Can the use of artificial intelligence allow us to create a utopian society?

Andre Hirschler

Abstract

The aim of this project is to discuss the developing field of artificial intelligence with regards to where future application and use of the technology might lead us in the future. This possible future is envisioned as a utopia which human society could aim to create.

The discussion focuses on current developments in the field of artificial intelligence as well as predicted breakthroughs as evidence of a possible intelligence explosion. I demonstrate that artificial intelligence is a very real and likely possibility in the near future. With the premise that such machinery will be available, the discussion then examines the possible uses of it to create a utopian state. I also show the difficulties in creating intelligent machines to the extent required of a utopia. I then examine the purpose and prospect of a utopia, and whether such a social state is achievable. I look at how such a state could, and would likely fail, even with the advent of intelligent machines.

I conclude by finding that given the current progress made by researchers, artificially intelligent machines are likely to become commonplace very soon. However, it rests with the fact that human society would have to be radically altered in order to achieve the state desired for a utopian society. Artificially intelligent entities could possibly help us achieve such a state, but without human social change it is unlikely, and so we cannot achieve a utopian society with the help of artificial intelligence.

Introduction

Artificial intelligence has had multiple breakthroughs recently, and it appears that there are new achievements by artificially intelligent machines each week. Since the widely publicised defeat of chess Grandmaster Gary Kasparov in 1997 by Deep Blue, an intelligent machine, it would appear that artificial intelligence research is on the rise. The past decade has seen greater breakthroughs than ever, with machines now competing in a wide range of both mental and physical activities. This has led to renewed speculation about the possibility of creating fully intelligent and possibly conscious machines, and if successful in that respect to what it would mean for human society.

Artificial intelligence is intelligence as exhibited by machines or programs which are commonly associated with human capabilities.[1] The main aspect of artificial intelligence is its ability to replicate human intelligence, and in some cases in surpass it with machine precision. Artificial intelligence is thus an incredibly radical and impactful technology. Its range of uses and the ability to apply intelligent programs to machines makes the technology useful in almost every aspect of society and in nearly every financial sector. Recent developments and research by companies such as Alphabet[2], Apple and other corporate giants have led to increased hope about the successful creation of intelligent machinery.

[1] The fact that artificial intelligence has had such varied successes in reproducing intelligent characteristics has led to many different interpretations of what is AI, an issue I address in my literature review. Some common descriptors of AI are cognitive functions, speech recognition and reason.

[2] Alphabet owns several subsidiaries which conduct research on artificial intelligence, such as Google X, DeepMind as well as other robotics companies.

Literature Review

What is Intelligence?

The first issue when considering artificial intelligence is the lack of a clear definition of intelligence. This lack of clarity not only applies to our understanding of our own, human intelligence, but passes to any "intelligent" machines as well. Given a lack of clarity on what would constitute intelligence, it becomes difficult to accurately assess any developments of AI (artificial intelligence).

If we were to compare it to Human Intelligence, it would include the ability to learn from experience, adapt to new environments and situations as well as manipulate those environments and also to be able to understand and handle abstract concepts[3]. Although fairly simple in definition, there are in fact four main areas of human intelligence studies, ranging from biological to mental studies of how intelligence would work (pp. 1, Human Intelligence). What this makes apparent is the lack of understanding of how our own intelligence functions as we have conflicting fields of thought about it.

The issue with intelligence regarding to Artificial Intelligence

An issue in the discussion of artificial intelligence therefore becomes the ability to truly define a machine or algorithm as intelligent due to our limited understanding of intelligence. We do not even yet have a complete understanding of human intelligence. Given that it is likely machine intelligence is altogether different to human intelligence, it allows the field of artificial intelligence research to expand and make breakthroughs, yet not be recognised as artificial intelligence. Data analysis on AI usage by Thomson Reuters[4] has shown that 57.9% of businesses using big data[5] employ AI solutions to deal with the data. It also demonstrates that the highest use of AI by 48% of the companies was for automated communications that give data and advice about potential decisions based on the data. (Thomson Reuters, 2015). What this means is that there is already extensive commercial use of forms of artificial intelligence in corporations. In fact, they seemingly require it in order to function as they do. These forms of AI are even involved in human communication, an area of AI considered to be the most difficult to achieve well. Yet it would seem that such programs, often nothing more than algorithms, are already extensively used in communication.

[3] This is the definition according to the Encyclopaedia Britannica entry. Entry was written by Robert J. Sternberg, currently a Professor of Human Development at Cornell University, but has previously served as Professor in many positions and fields of psychology at Yale, Tufts, Wyoming and Oklahoma. He has also served as President of the American Psychological Association.

[4] Thomson Reuters is a large multi-national data analysis and advice corporation. They provide a range of businesses in multiple fields analysis services of their company labour and data sectors. As such they would have the access to conduct conclusive and extensive research into the area.

[5] Big Data is an issue faced by many large corporations in recent years, as they lack the ability to effectively analyse and store their data as they receive so much of it on a daily basis.

Weak Artificial Intelligence

Such algorithms are considered weak AI, an artificial intelligence program that is effectively intelligent at a specific task, but not intelligent in the form a human is intelligent. Many of the current developments are focused on this area of intelligence, as it is simpler to create a weak AI program than an overall general intelligence AI. Nilsson[6] claims that the focus on this form of AI is because of a 'Concern for "respectability"' through success (Nilsson, 2009). This is because it is easier to successfully replicate intelligence in a specific and focused task, rather than intelligence similar to a human, in that it can learn to be capable in multiple fields. Much of the recent progress in AI is thanks to breakthroughs made with weak AI programs restoring faith in the field.

General Artificial Intelligence

The most promising aspect of AI however, is not soft intelligence, which can at most replace humans at specific jobs. The aspect with greatest potential use is what is considered general intelligence, or complete AI. This is an artificially intelligent machine with the same intelligence as a human, in the sense that it can learn and adapt, as well as understand given information or an environment. This is effectively the aim of AI, the development of a fully intelligent machine at a human level[7]. We effectively have the individual pieces, with machines considered to be artificially intelligent capable of a varied spectrum of intelligent task and thought processes. The difficulty comes in attempting to join them together.

The first clear incident of fully autonomous work was conducted by W. Grey Walter[8] with his autonomous robots Elsie and Elmer during the 1950's *Machina speculatrix* (S & B, 2003). These *Machina speculatrix* as he named them, were some of first autonomous robots constructed. They had simple touch and light sensors, and would move towards light sources and away from object they hit. However, as Walter changed the conditions, he began to notice odd behaviour. This was behaviour he had not expected, and as such he proclaimed it be limited intelligence, as the robots were reacting to the changes in the conditions. What this suggested was that intelligence in machines occurs when they display seemingly intelligent yet unpredictable actions.

The variations of Artificial Intelligence

The issue this caused was that critics claimed that Walter's alterations to the environment could have been designed to purposefully cause those supposedly randomised reactions. This can be more clearly seen in the possible objection to the

[6] Nils Nilsson was one of the founding fathers of artificial intelligence research. He has published several papers and books on the field of artificial intelligence, as well as worked at the Stanford AI Laboratory.
[7] However, it can be said that although this is the point to which the AI field leads, there are no projects attempting to create such a machine, but rather research attempting to create different aspects of intelligence for different purposes.
[8] W. Grey Walter was Neurophysiologist who also pioneered the field of cybernetics through his research into showing how small amounts of brain cells (e.g. small circuits) could show complex behaviour.

