Achieving Human-Level Intelligence through Integrated Systems and Research

Introduction to

This Special Issue

Nicholas Cassimatis, Erik T. Mueller,
and Patrick Henry Winston

■ This special issue is based on the premise that in order to achieve human-level artificial intelligence researchers will have to find ways to integrate insights from multiple computational frameworks and to exploit insights from other fields that study intelligence. Articles in this issue describe recent approaches for integrating algorithms and data structures from diverse subfields of AI. Much of this work incorporates insights from neuroscience, social and cognitive psychology or linguistics. The new applications and significant improvements to existing applications this work has enabled demonstrates the ability of integrated systems and research to continue progress towards human-level artificial intelligence.

In many subfields of artificial intelligence during the past several decades, there has been substantial progress that has resulted in significant near-term advances in theory and applications. However, we believe that progress towards human-level artificial intelligence and the applications it enables requires a deeper and more comprehensive understanding that cannot be achieved by studying individual areas in isolation. Two reasons, both involving integration, support this belief. First, many problems that human-level AIs must solve involve subproblems currently addressed by different subfields, often using very different computational methods. A human-level AI must either integrate, for example, backtracking search, partially observable Markov decision processes (POMDPs), logic theorem proving algorithms, productions systems, and neural networks, or it must be based on new, heretofore undiscovered computational methods that exhibit all of the best features of these computational methods. Second, the best currently existing example of human-level intelligence is of course the human being. We believe that the history of AI (and the articles in this special issue) demonstrate that insights into the mechanisms underlying human cognition can inform research towards human-level AI. To explore a more integrated approach to human-level AI, we helped organize the 2004 AAAI Fall Sympo-

sium on Achieving Human-Level Artificial Intelligence through Integrated Systems and Research.

The symposium and the articles in this issue are motivated by four recent trends that we believe present an opportunity to reinvigorate the field's focus on understanding and developing human-level intelligence. First, many new subfields of artificial intelligence are explicitly directed towards integration. For example, many in the robotics community (see, for instance, the article by Craig Schlenoff, Jim Albus, Elena Messina, Tony Barbera, Raj Madhavan, and Stephen Balakirsky in this issue) are attempting to integrate the complexity of behavior enabled by reasoning and planning algorithms with the robustness and flexibility of reactive robotic architectures. Second, many new application domains require systems either to behave at the level of humans or to understand the complexities of human behavior. For example, the goal of increasing the realism of synthetic characters in training simulators (see the article by William Swartout, Jonathan Gratch, Randall Hill, Eduard Hovy, Stacy Marsella, Jeff Rickel, and David Traum) creates a need for artificial agents that behave at a human level of intelligence. Third, many funding agencies have recently renewed their interest in systems that go beyond the level of intelligence afforded by individual computational methods. Many of the programs resulting from the Defense Advance Research Project Agency's (DARPA) recent focus on cognitive systems, for example, aim to achieve a level of reasoning, deliberation, and learning normally associated with human intelligence. Finally, the ever increasing amount of computational power available to the field enables large-scale integration efforts that were heretofore impractical.

The articles in this issue report research that capitalizes on these trends by integrating multiple computational methods and by taking inspiration from recent discoveries about human cognition. As the work reported in these articles demonstrates, this approach has led both to deeper insights into intelligent systems and, consequently, to significant qualitative advances in the applications they enable.

Architectures That Combine the Insights of Multiple Subfields of AI

Subfields in AI are often based on different computational methods that are difficult to reconcile. However, since, for example, POMDPs, case-based reasoners, and logic-theorem provers each have strengths not shared by the others, human-level AI researchers must often find ways to integrate these different methods into a single intelligent system. The special issue therefore contains a focus on architectures for achieving this integration.

Pat Langley provides an overview of cognitive architectures and their role in achieving human-level intelligence, in addition to a survey of past modeling efforts and how they relate to the present.

Randolph Jones and Robert Wray review common elements and differences among several popular agent architectures. Their work facilitates the evaluation of current agent architectures and provides a systematic framework within which to explore new ones.

