Available online at www.sciencedirect.com
SciVerse ScienceDirect
Int. J. Human-Computer Studies 71 (2013) 171–176
International Journal of Human-Computer Studies
www.elsevier.com/locate/ijhcs

From knowledge science to symbiosis science

Jeffrey M. Bradshaw

Florida Institute for Human and Machine Cognition (IHMC) 40 So. Alcaniz Street, Pensacola, FL 32571, USA

Accepted 25 September 2012
Available online 16 October 2012

Abstract

In the mid-1980s, Brian Gaines first developed a model to predict the trajectory of progress in human–computer relationships, including how the knowledge science research programme would naturally transform itself over time into something he called "symbiosis science." In this article, we reflect both on the extraordinary prescience of this model, and the contributions and challenges faced by researchers intent on progressive achievement toward the aspirations it inspires.

Keywords: Symbiosis science; BRETAM; Autonomous systems; Software agents; Personal assistants; Multi-agent systems; Human-agent-robot teamwork; Sensemaking; Wisdom

1. Introduction

Brian Gaines was always thinking decades ahead of the rest of us. His BRETAM diagram brilliantly predicted the trajectory of progress in human–computer relationships, including how the knowledge science research programme would naturally transform itself over time into something he called "symbiosis science" (see Gaines, 2013). The term "symbiosis" hearkens back to a 1960 article on *man-computer symbiosis* written by J.C.R. Licklider, the first director of the Information Processing Technology Office of the US Advanced Research Projects Agency—now DARPA (Licklider, 1960). In the ultimate form of such symbiosis, human capabilities would be transparently augmented by cognitive prostheses—computational systems that would leverage and extend human intellectual, perceptual, and collaborative capacities, just as a steam shovel is a sort of muscular prosthesis or eyeglasses are a sort of visual prosthesis (Ford et al., 1997; Ford, 1998; Hoffman et al., 2012).

This vision of symbiosis can be contrasted with early efforts in knowledge acquisition where our intelligent systems were somewhat like the disembodied brains shown in low-budget black-and-white science fiction movies: entities that ruled the world while floating in a glass jar tethered by wires.[1] While potentially rich in knowledge models and inferential power, their only direct experience of the world arrived through the impoverished modes of keyboard input and video display output. As a result these intelligent systems were virtually blind and helpless, having little they could realistically *learn about* and even less that they could directly *act upon*. As others in this special issue have observed, the rise of the Internet as the largest repository of knowledge on the planet has given intelligent systems immeasurably richer means to sense, learn, and interact with humans and with the myriad specialized interactive devices, sensors, and services on which people routinely rely.

However, this accumulation of human knowledge in machine interpretable form is only the beginning. Brian Gaines proposed four additional steps that would be necessary to bring the notion of symbiosis science to full fruition:

- the development of goal-directed autonomous knowledge-creating processes;
- the increasing coupling of knowledge processing entities in social networks;

E-mail address: jbradshaw@ihmc.us

[1]For a sampler of such movies, see http://grandoldmovies.wordpress.com/2011/08/20/there-ain't-nothing-like-a-brain-some-favorite-brain-movies/

1071-5819/$ - see front matter © 2012 Jeffrey M. Bradshaw. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ijhcs.2012.10.003

- the development of techniques to facilitate the synergy between human and computer knowledge processes;
- the synthesis of both into a unified system.

Let's look at progress on these steps in more detail.

2. The promise and problems of autonomous systems

Addressing the first step of developing "goal-directed autonomous knowledge-creating processes," one of Brian Gaines' students proposed in 1997 a conception of the future Internet as a "cyberorganism" consisting of "distributed intelligent agents," both human and software (Chen, 1997). Subsequently, proponents of the Semantic Web (Berners-Lee et al., 2001) envisioned that such agents would, as Mark Musen expresses it, "comb the Internet and would reason about user goals and how to achieve them" (Musen, 2013). In setting their sights on this goal, agent researchers abandoned the metaphor of the intelligent system qua disembodied brain and adopted the vision of software robots operating in a world of networked computing resources. In this change of metaphor, the research emphasis made an important shift from deliberation to doing, from reasoning to remote action.