Turing Test[9]. It became an aim to, rather than create true intelligence that could pass the test, design a machine specifically to pass it (S & B, 2003). As of yet the program named Eugene Goostman is the only successful program to have passed the test, having convinced 33% of judges that it was human[10] (BBC, 2014).This is why artificial intelligence becomes difficult to deal with, as are the machines intelligent in their ability to pass the test and do nothing else, or rather not intelligent due to specific programming? Goostman had the personality of a thirteen-year-old, and so questions a judge could possibly ask were limited, giving the program an advantage. If machines are thus considered intelligent due to their ability to fulfil a specific role, then we have arguably already achieved intelligence, and are now seeking a concept known as superintelligence. If not then it would mean that most of the current developments are simply small individual steps towards possible human level intelligence, or beyond that, of a single machine or program.

Current artificially intelligent machines are often role specific, and as such are often dismissed as capable programs rather than truly intelligent machines. As stated by Nick Bostrom[11] *"a lot of cutting edge AI has filtered into general applications, often without being called AI because once something becomes useful enough and common enough it's not labelled AI anymore."* (Bostrom, 2006). This is the notion that there is an underappreciation of the extensive use of AI programs already in field, as demonstrated by the (Thomson Reuters, 2015) research. This is again likely due to the lack of clarification as to what an artificial intelligence is. Experts such as professor Bostrom suggest that machine that can autonomously perform a task, any task, at its simplest level, is technically artificially intelligent.

The book 'Rise of the Robots' by Martin Ford[12] suggests a solution to this in the idea that for a machine to be intelligent it must be able to complete a predictable job, which it can learn how to do and be programmed for, but does not always have a single answer choice. These could be simple management jobs making the most productive decisions or even executive jobs analysing productivity in a workspace and effecting changes accordingly. This helps provide us with a guideline as to what we could consider intelligence, and consider the topic with that guideline in mind.

---

[9] The Turing Test was a method of testing a program to judge whether it is human or artificial, originally proposed by Alan Turing. The test was conducted by a human questioner of an invisible person or machine giving the answers. According to the responses, the questioner should be able to judge whether the responder is a machine or human. (Traiger, 2000)

[10] The program simulating a 13-year old boy conviced 33% of London Royal Society judges that it was human, enough to pass the 30% margin required of the Turing test. The fact other judges were not fooled does mean that perhaps the 33% were worse judges, and the testing itself was heavily criticised by others, such as Hugh Loebner, who devised his own form of Turing Test which Eugene failed.

[11] Nick Bostrom is a Professor of Philosophy at Oxford University. He is also the founder of the Oxford Martin Program on the Impacts of Future Technology and Founding Director of the Future of Humanity Institute at Oxford. He is one of the leading advocates in the field of A.I. developments, mostly concerned with its potential impact. He produced a book SuperIntelligence: Paths, Dangers, Strategies on the topic and is an advisor to the Future of Life Institute.

[12] Martin Ford is originally a founder of a Silicon Valley firm that develops software. In 2009 he was the first 21st century author to publish a book focusing on job loss from automation, and has continued that notion in Rise of the Robots, this time focusing on how "white collar" jobs could also be lost. His works have been supported by notable professors and academics, including Michael Osborne of Oxford Uni.

Thus when examining an issue of how we could harness the potential of AI, it seems clear that the initial uncertainty as to what artificial intelligence actually is makes it more complicated to understand any uses of AI, as we fail to understand its potential as a technology. The multiple fields which the AI covers also make it difficult to focus the topic. It seems clear that when considering AI as a threat, it is important to examine the separate impacts the different branches of artificial intelligence could have, and equally the possible collective impact it could have.

The beginnings of AI

It is surprising to think that the beginnings of artificial intelligence research appeared around the same time as computer development appeared. Yet artificial intelligence has developed at a slower rate, whereas computing capabilities and access has developed much faster and is in much more use than artificially intelligent technology. Thus since the work of the original pioneers of AI and computing, computing has become widespread thanks to its business utility, whereas AI remained constrained to impossible promises.

These pioneers[13] began their work in Dartmouth in 1956 (John, et al., 1955), at a conference specifically aimed to study the newly named field of artificial intelligence, a term coined by John McCarthy[14] in 1955. The conference lasted two months, and was comprised of a team of ten researchers who would attempt to test and create as many artificially intelligent programs as possible in those two months. This event would establish artificial intelligence research as a field. In those two short months' researchers created and developed many basic programs that became the foundations of AI research. Most important were the researchers that went on to work in the field over the next 'Golden Years' of AI from 1956-1974, during which astonishing breakthroughs were made very quickly. These ranged from programs that solved algebra to winning at checkers and learning English. It led to the researcher Herbert Simon[15] to predict that "machines will be capable, within twenty years, of doing any work man can do"[16].

The Artificial Intelligence Winter & its effects

This turned out to be wrong, and the initial optimism and underestimation by the pioneers in AI research meant the steady decline of AI as funding stopped due to lack of useful results. When the Science Research Council[17] approached Sir James

[13] The conference participants included John McCarthy, Marvin Minsky, Claude Shannon, Ray Solomonoff, Alan Newell, Herbert A. Simon, Nathaniel Rochester, Trenchard More, Oliver Selfridge and Arthur Samuel

[14] John McCarthy is considered one of the founders of the field of AI and also one of its greatest pioneers, being behind multiple advancements in the field throughout the late 1900's. He originally proposed the Dartmouth Conference and established its aims, as well as the establishment of a field of research called 'Artificial intelligence'.

[15] Herbert Simon was a distinguished scholar across many fields of computational technology, as well as Nobel laureate. He worked on several problem solving programs, and produced The Logic Theory machine along with Newell in 1956, often known as the first intelligent program due to its ability to reason and problem solve as a human.

[16] Simon was quoted to say this in 1965.

[17] At the time a UK council which provided public funding to research projects.

Lighthill[18] to produce a report on the field of AI, his subsequent findings of failure in the basic research led both the British and subsequent American governments to cut off funding. This period from 1974 to 1980 became known as the first AI winter[19].

The effects of the AI winter were the decline of breakthrough's in the fields and developments, and the general disinterest in AI for an extended period of time by comparison to other developing fields such as computing. This could provide a likely explanation as to why it is only recently that we have seen a resurfacing of the artificial intelligence field in terms of achievements. What it created was a situation in which researchers were able to procure more funding by mislabelling AI. As well, as explained by Herbert Simon, the field is finally producing successful machines and programs that are competing either commercially or in terms of breakthroughs with other major technological fields. This is also largely thanks to the AI winter, as researchers began investigating individual aspects of intelligence rather than a single, wholly intelligent machine. This led to breakthroughs such as the development of Deep Blue. It also means that in research terms there is perhaps less coverage and worry over research, as it is much less threatening to say you are developing a program to play chess or to have facial recognition, than it is to say you are recreating human intelligence in a machine.

The importance of Robots

It is also useful to consider at this point robots, a term often used to describe artificially intelligent physical machines, who have similar physical features and capabilities to humans. In a contemporary sense, a robot is effectively a slave. If we were to create an artificially intelligent robot, which would for all intents and purposes be alike to us mentally, it creates much ethical debate about what right we would have as their creators to use them as unpaid labour. Robots are the field of AI best known to the public judging by a survey (Dodd, et al., n.d.) conducted by research students[20], in which most people answered 'what they knew most about AI' as the development of robotics or robots in culture. The rise of robots and a popular stigma caused by the negative, often evil depiction of robots in popular culture has made robotics a seemingly essential field to showcase the development of AI. As AI improves so do the capabilities of robots, and it is robots who would be most useful in creating a utopian state, given their capabilities to fulfil human labour.

