



## Re: The Future

I'm a bit skeptical about this part. With a lot of hand waving, computer science has categorized problems according to their "difficulty" (how long it would take a computer to solve the given problem as a function of the size of the input), and two classes of difficulty they've come up with are P (which stands for "within Polynomial time") and NP (stands for "Non-deterministic Polynomial time").

Polynomial time means it's possible to write a polynomial where the variable is the size of the input, and the value of the whole polynomial is an upperbound on the number of computational "steps" (and I'm intentionally being vague about what constitutes a "step" here, but you may approximate it as being a single CPU instruction) that a "traditional" computer would have to perform to solve the problem. For example, sorting a list of numbers is within polynomial time, because given a list of  $N$  numbers, we can devise an algorithm that will sort the list within  $N^2$  steps. If we're sorting a list of a million numbers (and thus the "size" of the problem is said to be  $10^6$ ), it'll take at most  $10^{12}$  steps. If you have a computer which can perform a billion steps per second (e.g. a 1Ghz computer), you could solve the problem in  $10^3$  seconds, under 20 minutes.

Contrast this with Exponential-time problems (or EXP problems), where the time it would take to solve the problem can be expressed in the form of  $2^N$ , for example. Solving a problem of size  $10^6$  would take  $2^{1000000}$  steps. Even if you had a billion Ghz computer, you probably wouldn't be able to solve this problem before the universe collapsed upon itself.

NP problems lie somewhere between P and EXP. If we had some sort of "magical" non-deterministic computer, we could solve NP problems in polynomial time. But without these "magical" computers, it would take our "normal" computers exponential time. "Cracking encryption" is an example of an NP-problem: When some device claims it uses 1024 bits of encryption, cracking its encryption is a problem of size 1024, which would take  $2^{1024}$  steps, which again with our computers would likely take longer than the age of the universe: Google calculator won't let me compute  $2^{1024}$ , but  $2^{999} \approx 5 * 10^{300}$ ; a 100Ghz computer which can perform 100 billion ( $10^{11}$ ) steps per second would solve the problem in  $5 * 10^{289}$  seconds, or about  $10^{282}$  years.

It turns out that quantum computers have the properties of the above mentioned "magical" non-deterministic computers. That is, quantum computers can crack most encryption schemes that exist today within polynomial time.

So while it's true that a traditional computer can simulate a quantum computer, it would do so extremely slowly, and only feasibly for extremely small problems. For example, for an NP problem of size 5, the traditional computer would take  $2^5$  steps, or 32 steps, while a quantum computer would take perhaps  $5^2 = 25$  steps. Not a big difference. Essentially (and again, with a lot of handwaving), this means we could simulate a 5-neuron brain at real time speed (and in fact, I think AI researchers are currently

## Re: The Future

simulating 300 neuron brain at real time speed; recall that a computer is much faster than a brain, if we ignore the advantages of parallelism, which doesn't show up very much in a mere 5–neuron, or even 300 neuron brain). It becomes a big problem when you want to simulate a 100 billion neuron brain.

But as you mentioned above, it's not clear whether you really need "quantum magic" to simulate a human brain or not. If you don't, then merely having a billion–core CPU ought to be able to do it. If you do, then you'll probably need a quantum computer, and right now, quantum computer technology is still in its infancy, so it's not clear what kind of progress rate we can expect from it (e.g. will they too follow Moore's Law?).

And as you mention below, we might not need to actually fully simulate the brain in order to achieve Turing–test passing AI.

[...]

Kurzweil talks about simulating the brain at two levels and gives order of magnitude estimates for how much computer power will be needed.

On level is to simulate the brain at level function units. This requires somewhat less computations per second (cps) than simulating the neurons and all of their non–linearities. The later level of simulation would be need if you want to scan your barin an upload a copy to a computer.

Right now we have functional level simulation of some regions of the brain. There are probably several hundred of these functional regions.

Interesting, I hadn't heard anything about this area of research.

Right now the unused cycles of computers attached to the internet is probably more than enough for functional simulation.

Yes, but good luck trying to get all the computers attached to the internet to cooperative towards a single goal.

Around 2020 supercomputers will be powerful enough to do it and around 2025 a PC will have that much power. He predicts computers will start passing the Turing test around 2030 and by around 2037 full simulation of the neurons and their non–linearities will be possible.

Well, I'm predicting around 2047, so what you say sounds a tad

## Re: The Future

optimistic, but not entirely unreasonable.

[...]

The AI guys seem to want to do it so they can say "see we were right all along". Plus the idea of immortality and super intelligence with god-like powers is very seductive to guys who like to learn things and increase their knowledge. What worries me is that humans are a mixture of both good and evil potential and the million times faster processing will amplify both. Despite the rosey projections and optimistic scenarios I think nobody knows what the consequences will be. For every positive picture presented you can imagine an equally powerful negative picture. I think implementing it without some idea of what will happen is irresponsible. Technology always seems to result in some unintended consequences. As it gets more powerful those consequences could result in the end of us all.

The whole idea of "The Singularity"

([http://en.wikipedia.org/wiki/Technological\\_singularity](http://en.wikipedia.org/wiki/Technological_singularity)) is that you \*CANNOT\* have some idea of what will happen once the technology is implemented. And obviously, a lot of Sci-fi fiction play with this idea as well.

He says we evolved from bacteria and we did not exterminate them so super intelligent computers will not wipe us out.

