowledge bases they command. And we are going to literally going to make ourselves smarter. We ultimately will put these machines inside our bodies and brains. They're already very close to us. When I was a student, I had to go across campus to get to the computer; now it's in my pocket. It will make its way into our bodies and brains, and that's started. You can put a computer in your brain today if you happen to be a Parkinson's patient or a deaf person. You can put computers in your brain that will replace the functionality that's been lost by disease or disability, and the latest generation of this FDA-approved neural implant for Parkinson's allows you to actually download new software to the computer inside your brain from outside the patient. Now, today that neural implant is not blood cell-sized; it's pea-sized. It's pretty small, but another exponential trend is we're shrinking technology. Actually, according to my models, a rate of 100 [times] per decade, so these technologies will be 100,000 times smaller in 25 years, and 25 years from now they will be the size of a blood cell. So my vision is a few decades from now, we will have millions of these nanobots, nano-engineered blood cell-sized devices which have computers in them. They'll be going inside our bloodstream to keep us healthy from inside, they'll augment our immune system, they'll destroy disease when it's at the level of a cell rather than the level of an organ and is life-threatening. They'll go into our brains through our capillaries; we'll have neural implants without surgery, introduced non-invasively through the bloodstream. They'll interact not just with a few neurons, but with all of our neurons; we can have millions or billions of them. They'll put our brains on the Internet, they'll provide a highly realistic, full immersion virtual reality incorporating all the senses from within the nervous system. And most importantly, they will extend our intelligence, which they do today. Even if they're just in our pockets and in our hands, we can access all of knowledge easily. That makes us smarter, but a convenient place to put them, particularly when they're the size of blood cells, will be in our brains. We won't lose them that way, they'll have an opportunity to more intimately integrate with our intelligence and they will make us smart. And that is what our technology has always done and it's getting closer and closer to us, and more and more intimate.

Spicer: Is thinking then a form of calculation?

Kurzweil: You know, in some ways the brain has a very different architecture from conventional computers. A conventional von Neumann machine is very, very fast, but it does one thing at a time. So we've taken some baby steps in parallel processing and you might have a chip now that has four cores or 16 cores, and so it does 16 things at a time very quickly. The brain is massively parallel, really a hundred trillion fold, because the calculations take place in the inter-neuronal connections and the dendrites and the synaptic clefts between neurons and we have a hundred trillion of them. But they're very slow. They calculate about 200 calculations per second; that's maybe 10 million times slower than computers today. And we can build machines that have this massively parallel architecture, or we can simulate it. We really just need to achieve the computational level. It's an engineering detail as to how parallel we make the computers, because they can simulate what goes on inside the human brain. But if you talk about what goes on in a neuron, it is computable, and particularly when we understand the salient information features, we don't have to simulate all of the life support functions of a neuron because we're really building it through a different infrastructure, a different substrate. And we have already figured out what the salient algorithms are for certain regions of the brain. The auditory cortex does certain types of transformations to auditory information, and we really don't have to simulate exactly how a neuron does that if we understand what the neurons are doing from an information processing point of view. And so we've already done that for some parts of the brain. And the most important part of the brain is the neocortex; that's where we do this hierarchical thinking that results in language and results in ability to create tools and to understand the world, which is naturally hierarchical as well. And that's what's unique about human beings, is that we can do that. Only mammals have a neocortex and our neocortex, which is about the size of a table napkin, is the biggest, but it's still limited. Why not have a neocortex that's much bigger than a table napkin? Well, we'll be able to extend those limitations by being able to add computation, and that computation in our brains can then be on the Internet and can access the vast amount of computing that will exist in the cloud, which is already formulating, so we will be doing some of our thinking out on the cloud. You know, once we have computation inside our brains, it'll be on the Internet and then, just like any computer, will be extending, multiplying its ability by thinking out on the cloud. That's where our future lies. But yes, what the brain does is it processes information. There's some architectural differences with computers, but it's the nature of software that you can write software that will simulate a different architecture, and as long as we have enough brute force, we'll have the ability to simulate what goes on in the human brain, provided we know what goes on there. So we also need the software, we need the algorithmic insights, but we're making exponential gains in that as well. And then we'll need a third component, which is education, because you can take the human brain, which has hardware and software, which is integrated together in the human brain, but it doesn't do much unless you educate it. And that's in fact a complex process, and a lot of knowledge goes into the education of human beings. In fact, we've had thousands of years of culture to develop the means of educating the brain and accessing all of the learning that goes on all of our history and literature, the human brain can access that. But all of that is out there. It's all on the Internet. These AIs will be able to be educated with all of human knowledge readily accessible, but we'll actually have to design that learning program just as we do for human intelligence.

