

Ruiyang Ma

Dr. Berleant

IFSC 2200

10 December 2020

Computer Decision-making Under Ethical Dilemmas

On the night of March 18, 2018, a woman crossing a road in Arizona was hit by an Uber self-driving car and died. An unmanned car is a smart car that senses the road environment through the on-board sensor system, and the computer system automatically plans the driving route and controls the vehicle to reach the predetermined goal. Perhaps one day in the future, computers will not only be used as tools for calculation or negotiation but become mature decision-makers on important issues related to human well-being. People want to focus on decision-making issues. In this field, computers are most likely to affect and control people's lives. With the development of unmanned driving technology, facing the dilemma of computer ethics, it is necessary to consider whether the computer should make a decision.

With the continuous attention to unmanned driving technology, it may subvert the existing traffic mode in the future and bring great convenience to human travel. When the computer takes over the steering wheel from humans and faces the inevitable collision, ethical dilemmas may arise. This dilemma has caused people to rethink whether computers should make decisions. In a narrow sense, "decision making" is just an act of choice. The decision-making process is ignored. For example, a person is asked to choose any card in a card trick, so simply choosing a card constitutes a decision. Broadly speaking, "decision making" not only refers to decision-making but also process. Therefore, in a broad sense, making a decision may involve investigating a possible course of action, evaluating alternative strategies, and

choosing a course of action based on this investigation and evaluation (William 2018).

How should computer decision-making adopt ethical choices to resolve the conflict between ethics and law? We can analyze from the following aspects: the legitimacy of computer decision-making and the responsibility of computer decision-making; the level of computer decision-making ability; how to judge the ability of computer decision-making; the delegation of decision-making power to the computer and the ethical dilemmas that follow. In the reality that the decision-making ability of computers cannot exceed that of human beings, computer ethics relies on the cornerstone of human ethics. However, the existing human ethics has no good solutions to the unavoidable dilemma of collisions, and computer decision-making is temporarily unable to get out of the ethical dilemma.

Is the computer a decision-maker?

This issue has always been controversial. Opponents believe that computers may make narrow decisions, but they cannot make broad decisions. The computer is not a machine that can study and evaluate alternative strategies to select the course of action. In other words, computers can only make simple decisions like flipping a coin.

Moor (217-229) thinks this statement is wrong, and perhaps the computer can only do some simple calculations at first. However, computer activities can be used as a complex information analysis in many cases, leading to the process of selection behavior. In these cases, decision language is usually a very natural application.

For example, A. L. Samuel's classic checkers program. The computer uses the Samuel program not only to play checkers but also to improve its game through experience. The computer is considered a checkers player and is described as a decision-maker. When it's its turn, the computer must decide what action to take. And

to play checkers better, the computer must make decisions based on a complex decision-making process. As Samuel pointed out, there is currently no known algorithm that can guarantee victory in checkers, and a complete exploration of every possible path in a board game involves about 10^{40} move choices. If three choices can be made every millisecond, it still takes 10^{21} centuries to consider. The decision of a computer is different from that of a human. Because a person will take a few steps forward, the evaluation may produce a wide position, but the computer will directly evaluate the wide position. The computer calculates a broad position based on a polynomial, and each polynomial represents a parameter of the game. Through playing many games, the computer sometimes treats itself as an opponent and learns which parameters are important. By changing the weight of the parameters and trying different parameters, the evaluation mechanism of the computer becomes better and better. Although the computer played poorly at first, through adversarial learning, the computer's ability has been improved to the point where it can defeat human chess players.

Moor believes that the computer uses polynomials to decide to move the chess pieces, indicating that the computer is a decision-maker. In fact, computers and human players make decisions in the same way. Computers playing checkers can prove that computers can make decisions in a broad sense. The computer must analyze the situation, discover available operating steps, evaluate options, and select the course of action based on its information.

Some people still believe that decision-making must be conscious. People are aware of their many decisions, but there are also many decisions that people do not realize. For example, people make many complicated driving decisions while driving, and they may also think about other things at the same time, but do not realize that

they are making decisions. Since people make conscious and unconscious decisions differently, So Moor believes that consciousness is not the essence of decision-making

Finally, some people say that it is not computers that make decisions, but people who use computers. But this view confuses the ability to make decisions and the right to make decisions. The right to make a decision includes the right to decide under appropriate circumstances. Computer designers can also delegate the power to make decisions to computers. In fact, using a computer in this way is the same as using a computer as a decision-maker.

Moor believes that understanding computer behavior at some point can reduce human awareness of the potential impact of computers in human society. Entrusted decision-making power is entrusted management. Ultimately, the question is what aspects of human life the computer needs to control.

How is the decision-making ability of the computer?

If computers can make decisions at least in principle, then what kind of decisions a computer can do is still a question. Because computers are not limited to random, fixed, or arbitrary decisions. It seems that the decision-making of a computer is unlimited. However, the logical result clearly shows some limitations. If you accept Church's theory that the algorithmic computability of a function is equivalent to the computability of a Turing machine, then the limit of the Turing machine is the limit of the computer. Specifically, the results of the halting problem show that even a general Turing machine cannot make decisions effectively, because there is no universal Turing machine. It can determine whether each Turing machine will stop. The problem with this restriction is that it seems to apply to both humans and computers.

No one knows whether computers can match or exceed human capabilities in the decision-making process. Therefore, empiricism is advocated in computer decision-making. Moor advocates:

- 1) The ability level of a computer is essentially an empirical problem for a given intellectual activity.
- 2) Evidence can be collected to determine the computer's ability level for a given intellectual activity.
- 3) For most intellectual activities, it is still unknown whether computers will reach the human level someday.