It's more like we kill billions and billions of bacteria every second without even noticing it (via our immune system), because they're so low our radar, we don't even pay attention to them.

Hmm... suddenly, super-intelligent AI isn't sounding like such a good idea...

In fact they will be human and will value and respect their human heritage. He is optimistic about human values.

I'm not. It also depends on whether we program them to value and respect human heritage, whether or not we allow them to reprogram themselves, and whether or not we program them to not want to reprogram the part of themselves which values and respects human heritage.

If that last part confused you, consider this: Let's say you have the ability to reprogram yourself. Would you decide to reprogram yourself so that you could be a cold blooded killer, feeling no empathy, regret or

## Re: The Future

remorse for taking other people's life? Probably not. Why not? Because society has brainwashed you into thinking that reprogramming yourself to be a cold blooded killer is a "bad" thing.

But even if we program the AIs to like humans, there's tons of sci-fi fiction depicting bugs (either in the code, or even in the original design/requirement specifications themselves) that let the AIs kill humans anyway.

Take Asimov's three laws of robotics, for example. The first one is says that a robot may not allow a human to come to harm, even through inaction. Obviously, this an impossible requirement, as humans are continuously "coming to harm" (in the form of aging). So the AI might then decide to reformulate the rule as "minimize the amount of harm that any human will come to". From there, it may conclude that the best way to do that is to kill all humans. You'll have a sudden spike in harm in the immediate future, but then for the rest of eternity, you've guaranteed no harm will ever come to a human ever again.

Well we have not wiped out bacteria but we don't mind killing them by the millions or billions when they get in our way. We experiment on monkeys that share our genome to 95 to 99 percent.

Right; the way we handle animals with presumably inferior intelligence is mixed. Some we tend to kill with little remorse (so called "pests", such as rats, mosquitos, etc.) Others we try to protect, without disturbing (animals considered to be endangered). Some we keep as pets. Some we hunt (for food or for sport). How would we feel if AIs "domesticated" us, so that we were no longer able to survive in "the wild", but became entirely dependent on the AIs to take care of us? We may have a higher quality of life than what "wild" humans might have experienced, but we lose our independence, and become "second-class" citizens in a world dominated by superior life forms.

Human values have produced world wars, mass exterminations, genocide and many other horrible events. We often have good conscious intentions that are thwarted by subconscious programs. Often these subconscious programs are developed in childhood and they have a good purpose for us, but their methods are those that make sense for a child but are not appropriate behavior for an adult. What happens if these god-like super intelligences get in a power struggle to see who will be the head god?. What will the collateral damage be? What happens if the one million plus increase in speed results in mental illness or frustration and impatience with the slow responses of unenhanced humans? When they sneeze will we all catch cold? If they can modify their design and improve upon it at an ever increasing rate how long will they retain their "humanity". He projects that late in this century a single

## Re: The Future

computer will have trillions of times the processing power of 10 billion people. It does not take super intelligence to realize that mere humans with nuclear weapons and nanoweapons could be perceived by these AIs as a threat to their existence.

On the other hand, something with a trillion times the intelligence of humanity as a collective might not consider mere humans with nuclear and nano weapons to be a threat at all. Or at least, only as much as a threat as the common cold virus is to us. Yes, it's \*conceivable\* that someone might die from the common cold, but it's so low on our priority list of things to worry about... The AIs might have more important things to worry about than humans with bombs that can barely destroy a single planet, nevermind a galaxy... they might be more concerned with what the \*other\* AIs are up to.

He does not address this very deeply.

Understandably, as we're way out into wild-speculation-land here.

Most of his arguments are with those who say that what he is proposing cannot be done. He says the super intelligence will find a way to overcome any obstacle.

It really depends on what he means by this. On the one hand, we still face the same basic problems that animals face (how can I raise enough money/find food to survive? how can I claim territory/buy a house? how can I find a good mate? how can I raise my family successfully? etc.) only at a higher abstraction level. It seems that mere intelligence is not enough to solve these problems, because these problems arise from competition with other agents of comparable intelligence. On the other hand, presumably AIs don't need to worry about food or raising families (or do they?).

From another interpretation, we as humans have essentially "solved", at a superficial level at least, all the problems that animals were facing (how to satisfy basic survival needs), and we've moved on to problems that are probably inconceivable to the animals, like what to do about global warming, pollution, corrupt government, etc.

He does not mention much about trying to control it, although some of the references are to material that discusses steps we should take to make sure we have friendly relations with them. I suppose if we try to control it and it circumvents us then relations might not be very good.

## Re: The Future

That might not be the case. Think of park rangers to find an injured animal such as a wolf. They might try to help the wolf, and the wolf, not understanding what is going on, is aggressive and tries (futilely) to attack the rangers. The rangers are not upset with the wolf, or anything like that. Rather, they understand that the wolf does not have the intellectual capability to assess the situation, and is behaving from baser instincts. The AIs might have the same outlook towards us.

It might be amusing to picture an AI training other AIs into capturing, tagging, and releasing entire human societies back into the wild for further study. "Careful there, if you aggravate them too much, they may start launching missiles at you. It's not deadly, but it *\*is\** bothersome to deal with. You want to distract them here, and while they watch that point, quickly place the nano-RFID tags on them before they realise what happened, and then back off. We don't want to disturb their natural habitat."

– Oliver