

HRI Reading Group

@ Instituto Superior Técnico

Meeting #9 (11 May 2018)

Welcome!

Paper

Turkle, S., 2007. **Authenticity in the age of digital companions.** Interaction studies, 8(3), pp.501-517.

Ideas



Is an authentic agent conceivable?

Authenticity is...

Alignment between evocativeness and capability to evoke it evokes,

Genuineness seems to be stronger

Context- and task / reward dependent

Evocative but not authentic?

How do we relate with inauthentic agents?

But it seems odd to use these words to describe benchmarks in human-robot encounters, because we have traditionally reserved them for relationships in which all parties were capable of feeling them – that is, where all parties were people. With robots, people are acting out “both halves” of complex relationships, projecting the robot’s side as well as their own. Of course, we can also behave this way when interacting with people who refuse to engage with us, but people are at least capable of reciprocation.

What matters is the benefit on the user

Problem of expression of models that don't correspond to human models

Neural aspects



What should be the guidelines to develop these evocative/inauthentic agents?

Ethical dilemmas:

- Authenticity
- Human interactions will be lost
- Responsibility, accountability and agency: Lawmakers
- Discrimination due to current fears of AI

Guidelines:

- Identification should be required (expectation, avoid deception)
- Give control to the user
- Transparency and explainability in decision-making and the learning, but also in the state and its expression

**Who is bringing the
refreshment next week?**

HRI Reading Group

@ Instituto Superior Técnico

Meeting #10 (18 May 2018)