How to judge the decision-making ability of a computer?

To perfect the judgment of the nature of computer capabilities requires empirical investigation, but what kind of evidence is the most important? Because the ability of a computer refers to the ability to execute the program to a certain extent, the performance of the computer is important evidence for evaluating ability. In the process of computer decision making, two characteristics of performance are related to decision making. These two characteristics are the decision record and the justification for the decision. Moor hopes to distinguish these two extremes of decision-making.

About decision records, under clear criteria, every decision can be classified as correct or incorrect. For example, deciding which horse to bet on to win the race is a decision made under clear criteria. The horse will either win or not. After making many such decisions, the decision-maker will establish a clear record of correct or wrong decisions. In the decision-making process under clear standards, the reasons provided for the decision are usually not so important.

About the justification for the decision, if the decision record is not available or is not trusted for some reason, then the reason for the decision is very important when evaluating decisions under clear standards. Decisions made under fuzzy standards are difficult to classify as correct or incorrect. For many people, deciding which occupation to pursue is an example of making decisions under fuzzy standards. The vaguer the standards, the harder it is to establish a record of correct decisions.

Should the computer make a decision?

Joseph Weizenbaum believes that computers should not make decisions. He believes that computers can make judicial decisions as well as judgments of mental illness. Computers can flip coins in more complicated ways than most patients. But the crux of the problem is that computers should not be given such tasks. Although the computer may make the "correct" decision under certain circumstances, no one may be willing to accept it. Weizenbaum believes that computers are an outsider to human affairs. The basis for outsiders to make decisions is not commensurate with the environment in which they are made. But Moor believes that this view confuses the lack of information and the lack of decision-making ability. There may be good reasons not to grant outsiders the power to make certain decisions, but in principle, there is no reason to believe that an informed outsider is not a competent decision-maker.

Moor believes that Weizenbaum's examples of judicial decisions and mental disorders are all decision-making cases under fuzzy standards. The various reasons provided by the computer for its decision-making through close observation can be used to evaluate the decision-making ability of the computer in these fields. This makes the computer's ability to make decisions as good as or better than human

experts. From experience, this situation may not happen at all, but it is not an inevitable fact.

What Moor advocates is to use computer decision-making as a tool. For certain situations, people must decide whether to use computers to make decisions that will better accomplish people's goals. Within the scope of people's basic goals and values, people must use the experience to determine the ability of computers to make decisions. Computers should never determine what the basic goals and values of human beings should be. People want computers to work for people's purposes, and people want to prohibit computers from deciding to change these ultimate norms. To prohibit computers from making decisions on basic goals and values does not impose many restrictions on computer decisions. People's basic goals and values remain unchanged and people rarely change them. However, Moor believes that there is a reasonable concern that after the computerization of human society, human life will lose humanity. The root of this worry is not that computers do not have sufficient decision-making capabilities, nor that people are inherently evil. But maybe one day, computers will make most of the decisions better than humans to manipulate human society and greatly improve the quality of human life. The root of this concern is that the promotion of computerization should focus on the issue of responsibility. Computers are not humans. Even though computers are responsible for their decisions, they are not legally or morally responsible for their decisions. It is impossible for people to sue the computer. Therefore, human beings not only bear the initial responsibility but also continue to be responsible for improving their decision-making ability and values when computer decision-making problems arise.

The ethical dilemma faced by computer decision-making

With the popularity of driverless car technology, computers have taken over the right to drive cars from humans. Computers need to replace humans in the face of complex and changeable traffic conditions and make corresponding driving decisions based on traffic conditions. When faced with some inevitable collisions, ethical dilemmas may arise. These ethical dilemmas can be summarized into two typical problems: the tram problem and the tunnel problem.

Moral choice

When there is a gap in the legal judgment of computer decision-making, it is necessary to use moral choice to judge the correctness of the decision. Moral selection is the process of selecting possible moral viewpoints based on moral principles and conformity with moral principles when dealing with moral-related matters. Making ethical choices is a difficult and complicated thing. This choice is often accompanied by economic, professional, and social pressure. Sometimes these pressures challenge the moral principles or moral goals that people abide by. The complexity of ethical choices lies in the fact that in many cases there are multiple values and choices of different interests at the same time, and these competing values and interests must be chosen.

In computer decision-making, people need to think about the nature of computer power and how to show their nature. Secondly, given basic goals and values, why are computer decisions better than human decisions under certain conditions? Perhaps the responsibilities of human beings are easily destroyed by external pressure. As increasingly complex computer systems are embedded in social life, the pressure exerted on the computer field has gone beyond the purely technical level. In response,

people hope to get the "correct" answers to those moral questions. "Correct" is understood as "philosophical legitimacy or rootedness". This is undoubtedly an important field of philosophical analysis. But philosophy seldom tells people the relevance of philosophical positions that appear in practical calculations, let alone solve them effectively. This requires humans to give continuous feedback and corrections to computer decisions. But this is back to the ethical dilemma, what is right. This problem has still not been resolved. Computer scientists usually regard the basic ethics of human beings as the cornerstone for judging the legitimacy of computer decisions. The dilemma of human decision-making in ethical cognition has led to the dilemma of computer decision-making in ethics.

References

William, J. Rapaport. M. *Philosophy of Computer Science*. 2018.

Moor, J. H. *Are there decisions computers should never make?* Nature and System 1, 1979, 217–229.