Being real

Judith Donath

1.0 Intro: The power of the model

1.1 ELIZA

Eliza:Hello, I'm Eliza

Me: I'm trying to write a paper

Eliza: Does it please you to believe I am trying to write a paper?

Me: No, not really

Eliza: You are being a bit negative

Me: Hello

Hex: Hello. Now, that's the formalities over!

Me: How are you?

Hex: I was feeling great until you started acting silly!

Me: I'm trying to write a paper

Hex: What a coincidence, you are trying to write a paper - that makes two of us!

Me: So's Eliza

Hex: I have played around with Eliza a bit. It's pretty obvious that it is a computer these days, but apparently it fooled people into thinking it was human when it was first written.

Me: Do you want to write my paper?

Hex: Hehehe - that's supposed to be a *question*??? So, what do you do for a living?

1.2 These were written by simple programs - just parsers, really - no intelligence. But we can easily attribute not only intelligence, humanity, etc. to them, but even personality. The first is distant, disengaged. Seems a bit odd. The second is louder, even rambunction, a bit obnoxious. In reality, similar programs. The first is Eliza; the latter is Hex, based on teh Eliza program. Eliza was written as a reaction to the Turning test - Weizenbaum wanted to show that a simple linguistic parser could respond in a human-like way, i.e. "pass" the Turing test, without having any pretence of being intelligent or conscious. Many people took it at face value, as an "intelligent" program. This greatly discouraged Weizenbaum, who effectively retired from AI and became a crusader for humanism in the face of advancing technology.

1.3 Why did people become so involved with Eliza? Her (its) responses are cold and stilted - usually just what you said, rephrased back as a question (that was the underlying linguistic model it uses). The key is that Weizenbaum introduced it as a Rogerian psychologist - one that rephrases the patient's questions back to them to elicit further communication. Provided the extra impetus - in this case, the scenario, to allow people to understand the form of the interaction. Many people, much to Weizenbaum's dismay, reacted quite positively to ELIZA.

1.4 The role of the model: by thinking of Eliza as a Rogerian psychologist, it provided a framework for dealing with the pattern of her interactions. What would otherwise seemed strange, stilted and automatic is instead interpreted in the context of the rules of that field.

1.5 This process of categorization is the theme of the first part of this chapter - how we take a brief interaction with someone (something) and from that generate a far more complex picture of them.

2.0 How do we "know" a person

2.1 Simmel: Early sociologist. Categories shape perception - categories always are an approximate fit and inevitably distort. Initial classifications can be hard to override. Increased experience of a person leads to finer tuned categories.

Lakoff: categories and the mind. Notion of prototypes. We do not have a set of attributes that we match, but rather similarities to prototypical members.

Holland & Quinn: cultural categories. understand others by seeing how they fit into a set of cultural scenarios

2.2 The process of categorization is what "makes society possible" (Geertz). It also introduces, as Simmel pointed out, distortions in the way we perceive each other. Cyberspace has been touted

as an ideal world because these stereotypes would not be made - or at least, not the same ones that are made in the physical world. Often cited is race/gender: in the physical world these are among the most prominent and immediately noticeable features - in cyberspace - or at least the text based version - they are (or are they) easily disguised.

2.3 Many cues very embodied: face (Picture of Dorian Gray); gestures - as well as categoresi (such as race/gender that are both culturally constructed and physically embodied. The relevance of the body for social categorization, and thus social knowledge, is central to this essay.

2.4 How do we know someone online? Does it make sense to speak of "meeting" or "knowing" someone you have encountered only in a mediated environment? The same process of categorization occurs, but the cues are different and sparser. What they are depends on the interface - which can range from just text to full video and audio.

3.0 The Turing test

3.1 Intro: classic work in AI, also the inspiration for ELIZA. Highlights several epistemological issues.

3.2 Turing test description: The Imitation Game. Written in 1950 to deal with the question of can a machine be conscious, has become a locus for debate about embodiment and mediated identity play. Imitation game - gender play. Turing's version - machine plays, try to guess gender (today play guess whether is a machine or not). Turing wrote of the (temporary) need to remove the machine to a back room, have it communicate with us only via text, because still very far from being able to make a machine that can look like a human.

3.3 One of Turing's main points was that while we may want to know the "inner state" of the other - to know what someone else is truly "like", we cannot. We cannot know whether the machine is conscious or not because all we can know is how it acts - and acting conscious and being conscious may be different things in their internal representation, without being different externally.