[18] Sir James Lighthill was an English scholar in mathematics. He held multiple senior mathematics posts at universities across the country, including Imperial and Trinity College (Cambridge).

[19] AI winter is the term used to describe periods of time when funding for AI was particularly hard to find due to decreased appeal of the technology.

[20] Given that the project is a bachelor's degree report, there are some limitations to the extent of the project as it could only be conducted with a smaller group of people surveyed. However, the questions themselves are useful as many people wrote their insights into AI, a high percentage answered in relation to robots. The report was also conducted with a professor overseeing and advising, suggesting it should have been well-conducted and as such accurate.

Accelerated Development

The accelerated development of AI is linked to Moore's Law[21] (Intel, 2015) which proposed and later established that the number of transistors in a computer chip roughly doubled every 18-24 months. This theory has now been applied to technology as a whole in a very broad sense that our technological capabilities roughly double every 18-24 months. When compared to the rapid growth and development of AI in the past this does not seem like such an implausible application. Yet it is when considering AI in terms of current and future development that it seems reasonable to expect an increasing rate of breakthroughs and successes in the field.

If we compare the Logic Theorist program[22] (Encyclopaedia Britannica, n.d.), considered to be the first AI program, from 1955, to perhaps Google's Deepmind[23] (Rowan, 2015), a program that has demonstrated actual ability to learn and surpass humans in something as complex as Atari video games[24], we can see the relatively small time gap. The program was even able to develop a strategy in the game "breakout" that its original designers had not even considered, demonstrating that it soon learnt to understand a human design better than its designers. The difference from algorithm solving to human learning is an impressive leap in 60 years.

Even more significant, if we look at the language and photo recognition capabilities of programs such as Apple's Siri[25] (Apple Inc., n.d.) and Microsoft's Cortana[26] which can be found in mainstream products, the advancement is all the more impressive given how ordinary those programs have now become. Around 1.5 billion Apple products have been sold (Statistic Brain, 2015). Not all of these will have Siri, given the software was only introduced in 2012, but many of them will[27]. PC (Personal Computers running windows programming) sales following the release of Windows 10, the software containing Cortana also reached 75.7 million in 2015 (Gartner, 2016). Yet as Windows 10 was released free of charge to existing PC users, the

[21] Gordon E. Moore, co-founder and chairman of Intel Corporation, posited Moore's Law in 1965, as a standard to which a company should aim development, and as an observation of the development of computing speed or power.

[22] The logic theorist was a program designed by Newell, Shaw and Simon from 1955-56. The program could successfully prove theorems from the *Principia Mathematica* autonomously, the Mathematica being a study of the foundations of maths. In one case it even devised a more elegant proof of a theorem, surpassing humans.

[23] DeepMind is the company that developed what is considered the best learning algorithm to date. The algorithm it developed learnt to play thousands of game requiring various skills, and quickly mastered all of them in a matter of hours per game. This learning was done by feeding the computer limited information, such as instructions, then let the algorithm play and learn the best strategies.

[24] Atari are a series of arcade games, requiring skill and strategy to play. The games are as varied as kung-Fu fighting, racing and space flight games or as simple as pong, a paddle game.

[25] Siri is the personal virtual assistant developed by Apple for their products. A user can a have a light conversation with Siri, ask her questions and even tell her jokes.

[26] Cortana is Microsoft's version of Siri, sharing many of the same capabilities. However, Cortana was developed later, and Microsoft claim that Cortana has more personality than Siri, humanising the program more.

[27] I could not find the exact number of products sold that had Siri, but given the successes of the iPhone6 as the best-selling model, as well as iPad sales, it will be a considerable proportion, at least a billion, of products sold (laptops also have Siri) - http://bit.ly/19eBB1s ; http://bit.ly/1M5ygp8

number of Windows 10 users could be many hundred million higher[28] This indicates an already widespread presence of an artificially intelligent entity in personal and household use.

Cortana is the first widely developed office product to have a personality, Microsoft describing the program as 'Your clever new personal assistant' (Microsoft, n.d.); even I accidentally refer to Cortana as her. Cortana's ability to convincingly display gender is just another example of how developed these programs are becoming. Just as impressive, if not even more so, is Google Now, the mobile operating system developed by Google. Although Now doesn't present a personality, it's predictive and learning capabilities are what makes it stand out. Now will automatically track your movements to the point where it knows where you go at certain times during the day, such as to work in the morning, and then keeps you updated on your commute. It also analyses your search history, and based on the sort of websites you access or search for, provides you news based around that content. This pre-emptive, understanding program is perhaps most impressive given its ability to learn about you and what you do[29].

Thus accelerated development shows us how AI might quickly become a relevant and useful technology. It suggests that we will be further than we can imagine in 60 years, and perhaps by then will have developed a fully intelligent program. Early predictions about successes in the field of AI might have been wildly inaccurate, but they become more and more accurate as researchers set more realistic goals, and achieve them. When you couple the recent commercial successes of AI programs such as Cortana with the AI effect, it creates a situation where AI is introduced extensively into the public without them realising. This allows for public integration with AI without any possible stigma against such a possibly influential force.

The Artificial Intelligence effect

The AI effect was alluded to in my earlier quotation of Nick Bostrom. That is the general public and commercial reaction to the development of an artificial intelligence that it isn't actually intelligent, just an able simulation. Professor Drew McDermott[30] gives insight into the problem in his analysis of Deep Blue[31] following the computer's victory over Chess Grandmaster Kasparov in 1997. In a New York Times article (McDermott, 1997), McDermott argues that although Deep Blue is not intelligent in the sense of general intelligence, given its lack of depth, it is fallacious to claim it has no intelligence. McDermott tells us many 'commentators' said Deep Blue won by simply running through all possible moves. This view can be applied to any modern day program, and is the central argument against any AI development as actually being AI. This is the AI effect, the claim that an AI program is not truly

[28] I could not find any data on the number of downloads of Windows 10.

[29] For the program to fully function you would of course have to give it access to your search and location history – It is not an invasive entity.

[30] Drew McDermott is a computer science professor at Yale University. He has conducted research in multiple fields of AI and is considered an expert on the topic. He ran a class considering Deep Blue at Yale at the time of writing his Journal Article (How intelligent is Deep Blue?)

[31] Deep Blue was the first AI machine that managed to defeat a Chess Grandmaster in a six series game. Deep Blue's success made headlines that year, and it helped revitalise public perception of AI.

intelligent, but rather just very effective at emulating intelligence. McDermott suggests to us that this is no different to our own intelligence, given that we similarly function by the hundreds of messages fired through our neural networks. Simply because a computer's thought works differently does not mean that it isn't thought. In fact, it is the individual breakthroughs of machines which are dismissed as lacking in intelligence because they are lacking in breadth of capabilities that are likely the most worrying. This is because they can be dismissed as not intelligent, when in reality they can fulfil similar roles to the average human, only in most cases more effectively.

Machine Safe Jobs

This leads on to the idea of jobs which are safe from the technological revolution. Martin Ford argues that thus far, jobs which have replaced human workers with machines and those which are most threatened by further machine development can be deemed as "routine". Routine machines are now extremely varied, with both software and physical programs carrying out certain roles. Routine jobs can be characterised as jobs which consist of a repetitive action, things such as assembly line labour. The greatest labour replacement with robotic, intelligent machines has occurred in large factories. Amazon replaced its entire warehouse floor staff involved with the movement of goods with small autonomous robots[32] developed by Kiva industries.

