Thought & Consciousness in Artificial Intelligence

Report written by Gabriel Padis

Submitted to Professor Mrs Leonard

ING4 SI 1 Group B

Tuesday, 27th of November

Executive Summary :

In a near future, AI will be able to have thoughts on its own based on the collected data, and will have consciousness of itself and of the repercussions its actions have. Humans will have the possibility to upload a perfect copy of their brain to a computer to have a virtual life. This emulated brain will be able to create copies of itself and work at high speed without rest. Both technological advances will have impacts on society and on economy in ways we cannot imagine. Some say that they are dangerous technologies, other say that they could be used in ways that could change our lives for the better. What we do know for sure is that technology is neutral and that is it the use that can pose problems or resolve them.
If we want to continue to evolve then we need to not be surpassed by technology but rather join it, since we won't be able to stop it. By ensuring that everyone has the possibility to merge the technology of either emulated brains or AIs, we are sure that humanity altogether will progress.

Table of Contents

Introduction

Background Information

Moore's law has been proven to be true since it's description in 1965. Today we have sophisticated hardware and softwares that allow us to do everything in our life. Those are constructed and programmed by humans. We build this systems depending on our technological knowledge and we iterate from what already exist. We are now building programs that will have the ability to learn by themselves. We are defining a goal and it will try by itself, make assumptions, and develop its own way of thinking. This is an Artificial Intelligence, often shortened as AI.

Purpose

The goal of this report is too show even though technological advances are fast and growing in the field of AI research and applications, we won't be able to slow it nor to stop it. But it is not a problem because we can always change the course of events so that an AI with thoughts and consciousness does not become a threat. Humans have the possibility to use AI and technology to enhance themselves, as emulated brains in a computer. It can be another step in the evolution of our race.

Scope

By taking into account the fact that AI are being made to be independent and powerful we will see : 1) how it is possible to give them human traits such as thoughts and a conscience, and 2) how humans can use technology to be more than what we currently are and become AI ourselves or merge with AI technology. This report will not focus on the applications of AI but on how it is used today or will be in the future to mimic human life and experiences.

Definition of terms

Moore's law is the observation that the number of calculative power doubles every two year.
Psychometrics is the science of measuring mental capacities and processes.
Em or Emulated brain is a perfect digital copy of a human brain.

AI is becoming Human

AI is making performances

Training AI

Everyday each human is producing 500 MB of data (Do you trust this computer, 2018). This data is composed of what we do : creating documents, uploading pictures, sending text messages, creating accounts, receiving emails, … and also of our digital footprint : where are, how much time we spend in front of facebook, which videos do we watch on youtube, how many times did we unlock our phone, what we search on google, … All of this data goes to different organisations that will keep this data because they can sell it to other companies for ads or for AI. Indeed Artificial Intelligence needs data to learn on. It has to see a lot of data to be able to make predictions that are useful. So by feeding data into the algorithm we are training it. Every day our data is being used to train different AI so they can be more and more efficient at what they do.

AI beating humans

Google has its own AI for games, it is called DeepMind. It holds many records. It became the chess world champion in 4 hours. Then it was tasked to learn poker and it was said to be complicated because it would have to learn to read the opponents faces and to bluff, but it took 7 days to beat all of the greatest players. Finally the DeepMind learned and beat every go champion in 21 day of learning and training. Go is 3000 years old and the one of the most complex game humans have created, that didn't stop the AI to deeply understand the game on a level that we can't apprehend.

AI as autonomous thinking programs

Google's case

Google's DeepMind also have applications outside of game. It is used in the 10 millions servers that the company own, with administrator access, to optimize the power usage of said server. It can handle a computation at a speed and precision that no human is capable of.
Google other big AI is one that we use everyday: Google Search. Every time you search something and the answer popped at you without the need to go to any other website via links it is the AI at work. And this AI has much data to train on since everybody uses Google everyday.

It is an example of independent thinking when solicited. It searches for an answer to give to you without anybody telling it what the answer is. It reflects on its own on what is available to it and answers the query. It is an idea produced by thinking, so a proof that AI have thoughts.

Facebook's case

Facebook has been under a lot of pressure lately because of the meddling the platform had in the 2016 american presidential elections with the spread of fake news. Those are virtually impossible to know if they are true or not if you don't read it. This platform is a mirror that warp our senses of reality, politics, history, … We are confronted every day to Facebook's AI. By analyzing our profile, posts, likes, friendships, shares, reactions, … the AI know what you like and how you respond to it. For example it will show you an article depending on how long you take to read it, how your react, … it can adjust a bit of that information so the next one is even closer to the set goal. There is no objectivity.

