Reflections on Challenges and Promises of Mixed-Initiative Interaction

Eric Horvitz

Abstract

Research on mixed-initiative interaction and assistance is still in its infancy but is poised to blossom into a wellspring of innovation that promises to change the way we work with computing systems—and the way that computing systems work with us. I share reflections about the opportunities ahead for developing computational systems with the ability to engage people in a deeply collaborative manner, founded on their ability to support fluid mixed-initiative problem solving.

People have a remarkable ability to understand, communicate, and coordinate with one another to achieve mutual goals. Such collaborative intelligence sits at the veritable heart of human civilization. In the course of daily life, we assume and rely on a rich interleaving of efforts to achieve goals while immersed in shared context. We continue to engage one another in efficient, tightly woven collaborations, reasoning with remarkable efficiency about the beliefs, preferences, intentions, and skills of potential collaborators.

The inferences underlying successful collaborations typically stream in such an effortless and subconscious manner that we often fail to recognize the elegance and sophistication of these capabilities. The magic of human collaborative competency comes to the foreground with attempts to extend these skills to computational systems. Developing a better understanding of the core aspects of intelligence that enable people to collaborate with fluidity promises to enable new kinds of human—computer collaboration.

The nascent area of research on *mixed-initiative interaction* centers on developing methods that enable computing systems to support an efficient, natural interleaving of contributions by people and computers, aimed at converging on solutions to problems. In mixed-initiative interaction, people and computers take initiatives to contribute to solving a problem, achieving a goal, or coming to a joint understanding.

Conversational dialog is an oft-cited example of mixed-initiative interaction, referring to the ability of each participant in a dialog to take initiative to guide or add to a discussion. Endowing an automated dialog system with the ability to both take initiative ("What city do you wish a flight to?") and to also allow people to take conversational initiative ("Wait, I'd like to add a side trip.") can enhance the naturalness and effectiveness of dialog. However, mixed-initiative interaction extends beyond spoken conversations to include a broad spectrum of collaborative problem solving marked by an interleaving of contributions by different participants.

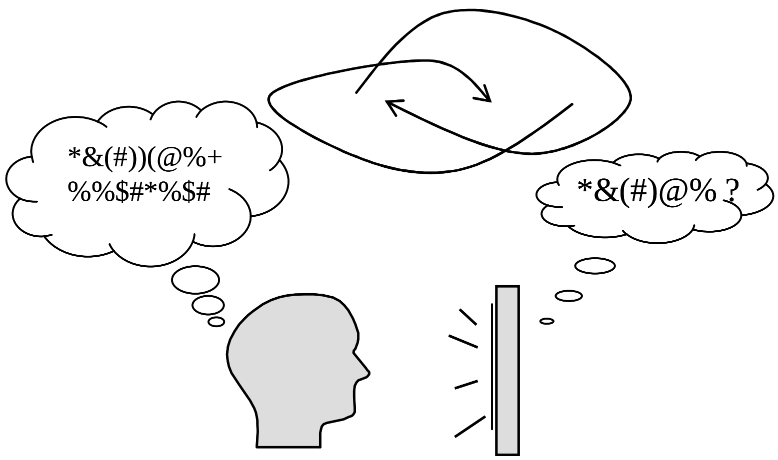

Figure 1. In pursuit of mutual understanding via grounding. There is much to be done on developing automated reasoning processes that perform efficient and effective *grounding* to develop a shared understanding among people and machines of context, beliefs, intentions, and preferences.

Mastering mixed-initiative interaction poses a constellation of fascinating challenges and opportunities for AI researchers. Figure 1 highlights the core challenge of seeking mutual understanding or *grounding* of *joint activity*. Joint activity describes the behavior displayed by people working together to solve a mutual goal. Participants in joint activity need to converge on some common understanding of beliefs about the setting, activity, goals, and the nature and timing of their individual contributions. Psychologists have referred to efforts to reach a mutual understanding or common ground on joint activity as the process of *grounding*. Challenges in grounding include the ongoing resolution of uncertainties about the focus of attention and comprehension of the participants, the nature of the problem to be solved, and about abilities and intentions to contribute to the solution in different ways.

Similar challenges of grounding are faced by people who work together to achieve goals, whether they are maneuvering an oversized piece of furniture through a doorway, docking a boat on a windy day, or are working toward understanding one another in a conversation. Effective collaboration among people relies on a sharing of context where there is a common view or "sense" of relevant aspects of the world in which the collaborators are jointly immersed, including a shared view of goals, intentions, abilities, and of causes and fluents. The automation of grounding by computing systems, via sensing, reasoning, and dialog about context and intentions, is a fundamental challenge for fluid, general mixed-initiative interaction.