Much of the early research on autonomous systems was motivated, not by cyber applications, but by situations in the physical world in which autonomous systems were required to "replace" human participation, thus minimizing the need for considering the human aspects of such solutions. For example, one of the earliest high-consequence applications of sophisticated agent technologies was in NASA's Remote Agent Architecture (RAA), designed to direct the activities of unmanned spacecraft engaged in distant planetary exploration (Muscettola et al., 1998). RAA was expressly designed for use in human-out-of-the-loop situations where response latencies in the transmission of round-trip control sequences from earth would have impaired a spacecraft's ability to respond to urgent problems or to take advantage of unexpected science opportunities.

Sadly, since those early days, most researchers in autonomous systems have continued to pursue their work in a technology-centric fashion, as if full autonomy—complete independence and self-sufficiency of each system—were the holy grail in every situation. Of course, there are problems like deep-space exploration where the goal of minimizing human involvement with autonomous systems can be argued effectively. However, reflection on the nature of human work reveals the shortsightedness of such a singular focus: What could be more troublesome to a group of individuals engaged in dynamic, fast-paced, real-world collaboration than a colleague who is perfectly able to perform tasks alone but lacks the skills required to coordinate his or her activities with those of others? Despite a widespread perception to the contrary, it should be noted that virtually all of the significant deployments of autonomous systems to date—e.g., military UAVs, NASA rovers, oil spill UUVs, and disaster inspection robots—have involved people in important roles, and that such involvement was not merely to make up for the current inadequacy of autonomous capabilities, but also because their jointly coordinated efforts with humans were—or should have been—intrinsically part of the mission planning and operations itself.

In view of the shortcomings of standalone autonomy for complex situations, interest has grown in the topic of "cooperative" or "collaborative" autonomy. Unfortunately, however, this research has a fundamental limitation—namely, that the kind of "collaboration" usually imagined encompasses solely the autonomous systems themselves, regrettably excluding the role of humans as potential collaborators. For example, the United States Department of Defense Unmanned Systems Roadmap stated the goal of pursuing "greater autonomy in order to improve the ability of unmanned systems to operate *independently* [i.e., without need for human intervention], either individually or collaboratively, to execute complex missions in a dynamic environment." Similar briefs have complained of the fact that because UxVs are not truly autonomous, their operation requires substantial input from remote operators. They ask whether additional research in cooperative autonomous behavior—referring to cooperation between the autonomous systems without any human element—could address this "problem."

3. Social machines and human–computer synergy

In contrast to such views, Brian Gaines never saw standalone agent autonomy as the end of the journey. He recognized that just as machine intelligence is hobbled without autonomy, so machine autonomy without sociality is reduced to mere autism. Thus, as a next step, he predicted "the increasing coupling of knowledge processing entities in social networks," a topic deftly summarized by Nigel Shadbolt in his discussion of "social machines" that embody new kinds of emergent and collective large-scale problem-solving by people who are supported by socially-contextualized machines (Shadbolt, 2013). My personal focus, however, has been primarily on the subsequent step in Brian Gaines' model, namely "the development of techniques to facilitate the synergy between human[s] and computer[s]," with the machines acting in the role of differently-abled teammates rather than of sophisticated tools.

Increased synergy between humans and autonomous systems as teammates requires a better understanding of how they become interdependent as part of joint activity. Regrettably, most methodologies for autonomous system design have not been formulated with a sufficient appreciation for the essential role of interdependence in joint human-machine activity (Johnson et al., 2010). While certain approaches to cooperative interaction between humans and machines have become widely known (e.g., dynamic function allocation, supervisory control, adaptive