Nicholas Cassimatis addresses the problem of the large array of procedures and knowledge that seem to be required for human-level artificial intelligence. He identifies a relatively small *cognitive substrate* of procedural and representational mechanisms that underlie most AI algorithms and knowledge representation formalisms. He argues that this substrate can greatly reduce the complexity of integrating many algorithms and diverse knowledge representations into one system.

The integration of reasoning and planning with perception and action is a difficult problem in robotics and other branches of AI. Craig Schlenoff et al. present an architecture to achieve this in the automated driving domain.

Richard Granger presents a framework for integrating the benefits of parallel neural hardware with more serial and symbolic processing in a way that is faithful to actual human neural processing and enables advances in perceptual processing.

The lifelike synthetic characters in William Swartout et al.'s training simulations are based on an architecture for combining computational methods used to model emotion, reasoning, and planning.

The architecture underlying Kenneth Forbus and Thomas Hinrichs's work is based on analogy and therefore a key part of integrative AI research because analogies enable computation in one domain to be used for another.

Inspiration from Actual Human Intelligence

Humans are the best example of human-level intelligence. Recent work in multiple branches of cognitive and neural science motivate new computational approaches to achieving human-level AI.

The computational architectures that enable the application advances in Granger's and in

Schlenoff et al.'s work are motivated by recent discoveries in neuroscience.

Swartout et al. are able to create more realistic simulated characters in training simulations by basing their work on a theory of how the human brain integrates emotion and cognition.

Cassimatis surveys developments in cognitive development, linguistics, and neuroscience that suggest a basic set of cognitive abilities underlie most of human cognition. He discusses how to convert these insights into a computational framework for achieving more integrated cognitive systems.

Langley, Jones, and Wray discuss cognitive and agent architectures that are inspired by and designed to conform to data on human cognition.

Human-Level AI Enables Profound Advances in the Power and Scope of Applications

Applications drive and validate the work presented in this special issue in three ways. First, many application domains require significant integration (for example, robots that must integrate reasoning, vision, and motion planning). Second, the ability to enable new or significantly improved applications is a motivation for achieving human-level AI. Finally, many applications (for example, training simulators) require humanlike behavior and thus drive the need for human-level intelligence. Applications discussed in this issue also include automated driving, robust machine perception, and companion cognitive systems.

The work reported in the articles in this special issue demonstrate that researchers can make substantial progress towards achieving human-level artificial intelligence by integrating the approaches of multiple subdisciplines in artificial intelligence and by drawing inspiration from the study of human intelligence.

Nicholas Cassimatis is an assistant professor of cognitive science at the Rensselaer Polytechnic Institute. He received his B.S. in mathematics from the Massachusetts Institute of Technology, his M.A. in psychology from Stanford University, and his Ph.D. from the MIT Media Laboratory. The overall goal of his research is to understand and reproduce in computers human-level intelligence. This has led him to do work in robotics, machine learning, cognitive development, physical reasoning, language understanding, and social reasoning.

Erik T. Mueller is a research staff member at the IBM Thomas J. Watson Research Center. He is the author of two books and creator of the ThoughtTreasure commonsense knowledge base and architecture for natural language processing. He received a B.S. in computer science and engineering from the Massachusetts Institute of Technology and an M.S. and Ph.D. in computer science from the University of California, Los Angeles. His research focuses on commonsense reasoning and its application to story understanding and intelligent user interfaces.

Patrick Henry Winston is a professor of computer science in the CSAIL Laboratory and Electrical Engineering and Computer Science Department at MIT. Winston received a B.S. and M.S. in electrical engineering and Ph.D. in computer science from the Massachusetts Institute of Technology. Winston's Genesis Group studies how vision, language, and motor faculties account for intelligence, integrating work from several allied fields, including not only AI, but also computer science, systems neuroscience, cognitive science, and linguistics. His EWall Group focuses on new ways to visualize information, making that information easier to comprehend, analyze, and communicate to others for learning and decision making. He is the author of many textbooks and edited volumes on AI and programming.