The Human-or-Machine Matter: Turing-Inspired Reflections on an Everyday Issue

(brief summary of paper)

David Harel and Assaf Marron

Department of Computer Science and Applied Mathematics
The Weizmann Institute of Science, Rehovot, 76100, Israel

In his seminal paper "Computing Machinery and Intelligence", Alan Turing introduced the "imitation game" as part of exploring the concept of machine intelligence. The Turing Test has since been the subject of much analysis, debate, refinement and extension. In this talk we sidestep the question of whether a particular machine can be labeled intelligent, or can be said to match human capabilities in a given context. Instead, we first draw attention to the seemingly simpler question a person may ask themselves in an everyday interaction: "Am I interacting with a human or with a machine?". We then shift the focus from seeking a method for eliciting the answer, and, rather, reflect upon the importance and significance of this Human-or-Machine question and the use one may make of a reliable answer thereto. Whereas Turing's original test is widely considered more of a thought experiment, the Human-or-Machine matter as discussed here has obvious practical relevance.

Among the issues we raise and discuss are:

1. What would the typical 'person on the street' say *are* the differences between machines and humans that are relevant to them, especially, when interacting with machines that mimic humans well. For example, are agent's emotions, learning abilities or common sense relevant factors?
2. How would such differences, real or perceived, affect such a person's interactions with the agent? In what way would a person's interaction with agents that are known to be machines that mimic humans (with high fidelity) be different from their interaction with agents that are known to be human? Would conversations be more matter-of-fact? Would a new kind of trust building evolve?
3. In normal, ethical, everyday situations, how would or should one discover the Human-or-Machine identity of an agent? Should this knowledge be always explicitly available, maybe by regulatory enforcement? Will there be protocols to elicit such knowledge from interactions that are narrowly focused on some business or personal issue? Are there ethical situations where people would prefer to *not* know if the agent is a human or a machine? How will human-to-human interactions change in a world where a growing portion of human-agent interactions are with human-like machines?

While it is still unclear if and when machines will be able to mimic human behavior with high fidelity in everyday contexts, we argue that near-term exploration of the issues raised here can contribute to improving methods for developing computerized systems, and may also lead to new insights into fundamental characteristics of human behavior.

The full paper is under review for publication. Also, see arXiv preprint here: https://arxiv.org/abs/2305.04312