However, he goes on to argue that "predictable[33]" should also not be considered machine safe. A sufficiently intelligent machine can predict what to do best in a situation better or as well as most human labourers. This would therefore affect not only 'blue-collar' jobs as the development of automated machinery has thus far, but also 'white-collar' jobs, which were previously believed to be risk free from machines due to the requirement of understanding and evaluation of situations.

Although it might seem irrelevant as to whether or not a robot can race a motorcycle, it brings about another fundamental set of questions. If a robot can learn to drive safely, as proven by Alphabet's driverless cars, what can it learn to do next?

Autonomous Vehicles

Driverless cars are one of the most publicised fields of artificial intelligence. There is little doubt that such a vehicle is intelligent given its ability to drive itself without help from a human controller. It is also the field that has achieved the greatest commercial as well as developmental success.

In the commercial sense, there are around 25 large car manufacturers that are currently developing and producing autonomous or semi-autonomous vehicles[34]. These include well known car manufacturers such as Mercedes and BMW, both of

[32] These robots are box-like, autonomous machines which can lift and carry up to 700 pounds.
[33] Predictable being a case where you can easily and accurately analyse what impact a decision would have and how you would apply it. E.g. any scenario where one could do a cost-benefit or other form of decision making analysis.
[34] Semi-autonomous vehicles are vehicles that still require a human operator for certain driving conditions.

which have unleashed future autonomous vehicle products[35]. Tesla cars already have certain autonomous features, such as parking and highway driving which have been demonstrated to function in road use.

The success had by these vehicles also suggests a high-level of reliability that is driving the industry forward. Google's driverless cars have only had 341 incidents in 682,895km of driving, only thirteen of which might have resulted in a crash (BBC, 2016). Despite this, public outcry such as strikes by teamster's unions have caused the policy issues over the allowance of autonomous vehicles in public[36]. Although it seems that autonomous vehicles are definitely becoming widespread, any major incidents associated with those vehicles could quickly halt development.

The Technological Utopia/Dystopia of Science Fiction

Given the uncertainties as to where the field is headed, and what could potentially be achieved, it is difficult to predict future impacts the development of either soft or hard AI could have. The often hopeful and unlikely claims made by previous developers or researches as to the success of their research and the likelihood of AI machines within a decade or less has made it difficult to use their estimates reliably. Thus when considering a probable, or even possible future, it is useful to examine the field of science fiction.

Science fiction has delved deep into the AI problem given its freedom to explore future scenarios. Some of the most famous scenarios include the works of Isaac Asimov (Asimov, 1953), (Asimov, 1950), the Terminator film series (Terminator, 1984) and other standalone films or works such as 2001: A Space Odyssey (2001: A Space Odyssey, 1968).

Dystopia

The Terminator franchise is exemplary of an AI dystopia. It demonstrates what could happen if a controlling artificially intelligent entity chose to eliminate the human race on the basis of their self-destructive nature. Such a machine has access to varied autonomous war machines that it can employ to the end of eliminating humans. Although the scenario has been dismissed as being ridiculous in scope and destruction, it does pose some questions as to what could and/or would machines be capable of doing if they act out of accordance with their programming.

More specific, and perhaps more believable examples are those given in (Ex Machina, 2015) and (2001: A Space Odyssey, 1968). These demonstrate how individual machines could act against human creators in order to achieve an intended goal. Both are particularly chilling examples as it demonstrates how a

[35] BMW has recently released its radical concept car BMW next 100 series, aimed at car development over the next 100 years - http://www.next100.bmw/en/index.html. Mercedes have also released projects such as the Mercedes-Benz Future Truck, a long-haul autonomous freight vehicle - https://www.mercedes-benz.com/en/mercedes-benz/innovation/the-long-haul-truck-of-the-future/.

[36] The teamster's unions, a union of freight transport drivers – long-haul truckers etc… have protested against the development of autonomous vehicles as it threatens to render them redundant and thus lose their jobs. Taxi cab drivers have and similar complaints.

machine could logically, and willingly kill a human being. Space Odyssey is effective at demonstrating how a misinterpretation by a controlling autonomous entity such as HAL[37] could be dangerous. The machine knows it is crucial to the completion of the mission, and so kills the humans attempting to shut in down as it has orders to complete the mission. Such a logical loop is difficult to program against. Ex Machina demonstrates how a superior intelligence could manipulate humans to its benefit. Also it raises some of the ethical issues involved with controlling what is apparently a conscious, intelligent being.

Utopia

The works of Isaac Asimov are the best example of an AI utopia. Humans live amongst robots, which are in fact integrated into the life of any human, either around them or with them. Asimov has justified such a peaceful society with his concept of 'The three laws of Robotics'[38] (Asimov, 1950). These laws allow the establishment of a society where people can always depend upon any robots. Although robots can be privately owned, they serve to the benefit of all. There are also no ethics to worry about as all robots are made subservient to humans, yet at the same time have an apparent lack of emotion and feeling.

[37] HAL is the intelligent on board ship computer that is tasked with control of the ship, and so in many ways dictates the success of the mission.

[38] The Three Laws:

1. A Robot may not injure a human being, or, through inaction, allow a human being to come to harm
2. A robot must obey the orders given it by human beings except where such orders conflict with the 1st Law.
3. A robot must protect its own existence as long as such protection does not conflict with the 1st or 2nd Law.

Discussion

Can we achieve an artificial intelligence Utopia?

Artificial Intelligence is here

There are now many clear examples of artificial intelligence, both in development and in society, that suggests that artificial intelligence is possible. The increasing successes that researchers in the field of artificial intelligence are achieving also means that artificial intelligence is beginning to have a more active role in society. However, previous difficulties and failures by artificial intelligence researchers given wild speculations means we must carefully consider whether artificial intelligence is actually a possibility. Once established that it is possible, we can then examine what use we can make of artificial intelligence to create a new and hopefully better society.

Driverless cars:

The driverless car is arguably the best example of how autonomous machinery can be used to improve society. If someone were to create a fully functioning autonomous vehicle, and if that vehicle was to work efficiently on roads, then it could very quickly become widespread due to the ease of legalising and developing said vehicle. Car accidents account for 32,675[39] annual deaths in the US, and there are many more impairments and disabilities caused by accidents. Most of these accidents are caused by the driver (E.g. Drunk driving, reckless driving, slow reactions) as opposed to faults in the cars themselves (Anderson, et al., 2014). The introduction of autonomous vehicles would drastically reduce these accidents through safer driving on the part of the vehicles.

In fact, the increased investment[40] into autonomous vehicles suggests an increased awareness as to their social and economic benefits as demand and consumption for them rises. This is best demonstrated by the growing numbers of autonomous vehicles being produced, with there now being over 25 companies working on autonomous (or semi-autonomous) vehicles. The BMW next 100 model plans to provide what it calls 'the Companion'[41], an autonomous entity that 'allows the vehicle to get to know its driver better' and so assist in driving. From this we can gather that not only would autonomous vehicles seek to integrate artificial intelligence in society but also create a harmony in between humans and machines.