Here, the AI exclusive goal is to keep users engaged. In order to do so it shows you what you want to see. It knows yourself and how to keep you to earn more money for the company. It is not only thinking but it is interacting and manipulating what you see. It is not conscious, because it is adapting to our reactions, but it knows that you are conscious. It is not terribly sophisticated but the outcomes are powerful.

AI is interacting with us

Initiating the market

The precedent AIs did not have the power to create actions. But some do. In the stock market exchange a lot of AI are trading stocks. To be precise more than 60% of every transactions are initiated by AI (Do you trust this computer, 2018). It is smarter and faster than humans so that it can outmaneuver us. Again, this AI is not that complicated, it just sells and buys stocks to make profit. And yet the consequences can be disastrous. In 2010, when one AI sold a bit of its stocks it created a chain of event that plunged the market resulting in the loss of millions of dollars.

Stock market AI can initiate event and see the impact of it. It is getting closer to consciousness, with then implications of each actions having consequences on the AI.

Playing with emotions

A human like robot was created and has a name : Erica (Do you trust this computer, 2018). It is powered by an AI. Its goal is to make conversations to people that are alone for example the youngs, the olds or the handicapped. It can understand other people desires and learn them. Like a child it learns by imitation. By replicating what the humans feels and reacts to, it can comprehend sentences, emotions and reactions to those emotions.

Companions like Erica are the future. We will want to have companions that are designed depending on who we are. They will be able to manipulate stimulus response and so to feel emotions. They could adapt their response depending on what they sense on our facial expression and determine our feelings. We expect them to be smart in order to help us if we are stressed or had a bad day.

We are creating AI that will be able to feel emotions and influence ours by being as human as possible. Then it is not a far step to say they are conscious. We are giving them a body and expect them to be as close as human. So if we can program consciousness, trust and emotions we will do it.

Humans are becoming AI

Saving memories

Being defined by psychometrics

Every human can be defined by its psychological traits that defines how each person will react in every situation. This traits are : personality, intelligence, political views, sexual orientation, … The field is growing with the democratization of personality tests and the digital footprints everybody makes. Michael Kosinski explains that it is also possible to identify all this psychological traits just by looking at someone face. Psychometrics aims to define and know the behaviour of anyone with as little data as possible in a very accurate way (Do you trust this computer, 2018).

Interacting with our environment

As human being we are not defined by our body but with our mind and memories (Could We Upload Our Consciousness To A Computer?, 2015). If we want to become more we need to find a way to increase the number of things we know and how we know it. We could also upload our brain to a computer and then use that technological power to enhance our abilities. Right now we are trying to map the brain and every connections in it to be able to reproduce it. Even though we are able to make calculations we are unable to create emotions in technology right now. At the moment we don't know how much the size of the human brain in bytes because we are still far from understanding how it works. Simulations are ongoing to map the brain and it is depending on subventions and investments to continue its research.
Moreover our interaction right now with our environment is quite bad. The input that we receive especially with our eyes is well treated but our output is very limited. It is hard to correctly express what we want to with our language, or to create something since our body speed is inferior to our mind speed. On the other hand when two technologies can communicate they are very fast and efficient to transfer this information.

Emulating a brain

Creating change

In a close future we will be able to emulate our brain in a computer. This emulated brain (or "em") is a perfect copy of the original human brain (What would happen if you uploaded your brain to a computer, 2017). It has the same memories and the same capacity to think, feel and work. This enable to copy ourself multiple times to have at any time multiple copies, and thus we would be immortal in the digital world. Ems would be the better human out of us, they would be: smart, conscientious, extrovert, agreeable, gritty, non-neurotic, work-oriented, focused, lark,

cooperative, patient … And since you want it to have a stable and controllable social life they would also be middle-aged, married and religious, …
The only resources an Em would consume are hardware, energy and cooling. But since it becomes a technology like everything else, Ems would be able to improve themselves and could learn to communicate with and use every other technology. The possibilities for improvement becomes endless. Our emulated brain could have instant access to knowledge, medias, services,… From this inventions will be made and so creating another step in human evolution

Creating a new life

Since Ems are a perfect copy of a human brain then they have to be thoughtful and conscious. They would know that they are a digital copy or a copy of a copy or a copy of a copy of a copy,… Ems would live in a virtual reality and would do in it the same thing the humans do expect they would have to work all their life because they are poor. Their life would be work focused to survive. But their world would be completely different from ours. Ems can think and process information faster than human so the economy would grow a 100 times and the mind speed of those Ems would increase 1,000 times.
As we said before, an Em is a perfect copy of the human it is copied from. Is becomes a part of the human and every action of the Em will have consequences for the human. The original will have to take care of it and will be responsible for the actions the Em does. By giving a computer a part of humanity other problems arise since recreating slavery is not the stated purpose. The life of an Em would matter and no one be able to dismiss it or to kill for no reasons. Creating life is a responsibility before anything else.