Figure 2 highlights in schematic manner several challenging problems with endowing computing systems with mixed-initiative problem-solving skills. In the general case, opportunities for taking initiative to assist with problem solving may come in the absence of explicit signals from a computer user. Thus, research challenges for mixed-initiative interaction include providing systems with the abilities to recognize problem-solving opportunities, including opportunities outside the scope of someone's current focus of attention, and to understand where automated capabilities might complement human skills in solving the problems in a useful and desirable manner.

In addition to recognizing opportunities for solving problems, mixed-initiative systems may benefit from skills that enable them to decompose problems into sets of subproblems (α and β in the figure), and to consider how people and machines might each contribute in symphony or sequentially to solving the subproblems. After solving one or more subproblems, and observing the effort by a human partner on solving other subproblems, a mixed-initiative system might also contribute by helping to weave together the results of problem solving into larger solutions.

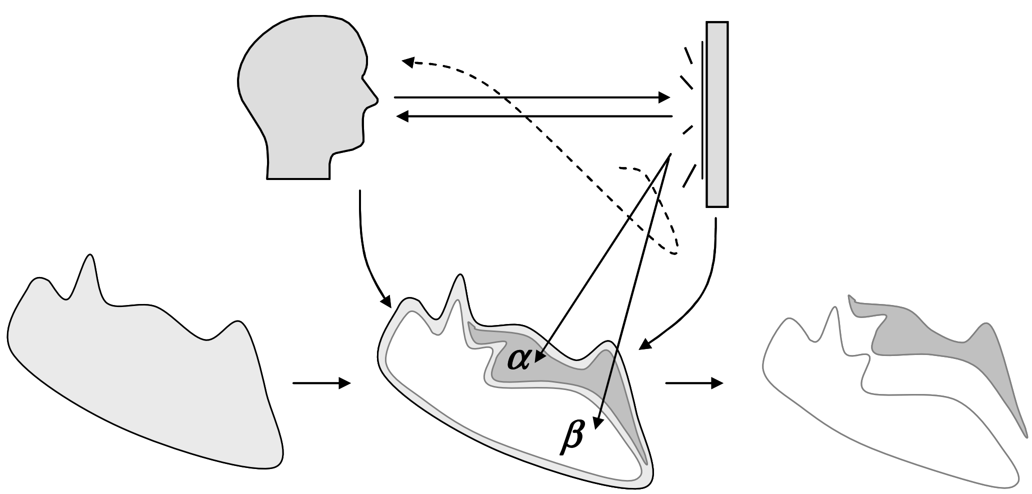

Figure 2. Challenges of mixed-initiative interaction. Mixed-initiative interaction challenges includes recognition of relevant problems, decomposition of the problems into subproblems, identification of subproblems that might be best solved via automation, solution of the subproblems, integration of human and machine contributions, and the ongoing communication and coordination about this reasoning and problem solving.

Such automation of mixed-initiative collaboration might rely on scripted plans, executed at particular points in an interaction within a relatively self-centered vacuum. However, a longer-term dream for human-computer interaction is one where quick-paced sensing, reasoning, and reacting supports an elegant problem-solving dance among parties, where the nature and timing of human and machine contributions are coordinated carefully.

Mixed-initiative assistants may often face inescapable uncertainties about human goals, about the accuracy and complementarity of computed solutions, and the overall desirability of intervening at different times. Thus, valuable mixed-initiative behavior may depend critically on machinery for making decisions under uncertainty, taking into consideration human preferences about collaboration. These methods can endow a mixed-initiative system with the ability to continue to weigh the expected costs and benefits of alternative actions (or inaction), and also to consider when to pause to better understand a situation via dialog or additional sensing.

Achieving fluid collaborations will often require efficient signaling between people and computers about the changing focus of attention and proposed contributions (highlighted by the dashed curve in Figure 2), as well as such important details as the degree of understanding or confusion about a situation, the status of problem solving, and the overall progression of the collaboration. There are opportunities to formulate sets of gestural, verbal, auditory, and graphical cues or richer languages for coordinating problem solving within specific domains or to serve as cross-application conventions. Signaling strategies might be informed by the natural, subtle coordinative signals about initiative, contribution, and comprehension that people employ when they converse or collaborate in other ways with one another.