Research suggests consumers will be comfortable with fully autonomous vehicles and that they would also be willing to see them widely introduced. The (Accenture, 2011) survey found that around half of consumers would be comfortable using an

[39] These figures are from the National Highway Traffic Safety Administration's 2014 Fatal Analysis Reporting. They are based on figures from countrywide (US) fatalities caused by driving.
[40] Industry estimates hold that the industry could be worth $87Billion by 2020 (Accenture, 2014)
[41] This is taken from the next 100 overview given on the BMW website.
https://www.bmwgroup.com/en/next100/brandvisions.html

autonomous vehicle, mostly the younger half. This suggests that future generations should be increasingly comfortable with the use of autonomous technology.

The economic benefits of driverless cars are also crucial to consider. The benefits could mean that not only consumers will push for their integration, but also governments and firms. Major industries such as transportation and line trucking[42] could be replaced by autonomous vehicles as a means of cutting costs. General Motors experts believe driverless trucks cold allow for a 40% cost reduction to the freight transport industry[43]. This provides an incentive for production and development of autonomous vehicles, already demonstrated through the current and planned production of autonomous vehicles[44]. Provided the vehicles work as efficiently as promised[45], Google's driverless cars had 341 incident reports[46] (BBC, 2016) in which the driver was forced to take control, then they could become widespread.

What this leads to is an acceptance of artificially intelligent machinery into the daily life of a society. The integration of a technology such as autonomous vehicles is demonstrative of how we could achieve a form of utopia using intelligent machinery. It demonstrates a willingness to transfer a very large labour and economic sector (transport) to an autonomous workforce. If successful, requirement for such a transfer could then be made apparent in other labour and social sectors. This would signify a shift towards dependence on autonomous systems, a requirement for the achievement of a utopian state.

The service sector:

Driverless cars are also important as they are able to enter the service sector, an area often high in human interaction. If we were to have automated taxis, then people would suddenly find themselves in constant contact with automated machinery. In many ways people are already in constant contact with artificial intelligence. Around 1.5 billion Apple products have been sold, mostly iPhones. These contain Siri, intended to be a personal assistant, another role fulfilled by humans requiring high levels of human interaction. The high levels of customer satisfaction that Apple[47] has attained with devices containing Siri also suggests consumer comfort with the program.

A high level of comfort with intelligent programs is required for a utopian society based around artificial intelligence. The possibility of autonomous machines permeating into the service sector is an indication of how willing society is to produce

[42] The highway transportation of loads by cargo trucks – a several trillion-dollar industry

[43] This according to former GM Executive Larry Burns.

[44] 4/10 vehicles could be autonomous, according to Litman, T. 'Autonomous Vehicle Implementation Predictions'

[45] The general idea of driverless vehicles is that they are accident free

[46] However, these numbers can actually be regarded as successful given that only 13 of those incidents might have resulted in a crash, the rest were just precautionary measures by the software.

[47] All Apple devices (Devices excluding laptops and desktops) have an above 90% consumer satisfaction – (Apple Insider) - http://bit.ly/1gm2sT2

and consume artificially intelligent machines. The production is best indicated by the many private[48] enterprises focused on developing new autonomous machinery.

A particularly niche product that is exemplary of this is the device produced by Momentum Machines, their gourmet hamburger maker[49] (Ford, 2015). Vardakostas[50] stated that the machine was intended 'to completely obviate them (employees of restaurants where the machine would be employed)'[51]. Here we can begin to see the mixed impact that such products would have on society. They would be willingly integrated because it would nearly eliminate the $9-billion-dollar labour cost for burger production (Momentum Machines, 2012), which would allow for the production of 'Gourmet'[52] burgers by the machine for fast food prices. This would increase social welfare in two ways; the first is the use of high quality ingredients, and the second is the cheap availability of these burgers. The high quality ingredients reduce the weight gain incurred globally by fast food burger producers such as McDonalds, improving social health, largely of lower income consumers of fast food. The cheap availability of such burgers means lower income classes have access to higher quality, tastier food, increasing happiness.

Yet such products are intended, as stated by Vardakostas, to put millions into unemployment. The same applies with autonomous vehicles, the use of such vehicles for trucking and taxi services would also put many thousands[53] into unemployment. This would not merely be a temporary unemployment, but rather the displacement of a whole skill set from work. Unless drivers were willing and able to learn a new skillset, they would be unable to find new employment due to the saturation of the market for drivers by autonomous vehicles. Similarly, fast food employees[54] are placed under threat by products such as Momentum Machines. Both of these major service industries could be displaced. This would cause drastic inequality in a society, suggesting a move towards a dystopian rather than utopian state.

'Blue-collar' labour:

Similar problems arise when considering what is considered the 'blue-collar' labour sector. These are often assembly line, low-skilled labour jobs. This is one of the biggest global employment sectors with 714,704 (Thousand) employed globally[55] (International Labour Organisation, 2014). The creation of an automated labour force

[48] In actuality, many of these private entities are now owned by umbrella corporations such as Alphabet or Apple.
[49] This is a 'gourmet' burger making machine capable of producing a burger every 10 seconds. Not only that, but the machine can custom make each burger, including how the meat is cooked, what meat it uses, toppings, bread toasted etc.
[50] Vardakostsas is the co-founder of Momentum Machines.
[51] Quoted from 'Hamburgers, Coffee, Guitars, and Cars: A report from Lemnos Labs' – Wade Roush
[52] Gourmet in this sense as being high-quality, fresh ingredients and well-cooked
[53] There are 1,005,000 taxi drivers in UK employment as of 2016 - http://bit.ly/1PJV5tb (National Careers Service)
[54] McDonalds alone employs 1.8 million workers globally (Ford, 2015)
[55] This according to the International Labour Organisation, a United Nations agency aimed at establishing labour standards and collating global labour information - http://bit.ly/1lpbOHU

could potentially displace a majority proportion of the world's workforce from their jobs.

Such an automated labour force would not be difficult to design. We already have large industrial processing machines in use in the automated assembly lines, and new machines such as the Kiva industries packing and Delta's Assembly robot. The automation that has already occurred has done economic wonders, an example being the U.S. textile industry. The industry lost around ¾ of domestic employment in textiles between 1990 and 2012. However, U.S. textile exports recently surged by a $27billion value, largely thanks to automation (Ford, 2015). Only highly efficient automated technology can compete with low-wage workers in highly industrial countries such as China. This allows nations like the U.S. to fully automate their low income manufacturing classes. The low 10% employment in the sector (International Labour Organisation, 2014) also means that it will have less of an impact on the workforce.

This would likely even be beneficial for the 10%. Although automation would mean an elimination of their jobs, it would allow for decreased price of goods for a substantial decrease in the cost of production. This in turn decreases the costs of living, giving consumers an increased expendable income that they might spend on luxury goods or the service sector, expanding the demand for labour in sectors other than assembly line workforce. Thus the 10% could be absorbed into the middle class through higher paying jobs and decreased living costs, effectively eliminating lower classes.

This sort of labour migration is not uncommon. The industrial revolution had a similar impact on the British economy, as the effective automation of many services put thousands out of work. History shows that those thousands managed to survive, and eventually the economy provided new labour sectors for them to join. From this we can see how a forced labour migration is not necessarily a negative thing. We can also see how the automation of a 'blue-collar' workforce is possible, and in many ways could be beneficial to a society.

