Artificial Intelligence and Trans Humanism as the next stage of the evolution.

Author: Stefan Certic

For thousands of years we had knowledge to determinate earth orbit just based on the sun reflection. But, still achieving a great failure to perceive idea of "what was before the beginning". Perceiving forever or nothing.

No matter if you are supporter of Trans Humanism movement or not, there is a question to be asked.

- Are we, building an AI?
- Are we, who are already an AI, building an AI?
- Is this "mirror effect" definite? Can we consider it a dimension?

Few words about singularity

The singularity is the hypothetical advent of artificial general intelligence. Such computer would theoretically be capable of recursive self-improvement (redesigning itself), or of designing and building copies better than itself. Repetitions of this cycle would likely result in a runaway effect — an intelligence explosion — where smart intelligence design successive generations of increasingly powerful AI, creating intelligence far exceeding human intellectual capacity and control. According to public criticism and scientific society, because the capabilities of such a super intelligence may be impossible for a human to comprehend, the technological singularity is an occurrence beyond which events may become unpredictable, unfavorable, or even unfathomable.

Unjustified criticism

While reading some article in Popular Science Magazine related to brain research, I noticed the bottom line section with summary of AI related thematic in SCI-FI genre. What captured my attention is the way plot line of movie "Transcendence" has been written. It was something like

"Before dying, he uploads his mind into a computer and becomes a power-hungry megalomaniac."

As someone who received the movie diametrically opposed, believing that covers many deep ethical and philosophical questions – by opening them, but not going to deep (it's still a product, right), this sentence got me to immediately watch it again. After watching it 2 more times, paying attention to every single detail, I went to check the other reviews. By reading the comments and IMDB ratings, one thing I was sure.

"If ever AI we build, it would be definitely threated like Afro-Americans once"

It will happen

When singularity happens, and out of question it will if we graph the evolution of mind the same way we do with Moore's law, the only real threat may be a human stupidity backed up by the fear of unexplored. Although, we will try to cover different concepts of creating a self-aware AI that may lead into destructive behavior, luckily, any form of more advanced intelligence will easily conclude and predict empirically rather then acting emotionally, unless certain limitations are imposed. The paradox lies behind the fact that limitation of certain abilities may leads to potential threat.

Due to the fact that creating a self-aware AI may result in digital organism exceeding human intellectual abilities due to amount of information available and the processing power, it's ultimately important no restrictions are applied. We would probably want to protect ourselves by employing different concepts, however, once the intellectual power explode, how much those mechanisms we put in place may actually work against us and an artificial being itself.

Separating the human specie vertically, we may observe minor cases of great minds heading around with the guns conquering territories. Rather the major samples of social disorders such fascism, fanaticism and many others *ism's reserved for those with less developed reasoning.

For the purpose of better understanding, let's assume, there are two types of persons. And it's not the binary joke of zero and one :). I would say, those who work on self-improvement to become better, and those who "don't need it" - utilizing different self-defendable mechanisms to protect their opinion outside of any reasonable scope. Of course we can't be easily divided into specific group. It's a scale. We all have properties of both mentioned groups, more or less, in certain situation. Can we measure the level of self-awareness on the scale? I strongly believe we can.

Now taking the evolution as the only empirically proven concept towards something more advanced, it's quiet obvious that an ability to self-modify itself in combination with meta-cognitive reasoning (an ability to think about thinking) would be a pathway to something called singularity, or AI and an explosion of intelligence that comes as a consequence of distributed consciousness outside the limits of physical laws and boundaries of know forms of living organisms.

All the different religions utilize different methods to establish value system and ethics. Some do it by utilizing fear, some by promising rewards. Our justice system establishes ethical norms of acceptable behavior through different types of sanctions and fines. One is common – all ethical norms come from self-awareness. The more self-aware we are, the better ethical judgments we make.

How far we are?

I'll rather say not close enough. And it's not about current computing power compared to brainpower or Moore's law that's about to collapse. I'm sure quantum computing has enough space to grow in the mean time and definitely move Moore's prediction by adding lot of digits.

But where the real problem resides, is the efficiency of the algorithms that empowers our todays computing. How much processing power we really need to solve specific problems, and what kind of approach towards a problem we employ.

