Minds The Morality of Mechanical

Yochanan Bigman

On March 18th 2018 an Uber self-driving car killed a cyclist who was crossing the street. As we adopt to the increasing presence of automated robots playing significant roles in our society, it is important to understand how people think about robots making moral decisions. In my research I examine how people's attitudes towards robots making moral decisions are shaped by the mental states people attribute to robots and to humans.

In my talk I will discuss three questions regarding robot morality. The first is do people think that robots *should* make moral decisions. I find that people want moral decision makers to have the ability to think but also to feel, and since robots are perceived as being unable to feel, people do not want them to make moral decisions. The second is when robots would have to make moral decisions, such as in the case of self-driving cars, *how* do people want them to do that. I show that, in contrast to other published work, people want robots not to discriminate between humans according to age, gender and social status. The third is how people would respond *if* robots behave immorally, such as exhibit gender or race discrimination. I show that because robots are not seen as possessing a negative (sexist or racist) motivation, people are less outraged when robots discriminate than when humans do.

In addition to understanding people's attitude towards robot morality, my research reveals basic moral cognition processes. I show that people view emotions as necessary for moral decision making and highlight how the attribution of mental states shapes moral judgment.