What constitutes a Utopian

Traditional ideal of Utopia:

The traditional definition of Utopia is 'Involving, based or founded on, imaginary or chimerical perfection; impossibly ideal, visionary' (Oxford English Dictionary, n.d.), and the inherent nature of a Utopia appears to be the idea that it is impossible to achieve. The traditional ideal is intended to be impossible to achieve, and given the evidence for artificially intelligent technology being developed that could in fact allow us to attain such a state, this notion is useless to us. Thomas More, who originally coined the term, describes it as a 'fruitful and pleasant work of the best state of public well' (More, 1516). The idea of fruitfulness and work suggests a highly productive human society. Yet this again is inadequate when applied to an artificial intelligence utopia given the notion of productivity. We can assume that in a perfect machine intelligence Utopia, work would be optional for humans given the

intelligence of machines to produce what is required. Thus if anything we would be creating a society where no one really does anything, and humanity would reach its lowest level of productivity in years.

Yet perhaps Utopia as a place in the traditional sense is not what we would achieve, but rather a utopian society that can only feasibly be achieved through a combination of human ingenuity in machine intelligence and then the use of that intelligence to support said utopian society. These conditions are impossible given our current social and economic world structure. Resources are arguably too scarce in our world, preventing full equality through uneven distribution. The similar requirement for a labour force to be employed as cheaply[56] as possible also means that inequality is inevitable in our current economic productive system. The creation of artificial intelligence would give us a vastly superior labour force that we could then employ for the betterment of society through cheap and efficient production.

An artificial intelligence Utopia:

Given that the original notion of Utopia was defined long before artificial intelligence could even be imagined, I would argue that an artificial intelligence utopia must therefore be different to a standard Utopia, given the contributions artificial intelligence can make to a society. Thus the main aspects of a traditional Utopia relevant to this discussion are the aspects of perfection and welfare state that a utopia implies.

A new term we could apply to a utopia accounting for artificial intelligence could be:

'A state of social and financial equality of every human individual, in which they are at liberty to conduct themselves as they please, within legal restrictions.'

We can assume, given the possibilities that artificially intelligent machines could grant, that we could very likely achieve a utopia in this form. As previously discussed, artificial intelligence could likely pervade every aspect of society, and as such replace humanity as a whole. However, the restriction of artificial intelligence to use as a tool would effectively allow us to create a whole new artificial human race to provide for the actual human race, given their lack of necessity for food and rest and other human shortcomings when it comes to labour. Thus we could remove a human requirement for labour entirely, provided we created machines to run other machines and to repair each other. Arguably only this would allow a wholly equal social state given that most social and class divides can be traced to an economic basis. Also, every person would in theory be financially equal given that the removal of job

[56] In a capitalist, profit driven market, lower tiers of labour are employed cheaply due to low skill requirements for the jobs, as well as high supply of labour in large economies. Higher wages would mean lower revenue for firms.

differences would eliminate income classes. Rather wealth[57] would ideally be equally distributed by an incorruptible intelligence entity[58].

Such a state would of course require an almost complete "submission" of human individuals to intelligent programs. In order to achieve equality, all individuals within the state would have to subject themselves to said incorruptible entity to accurately allot wealth. The wealth would be gained from the effective production of independent machines working to production measures set by popular demand from the individuals. Similarly, in order to ensure complete equality there must presumably a non-corruptible governing body that would effectively be above the rest of humanity. If any human person were allowed to do this, they would be breaching the set state of equality and it would therefore have to be an autonomous intelligence. Only such an intelligence could effectively lead to a society without any prejudice, and also presumably without being able to be corrupted into dividing the products of the society unfairly.

Synergy – Can robots and humans co-exist:

This begs the question of co-existence between humans and robots. Given the apparent requirement to be utterly dependant on and trustful of artificially intelligent machines, humanity would have to co-exist with artificially intelligent entities. The only method of achieving said utopia would be if there could be some level of co-operation and co-existence between humans and artificially intelligent machines.

Co-operation is a difficult term to use as a description, as it allows for so many different levels of interpretation. Is it simple co-operation if a human buys something from a machine seller, or is that just co-existence. Would co-operation have to be a more complex activity, such as a machine and a researcher working in unison through a project, and their joint ideas form the end product?

Many problems would arise if such an extent of co-operation were to be necessary. What would we do if some researchers are happy to work with artificial intelligence, but others refuse and thus hamper research progress. An example would be the job encroachment currently caused by the use of artificial intelligence in fields such as medicine[59] with IBM's Watson. Several news agencies[60] reported (in 2014) of the astounding 12 million annual misdiagnosis (US average). Employment of machines such as Watson would reduce errors, and most importantly allow for Doctors to have more time with patients administering treatment. However, all this change would be predicated on doctors' acceptance of artificially intelligent machines, and willingness to rely on the judgement they give.

[57] Ideally, there would be no real wealth, but rather an allotment of scarce resources to individuals, the value of which could possibly be based upon production costs.

[58] This would have to be a machine entity. It is said absolute power corrupts absolutely, and so a human cannot fulfil such a role due to their susceptibility to corruption, as well as inability to meet the demanding requirements of that role.

[59] IBM's Watson is used in this field. It analyses medical data available to form an evidenced based patient diagnosis.

[60] Including; Reuters, Forbes and CBS for US statistics

If we consider societies most often explored in science fiction novels such as the work of Isaac Asimov, it gives us an indication of how humans could possibly interact with machines.

Sustainability

Can we overcome mistrust of intelligent machinery?

Such machine interaction is also essential to the function of an artificial intelligence utopia. Humanity would have to be comfortable with the regular interaction with robot entities.

AI permeation into everything:

A forced labour migration might not necessarily be a negative, but it is certainly a requirement. If we successfully replace “predictable” and repetitive assembly line labour with autonomous machines, it would mean the elimination of a massive job market. This would be alright provided workers could migrate into higher paid jobs, but not if those other jobs were also taken by autonomous entities.

The current development of autonomous machinery suggests that even highly skilled ‘white collar’ jobs could be replaced with machines. This would mean a saturation across the whole of the labour market, rather than in just specific sectors.

Such a saturation would be unsustainable. Not only would it cause widespread unemployment and so extreme inequality, but even those who profit from such automation could suffer as they no longer have any consumers for their production. A permeation of artificial intelligence into the labour market would have to be very carefully controlled in order to prevent an economic collapse. With such control we could move products into a form of public distribution system I discussed under the notion of an artificial intelligence utopia. However, a failure to control the movement could lead us towards dystopia.

Psychological impacts on humanity

Superficiality:

Perhaps the greatest issue in sustaining such a utopia is the psychological impact that it would have on humans. Humans arguably require some purpose to keep us going[61]. If we effectively create a society in which humans are no longer actually necessary aside from consuming goods produced by machines, what would be the possible psychological impact on society?

We would technically be reduced to a wholly superficial society. This implies an absence of anything deeper to the society, anything that we could consider to make us human.

[61] I am unwilling to delve into mental illness, as that would be a whole field in itself. However, I will add that there are many new considerations of mental illness as an epigenetic disorder rather than environmental. Your environment and upbringing might trigger a genetic disorder, but is not usually the sole cause. This means that artificial intelligence might not have as great an impact on mental illness as I suggest - http://bit.ly/1pQ5TAS

Purpose is arguably quintessential to human life. Without purpose there really is little meaning to life. Yet such a utopia would essentially rob you of any purpose, aside from controlling the machines to give you your desires. Self-indulgence does not seem like an effective basis for a society. What you would end up creating was a wholly selfish and useless society, who would care nothing for self-improvement given the lack of necessity for such a thing. It is no small thing to make humanity redundant, but there is nothing like artificial intelligence for achieving this, and indeed must achieve this in order to create utopia.