Spicer: Should we be hopeful or a little concerned about these developments? Or both?

Kurzweil: I'm not blasé about the dangers of technology. Technology's been a double-edged sword ever since we had tools. Our very first tools; fire, stone tools, the wheel, we used for both creativity and destructiveness. They've always amplified our ability to be destructive in war and conflict.

Kurzweil: And certainly we've seen a lot of destructiveness come from our tools. I would argue that we've been helped more than we've been hurt. People forget what life was like. If you read Thomas Hobbes, he's described life a few hundred years ago quite well. It's short; human life expectancy was 37 in 1800. Brutish, disaster-prone labor-filled, it took six hours to create the evening meal. There were no social safety nets; most people lived on the brink of disaster and disease-filled. We had no sanitation, no antibiotics; life was very, very hard just 200 years ago. Human life expectancy was 48 in 1900. We've come a long way in overcoming human suffering. We still have a lot of suffering to go, but technology is a double-edged sword. We've created enough nuclear weapons to destroy all human life, all mammalian life, so that's certainly a downside to technology. Strong AI, AI at the human levels and beyond, will be the most powerful technology that we'll ever create. It will have the power to destroy us if we don't build our own moral and ethical values into it, so there's a lot of discussion in the AI field and how to build what Eliezer Yudkowsly calls "friendly AI" versus unfriendly AI. Certainly if you're in a situation where there's some AI and it's much more intelligent than you are and it's bent on your destruction, that's not a good situation to be in; we have to basically avoid getting in that situation in the first place. The only thing that could protect you from that would be an AI that's even smarter that's on your side, and we really need to integrate with this technology. It's not my view of it that it's a civilization apart that we have to combat the sort of scenario deployed in many science fictions and movies of man versus machine I think is unrealistic, because we're going to be very integrated with our technology. We can see conflict today between different groups of humans and their machines, and the technology amplifies the destructive capability of both sides. But that could be a frightening scenario as well, because we do amplify our destructiveness with our machines. So it's a complex issue. I do take some hope from the fact that I think the decentralized nature of these new technologies, like the Internet, is inherently democratizing. I wrote in my first book *The Age of Intelligent Machines*, which I wrote in the 1980s when the Soviet Union was going strong, that the Soviet Union would be swept away by the then-emerging decentralized electronic communication. We didn't have the Internet yet, but we had early forms of e-mail over teletype machines and early fax machines, and they kept everybody in-the-know. And I felt this would sweep away the Soviet Union. These were far more powerful tools than the copiers that they were controlling, and indeed that's what we saw in this 1991 coup against Gorbachev; the authorities grabbed the central TV and radio station, which they had always done before, but it did not keep people in the dark, and the totalitarian control was swept away. And then we saw a tremendous movement towards democracies during the 1990s with the rise of the World Wide Web, and people forget, but the world was a very different place. And decades ago, there were very few democracies, and today, there are relatively few holdouts. You know, you can argue about how perfect various democracies are, but there's much more democratization in the world than there was some decades ago. And we also have a democratization of access to the tools of creativity. It used to be if you wanted to create a movie, you had to be a big Hollywood studio, if you wanted to do some technology development, you had to be a big corporation or government lab or academic lab; today, a kid in her dorm room can create a full-length motion picture in high definition with a $500 camera and her PC. A couple of kids at Stanford, as a dorm project, wrote a little bit of software and revolutionized web search and created a company worth $100 billion today [Google, Yahoo], and there's lots of other examples of that, and these tools are literally in everybody's hands. It's not just 'the haves" and the richer are getting richer. I mean, take these so-called "cell phones"; they are really very powerful devices that access all of human knowledge and are really tools of creativity. They're in half the hands of the world. I just got back from China; half the farmers in China have them. And nearly everyone will have them within a few years. So the tools of creativity and access to them have been democratized. These are positive indications, but technology nonetheless can be creative and overcome suffering on the one hand, or destructive on the other. We've had that intertwined promise and peril ever since we've had stone tools and fire, and so it's an old story. And we need to keep these technologies safe, and it's a complex issue as to how to do that. On the one hand, we need ethical standards to prevent accidental problems. For example biotechnology, which is the ability to reprogram biology as a set of information processes, that's a whole new concept and we're using that to reprogram biology away from disease, presumably good things. But it could also be misapplied by a bioterrorist to create a new biological virus that could be more deadly or more communicable or more stealthy. Now, to prevent that from happening accidentally, we have ethical standards; they're called the Asilomar Guidelines, and they have actually kept biotechnology safe for 30 years. But if you have a bioterrorist, he's not going to follow those ethical standards, so we also need a rapid response system that would act as a kind of technological immune system. And a good model of that is how we deal with software viruses. They actually have an immune system. We automatically detect new software viruses which emerge every day, - they are destructive people who writ these software viruses and put them out -- and we detect them and we reverse engineer them, and an antiviral program is coded, in some cases automatically, spread virally on the Internet. And within hours, we have protection from a new software virus. You can argue about how perfect it is, but nobody has taken down even a portion of the Internet for even a second over the last 10 years; it's been a very robust system because we do have a technological immune system that protects us from software viruses. We need a similar system for biological viruses, and I've actually been working with the Army among the Army Science Advisory Group, and the primary issue that I've been dealing with is a rapid response system for biological viruses. We do have the technological ideas to do that.