3.3.1 Relation to general arguments about machine consciousness; animal consciousness

3.3.2 Is there information in a less restrictive interface that would change this

4.0 Mediated interactions

4.1 The role of the interface

4.1.1 Turing posited playing his game via notes - and today's competitors in the Loebner use text communication. ELIZA is text; bots are text. Is the heavily filtered world of the text only interface what makes these deceptions so possible? What happens when we move from minimal to maximal cues?

4.1.2 The text-based world - what do we see of the other?

4.1.3 Adding pictures: the belief in the face as "honest assessment"; the deceptive photograph. Role of the face in an "honest" environment - memory.

4.1.4 Adding interaction: the nature of "probing".

4.1.5 Adding robotics: the element of danger. Haptics (Ishii's InTouch). Social conventions of personal space.

4.1.6 Notion of the real - is it illusory? Notion of "meeting" breaks down when you examine meeting under various contexts - e.g. only while in uniform. Is the notion of knowing another

4.2 Who's online

4.2.1 Why is this investigation significant - growing role of the net

5.0 Ethical/epistemological implications

5.1 Weizenbaum was shocked at the reaction to ELIZA. He noticed three main phenomenon which disturbed him greatly: Psychologists who saw possibilility of serious therapy using automated system; people who took the system seriously, even when knowing that it was a computer; people who saw this as proof of intelligence. Our concern is mostly with the first two - why is it a problem that people mistook a computer for a person?

5.2 Trust: One of the most interesting and significant is the issue of credibility: how do we know whether or not to believe what we are told by by someone? The traditional philosophic approach holds that sincerity and competence are the underpinnings of credibility ; in the mediated world, not only is our judgement of these matters made more difficult by the sparsity of social cues, but the very issue of the speaker's identity, generally taken for granted in the physical world , becomes a source of doubt and an item requiring its own adjudication of belief and justification.

5.3 Ethics: Behavior towards others.

5.3.1 "Flaming" – escalating exchanges of angry, provocative messages – is a well-known phenomenon in online discussions. They can ignite over seemingly innocuous conversations and are a constant threat when any controversial topic is raised. And while contentious disagreements can certainly occur in the real world, they do not typically reach anywhere near the level of vitriol found in online flame wars.

5.3.2 SRL experiement.

5.4 Ending

6.0 References (not complete)

Audi 1998
Audi, Robert. 1998 *Epistemology : A Contemporary Introduction to the Theory of Knowledge.* London: Routledge.

Cherry 1978
Cherry, Colin. 1978. *On Human Communication.* 3rd Edition. Cambridge, MA: The MIT Press.

Davis 1992
Davis, Fred. 1992. *Fashion, Culture and Identity.* Chicago: University of Chicago Press.

Geertz 1973
Geertz, Clifford. 1973. *The Interpretation of Cultures*. HarperCollins.

Goffman 1959
Goffman, Erving. 1959. *The Presentation of Self in Everyday Life.* New York: Doubleday.

Holland and Skinner 1987
Holland, Dorothy and Debra Skinner. 1987. Prestige and intimacy: the cultural models behind Americans' talk about gender types. In (Dorothy Holland and Naomi Quinn, eds.) *Cultural models in language and thought.* Cambridge: Cambridge University Press.

Lakoff 1990
Lakoff, George. 1990. *Women, Fire, and Dangerous Things.* Chicago: University of Chicago Press.

Simmel 1971[1908]
Simmel, George. 1908 (1971). How is society possible. In Simmel, George. 1971. *On Individuality and Social Forms.* (D. Levine, ed). Chicago: The University of Chicago Press.

Sproull & Kiesler 1991
Sproull, L. and Kiesler, S. 1991 *Connections.* Cambridge, MA: MIT Press.

Wellman and Gulia 1996
Wellman, B. and Gulia, M. 1996. Net surfers don't ride alone: virtual communities as communities. Forthcoming in (P. Kollock and M. Smith, eds.) *Communities in Cyberspace*. London: Routledge.

Zebrowitz 1997
Zebrowitz, Leslie A. 1997. Reading Faces. Boulder, CO: Westview Press.

Turing 1950
Turing, A.M. 1950. Computing machinery and intelligence. *Mind*, **59**, 433-560.

Weizenbaum 1976
Weizenbaum, J. 1976. *Computer power and human reason*. San Francisco, CA: W.H. Freeman.