Could this actually improve mental health?

Recent medical studies into the root of mental illness have suggested it lies somewhere in our genetic make-up (Weir, 2012). This means that the impact caused by major social changes might not actually have a negative effect on human mental health, and so can be disregarded as a problem in achieving a utopian state.

If anything, such a state would improve mental health. Anxiety, stress and social pressures are the leading environmental causes of mental illness (Animals in Research, n.d.). Creating a work-free, equal society would considerably reduce these factors, perhaps eliminate them given the removal of job and social pressures, and leaving the sole cause of disorders as genetics.

Decline into Dystopia

Despite the possibility of improving mental health, there is still a contradiction in human nature required to achieve utopia, as humanity must make themselves redundant. But if humanity feels redundant, they can never truly achieve the psychological happiness required for it to be a utopia. Thus humanity would become trapped in an undesirable situation were we to reach such a state, and therefore become a dystopian state.

Zeroth Law:

The Zeroth law is likely one of the greatest difficulties in creating a utopia through artificial intelligence, as it would cause the most unpredictable difficulty. The Zeroth is a formulation that intelligent robots could arrive at. It is derived from Isaac Asimov's stories, and Asimov's application of the three laws to all robots. It is a law that arises from Asimov's first law that prevents robots from allowing humans to come to harm. However, there is an issue when considering how self-harming humanity can be to itself. Current examples such as pollution and war can be considered, according to the Zeroth Law, as harm caused to humans. Thus the logical process to a machine therefore becomes that humans are harming other humans.

What this means is that you could potentially end up with a situation similar to that depicted in the Terminator film franchise. Say a machine decides that humanity must be protected from itself. It would excuse its harming of humans as beneficial to humanity through a long-term consideration of harm rather than short term. In fact, you might even have machines controlling humans in order to allow no harm whatsoever. By control I mean potentially guiding their actions so that they take no

actions to cause them harm. Perhaps even a containment of humanity in order to better safeguard them.

Most likely you have a situation like in (2001: A Space Odyssey, 1968). This is a machine too heavily prioritizing its program instructions over other equally important factors, such as human life. Professor Bostrom suggests an even simpler form of Zeroth. Imagine you program a machine to produce paperclips (Bostrom, 2014). You fail to account for the efficiency of the machine, and before you can prevent it, the machine has produced millions of unwanted paperclips, wasting tons of resources. Such problems seem negligent, but say this were to occur in many different instances simultaneously, you would suddenly be faced with a massive waste of resources and suffer the consequences of those shortages.

The most important point to consider when considering the problems caused by the Zeroth law is the inability to program against it. Even more computer programs have bugs, instances where they work ineffectively. It is much more difficult to find bugs and program flaws in the complex software involved in artificial intelligence. The whole point of the Zeroth Law is that it arises from an unpredictable interpretation of a program. Given the nature of artificial intelligence as unpredictable, it means it is impossible to safeguard against it.

Military Development:

However, the Zeroth law motion that a robot overlord could eliminate humanity seems an unlikely occurrence. Skynet has long been dismissed as an impossible science fiction fantasy, given that we could easily reprogram machines that function against humans. Also, in a utopian society humanity would presumably be in control, power being a desirable function. Thus it seems unlikely for a society to ever hand over complete economic, and to an extent social, control to a machine as occurs both in the Terminator franchise and I, Robot. If a human is in control they could prevent any such occurrences.

However, any misinterpretation of instructions could potentially be severe, given the likely capabilities artificial intelligence would have. Much of the current funding for artificial intelligence is given by institutions such as DARPA (Defence Advanced Research Projects Agency), a branch of the US military. Even academic and financial institutions help develop technology for use in modern combat, such as Boston Dynamics development of BigDog[62]. Thus there is a definite movement in the development of artificial intelligence towards military machines.

A utopian society is peaceful. It seems like a contradiction if the society is both peaceful, yet generates and has available a large number of military robots. That almost seems to beg the temptation of a military coup or revolution at the first instance of dissatisfaction. Also, it allows for different sectors of AI development to occur. The creation of a military driven sector of artificial intelligence means there is less potential for development in other sectors given a lower allocation of funds. Not

[62] BigDog is a military support unit developed for field use. Although eventually discontinued due to noise, and despite not being an offensive machine, it demonstrated the time and money invested into development of military machines.

only that, but there is much opposition to the use of automated vehicles such as drones in war. The further development of autonomous war machines could lead to public outcry against development, and the scrutiny could very easily be turned to artificial intelligence as a whole, hindering development.

Ethics

The ethical difficulty of artificial intelligence:

One of the difficulties achieving complete artificial intelligence would cause is the status of said intelligent machinery in our society. There is some conflict as to whether a machine that would technically have human understanding and intelligence would therefore be counted as human within the extent of social status and morality.

Let us consider morality first, as we can then apply morality to any formation of social status. Most morality is concerned with interactions between sapient persons within a society in order to achieve the best state of welfare through a set of laws that can be applied universally. True, the notion of universality is not constant to all the moral laws, but given the requirements of a Utopia as a societal best state of welfare, we can consider society to account for all persons within that society, and thus a moral law must too.

The difficulty that arises is if a seemingly sentient machine can ever be considered a person, and as such be deserving of equal rights and ethics being applied to them as are applied to humans, should they receive such rights? For all intents and purposes, machines would be built primarily to be used as tools. Therefore, in granting them rights, we are reducing their usefulness to us as rights would grant them certain freedoms that would make them less efficient. An example might be that a machine could develop preferences or hobbies, and would require time off work for itself to pursue those. This could prove detrimental to the utopian nature of the society, as it would then create a divisive society in between machines and their human makers.

How to solve the ethical situation:

The requirement to resolve the situation would therefore be a necessity to build any artificially intelligent machines that are to a large extent dumb. This however, creates a whole new slew of ethical issues, given that we would effectively be removing or preventing potentially intelligent beings from having intelligence. This would also mean that we are socially excluding certain potential social groups (Intelligent, conscious machines), and thus undoing that which makes a Utopia.

Conclusion:

In conclusion, it seems apparent that we could, given the current developments being made in the field of artificial intelligence, achieve a state resembling a utopian society. The current rate and acceleration of developments all suggest a general movement towards more effective, more 'intelligent' software which would allow us to build machinery capable of an economic revolution, and through that process a social revolution.

However, the clear requirements of equality and lack of a corrupting entity, and many of the predicates upon which such a society would be based appear contradictory in nature to human behaviour and social interactions of today. Thus, there would need to be a psychological upheaval and a radical change of human social interaction in order to allow such a society to occur. Although this could arguably occur over time, at the pace with which the technological revolution might occur, the apparent inability of changes in basic human social instinct through history suggest that so radical a change in human nature is likely impossible. From the time of the earliest civilisations of the Babylonians and the Egyptians, there has been evidence of social inequality and instability. Slavery was a staple of human civilisation until the 19th century. It seems that human society is thus built upon layers of inequality and social differences which would be difficult to undo in order to create a totally egalitarian state required of a utopia. The world has seen unprecedented growth and accumulation of wealth as it crossed into the 21st century, yet inequality is still a major social and global issue. Wealth has accumulated into the hands of a small percentage of the global population, and there is little reason to think that the same technology that would help us achieve a utopian state would not also be controlled by them.

