Midterm

Manoj Vasa

March 6, 2017

PHI 3370-001

Computers, Ethics & Society

1. Artificial Intelligence, it is claimed, has its foundation in the Scientific Modern conception of mind. That is, given the mechanical view of mind argued for, in albeit different ways, of Descartes, and Hobbes, it is claimed that thinking being like computation, offers an appropriate model for claiming non-human devices can think, in a similar fashion as human beings. Further supporting this claim is the idea advanced by David Hume, that when we consider the processes of thinking it is not necessary to make any claims about the nature of the entity that is thinking. The question to be asked here is: is such a model of thinking sufficient to claim the reality of artificial intelligence? In your answer you may wish to draw on the relevance, or irrelevance of the Turing Test, which claims that singular feature of thinking, is the use of language. That is, some would claim, that there is an isomorphism between human thinking and machine thinking.

A machine has the incredible power to make extensive calculations that seem rather very difficult for a human being to do in just that amount of time. A machine may exhibit characteristics of human conversation, but its process of expression is quite unique from that of a human being. Turing addresses several objections in his paper on the particular topic. Some of these objections do portray very convincing arguments against the ability of machines to think just as humans. Something that is unique to computer is their entire existence and calculations are determined beforehand rather than the spontaneous unpredictability that humans express. In other words, computers are given directions for their actions that determine their behaviors. Such arguments convince me that machines do not think.

The Turing test has portrayed an interesting situation that truly questions the ability of machines to think. Such a test leaves the interrogator in inquisition as to which responses are coming from a human and which from a computer. As Turing addressed in his objections, however, the mathematical probability of the responses of the machine to be appropriate does define if the machine is thinking or not. This probability expresses the concern that the machine may respond incorrectly or not respond at all given particular questions. Given a question the machine may match keywords and not be able to make sense of the meaning of the sentence before replying back. Just the use of this language clearly doesn’t imply that the machine is thinking and understands the contents of the conversation. John Searle proposes the Chinese-room argument to counter the proposition that machines can think like humans and understand these conversations. With this argument, it can be conveyed that the machine just matches keywords and phrases to reply back, rather than actually understanding what it’s expressing. Through this argument, I propose that a machine doesn’t understand and think, but performs algorithms to match receiving words and phrases with responses. Such algorithms are clearly defined by humans and the machine may not deviate from it to express any understanding of the conversation.

Such a program that provides as a prime example to this set of algorithms is the ELIZA therapeutic software. This program was designed as a therapeutic conversation program that takes human responses and stores this data in memory to match with its own keywords and phrases to develop an output appropriate to the user’s statement. This is an algorithm. Such an algorithm was defined and constructed by a human being. The machine itself doesn’t consciously perform these actions but it only follows a set of given directions. A script given to the computer determines what questions to ask and what keywords to store. Machines only do what they are programmed to do. Just as how a human can translate a foreign word from that language to English with a dictionary without understanding anything, computers exist and converse without meaning. The algorithms are defined by humans and the responses are nothing more than expected.

Nonetheless, computers don’t convey a process of thought and understanding. They just follow directions and algorithms defined by humans. From this perspective, computers may not be able to respond to unexpected input or input that is out of their scope. Their instructions and predetermined algorithms are what determine computers. Through this argument I deny the reality of AI. The “thinking” done by machines is not like the thinking done by humans. The process of AI follows a set of algorithms defined by humans and responds nonsensically when given unexpected input. These outputs by AI provide the results of algorithms and instructions defined for their specific intended purpose. When stepping out of this scope of knowledge the AI begins responding with incoherent output. Nonetheless, they keep following their predetermined list of instructions without deviation.

2. What might you see as an overriding ethical issue associated with the possible further development of Artificial Intelligence? That is, is there something unique and different about Artificial Intelligence which may demand that new ethical frameworks need be developed? You may wish to consider the use of Artificial Intelligence in terms of military usage, or economic management on a global scale.

The question of developing and advancing artificial intelligence concerns the idea of rationality and morality. The implementation of an AI with self-learning capabilities will advance its own knowledge and develop new pathways of rationality within its own reasoning. This may provide some conflict. Such conflicts involve the extent of the AI’s rationality and the defined boundaries of its own morality. Such questions may raise concern as to what the true capabilities of an AI might be and what they may learn.

AI has the ability to self-learn and apply this newly learned information. However, what it learns is essentially dependent on human actions, thoughts, and other information available to it. This information varies drastically from one sector to another. Conflicting opinions expressed by humans both in words and text may contradict the existing knowledge the AI is learning from. This severely questions if the AI is learning rational or irrational thought that may be expressed in a multitude of domains. The AI has to learn before being able to rationalize with humans and perform economic functions. The economics of the world are performed through very extensive and careful calculation. This feature is simple for the AI to grasp. However, the political aspect of these economic interactions may lead to ambiguous results. This entirely depends on what the AI learns and concludes as useful information that defines rational thought. The information available to an AI is what makes it a primary threat. The interactions between humans may be too complex as the ideals and moral expressed by humans conflict and contradict each other. One ethical framework concerning this discussion is that the AI should be given the information that it is meant to learn and use to perform calculations and interact with other systems rationally. The information revealed to an AI can drastically affect its direction in the crucial decisions it has to take in the future.