As an example of coordinative cues, a mixed-initiative system operating in a desktop setting might communicate its assessment of the status of grounding with human collaborators via a graphic that shifts continuously from a green glow when comprehension is good, to yellow to show some confusion, and to red when understanding and joint activity is likely failing. Such signals from mixed-initiative systems could become as familiar as confirming nods, or knitted eyebrows and confused squints, from human collaborators. Beyond providing general indications of comprehension, more complex spatiotemporal patterns of cues might support fast-paced volleys of contributions from people and machines. To whet the imagination, consider the prospect of one day seeing—when peeking over the shoulder of someone engaged in a mixed-initiative session—collaborative signaling as dancing sparks of light of different colors and intensities, surrounding, filling, and highlighting representations of problems and problem solving, sharing a stream of information between the computer and user about proposals and acceptances of contributions, and indications of attention, competencies, comprehension, and progress.

Mixed-initiative systems promise to qualitatively change how it feels to work with computers. Jumping off the desktop, mixed-initiative assistants promise to weave together computational and human intelligence in the course of daily activities in ways that could significantly enhance the quality of life for both healthy and impaired people, performing such tasks as helping people to remember things they might likely forget, addressing unplanned difficulties, looking out for surprises, taking advantage of opportunities that come along, and assisting with the achievement of acute needs and long-term goals.

New possibilities for mixed-initiative systems, such as applications aimed at augmenting native human intelligence in a graceful manner, may be enabled with advances in sensing, learning, and reasoning about human cognition. General and personalized models of the operation of human attention, memory, and judgment, including such subtleties as the timing of cognitive processes, will likely be important for success.

Flowing more deeply into the world, principles of mixed-initiative interaction promise to enable new forms of tightly synchronized collaborations of people and robotic systems on physical challenges in the world. Mixed-initiative robotic systems with exquisite skills at sensing and effecting might one day work hand-in-hand with people, in parallel with human efforts or in highly coordinated exchanges of actions with people to help bolster, balance, guide, position, cut, and shape objects in the world.

Beyond leading to new kinds of collaborations between people and computers, insights about automation of mixed-initiative interaction could spawn new kinds of capabilities and applications. For example, advances in our understanding of collaborative intelligence can be expected to support more effective cooperation among autonomous systems, and thus enable new forms of computational teamwork.

Advances in mixed-initiative problem solving will be important in enabling long-imagined scenarios where one person or just a few people coordinate larger numbers of semi-autonomous systems. Today, the ratio of people to semi-autonomous systems in operational environments is best characterized as many-to-one; critical semi-autonomous robotic systems, such as unmanned aerial vehicles, are managed in real time by teams of people. Providing semi-autonomous systems with such skills as the ability to sense, infer, and understand the current and future status of the attentional focus and cognitive load of human operators, and to coordinate amongst themselves on the timing and nature of requests for guidance from people, will reduce the numbers of people required to manage constellations of semi-autonomous systems.

In another role, systems with the ability to observe and reason *about* mixed-initiative interaction among human collaborators may find diverse uses. Methods for automated understanding of joint activity and grounding—developed in the course of research on mixed-initiative systems—could be deployed in vigilant systems that look out for human safety in high-stakes situations that rely on collaboration. Consider, as an example, the prospect of deploying *mixed-initiative monitoring systems* to track the conversations and overall joint activity of pilots and air-traffic controllers. Such systems could be tasked with reasoning behind the scenes about world state, beliefs, and intentions, and with alerting people or delaying the progression of plans given the detection of a potentially costly failure of mutual understanding. A motivating and heartbreaking example is the catastrophic breakdown of mutual understanding in the largest aviation accident to date, at Tenerife, Canary Islands in 1977. The disaster highlights a number of intriguing challenges in reasoning about beliefs, intentions, and interactions among multiple participants in a high-stakes collaboration. Readers may find it an engaging exercise to review the transmissions, cockpit recordings, and the overall cascade of events, and to reflect about the challenges with designing a mixed-initiative understanding system that could have averted this catastrophe, thus saving the lives of 583 people.

It is exciting to see growing interest and an acceleration of research on mixed-initiative interaction for conversation, problem solving, and assistance. Research on methods and machinery for supporting fluid mixed-initiative interaction, whether focused on specific problems or on tackling general principles, promises to generate insights about collaborative intelligence. Advances will have numerous influences on the way that people and computing systems interact, and will undoubtedly lead to new applications of automated reasoning. We are separated from such advances by hard theoretical and practical problems. It is up to the artificial intelligence research community to tackle these challenges. As demonstrated by recent and forthcoming work, we are clearly on our way.

Bio

Eric Horvitz is principal researcher and research area manager at Microsoft Research. He pursues basic and applied research in machine learning and reasoning, information retrieval, and human-computer interaction. He is president-elect of the American Association for Artificial Intelligence. More information can be found at http://research.microsoft.com/~horvitz.

In: *AAAI Magazine* 28, Special Issue on Mixed-Initiative Assistants (2007).