Furthermore, the difficulty in maintaining the stability of such a society would present even further difficulty in our achievement of a utopia. It seems illogical to consider achievement of a utopian state as temporary, but should rather be understood to be as permanent a state as possible. The myriad ways in which such a state might be undermined as presented in my discussion such as psychological difficulties caused by a superficial state and technical problems that would be caused the Zeroth Law as well as the difficulties in maintaining it both in long and short term imply that it is by nature unsustainable.

Perhaps, given plenty of time and upheaval, such a state might be achieved. This would require technology which contemporary evidence of failures in promised technological achievements with artificial intelligence suggests is unobtainable, even with breakthroughs occurring as they are now. Also a contemporary view of human nature and social interaction would suggest that an artificial intelligence utopia is impossible. At best, we might hope to achieve a capitalist society in which humans still govern, but there is an elimination of the greater proportion of the lower-labour intensive classes, thus leading to a more equal and productive society. The greatest utopia humans could achieve might simply be a world in which there is greater wealth from machine labour and production, so that those in an economically poor state might be rich by comparison to those who are poor today.

Evaluation

Having completed my dissertation, I have very mixed feelings about it. I am satisfied with the final project, having completed extensive pieces of writing both for the discussion and other aspects of my dissertation. However, I worry that my difficulties with keeping to guidelines and the delay in producing my work meant that it is often not as coherent or as high quality as it ought to be.

As for the project topic itself, I am very satisfied with my research into it and the areas I discussed on it. The area of artificial intelligence has a lot of research and writing, so if anything sources were too easy to find, so much that there was an almost insurmountable amount of material on the topic. I learnt to filter out the relevant and important sources, as well as what made for an interesting discussion point.

I felt my content was effective, mostly where there was substantial research for it. This concerned area of current developments and impacts, as well as how past successes in the field could demonstrate future success. However, I am slightly less confident in the more speculative aspects of my discussion. Although I enjoyed writing those areas of discussion more, as it was there where my own thinking largely came through, they were largely unsubstantiated. This means they can be considered less seriously than other aspects of my discussion as there is less evidence supporting them, due to the little speculation that has been made into the topic. I did feel that the topic was worth discussing. There appears to be hesitation in the scientific community to speculate to wildly as to where artificial intelligence could take us, especially following the events of the ‘artificial intelligence winter’ in the late 20th century.

With more time, I would however be more willing to analyse the current impacts artificially intelligent products our having on society, both economically and on a more personal level. There is little research done into consumer reactions to AI products, or sentiment about development of such technology. With the technology progressing as the rate it is, it would be interesting to see how society reacts to it.

Unfortunately, the field of artificial intelligence is so vast, and it incorporates so many aspects of technological and social developments that it would be impossible to fully evaluate where we might go with it. There are so many developments in such a wide range of fields that you would have to be an immense polymath to study them all.

Bibliography

2001: A Space Odyssey. 1968. [Film] Directed by Stanley Kubrick. s.l.: Metro Goldwyn Mayer.

Accenture, 2011. *Consumers in US and UK,* s.l.: Accenture.

Accenture, 2014. *Realising the benefits of autonomous vehicles,* s.l.: Accenture.

Anderson, J. M. et al., 2014. Autonomous Vehicle Technology: A Guide for Policymakers. *Rand Corporation.*

Animals in Research, n.d. *Understanding the root cause of anxiety and depression,* s.l.: s.n.

Apple Inc., n.d. *Apple.* [Online]
Available at: http://www.apple.com/uk/ios/siri/

Asimov, I., 1950. *I, Robot.* s.l.:Gnome Press.

Asimov, I., 1953. *The Caves of Steel.* s.l.:Doubleday.

BBC, 2014. Computer AI passes Turing test in 'world first'. *BBC.*

BBC, 2016. Google's self-drive cars had to be stopped from crashing. *BBC.*

Bostrom, N., 2006. *CNN on Artificial Intelligence,* s.l.: CNN.

Bostrom, N., 2014. *SuperIntelligence: Paths, Dangers, Strategies.* Oxford: Oxford University Press.

Dodd, M., Grant, A. & Serugawi, L., n.d. *Artificial Intelligence through the eyes of the public,* s.l.: Worcester Polytechnic Institute.

Economist, n.d.

Encyclopaedia Britannica, n.d. *Encyclopaedia Britannica.* [Online]
Available at: www.britannica.com

Encyclopedia Britannica, n.d. *www.britannica.com.* [Online].

Encylopaedia Britannica, n.d. *Encyclopaedia Britannica.* [Online]
Available at: www.britannica.com

Ex Machina. 2015. [Film] s.l.: s.n.

Financial Times, n.d.

Ford, M., 2015. *Rise of the Robots: Technology and the Threat of a Jobless Future.* s.l.:Basic Books.

Gartner, 2016. *Gartner Says Worldwide PC Shipments Declined 8.3 Percent in Fourth Quarter of 2015.* s.l.:s.n.

Intel, 2015. *Intel.* [Online]
Available at: http://www.intel.com/content/www/us/en/silicon-innovations/moores-law-technology.html

International Labour Organisation, 2014. *Global Employment Trends 2014: supporting data sets,* s.l.: International Labour Organisation.

John, M., Minsky, M. L., N., R. & Bell, S., 1955. *A PROPOSAL FOR THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE.* s.l., s.n.

McCarthy, J., 2007. *What is Artificial Intelligence?.* [Online].

McDermott, D., 1997. How Intelligent is Deep Blue. *NY Times.*

Microsoft, n.d. *Windows Microsoft.* [Online]
Available at: http://windows.microsoft.com/en-gb/windows-10/getstarted-what-is-cortana

Momentum Machines, 2012. *Momentum Machines.* [Online]
Available at: http://momentummachines.com

More, T., 1516. *Utopia.* s.l.:s.n.

Nilsson, N., 2009 . *Quote from Nils Nilsson.* s.l.:s.n.

Office of National Statistics, n.d. [Online]
Available at: www.ons.gov.uk

Oxford English Dictionary, n.d. *Oxford English Dictionary.* Oxford: Oxford Press.

Reuters , n.d.

Rotman, D., 2013. How Technology Is Destroying Jobs. *MIT Technology Review.*

Rotman, D., 2014. Technology and Inequality. *MIT Technology Review.*

Rotman, D., 2015. Who will own the Robots?. *MIT Technological Review*.

Roush, W., 2012. In: *Hamburgers, Coffee, Guitars and Cars: A report from Lemnos Labs.* s.l.:Xconomy.

Rowan, D., 2015. *Wired.* [Online]
Available at: http://www.wired.co.uk/magazine/archive/2015/07/features/deepmind

S, H. & B, H., 2003. *Introducing Artificial Intelligence.* s.l.:Icon Books.

Shu, C., 2014. *TechCrunch.* [Online]
Available at: http://techcrunch.com/2014/01/26/google-deepmind/

Statistic Brain, 2015. *Apple Computer Company Statistics.* s.l.:s.n.

Szondy, D., 2012. *Gizmag.com.* [Online]
Available at: http://www.gizmag.com/hamburger-machine/25159/

Terminator. 1984. [Film] Directed by James Cameron. s.l.: Cinema 84 Pacific Western.

Thomson Reuters, 2015. *Learning from data: building brains and using them,* s.l.: Thomson Reuters.

Traiger, S., 2000. *Making the right indetification in the Turing Test.* s.l.:s.n.

Weir, K., 2012. The roots of mental illness. *American PSychological Association,* Volume 43.

Zuckerberg, M., 2016. *Facebook Post - I want to make Jarvis by 2017.* Palo Alto: s.n.