Available online at www.sciencedirect.com
ScienceDirect
Artificial Intelligence 170 (2006) 1256–1258
Artificial Intelligence
www.elsevier.com/locate/artint

Shifting viewpoints: Artificial intelligence and human–computer interaction

Terry Winograd

Stanford University, 353 Serra Mall, Stanford, CA, USA

Available online 1 November 2006

Abstract

The AI and HCI communities have often been characterized as having opposing views of how humans and computers should interact. As both of them evolve, there is a deeper contrast that cuts across these communities, in how researchers conceive the relationship between knowledge and design. By examining the rationalistic and design orientations underlying bodies of work in both disciplines, we highlight relevant differences and possibilities for effective interaction with computers.

Keywords: HCI; AI; Human–computer interaction; Rationalistic; Design; Symbolic; Phenomenology

1. AI and HCI

In his engaging account of the early history of personal computing in Silicon Valley, Markoff [7] describes a confrontation of views between pioneering researchers. In an interview about the book, he stated it in simple terms:

> McCarthy's work was really right in the center of the computer science world at that time. Before arriving at Stanford in 1962, he had already invented the LISP programming language, which became the standard for artificial intelligence researchers, and he'd pioneered the time-shared operating systems that became the foundation of interactive computing. When he set up his research center his idea was basically to come up with a working artificial intelligence in the decade—a kind of "superbrain". ... But [Douglas] Engelbart was philosophically opposed to that approach. His idea, which he called "augmentation", was really not about replacing the human in the loop but instead to use computing help to augment the human. ... So a philosophical divide developed between Engelbart and McCarthy. [8]

This decades-old contrast between artificial intelligence and human-computer interaction was the topic of a recent retrospective by Grudin [4]. He describes the alternating cycles in which one approach was ascendant while the other suffered a "winter" in terms of attracting interest and funding from scientific agencies. The clear implication is that these are two different dreams competing for the hearts (and pocketbooks) of those with the resources to pursue them.

On another level, there have long been debates, most notably represented by Shneiderman and Maes [10], about whether AI should be a primary metaphor in the human interface to computers. Should we expect to communicate

E-mail address: winograd@cs.stanford.edu (T. Winograd).

0004-3702/$ – see front matter
doi:10.1016/j.artint.2006.10.011

with the computer in the same way we would to another human, as persuasively portrayed in Apple's early vision of the Knowledge Navigator [3]? Or are there practical and even philosophical objections to encouraging people to attribute human attributes and abilities to their computers?

My own history can be seen as reflecting a shift across this divide, starting with work in the core areas of AI [11] and then rejecting an AI approach [12] and moving to HCI [14]. Like many others in HCI, I have at times seen this as a battle between competing philosophies of what it is most effective to do with computers.

It would be easy at this point in history to simply take a conciliatory view of the situation and say "You're both right." As computers become more a part of everyday life, there are clearly places where AI can be applied and places where other styles of interaction are more appropriate. AI is a key technology in many interfaces, such as the recommender interfaces of on-line stores. Even in the Knowledge Navigator vision of twenty years ago, the enactment's dynamic appeal came from the natural dialogue with Phil, the AI agent, combined with direct gestural manipulation of maps, graphics, and other on-screen objects. And as to funding, it doesn't have to be an either/or. The battles for whether AI or HCI should get research support may be better fought on the level of encouraging broader support of innovative computing methods and interactions of all kinds.

But there is a deeper division, which is obscured by the simple opposition of AI and HCI: a contrast between two different approaches to how we understand people and create technologies for their benefit.

2. The rationalistic and design approaches

The first approach, which I will call "rationalistic" (see discussion of this term in [12]) aspires to model people as cognitive machines, whose internal mechanisms parallel those we build into digital computers. The clearest expression of this view is Newell and Simon's Physical Symbol System Hypothesis [9], which influenced a generation of researchers both in AI and in HCI. It was the theoretical expression of what Haugeland [5] calls "Good Old Fashioned AI" (GOFAI), which was the dominant paradigm in the era of high expectations for the near-term creation of human-like AI. Newell's conception was also the key inspiration for the founding text on HCI as a discipline of cognitive engineering [1], which remains influential in the HCI community today. A quick glance at the papers in the annual ACM SigCHI conference shows many papers that address interaction problems from an empirical quantitative perspective.

The key assumptions of the rationalistic approach are that the essential aspects of thought can be captured in a formal symbolic representation. Whether or not it corresponds directly to a formal logic, it operates like a logic in that well-defined algorithmic rules can be applied to models (processes and knowledge) in the form of symbol structures. Armed with this logic, we can create intelligent programs and we can design systems that optimize human interaction.

The second approach is harder to label. It has affinity to those who call their approach "phenomenological, "constructivist", and "ecological", and I will refer to it as a "design" approach. In the design approach, the focus is not on modeling intelligent internal workings, but on the interactions between a person and the enveloping environment. Along with this shift of focus goes a shift in the kind of understanding that is pursued.

