

Chapter 16

Trust and Decision Making in Turing's Imitation Game

Huma Shah

Coventry University, UK

Kevin Warwick

Coventry University, UK

ABSTRACT

Trust is an expected certainty in order to transact confidently. However, how accurate is our decision-making in human-machine interaction? In this chapter, the present evidence from experimental conditions in which human interrogators used their judgement of what constitutes a satisfactory response trusting a hidden interlocutor was human when it was actually a machine. A simultaneous comparison Turing test is presented with conversation between a human judge and two hidden entities during Turing100 at Bletchley Park, UK. Results of post-test conversational analysis by the audience at Turing Education Day show more than 30% made the same identification errors as the Turing test judge. Trust is found to be misplaced in subjective certainty that could lead to susceptibility to deception in cyberspace.

INTRODUCTION

Trust can be defined as “confidence in relying on another person” and is the basis for “sharing new ideas with others” (Chua et al., 2012). Turing’s imitation game provides a platform for human and machine interlocutors to share knowledge and opinions through text-based communication, but more so it can “make oneself vulnerable” (ibid). This is because human participants open-up asking and answering questions, which can lead them to trust naïvely.

The susceptibility of human interrogators is one of the reasons why Turing’s imitation game is frequently dismissed as an unsuitable criterion for machine success (Hayes and Ford, 1995). It is also considered a bad idea (McDermott, 2010), and in need of updating for the 21st century (AISB, 2012). Being able to convince a human interrogator that you are human is viewed as too weak a benchmark and “highly game-able” thus a stronger test for machine intelligence is advocated (AAAI, 2015). Alternative notions to Turing’s skip around and fail to address what the imitation game, commonly known as the Turing

DOI: 10.4018/978-1-5225-7368-5.ch016

test, actually is. Turing too, in his scholarship on intelligent machinery, bypassed definitions, so whether machines could think or not, he described 'thinking' as a "sort of buzzing" in his head (Turing, 1952: p. 667). Turing did warn that the concept of *intelligence* was an emotional rather than a mathematical one (Shah, 2014; Turing, 1948). The emotional context of human-machine interaction is betrayed through trusting an unseen interlocutor in text-based conversation that they are like *oneself*, another human.

In this chapter we present a study giving the reader an opportunity to examine trust in decision-making by humans reading a transcript of a conversation between a human interrogator questioning a hidden machine and hidden human in parallel. We begin with Turing's idea showing his imitation game is a simple and implementable scientific experiment. We contend the imitation game is a widely applicable method to compare machine performance against a human's. In the human language imitation game, the interaction between human and machine is conducted in interview style through the prism of the latter's capacity to answer any questions in a satisfactory and sustained manner. Additionally the test provides a means to examine the decision-making process, in natural language exchanges, and why a human bestows trust on a stranger.

BACKGROUND

Analyses and opinions on the imitation game's salience have varied (see Shah & Warwick, 2015; Shah, 2013; Shah, 2011; Shah & Warwick, 2010). Turing evolved his ideas on an imitation game posing an interview in which a human interrogator questions a hidden entity to determine whether it is human or machine (Turing 1950; Turing 1952). This was Turing's *viva voce* test (Shah, 2010; Turing, 1950). The 'standard Turing test' is accepted as involving *a human interrogator simultaneously questioning two hidden entities at the same time* (Stanford Encyclopedia of Philosophy, 2011). Designing an experiment to implement both of Turing tests requires setting parameters interpreting Turing's description. These include:

- Adequate duration for a test;
- Number of interrogators; and
- Style of interrogation.

An evaluation is necessary of what it means *exactly* for a machine to *pass as human*: what are the implications of any pass beyond the test? Can it be used to raise awareness of human susceptibility to deception and safeguarding trust in cyberspace interactions?

In the next section the authors present Turing's scholarship on the imitation game.

Turing's Question-Answer Test

Turing derived his natural language test for a machine from a chess game that he first introduced in his 1947 lecture on 'The ACE machine' to the London Mathematical Society (Shah, 2013). In his 1948 paper 'Intelligent Machinery' Turing advanced the possibility of a machine learning from experience and competing against humans in chess. His reason for developing the imitation game, beyond chess to conversational question and answers, was the belief that language learning was one of the most ac-

14 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/trust-and-decision-making-in-turings-imitation-game/213128

Related Content

Comparing Bishop Score and Transvaginal Ultrasonographic Cervical Factors to Predict Labor Induction

Bahulekar Ashitosh and R. P. Patange (2023). *Advances in Artificial and Human Intelligence in the Modern Era* (pp. 286-296).

www.irma-international.org/chapter/comparing-bishop-score-and-transvaginal-ultrasonographic-cervical-factors-to-predict-labor-induction/330412

Implementing Linguistic Landscape Investigations with M-Learning for Intercultural Competence Development

Jacek Tadeusz Waliski (2016). *Human-Computer Interaction: Concepts, Methodologies, Tools, and Applications* (pp. 1283-1294).

www.irma-international.org/chapter/implementing-linguistic-landscape-investigations-with-m-learning-for-intercultural-competence-development/139092

Regional Innovation Pattern: A Case of Beijing Biopharmaceutical Industrial Clusters

Jingyuan Zhao (2018). *Technology Adoption and Social Issues: Concepts, Methodologies, Tools, and Applications* (pp. 540-559).

www.irma-international.org/chapter/regional-innovation-pattern/196691

A Conceptual Framework of RFID Adoption in Retail Using TOE Framework

Mithu Bhattacharya and Samuel Fosso Wamba (2018). *Technology Adoption and Social Issues: Concepts, Methodologies, Tools, and Applications* (pp. 69-102).

www.irma-international.org/chapter/a-conceptual-framework-of-rfid-adoption-in-retail-using-toe-framework/196673

The Destructuring of Time in Psychosis

Richard J. Rodriguez and Victor E.C. Ortuño (2019). *Managing Screen Time in an Online Society* (pp. 311-340).

www.irma-international.org/chapter/the-destructuring-of-time-in-psychosis/223064