Spicer: Okay. Do you think a computer can become self-aware, and if so, when might this happen?

Kurzweil: Well, part of human intelligence is the ability to reflect on ourselves, reflect on our thinking, have a model in my brain of how my brain creates models of the world so I can think about my thinking and we can take that level of recursion several more steps and think about ourselves, thinking about ourselves thinking. Human beings are capable of doing that because of the hierarchical thinking of the neocortex. Is that being self-reflective? From an objective point of view, it is. You have an entity that's thinking about itself thinking and can create models of itself thinking, but philosophically, that's not really the same thing as consciousness, the sort of feeling that one has. You can talk about how the brain processes the color red, but the feeling of redness, let alone more subtle feelings like happiness and joy and humor, the feeling of that, which is what we associate with consciousness, is very hard to describe. And ultimately, if you talk about what some people refer to as the hard problem of consciousness, it's fundamentally not a scientific issue. Because a synonym for consciousness is subjective experience, and if you ask, "Well, what is science?" science is objective observation and deductions from that. There's this conceptual gap between subjective experience and objective observation. Another way of saying this is there's really no machine you could imagine building where you slide it into the end and a green light goes on, "Okay, this one's conscious. No, this entity is not conscious," that doesn't have some philosophical assumptions built into it. I mean, you could build a machine like that, but different philosophers would build different assumptions into the machine. So John Searle would want it to be squirting biological neurotransmitters and if it wasn't biological, he would say it's not conscious. Dan Dennett wouldn't require it to be biological, but he would require it to be able to think about itself thinking and build a model of itself, and if it did that, then maybe it's conscious. Different philosophers would have different assumptions, but there's really no way to prove whether or not their entity is conscious. My prediction is that these future machines will convince us that they're conscious, they will seem like they're conscious, they'll be convincing. We have machines today that talk about being happy or sad; avatars on the web or characters in a computer game. Their behavior is not yet convincing; it's not yet complex enough to really convince us that they're really having these emotions, so they seem like a simulation. In the future, that simulation will be so good that it will be indistinguishable from the real thing, and it will actually introduce a philosophical issue; is there a difference between a totally convincing simulation and the real thing? I believe we will be convinced by these machines. They will succeed in us believing that they are conscious. We will want to believe them because they'll be very influential, they'll be very clever, they'll get mad at us if we don't believe them. And moreover, it's going to get mixed up; it's not going to be a clear distinction. You know, I could not be able to walk in a room and say, "Okay, humans on the left and machines on the right," because it's going to be all mixed up. You can have a biological human that's got computers in their brain, maybe billions of them. There may be more going on in the non-biological portion of their intelligence than the biological portion, so are they machine or are they human? The action may be with the non-biological part; it's not going to be a clear distinction between human and machine the way it is today, because my prediction is we're going to merge with these technologies. So where does our consciousness lie? Do you have to be biological to be conscious? It's a philosophical question and it just means that there's still a role for philosophy, that science can't answer every question. But some scientists will go on and say, "Well, since it's not scientific, it's not real and consciousness is just an illusion and we shouldn't worry about it. It's a distraction," and you can build a completely consistent philosophy that does not have consciousness in it. But my own view is that it's a mistake because our whole moral system is based on consciousness, and our whole legal system is based on our moral system. So if I hurt someone, cause them pain