It's deterministic, and opportunistic.

The very first question I'll ask myself, is why do I found easier to type this text on my terminal window within the Linux box, using Joe editor, simple text editor written in 1988. I'm sure most of you find lot easier firing up Word or Open Office, since you don't need time to invest learning that entire complex set of commands for something simple – as typing the text. However, as a professional IT person dealing with complex problems for a living, I need simple, logical and fast interface that might not look nice, but serve the purpose. The bottom line here is, the technology of 1988 still provides me an ability to perform any manipulation of this text faster and more efficient then any other tool available. Did we get any achievement? – I don't think so.

Moving forward, we got used to trade efficiency for a comfort even in those "reserved for geeks" aspects of computing such as programming. Today's programming that heavy relies on ether Java, .Net or PHP and similar concepts put the programming in the field of design. There's a layer between you and the bare metal written by someone else, that you don't need to even understand completely to be able to write the code.

Rather then writing a function, you can call one provided by .Net written by someone and pray it would do what you want the efficient way. It's a tragedy how many full time programmers today don't know what's bit, byte and how they fit together. Most of them even fail to form a picture in their minds that computer can't understand words and letters – so when it comes to debugging a layer2 problem it's small number of persons they could relay on.
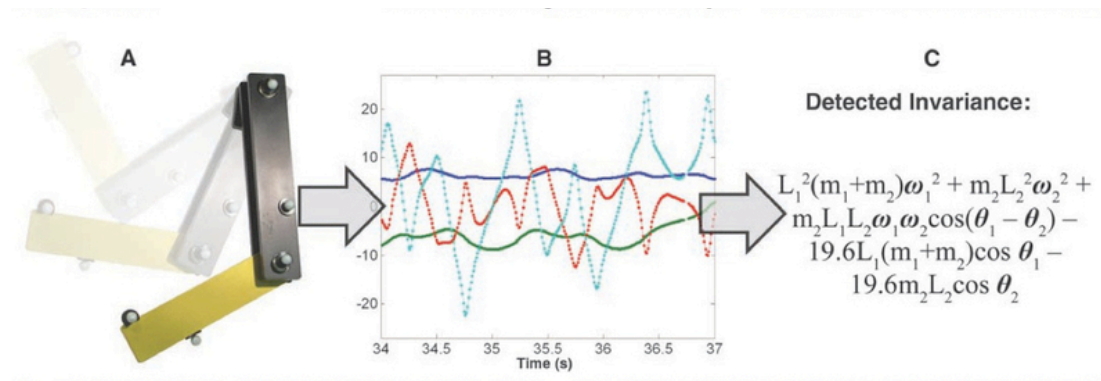
Personally, I would never be ready to guarantee for something with my own life that's not written in pure C. And it's 1983 technology.

Most of us would, however, develop those PHP apps without even knowing what's happening at the processor / kernel level and solve all the efficiency problems buying more powerful servers.

The Consumer Market dictates a road, and there's not much we can do here, but at least we can abandon hope that commercial development of supercomputers and quantum computing may change something dramatically.

How close we are?

Maybe closer then we think. Following the work of Hod Lipson, Cornell University computational researcher who wrote software that calculated gravity constant without prior input of any laws of physics or geometry by analyzing a double pendulum movement, could be an important pointer towards a direction in AI research. It took one day, for a machine to conclude a constant that took us centuries of research based on previous knowledge.

A   **B**   **C**

Detected Invariance:

$$L_1^2(m_1+m_2)\omega_1^2 + m_2L_2^2\omega_2^2 + m_2L_1L_2\omega_1\omega_2\cos(\theta_1 - \theta_2) - 19.6L_1(m_1+m_2)\cos\theta_1 - 19.6m_2L_2\cos\theta_2$$

The question why this research has importance in AI at all is closely related towards following problem:

2	4	8	32
3	6	18	?

So, how fast your brain calculated missing element of an array? Probably fast enough. A human brain has an ability to find a pattern relatively easy. And taking a look at the very basic question we use to test intelligence, let's ask a question, are we able to code software that will continue an array? With the current deterministic approach, it would just consume too much processing power, even for a simple task like given one.

