Artificial Intelligence and Intelligent Systems

Pat Langley
Computational Learning Laboratory
Center for the Study of Language and Information
Stanford University, Stanford, CA 94305

John E. Laird
Computer Science and Engineering
The University of Michigan
Ann Arbor, MI 48109-2121

The original vision of AI was to develop intelligent artifacts with the same broad range of capabilities as we observe in humans. None of the field's early programs achieved this lofty goal, but the tone of many seminal papers makes this motivation very clear. When we entered graduate school at Carnegie Mellon University in the mid 1970s, this aim was a central tenet of many researchers in the field, and it was still important in many circles when we became professors in the mid 1980s. At that time, AI was still generally viewed as a single field with a common set of goals.

By the late 1980s, that situation had started to change. Subfields like machine learning, knowledge representation, and planning began to break away from AI, establishing their own conferences, journals, and criteria for progress. One of us served as an active proponent of such developments in the area of machine learning, which launched one of the first specialized journals and which played a leading role in introducing careful experimental evaluation. To researchers who were involved in these movements, these changes seemed necessary at the time for advancing the parent field.

However, the down side of this speciation was that students began to identify more with their subfield than with AI in general. They began to focus their energies on solving component problems, like supervised learning or constraint satisfaction, with little concern for how their results might be used in the context of larger AI systems. Over the past 20 years, this trend has continued unabated. Clearly, it has produced considerable technical progress within each of AI's subfields, but it has also led to a narrowness of vision among many otherwise excellent researchers.

Today, many AI practitioners consider their main affiliation to be not with AI itself but with their subfield, and their primary conference is not AAAI but one of the specialized meetings. In many cases, they lack the training to understand results in other areas or even to appreciate their goals, an effect that is exacerbated by the specialized jargons that have emerged. In fact, graduate education in AI subfields has become so specialized that the only common knowledge concerns algorithmic complexity. This trend has reached its fullest development in Carnegie Mellon's new machine learning department, which requires its graduate students to take many statistics classes but not a single course in artificial intelligence.

This change in affiliation has been accompanied with a shift in the level of problems tackled. Each subfield has defined its own set of well-defined tasks on which it measures progress. For instance, machine learning now deals almost exclusively with supervised learning for classification and with learning from delayed reward for reactive control, with other subfields having analogous problems. Despite their advantages for purposes of evaluation and, in some cases, their practical usefulness (e.g., for corporate data mining), they tell us little about the nature of intelligence.

Computer science departments, which have become the typical home for AI research, often divide themselves into areas such as systems and algorithms. When we entered the field, most of the AI researchers we knew would have said they were more like systems people than algorithms people, whereas today most would take exactly the opposite position. This reflects the current concern with component mechanisms rather than with the manner in which they work together to support intelligent behavior. We believe this attitude must change if we hope to develop intelligent artifacts with the same scope as humans.

Despite these developments, there have been some important countertrends. In particular, research on cognitive architectures (Newell, 1990; Langley et al., in press) deals with systems-level accounts of intelligent behavior. Launched in the late 1970s, this paradigm has continued through the present and it has produced some impressive examples of intelligent agents that operate in complex environments. For instance, Tambe et al. (1995) report a simulated fighter pilot, implemented within the Soar framework, that incorporates substantial knowledge about flying missions and that has been used repeatedly in large-scale military training exercises. Similarly, Trafton et al. (2005) describe an ACT-R system which controls a mobile robot that converses and interacts with humans in building environments having obstacles and occlusion.

There has also been a growing interest in interactive computer games, which constitute a promising area in which to study cognitive systems. As Laird and van Lent (2000) have argued, such games provide realistic environments that let one address many of the integration issues that arise in developing human-level AI systems. Moreover, they hold great attraction for many students and avoid the hardware challenges associated with robotics research. The recently launched conference on AI and Interactive Digital Entertainment (http://www.aiide.org/) builds on this insight, and it definitely constitutes a move in the right direction.

Intelligent tutoring systems also have many advantages for the study of integrated cognitive systems. They require not only knowledge about the target domain, but also the ability to infer student knowledge states and to guide tutorial dialogues that achieve instructional goals. Some recent tutoring systems (e.g., Rickel et al., 2002) that deal with emotionally-charged social situations are especially compelling and raise intriguing issues at the intersection of cognition and affect. Research on such integrated agents fulfills the spirit of AI's original vision for intelligent systems.

Recently, other evidence has emerged of increased interest in combining the pieces of our fragmented field. DARPA IPTO has funded a number of innovative programs that explicitly involve work on cognitive systems (Brachman & Lemnios, 2002) and that involve integration of mechanisms which are generally studied in isolation. Two other positive signs are the newly established AAAI conference track on integrated systems and a recent Fall Symposium (Cassimatis et al., 2006) on achieving human-level intelligence through research at the integrated systems level.

We are not recommending that AI abandon its research specialties, which have produced undeniable scientific progress. Rather, we are encouraging them to take the broader view of considering problems that require integration across specialties. Working together, they should direct their efforts toward building intelligent agents that incorporate a broad range of capabilities, use knowledge from a variety of sources (preprogrammed, learned from experience, and acquired from other agents), and operate in complex dynamic environments that cannot be fully predicted in advance.

We need many more efforts along these lines to tackle AI's original dream of building complete intelligent systems. In parallel, we must work more diligently at communicating that vision to younger researchers who may not even realize its historical importance to our field. Universities should offer more courses that address the issues that arise in building integrated agents and that give students experience with developing them. More generally, we must convince young scientists that research at the systems level is key to understanding the nature of the mind.

In summary, the original aim of AI was to understand the nature of intelligence by constructing artifacts that exhibited the same breadth and depth of cognition as humans. The fragmentation of AI into specialized subfields has produced powerful component methods, but it has also distracted many researchers and educators from the field's initial vision. The focus on standalone algorithms has helped us understand the elements of cognition, but it has told us little about how they might work together. The fiftieth anniversary of AI's establishment is an appropriate time to begin reversing that trend and to redirect our energies toward the construction of integrated cognitive systems that exhibit broad, human-level intelligence.

References

Brachman, R., & Lemnios, Z. (2002). DARPA's new cognitive systems vision. *Computing Research News*, *14*, 1.

Cassimatis, N. L., Mueller, E. K., & Winston, P. H. (2006). Achieving human-level intelligence through integrated systems and research. *AI Magazine*, *27*, 12–14.

Laird, J. E., & van Lent, M. (2000). Interactive computer games: Human-level AI's killer application. *Proceedings of the Seventeenth National Conference on Artificial Intelligence* (pp. 1171–1178). Austin, TX: AAAI Press.

Langley, P., Laird, J. E., & Rogers, S. (in press). Cognitive architectures: Research issues and challenges. *Cognitive Systems Research*.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Rickel, J., Marsella, S., Gratch, J., Hill, R., Traum, D., & Swartout, B. (2002). Towards a new generation of virtual humans for interactive experiences. *IEEE Intelligent Systems*, July/August, 32–38

Tambe, M., Johnson, W. L., Jones, R. M., Koss, F., Laird, J. E., Rosenbloom, P. S., & Schwamb, K. B. (1995). Intelligent agents for interactive simulation environments. *AI Magazine*, *16*, 15–39.

Trafton, J. G., Cassimatis, N. L., Bugajska, M., Brock, D., Mintz, F., & Schultz, A. (2005). Enabling effective human-robot interaction using perspective-taking in robots. *IEEE Transactions on Systems, Man and Cybernetics*, *25*, 460–470.