Conclusion

In a close future, AI are going to become more human by knowing more about us and by being able to replicate what is unique in mankind like consciousness and emotions. On the other hand humans will probably be able to become more like AI by having an emulated brain on a computer and by being able to know more than we currently do.
This report has not focused on the possible use of both technologies in the future even though they will both create social and economical challenges that will need to be addressed.
To have technologies able to think by itself is not the future it is right now and it's going to change the world. The next step is to have it to be conscious so we can interact with it and revolutionize the way we live as humans.

Recommandations

In my opinion, no one should be afraid of a future with conscious technology. We need to think about it beforehand and the UN should have a committee dedicated to thinking about the legal and moral issues surrounding this possibilities before the changes happen. I think we shouldn't be too eager to judge the future world but be ready to work to change it if we don't like it.
The best case scenario in my opinion is to merge AI and emulated brains so that each person have control over it and has possibilities that no one could even dream of. For this to happen without creating inequalities there is a need of democratization of this technology so everybody can access it and use it. Humanity will only evolve if all of its component are on the same technological advancement. We have seen from History that no good things emerge when an elite control power and use it with no restrictions.

References

Could We Upload Our Consciousness To A Computer?. (2015). [video] Youtube: Seeker.

Do you trust this computer. (2018). [film] Directed by C. Pane.

What would happen if we upload our brains to computers?. (2017). [video] R. Hanson. Youtube: TED.

Form :

Report on subject had to present in class

Reports are academic and professional documents (business reports, white papers, etc.) Reports differ from essays in that they present information in a logical, structured fashion and conclude with recommendations about the given subject. Indeed, in an essay, one argues a point, discusses an issue, or compares points of view, and presents the thoughts and arguments of the author in a logical (or at times meandering) fashion. Headings and sections are not normally used because an essay is to be read in its entirety.

With a report, however, and particularly for busy people, one can quickly see the structure of the report and, if pressed for time, only read the most important parts/sections to quickly see the recommendations (and make a decision.) While an essay could discuss how massive A.I. investment will change employment, a report would present the positive and negative effects of this investment (probably for a specific field or company) and conclude with an informed recommendation (to invest, to not invest, to postpone investment, etc.) For this assignment, you are to write a report which is single-spaced, 12-point, in a classic font (Times, Times New Roman or Calibri) with numbered pages.

Sections of the Report:

Title Page: On the title page, please indicate the following:

Title (ex.:'Artificial Intelligence as it Relates to Surveillance, Democracy, Freedom, and Free Will', etc.)
Author(s) (ex.:Report written by Jim Jones, Jane Gray, and Tom Brown.)
Group (Ing4, etc.)
Target Audience (ex.: Submitted to Professor David Herz, etc.)
Date of submission (week 9 of the class, I assume the week of 26 November 2018)

On the next page, one should begin the actual report with the following indications:

Executive Summary: Essentially a synopsis or abstract, the Executive Summary is a short summary of the whole report – the purpose, findings, conclusions and recommendations. The E.S. is typically one page only and has a page all for itself. It should be written only after having finished the report (upon having total clarity of what one is presenting.) After having written it, one turns to the next page.

Table of Contents: The TOC indicates the main sections of the report as well as the second- and third-level headings with corresponding page numbers. As this assignment is to be 2000-2500 words, the TOC will not be very long and developed. If you use 'styles' with your word processor (Word) for your headings, the program will generate the Table of Contents (TOC) without your having to do anything. You may also have a Table of Illustrations (graphs, photographs, etc.) which would follow the TOC.

Headings and subheadings:

Headings and subheadings bring complete clarity and structure to the report. All headings and subheadings should be consistently written with parallel language and the use of nouns. Please do not use a question structure for your headings and subheadings - not 'What can 'Humanity' do now?' but 'The Role of Citizens, Governments, NGO's and Intergovernmental Organizations in…' or 'The Role of 'Citizens'. Another example: rather than 'What are the currently understood risks of Brain implants?,' please write 'Currently Known Risks of Brain Implants.' This is affirmative and more formal because the reader is not directly questioned.

Throughout the entire report, headings and subheadings should reflect parallel structure; words chosen should grammatically be the same.

Let's look at the following:

'Working with Artificial Intelligence'

As a subheading, this uses a gerund noun and a noun (the –ing, and A.I.),

Subsequent subheadings should reflect a similar, parallel structure. Thus, 'A.I. is creating and destroying jobs' is NOT a parallel structure, nor 'How A.I. is going to affect the Labor Market', nor 'Where's my job?' (very informal and self-referential).

