

## Robotics: Some Laws

for CSE 473: Introduction to Artificial Intelligence, Autumn 2018

S. Tanimoto

Paul G. Allen School of Computer Science and Engineering

University of Washington



## Outline

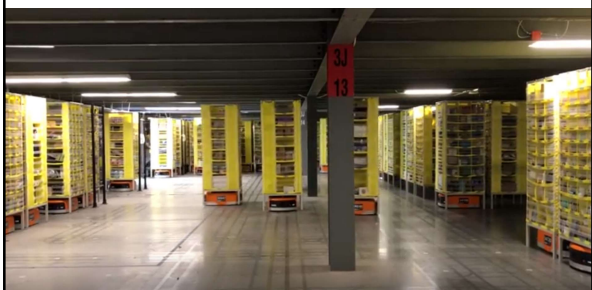
What are we afraid of?

\*Asimov's Laws

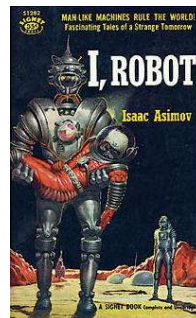
More recent laws of robotics

What else might happen?

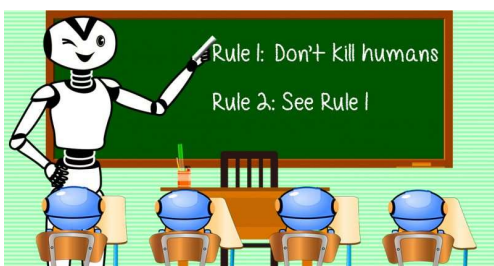
## Robots in an Amazon Fulfillment Center



## Are Humans Safe with Robots?



## Laws of Robotics ???



## Isaac Asimov's Laws of Robotics !!!

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 to it by human beings, except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.



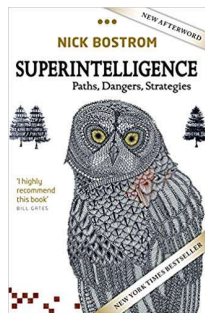


- **EPSRC's Laws of Robotics:**
- Robots should not be designed as weapons, except for national security reasons
- Robots should be designed and operated to comply with existing law, including privacy
- Robots are products: as with other products, they should be designed to be safe and secure
- Robots are manufactured artifacts: the illusion of emotions and intent should not be used to exploit vulnerable users
- It should be possible to find out who is responsible for any robot

Engineering and Physical Sciences Research Council (UK)

## What Else May Go Wrong?

- Even without "robots", A.I. could become malevolent.
- (e.g., hacking + a.i. = bad news)
- "superintelligence" could be good or bad.



## Definition of "Superintelligence"

By a "superintelligence" we mean an intellect that is much smarter than the best human brains in practically every field, including scientific creativity, general wisdom and social skills.

This definition leaves open how the superintelligence is implemented: it could be a digital computer, an ensemble of networked computers, cultured cortical tissue or what have you. It also leaves open whether the superintelligence is conscious and has subjective experiences.

--Nick Bostrom (Oxford University, U.K.)

## Conclusions

- Both robotics, and A.I. more generally, can affect human individuals and humanity broadly.
- It's important for anyone involved in creating or using AI to be aware of ethical issues and dangers.
- Understanding systems of laws such as Asimov's, can be helpful, even if nobody is enforcing them yet.