These decisions are very critical and sensitive when referenced to in the military domain of technology. A military technology programmed to do a set of tasks with clearly just mechanical processes without the addition of thought can perform an action without a problem. However, a self-learning AI has to think and evaluate its decisions in a battlefield. The rationality and morality of an AI is heavily dependent on the ideals and morals it is exposed to previously. The actions following these morals may be drastically different due to lack of emotion. Therefore, AI in military shouldn’t be exposed to lethal operations, but be used in calculations and processing leading to the formal executions done by humans. Trusting an AI to follow the morality and rationality of humans is a problematic statement. This is due to the fact that the idea, morals, and knowledge that is available to AI comes primarily from human knowledge and actions. Military technology as of now that promote drones to jam communication signals and target enemies in a war zone is already a dangerous application. Reconnaissance and signal interference is acceptable, but targeting by drones is somewhat questionable. These targeting systems may target civilians and fire without question. The morbid possibility of AI to perform such decisions in a war zone may lead to civilian casualties that the AI may reason to be perfectly normal.

Nonetheless, this is a question of rationality and morality. The actions of an AI really depend on what the AI learns and comprehends to be common human knowledge. The calculation of numbers by an AI is a trivial matter as opposed to the calculation of emotions. With the vast amount of knowledge available to AI that contradicts many other pieces of information available, it seems like a questioning matter to consider how an AI would react to human spontaneity. The AI learns from humans and sources online before creating its own rational line of thinking and morality. This is concerning given the drastic variation of human thought that exists online and in the real world.

3. It can be claimed that there are three ethical central views which have developed in concert with the scientific world view; they are: Contractualism, Utilitarianism, and Deontology, or Kantianism. In addition, there is the ancient view of Virtue Ethics which may be of interest as well. In your essay you may wish to reference the critique or support of the idea that the “new-wine” of computer information problems, especially in regard to Artificial Intelligence, is problematic.

The ethical view of Utilitarianism follows the idea of the greater good for the greater number. This ethical framework works very well within society. The greater the number of people that are happy, the less likely a rebellion will occur. In this society the action is dependent on what maximizes utility of the society. Such a utility is the well-being of the people and animals in the community. This maximization of utility assists in only executing actions and decisions that have the most positive effect on the greatest number of people. This way a negative minority won’t have an undesirable influence after every decision made.

A decision that can take place is the development of the economic organization of a society. What brings about the most happiness is what the community favors as a whole. Such a decision can be low taxes or higher wages for the people. Usually these positive actions provide happiness rather than the distaste that would spread with the nefarious act of raising taxes. However, to keep the whole society content with the architectural and recreational development of the society, some demands must be set. Even with such demands as taxes and raised prices to account for inflation, a community may benefit from the developments and services offered from their community with this change. The increased costs allow for the state or the country to offer more services that assist in bettering the conditions for all people of the state or the country. Utilitarianism offers a way to raise demands and provide benefits to maximize the utility of well-being through a whole society. Such an ethical framework is also only possible given the moral responsibility of the community as a whole. This moral responsibility prevents ominous acts to be carried out and preserves the well-being of the society. This moral compass is what an AI might find some trouble with.

An AI’s calculation of the utility of well-being may differ from that of human beings. This rather vaguely defined factor of Utilitarianism may conflict with the processing of an AI. For example, the process to improving the quality of life of many citizens by making government services more available may require drastic changes in the economic infrastructure. However, an AI may take many more drastic measures that may lead to a desired end result in terms of happiness, but the path there seems to be troubling and conflicting with the society as a whole. Such measure may include drastic raise in taxes that would help the government to raise budgets for better causes in the long term, but leaves citizens poor and unhappy in the short term. Another scenario may be the directing of transit routes. The optimal route to a destination calculated by a human may be very different than the one calculated by an AI. This approaches the travelling salesman problem. There exist numerous possible routes, but the criteria an AI uses may enlist an unfavorable route to the majority of the commuters. The AI might choose such a route that is shortest in terms of whole bus route but may reach each destination in an inefficient manner so it may load too many people in the bus and have people wait longer than they need to in order to reach their destination. The question resides in how the AI perceive the happiness of the greater number.

A very crucial application of AI is its involvement in the military. Enlisting a truly self-learning and aware AI in the military may provide for some problems if it was designed to use this ethical approach. The AI may just serve for the well-being of the citizens of the country it was deployed from and indiscriminately carry orders and execute collateral damage only to serve for the happiness of its home. These actions are rational to the AI that was designed with the ethical framework of utilitarianism. These drastic, violent actions are rational to the AI since it maximizes the utility of well-being for the people of its country. Through this logic the AI could keep carrying out such orders and actions unquestionably to keep its people happy and safe. This clearly is a conflicting demonstration. Such an action should be avoided due to its grave consequences. The AI could serve to be very problematic in the field because of this ethical framework. The morality of its actions may be skewed due to ensuring the happiness of one group but disregarding that of another group of people.