

Current challenges in *Ethics of AI* or: old wine in new bottles

Leonie Seng

First main thesis

➔ **Current discussions on *moral* development of *intelligent machines/ robots/ algorithms/ programs* in accordance to certain *ethical* criteria is old wine (in new bottles).**

- Reason: Answering questions concerning development of *moral machines* requires first of all the *systematic* development of concrete *concepts* of human beings and their ideas of certain *moral values*.
- This has been part of most considerations since the very beginning of philosophical investigations (philosophical *anthropology*, *humanism*, *ethics*, *meta-ethics*...) → *new challenges* of AI = **old wine**.



Second main thesis

➔ **Currently most ethical approaches tend to a form of *general Ethics of AI*. Such approaches are not very promising.**

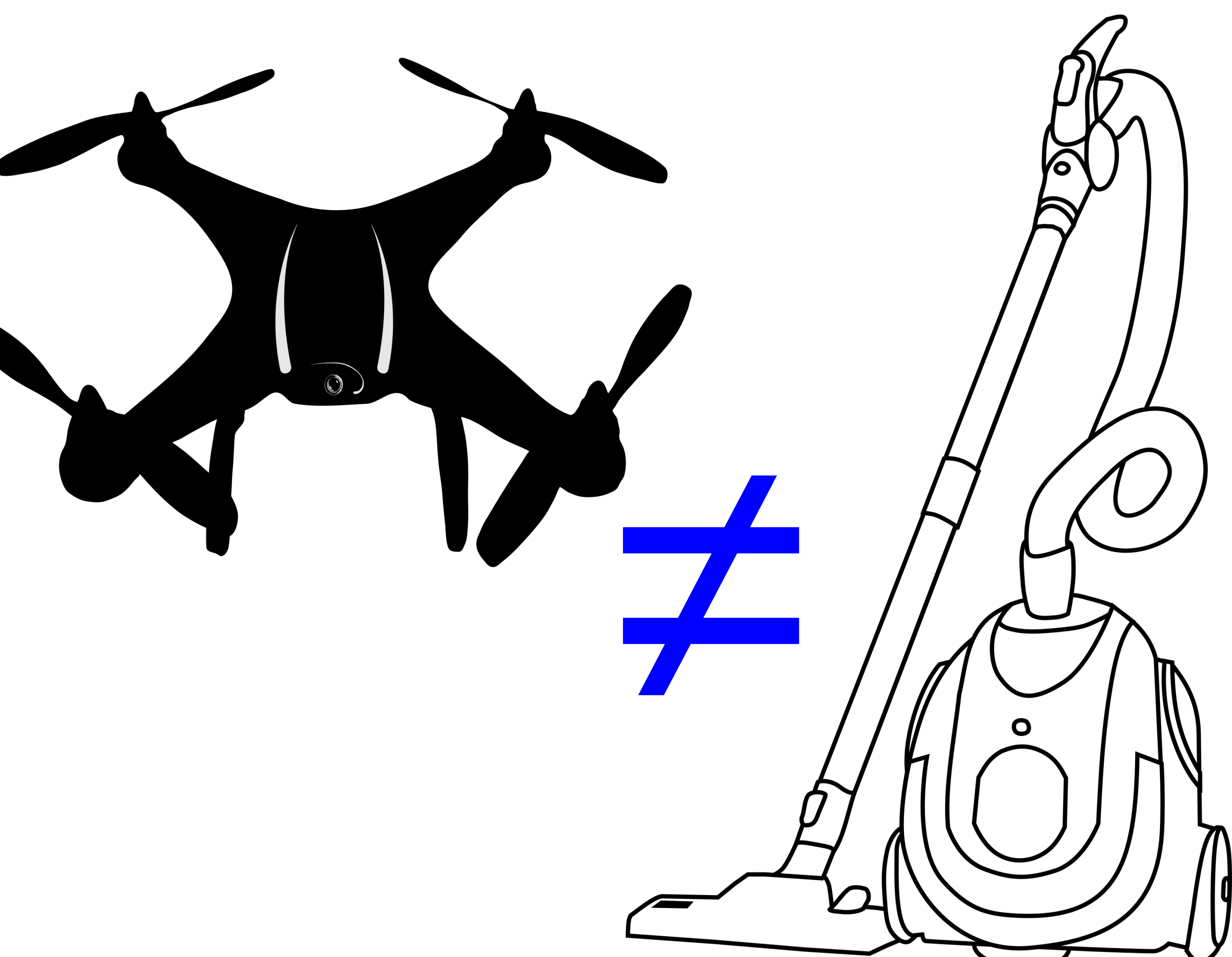
- Reasons: talking generally about *Ethics of AI* → *general* and hence superficial results (see Issac Asimov's *Three Laws of Robotics* →)
- general rules or laws such as *physical integrity* of humans = very impractical when developing *intelligent machines/ programs*, because *too vague*
- As Asimov's science-fiction stories (among others) show, laws/ goals/ rules can come into conflict
- moreover tendency : *learning* systems = spontaneous reactions

Asimov's Three 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 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 Laws."
- (Asimov 1940)

Problems

1. if moral rules are more concrete → only specific statements for certain domains
2. *utilitarian* machine ethics contradicts current technological developments (see above)
3. discussing *insoluble dilemma situations* = useless qua concept! (cf. among many others Birnbacher and Birnbacher 2016; Seng 2017)



Solutions/ Conclusion

1. pluralist world → not one general answer or definition can be considered as unique standard → moral values differ between different countries and cultures → **The development of military drones is much more delicate as the development of robot vacuum cleaners!**
2. installing ethical supervisory bodies in every department where forms of artificial intelligence/ androids/ robots are developed is necessary – even more, as sometimes *names* differ from the *content* (e. g. in the *Ethik-Kommission des BMVI* (2017) hardly any ethicists take part!)