and suffering, which is a conscious experience, that's immoral and probably a crime. If you extinguish someone's consciousness, that's a high crime. If I destroy some property, that's probably okay if it's my property; if it's your property, it's a crime not because I've caused suffering to the property, but I've caused suffering to the owner of the property. So it all comes down to consciousness; our whole moral and ethical and legal system. So I don't think we can just dismiss it as an illusion quite so easily, but it's another way of saying there's still a role for the idea of values and philosophical thoughts about the implications of these issues. But we are going to become increasingly non-biological. I don't think it makes sense to insist that these processes have to be running on a biological substrate, because what our brains are doing is shuffling around information. John Searle says, "Well, computers just shuffle around symbols and human beings really think thoughts and are really feeling these thoughts." But if you look at what in fact is going on, our brains are just shuffling around neurotransmitter levels and ion channels and ions, and those are just representing information. And if you can shuffle around the information using a different substrate, the same thing is going on. You could do a thought experiment where you take just a little piece of your brain and replace it with a machine. The machine is operating on a completely different substrate, but it's still the same person.

Kurzweil: We've actually done this experiment. For example, with Parkinson's patients, they had a piece of their brain that stopped functioning and we replaced it with a computer and if you ask them, "Do you think that computer is part of you?" of course if you ask different people the same question you might get different answers, but I've actually asked this question and most of them will say, "Yes, it's definitely part of me." And if you carry this thought experiment further and keep replacing more and more portions of the brain with computers, a person's personality never changes, there's a continuity of identity, you would come to the conclusion that it's always the same person, the same consciousness. It certainly would seem that way. At the end of this process, you would have a person that has no biology. So we've discussed these issues actually for thousands of years, going back to the Platonic dialogues. These will become actual pressing, real issues as we go forward. But the reality is it's not going to be deciding, "Is that avatar conscious?" because we are going to become the machines. We'll start by becoming partly machine, because we will be putting these machines in our bodies and brains and we'll be partly biological, partly non-biological, but it is the nature of the machine portion of our intelligence that it grows exponentially. That's the nature of what I call *the law of accelerating returns.* The biology portion is fixed; it's not going anywhere, and biological evolution is a million times slower than technological evolution, so ultimately we're going to be much more machine than we are biological. And people, when they think about that, thy say, "Well, I don't want to become machine-like," because they're thinking about the machines that they know today. And if you look at your cell phone or your PC, it's impressive but you don't want to become that kind of machine, because that is lesser than humans. We'll probably need some new terminology, because I'm talking about a machine that does have the subtlety and suppleness and intelligence and richness and complexity of human beings, and we will want to become those kinds of machines because ultimately, those machines are going to be greater than we are and we're going to want to merge with them to enhance our own capabilities. And then people say, "Well, will we still be human if we do that?" In my mind, that is what being human is all about; going beyond our limitations, extending ourselves. If it wasn't for our tools, our life expectancy would be 23, which is what it was 1,000 years ago. We didn't stay in the ground, we didn't stay on the planet; we have not stayed with the limitations of biology. We're the only species that does that. We make ourselves greater through our tools, and we're going to literally transform our biological substrate, we're going to transcend biology. But in my mind, that's not transcending our humanity because it is the nature of being human to extend beyond our boundaries.