A key part of the difference is in the role of formal modeling and explanation. In design, we often work in areas of human interpretations and behaviors for which we do not have predictive models. The question of "Does it work?" is not approached as a calculation before construction, but as an iterative process of prototype testing and refinement. David Kelley, founder of the design firm IDEO and the leader of Stanford's Institute for Design (dschool.stanford.edu) is often cited for his statement that "Enlightened trial and error outperforms the planning of flawless intellect." This is not a statement against intellect, but rather an acknowledgment of the limitations of knowing and modeling the complexities of the real human world.

Within the HCI community, there has been an ongoing progression away from Newell's cognitive engineering approach, towards a design-oriented stance that draws much more from the experience of other design disciplines [14]. Today's CHI community is engaged in ongoing debates about the roles of analysis and design in creating new interfaces and understanding the way people interact with them (see, for example, [15]). Of course "design" is interpreted by different people in many related but non-identical ways. A quick web search for the terms HCI and Design gives a feel for the diversity and liveliness of the discussion.

Over the same period, AI has followed an analogous trajectory. We have seen the general abandonment of GOFAI and the ascendance of statistical, embodied, and constructivist approaches. It is impossible to do justice to the important concepts and differences among these in a short comment, but a thread runs through them that resonates with the

design approach in HCI. Rather than basing AI competence on a logical representation of the setting and the agent's knowledge, there is an interplay between general adaptive mechanisms and world experience, which leads over time to intelligent behaviors, often as the outcome of extensive examples and training.

Many critics of AI, beginning with Dreyfus [2] have emphasized the inseparability of human thought and human physical embodiment. Thinking like a human is not just a matter of the right brain architecture or gene sequence, but of a complex developmental interaction between the whole organism and its environment, including other people. He and others have argued that for a device to have human-like intelligence would require its experiencing something akin to human development and upbringing. The rationalistic response to this is to replicate the end-result of physical world experience with a collection of propositions, as in CYC [6], a project which has failed to live up to its expectations after many years of effort (see [13]).

Of course there is a major role for a rationalistic approach in creating and understanding the adaptive mechanisms. Work in statistical language understanding, neural networks, or machine learning is based on deep analysis and quantitative models of the different mechanisms and techniques that form the basis for adaptation. But the researcher is not required (or able) to explicitly represent the knowledge or rules of behavior for the intelligent system. In the end, "Enlightened trial and error outperforms the planning of flawless intellect."

3. Conclusion

So if we agree that the underlying divide is not between HCI and AI, but between rationalistic and design approaches, what is the new bottom line? In spite of my obvious affinity to the design perspective, I would be foolish to take an either/or stance in this case either. The increasing influence of the design approach was a response to an over-valuing of the rationalistic approach that grew up around the advances in pure science and technology of the mid 20th century, along with attempts to push the computer metaphor onto all of reality. It was a needed corrective, but like all correctives it doesn't negate its opposite.

In our interdisciplinary design program we talk about the need to create "T-shaped" ways of understanding and working. The vertical bar of the T is a deep analytic understanding of a scientific or technological domain. The crossbar is the ability to bring design thinking to problems in a holistic way, recognizing the limitations of analysis and developing the ability to work effectively in problem areas that carry the unavoidable complexity and messiness of the human situation. There is a need for T-shaped AI and T-shaped HCI as well. We may be surprised at how closely they end up resembling one another.

References

[1] S.K. Card, T.P. Moran, A. Newell, The Psychology of Human Computer Interaction, Lawrence Erlbaum, 1983.
[2] H. Dreyfus, What Computers Can't Do: The Limits of Artificial Intelligence, MIT Press, 1972.
[3] H. Dubberly, D. Mitch, The Knowledge Navigator, Apple Computer, Inc., video, 1987.
[4] J. Grudin, Turing maturing: the separation of artificial intelligence and human–computer interaction, Interactions 13 (5) (September–October 2006) 54–57.
[5] J. Haugeland, Artificial Intelligence: The Very Idea, Cambridge, MA, Bradford/MIT Press, 1985.
[6] D. Lenat, Cyc: A large-scale investment in knowledge infrastructure, Communications of the ACM 38 (11) (November 1995).
[7] J. Markoff, What the Dormouse Said: How the Sixties Counterculture Shaped the Personal Computer Industry, Viking, New York, 2005.
[8] J. Markoff, Interview, Ubiquity 6 (29) (August 9–16, 2005).
[9] A. Newell, H.A. Simon, Computer science as empirical enquiry: Symbols and search, Communications of ACM 19 (3) (1976) 113–126.
[10] B. Shneiderman, P. Maes, Direct manipulation vs. interface agents, Interactions 4 (6) (November–December 1997) 42–61.
[11] T. Winograd, Understanding Natural Language, Academic Press, New York, 1972.
[12] T. Winograd, F. Flores, Understanding Computers and Cognition, Ablex, 1986, Addison Wesley, Reading, MA, 1987.
[13] T. Winograd, Thinking machines: Can there be? Are we?, in: J. Sheehan, M. Sosna (Eds.), The Boundaries of Humanity: Humans, Animals, Machines, University of California Press, Berkeley, CA, 1991, pp. 198–223.
[14] T. Winograd, Bringing Design to Software, Addison Wesley, Reading, MA, 1996.
[15] J. Zimmerman, S. Evenson, K. Baumann, F.H. Joanneum, P. Purgathofer, in: Workshop on the Relationship between Design and HCI, CHI '04 extended abstracts, 2004, pp. 1741–1742.