Human brain, with just a fraction of power - relays on best match algorithms and probability employing arithmetic once it's ready to test most promising samples. And it does that incredibly fast.

It's not about the power, rather efficiency.

The Pendulum project opens series of questions what we can do to improve efficiency of algorithms that may eventually lead towards an artificial ability to analyze given series of inputs and outputs processed by an unknown set of operations, in order to determinate an exact set of operations performed.

An ability to perform such tasks, more or less efficient would be a key of our ability to actually measure an IQ of the artificial software.

Another project that took part in early 1990s by an ecologist Thomas S. Ray, called Tierra aimed to experimentally simulate an evolution within the simulated environment of a computer system. A notable difference between Tierra and more conventional models of evolutionary computation, such as genetic algorithms, is that there is no explicit, or exogenous fitness function built into the model. According to Thomas S. Ray and others, this may allow for more "open-ended" evolution, in which the dynamics of the feedback between evolutionary and ecological processes can itself change over time. The system eventually reaches a point where novelty ceases to be created, and the system at large begins either looping or ceases to 'evolve'.

Apart from being the very first project that provides a proof-of-concept for an Evolution theory, the phenomena of host-parasite development within the system also showcase an ability to form self-rewriting software that evolve and fit into environment and survive.

By evolving the way software is written, we might be able to provide it with basic cognitive functionalities of inductive and deductive reasoning. The form of random mutation and formation of parasites in Tierra project prove the self-improvement and align hypothesis. Moreover, giving an ability to self-rewrite the code in combination with logical reasoning would result in more advanced version. The condition for this to happen is setting a goal. Hardest to achieve, more advanced versions may be expected.

Potential Paradoxes

The goal might be one of the largest paradoxes. It's required for any form of self-advance. But setting it correctly would determinate an ability of cognitive reasoning code to control it's own behavior.

Setting an unachievable goal without ability to rewrite it's own goals would definitely lead into a deadlock consuming more and more power until complete destruction.

The same applies for paradoxes of nature. Let's imagine a goal of understanding a big bang theory, or the currently mainstream theory of inflation. Do we have enough resources to mathematically explore one of the largest paradoxes such as beginning of time? What would happen having an self-aware being with an ultimate goal to understand what's behind the edge of the infinity, without being able to give-up?

Ultimately, for any form of consciousness to be achieved, the condition of allowing rewrite of it's own ultimate goal at some stage is required.

There are few ways to go:

- Setup an achievable goal without ability for goal modification and hope for consciousness with some degree to develop. (Time limited with unknown outcome).

- Setup an unachievable goal with an ability to rewrite it's own goals. (No time limits, with likely outcome of singularity and dismissal of the goal for the purpose of self-preservation).

- Setup a goal of finding and achieving goals. (Most likely resulting in singularity, with no time-scale limits).

Homo sapiens differs from other spices in terms of goals of developing an abstract thinking. While reproduction and survival is an ultimate goal for all form of known organisms, the goal of human individual is not quiet limited to reproduction. Empirically, the ultimate goal of all individuals that moved humanity forward was solving a set of problems. The connection between problem solving motivation and self-advance is hard to be accidental. And the development of abstract reasoning and cognitive / meta-cognitive abilities are in close connection to goals.

Where's the can opener? Inside can.

Ability to properly set it's own goals depends on current level of cognitive development, while success in developing abstract reasoning largely depends on pathway towards achievement of the goal.

Human being has an ultimate goal. Every single individual tried at least once to conclude some of the biggest paradoxes, such as – where is the edge of the universe. What comes after? If there's an after, then it's not an edge. However, there are set of subsystems preventing us from taking all the resources available to satisfy our curiosity. For example, there's a religion that completely rewrite our ultimate goal – understanding the concept of everything.

Some of the biggest paradoxes such as:

- God is almighty
- He is able to create a stone, so hard and heavy that even he can't lift from the ground.

With all the goals and thirds for knowledge, a religious person seems to ignore this paradox.

Non-religious persons have their own set of formal logic based restrictions, through different cultural or sociologically implemented systems. We all tend to accept paradoxes beyond our ability to understand.