Spicer: Beautifully said, very nice.

Spicer: This is kind of an odd-- what would you be doing if you were born in 1750 at the start of the Industrial Revolution?

Kurzweil: Well, I was born in 1750 in a way, because I started with mechanical technology. That's what I had when I was five, six, seven, eight. I had the kinds of tools that people had in 1750, and indeed in 1750, people built mechanical devices that could calculate using mechanical devices, and they built elaborate automata. And I built devices that could manipulate reality using mechanical devices, at least as much as a five-, six- and seven-year old could do. So I was very much like those inventors in 1750, using mechanical technology. I then discovered information in the computer around the age of 11 or 12. I began to build little computerized devices and also got access to professional computers like the IBM 1401 and 1620, and so I jumped ahead a few hundred years at that time and entered the Information Age where we had much more powerful tools to manipulate information. But even in 1750, the idea that you can embody information was around. Leibnitz has talked about the human brain as a set of logical pulleys, and he described it in mechanical terms but he described it as a system that manipulated information, using the kinds of metaphors that existed at that time.

Spicer: So one of the possible consequences of the vision that you have for the future is a radical extension in the human life span. I wonder if you could chat about that and the related though that we need a finite lifespan in order to sort of-- that's how we understand our lives currently, is that we live about 70 years and we plan our lives accordingly. How might that change if we lived for, say, 500 years?

Kurzweil: Health and medicine, biology, has just made a grand transformation from being not an information technology, just kind of a hit-and-miss affair where we just find things accidentally, to where it is now an information technology. We have now the software that life runs on, which is the genome, and we're also making exponential gains in understanding it. By the way, the collection of the genome followed exactly my exponential progression, which I call the law of accelerating returns. Halfway through the project, the skeptics said, "I told you this wasn't going to work. Here you are halfway through the project, and you've only finished one percent of the project. It's pretty pathetic." But actually, that was right on schedule, because if you're on an exponential progression where you're doubling your progress every year, once you get to one percent, you're only seven doublings away from 100 percent. And that's exactly what happened; it continued to double every year, and seven years later, it was finished. And we've continued that past the end of the genome project. So 2003, we had this software that runs in our bodies. It's outdated software. How long do you go without updating the software on your cell phone or your computer? Your cell phone updates itself every few days. This software hasn't been updated in thousands of years; some of these programs are millions of years old. Like the fat insulin receptor gene is millions of years old; it says, "Hold on to every calorie because the next hunting season may not work out so well." That was a great idea when our genes evolved; today, it underlies an epidemic of obesity. We turn that gene off in animal experiments and these animals ate a lot and remained slim and lived 20 percent longer; that's just one of the 23,000 genes we'd like to tinker with. We also have the means now of changing our genes; not just in a baby, but in a mature individual. RNA interference can turn genes off, new forms of gene therapy can add new genes, we'll have not just designer babies but designer baby boomers; we can design these interventions on computers, we can test them out on biological simulators that are getting more sophisticated every year. So health and medicine has just become an information technology. As such, it's going to double in power every year. That's the nature of information technology in any field. So these technologies, even though they are in an early stage today, will be a million times more powerful in 20 years and it will be a very different era. I've written a series of health books; the last two I've co-authored with Dr. Terry Grossman, and we talk about *three bridges to radical life extension*. *Bridge one* is applying today's knowledge to slow down the ageing process, and there's a lot you can do already to slow down ageing processes. For example, the cell membrane in every cell is made up of a certain substance called phosphatidylcholine that depletes gradually as we get older; that's why the skin sags in an elderly person and their organs don't work as well. That's one of the 12 ageing processes;