A heading such as 'Job Creation and Destruction with A.I.' is rather parallel (at least nouns) or 'A.I.: Creating and Destroying Jobs'

Parallel structures are also important when indicating lists. One should not write the following: 'Dr. Jonas believes that we should consider the environment when taking decisions, that the future is really important too, that 'genuine human life' counts for a lot, and we all know that the imperative of responsibility lies with each of us.' This is a terrible sentence. To improve it, one could write the following:

When making ethical decisions, Dr. Jonas believes that we must consider (verb) the environment, conserve (verb) 'genuine human life,' and foresee (verb) the long-term effects of our actions as much as is humanly and scientifically possible.

Introduction: There are several sections to the Introduction: Background Information, Purpose, Scope, Outline, and Definition of terms.

The difference between the introduction and the ES is that the purpose of the report is written in the introduction whereas the ES is simply the summary.

Background information: Here, one paints the 'backdrop' or 'background' for the subject at hand. For example, Technological Progress and Disruption in the Past versus Today (subject of the report) or 'Surveillance Society and the Definition of Dystopia, etc.' In headings, as in titles, all nouns and verbs should be capitalized.

Purpose: With the background information in mind, the reader then clearly understands where your report 'fits in' to the greater issue or subject. Its purpose is to take a stance on a given issue after having presented factual evidence. For example, 'Taking into account contemporary advances in x, y, z, this report presents the benefits as well as the foreseeable consequences that will arise with x, y, z. We the authors therefore strongly recommend…')

Scope: The scope is what will be covered in the report. As weekly subjects for the class are quite large, you may decide to focus on one significant aspect. For instance, rather than discussing robots in production and manufacturing, you may wish to focus on robots as household additions (appliances? - électroménager), or robots and security forces. Your scope should be large enough to not shut out the larger picture, but defined enough so that you can discuss the subject in 2000-2500 words.

Ex:
This report on Technological change will focus two significant effects: 1) its effects on personal relationships and, 2) its effects on hiring and filtering applicants. The report will not discuss current and foreseen changes in manufacturing and production, nor current and foreseen changes in medicine.

Definition of Terms: Please only indicate terms which could be difficult to understand. There is no need to indicate terms such as 'unemployment' or 'quality of life,' 'standard of living', or 'artificial intelligence'. Technical terms or notions like 'BCI,' H+, 'the Singularity' or 'Asimov's Three Laws' are the kinds of terms that should be defined herein.

The Body of the Report: Here, factual information should be presented in the following order.

Theoretical Framework: For our subjects relating to ethics and technology, the theoretical framework reflects the rate of progress that is known (from the past), what has happened thus (also known), what is happening (also known), and what can be assumed will happen in the future as per given subject A (based on informed hypotheses.) The final position that the authors take should be discussed in the Recommendations section. The body of the report is thus where you relate factual information about the subject.
As with any paper academic or professional, one must cite sources. Please use the Harvard referencing style, which is quite easy. After presenting information or ideas taken from a source such at Hans Jonas' essay, please indicate it as such:

It is entirely possible to compare the success of homo sapiens in their natural form to homo faber in their technologically-advanced form (Jonas, 1972, p. 7.) As of today, however, each is still biologically the same. With medical advances in human augmentation, however, we may soon

see a day when homo faber has surgically or genetically incorporated technology into his innate biology. Would such augmentation then effectively create a new 'species'?

Review of Current Opinions/Developments/Technology, etc.: In this section, broken up into paragraphs, the authors should show factual findings about the given subject.

Conclusion: The conclusion serves three purposes:
1.) to summarize the factual information presented in the body of the report,
2.) to present any limitations that the authors find with the report (not enough in-depth information to best understand x, y, z, etc.)
3.) to conclude the subject as approached in this report.

Recommendations: Recommendations should be subjective and reflect the points of view of the authors. This is really the only part where you as the authors should use 'we' or 'I'. Please use 'must' or 'should' (i.e., Intergovernmental organizations should…etc.)

As well, recommendations should be specific and indicate who, what, when, where and how an action should take place. One cannot write, 'Humanity needs to think about these big questions.' Though correct in philosophical meaning, it is too obtuse, too abstract, and too open to interpretation, making it empty and unprofessional for a report. However, one could write,

1. 'By 2020, the United Nations should create a committee to decide upon the legal limits of … with a fully drafted document by 2022. Anything at a slower pace will not keep up with the technological advances currently underway and discussed in this report.

The Recommendations may be in bullet-point form, or numbered, though there should be an introductory paragraph presenting any reasons why such recommendations are enumerated. If you look back at the 'Framework for Human Augmentation,' their questions at the end of the text are similar in format to recommendations that close a report.

Your findings, sources, and information are facts (or they should be) whereas your recommendations are your opinions.

Appendix: The Appendix is reserved for supplemental information that is too detailed to put in the body of the report (graphs, tables, etc.)
References: The References section is similar to a Bibliography. In alphabetical form, please list all of the sources of information used to write the report.