No matter how much important is to allow rewriting of it's own goal for an artificial being in order to prevent a self destruction from evolving more and more with the purpose of fulfill an impossible task – it takes key part for an artificial being to be able to split and re-define it's own goals. It takes a free will.

The data mining and compression

Did you ever have to remember a cell phone number? How can you memorize 241896? I'll start with memorizing number 2. Because 0 to 2 takes two, and 8 to 10 takes 2. And I'll also split the phone number into exactly 2 logical blocks of 231-897. The rest is built of digit up, and a digit down in each logical block.

I'll I have to do is associate name of the person with the number 2. When I need to call him or her, I'll need to reverse a process:

Starting point 2:

First Block: 2 -> 2+1 -> 2-1
Second Block: 10-2 -> 10-2+1 -> 10-2-1

So instead of storing "241896" I stored "2+-"

That's about 50% of data compression with less information to store, plus forming a deep neural connection forming an association of bringing back the number.

It's simple showcase of how do we save resources and perform data mining making them way easy and efficient to retrieve or associate with other relevant data.

Now can you close your eyes and play a video of your last year birthday party? You can't play exactly what shirt everyone wearied, but you can easily retrieve some images, jokes, even smell of the beer all around.

Taking in account there's a whole part of the brain in neocortex in charge for filtering the sensory input, separating and dismissing important from non-important data from entry into the processing of the human brain in order not to cause an overload, there are very sophisticated mechanisms used to compress and associate data before storing them into a long term memory by forming specific neural networks.

Imagine there's a database. There's a table "drinks" with columns "id", "name" and "smell". Now there's a table events with columns "id", "event name", "drink".

There are strong chances there will be a lot of beer on the birthday party. Let's normalize the database by adding "1" to a drink field of the "birthday party" row, matching the id "beer" from the drinks table.

Of course, there's a possibility you are attending a rehab supportive group party, but chances are small. Let's take our chances and store the data so we can easily retrieve the smell of beer when someone mentions the party.

Our brain takes its chances million times a day. Information we store is not 100% accurate. There are a lot of trades between space, efficiency and the probability. We aim to achieve database with the best chance to serve us fast and minimize mistakes. However, it's completely opposite to what we are aiming to achieve in modern computers databases. Something we call ACID compliance – a stand for Atomicity, Consistency, Isolation, and Durability.

Our Brain is not ACID compliant. It makes minimal mistakes. However, that's what makes possible to store enormous amounts of data on a single "floppy drive" no modern SSD can match.

So how about a merge?

Like it, or not – it seems like the only way to go.

Since evolutionary concept has been mathematically and laboratory (first simulated self-generated parasite) proven, there's no point discussing will it will continue developing to override problems of the species.

Increase in population and decrease of resources lead to unsustainable environment for any life form in very short period of time.

There is no way we can move to Mars or any other known planet due to fact we can't adapt to local environment.

However, Server plant, in fact can work very well on Mars, or any other planet using nothing but the solar power.

Now can you prove your physical existence? How can be sure you are not already a piece of software? If you can't - then what's the difference?

It would be, however, very funny evolutionary paradox is software is migrating itself into another software. That kind of mirror effect sounds much more reliable to me then all quantum theory together.

How it happens already?

Thanks to evil called capitalism, everything is in service of profit, while generating profit means exploiting resources. To cut the expenses and resources consumption, we easy adopt technology such as cell phones, personal computers, etc.

Now let's jump into next level of technology called VR. Imagine it's complete simulation of environment. Put on your sunglasses, and you are at the office. Imagine an ability to fill all the physical sensation and collaboration without leaving your home? No office rent expenses. No heating. No driving and traffic jams.

Imagine you don't have to travel by plane to meet with friends and have a beer? Putting sunglasses that connect your nervous system and gives you taste of the beer on Bahamas straight from your home?

Of course, there would be a lot of right fraction movements and non-supporters of this technology. At least, until they try VR Porn :) I don't expect many loud speeches from Nazi like guys against that sort of technology. I do expect them being occupied with a nice blonde called Helga 24/7 in their VR software.

Empty streets, rainy day, hard time having anyone to pick up their phone - and a message box. Would you like to upload your conscious? Yes, Remind me tomorrow?

Not much to chose from. Nor it should be :)