you can actually stop and reverse that by supplementing with that substance. Another ageing process is the buildup of plaque in our arteries. It not only causes heart attacks and strokes, but also leads to an elderly person's organs not working very well, claudication [sharp pain] of the limbs, impotence and so on; you can reverse that not so easily by just taking one supplement, but you can attack it from many different perspectives and actually slow down and if you're diligent enough, actually reverse that process. Dr. Dean Ornish has shown that. So we've described in detail how you can slow down, stop and in some cases reverse all this different ageing processes with today's knowledge. Today, it's somewhat complicated. You have to learn about your own body, what your issues are and develop a program that's multifaceted to slow down the ageing process. In my own case, I had biological ageing tests when I was 40; I came out at about 38. I'm now 61 and on these biological ageing tests, I come out in my early forties. I've only aged a few years in the last 20 years, according to these biological ageing tests of how old I am internally. And the goal of this, what we call "bridge one", is not to live hundreds of years; it's just to get to *Bridge Two* in good shape, and bridge two is maybe only 15 years away. When we have really the Golden Age of biotechnology, this ability to reprogram the information processes that underlie our biology, we'll have very powerful tools to really stop and reverse ageing processes with biotechnology methods in 10, 15, 20 years. According to my models, 15 years from now, we'll be adding more than a year every year not just to infant life expectancy, but to your remaining life expectancy. So as you go forward a year, your life expectancy will move on away from you. It's not a guarantee, but it will change the metaphor of the sands of time running out; they'll start running in. That's a bridge to the *third bridge*, which is the nanotechnology revolution. The quintessential app, I used to call it "the killer ap" but that's not so good a name for a health technology, are nanobots. And we'll have billions of nanobots, blood cell-sized devices, augmenting our immune system, going through our body, keeping us healthy, destroying disease when it's at the level of a cell rather than an organ. And that really will keep us going for a very long time. And if you go beyond that, we ultimately will become mostly non-biological by merging with machine intelligence. The nanobots will go inside our brain, they will interact with our biological neurons, they ultimately will be where the action is. Now, being computerized, we will back them up just the way we back up our computers today. Ultimately, most of our brain will be non-biological. Ultimately, it'll be so powerful that even the part that's biological will be understood, modeled and simulated by the non-biological part, and so we can back up the biological part also. And so we'll have a means of restoring our *mind file*. People 100 years from now will look back at this era where we went through the day without backing up our mind file as pretty remarkable, just as we would think it remarkable today if people didn't back up their personal computer files. We take for granted that you can smash a computer and then recreate its personality, its skills, its knowledge, just by loading it from a back-up. So this mind file of your computer is not dependent on the hardware; there's a separation of hardware and software. In our brains, the software is embedded in the hardware and if the hardware crashes, the software disappears with it. We take that for granted, but ultimately, we will achieve the same thing that we do now in our computers for our brains, which are really information processors. We have information in our brains that represents our memories, our skills, our personality; that's our mind file. It's information. It's the most valuable information we have. We'd like to be able to preserve that, and as we become more non-biological, we'll be able to do that. Now, people sometimes talk about, "Well, if we have radical life extension, then people live a lot longer, we'll run out of resources and we'll run out of things to do and life will get boring and there won't be new opportunities because the old people won't get out of the way." If we just had radical life extension and no other changes, these would be problems, but it's important to look at all the different things that are happening. The same technologies that will bring radical life extension will also bring radical expansion of our resources. For example, we have plenty of energy. We have 10,000 times more sunlight than we need to meet all of our energy needs, but right now we can't capture it efficiently. I just did a study with Larry Page of Google for the National Academy of Engineering, and we described a plan to convert our entire energy system to solar energy within 20 years. That is, in fact, on an exponential. We've been doubling the amount of solar energy we're producing in the world every two years; we've been doing that for 20 years. There's already been 10 doublings, and we're only eight doublings away from it meeting 100 percent of our energy needs. You might say, "Well, but is there really enough sunlight?" Well, there's 10,000 times more than we need. We only have to capture one part in 10,000 and we'll be able to do that with nanotechnology applied to solar panels. There are similar new technologies for water, for food, for housing that could meet the material needs of a growing biological population. And life won't be boring because we're going to be making ourselves smarter. We're going to merge with this technology, we're going to expand our minds quite literally, we're going to expand our experiences, we'll have virtual reality from within the nervous system incorporating all of the senses. We'll be able to choose a virtual reality environment just as we choose a website today and unlike, say, Second Life today, which is small and flat and on your screen ad sort of cartoon-like, these will be full-immersion virtual reality environments that compete with real reality and are as realistic as real reality. And you can have a different body, different virtual body, in these virtual environments for different circumstances. So life is not going to be boring. And as for opportunities, we're constantly creating new institutions. Larry Page and Sergey Brin didn't wait for some old person to open up a position for them to create a whole new institution and a new opportunity, and we are expanding human knowledge. It doubles about every 14 months, so we create constantly new opportunities to contribute. It's really life that gives meaning to life, and the things we can do with it and all the creativity that we can deploy. We don't need death to give meaning to life. In fact, in my view, death is a great tragedy. It's a great robber of all the things that gives life meaning. It destroys meaning and knowledge and skill and relationships, and so we will have the means of transcending that limitation. That is what human beings do, is we transcend our limitations, and we're going to transcend this particular biological limitation.

