Responsibility in the Age of Autonomous Machines

IEEE ARSO 2016
July 8, 2016
Shanghai

Minao Kukita
Graduate School of Information Science
Nagoya University
A fatal car accident by Tesla Model S while the driver used autopilot.

We learned yesterday evening that NHTSA is opening a preliminary evaluation into the performance of Autopilot during a recent fatal crash that occurred in a Model S. This is the first known fatality in just over 130 million miles where Autopilot was activated. Among all vehicles in the US, there is a fatality every 94 million miles. Worldwide, there is a fatality approximately every 60 million miles. It is important to emphasize that the NHTSA action is simply a preliminary evaluation to determine whether the system worked according to expectations.

Following our standard practice, Tesla informed NHTSA about the incident immediately after it
According to Tesla

• It is the first fatal accident after 130 million miles with autopilot activated, compared to the US average of one fatal accident every 94 million miles.

• Neither the driver (of Model S) nor the car didn’t notice the white body of the trailer against the brightly lit sky.

• Tesla’s autopilot system is “public beta phase” and they require the driver to maintain control and responsibility for the vehicle while the autopilot is activated.
Who is responsible?

- Tesla? who is selling a car still in “beta phase.”
- The driver? who was informed of the imperfection of the system.

- The question would be harder and far more serious if the victim were pedestrian or other drivers, not the driver of the autonomous car.
The concept of “responsibility”

• “You are responsible” may mean either
  – that you are **obliged** to make sure that some good thing will happen, or to prevent some bad thing will not,
  – that you are **among the major factors** which caused some (bad) thing,
  – or that you are **to blame**.
The concept of “responsibility”

- Traditionally, **only human adult can be held responsible** who is able to predict results of one’s action, to make autonomous decisions, and to control the course of one’s action.
Responsibility gap

• As for a highly complex and autonomous system, it is difficult to identify who is responsible for its behaviour, because…
  – It is difficult for anyone to predict the behaviours of the system.
  – Too many people are involved in its development and deployment.
  – Accidents will occur without any obvious fault of anyone.
  – There will be no one who is controlling it.
The function of “responsibility”

• The responsibility is embedded in our practice in which people are **obliged** to do something, and if one fails the obligation and **does some harm** to others, one will be **blamed** for that.
The function of “responsibility”

• Having this concept, we encourage ourselves to do good to the society and discourage ourselves from doing harm.

• The concept of responsibility seems to have developed because of this function.

• The responsibility gap will make it hard for the concept to fulfill the function.
Emotions

• Strong emotions, e.g. anger or guilt, accompany the claim “You are responsible”.

• Recent results in psychology reveal that these **emotions are key to our morality**; they are important factors driving people to moral behaviours.

• However in some situations, **emotions get in the way to rational, really beneficial behaviours**.

Cf. Joshua Greene, *Moral Tribes*
Necessity for revision of the concept

- It can be useless to follow our emotions and try to identify the culprit that is responsible for the accident caused by a complex, autonomous system.
- Given the original function of the notion, we have to revise it so that it will be useful in a society where humans and autonomous machines coexist.
Suggestions

• Put more emphasis on **social decision making** concerning an autonomous system than on individual persons involved in its development or its usage.

• **Ethical and democratic risk analysis** involving various stakeholders is necessary in the process of the decision making.
Suggestions

• Do not invest too much in manhunt, and instead do our best to improve the system, though it will be difficult not to follow our deeply wired emotions.

• Transparency and data-sharing is crucial. Do not allow huge or hasty companies to monopolise the data.

• The duty of compensation should be distributed to beneficiaries, including users, manufacturers, and the society itself.
Thank you!