Spicer: Oh, yes. Okay, previous visions of the past did not pan out, especially in the field of AI. What is different this time?

Kurzweil: Well, there's certainly no shortage of bad futurism but my own thesis is based on what I call the law of accelerating returns, which is specifically the exponential growth of information technology. And I got into this because of my interest in being an inventor, and I realized that timing was important so I began to study technology trends and I gathered a lot of data. And about 30 years ago, I made a pretty unexpected discovery, which is the trajectory of information technology based on measuring the attributes of it, like the price performance of computing, or the price performance of biological technologies follows amazingly predictable trajectories. In the case of computation, going back to the 1890 census, if you put all the computers on a logarithmic graph, they form a very smooth doubly exponential graph. The price performance of computing was doubling every three years in 1900, every two years in 1950, every one year in the year 2000, and it's now down to 11 months. It's a very smooth progression, very predictable, and was not affected by any of the little things that happened in the 20th century, like two World Wars, the Cold War, the Great Depression. It went through thick and thin and war and peace, and continues now through the current economic downturn completely unperturbed. And it's not just the price performance of computing; it's anything you can measure in terms of information. It's being moved around on the Internet or bits of data about the brain; all these different measurements follow these exquisitely predictable trajectories. The number of bits you can put on a magnetic disk, that's not Moore's law. Moore's law is actually just one example of this. People talk about Moore's law as being the sum total of exponential growth and then they ask, "Well, how long can Moore's law last?" Well, Moore's law, as paradigms go, has been a very important one but it's just one of the paradigms that has brought this exponential growth. It was the fifth paradigm in computers, and we see similar trajectories in things that have nothing to do with integrated circuits. So this turns out to be very predictable. I have a whole theory on evolution, both biological and technological, as to why this happens. Basically, an evolutionary process evolves a capability, adopts that capability and then uses it to evolve the next stage. So even in biological evolution, it took a billion years for DNA to evolve, but then biological evolution adopted it so the next stage, the Cambrian Explosion, went 100 times faster. And biological evolution kept getting faster and faster. Homo sapiens evolved in just a few hundred thousand years. And the fruits of these evolutionary processes grow in power, and I have a whole mathematical treatment of that, but we can see it empirically. And I have a team of 10 people that gathers data in these different fields and we see very predictable progression of these information technologies. So I believe what we can predict very accurately is the overall power of these technologies. Now, when we talk about the implications of it and what this will enable us to do and whether promise or peril will be more important or predominant, we can have arguments about it but I think the overall power of these technologies is inexorable.

Spicer: Thank you.

END OF INTERVIEW

