Discovery Systems Check Their Own Facts

James Figueroa

Scientists have long sought an AI system that completely automates the discovery process: a system that forms its own hypotheses based on raw data, runs tests, and makes adjustments based on results, much like human scientists in a laboratory.

Early AI projects such as Automated Mathematician (AM) could fulfill only part of that idea, making calculations from large data sets but stopping short at the evaluation step. Decades after AM set the stage for discovery systems with its mathematical concepts in the 1970s, two recent university projects have shown greater promise for completely automated science.

Reports on both projects appeared in the 3 April Science (http://sciencenow.sciencemag.org/cgi/content/full/2009/402/1). One is a Cornell University system that identified classic physics concepts such as Hamiltonian mechanics solely on the basis of data observed in physical systems, including a double pendulum. The other is a British “robot scientist” called Adam that discovered new genomics information from baker’s yeast. A commentary from two computer science professors, Columbia University’s David Waltz, an IS advisory board member, and the University of Pittsburgh’s Bruce Buchanan, accompanied the reports.

“This is really carrying out the full cycle of hypothesis evaluation, test generation, and then completing again in a circle,” Waltz said in a Science podcast (www.sciencemag.org/cgi/content/full/sci;324/5923/DC1), stressing that the projects “close the loop” on discovery systems. “These programs can operate not just from a start to some sort of completion, and then you look at the results—they actually can operate essentially infinitely.”

The leap could prove key to helping scientists contend with volumes of data that would overwhelm traditional analysis techniques. Waltz and Buchanan see the projects as important breakthroughs during a time when supercomputers and the Internet have enabled new levels of automation.

“Overall, science is changing the way it’s done, in part because we have vastly more data that we can see and gather and vastly more powerful machines that we can work on it with, and much more storage,” Waltz said.

Deducing Centuries of Science

The Cornell University project developed from an effort to model biological systems. Computer science professor Hod Lipson and doctoral student Michael Schmidt found that their efforts to build explicit and differential equations broke down when they tried to model implicit or invariant relationships. So they set out to develop a system that could meet the challenge.

The program they created uses motion-tracking technology to collect data from physical systems commonly found in physics laboratories, including harmonic oscillators and double pendulums (see Figure 1). Without any information about physics, kinematics, or geometry, the program automatically detected...
several natural laws on the basis of its computations, capturing centuries of mathematical advancement within roughly a day’s time.

Lipson and Schmidt expect the breakthrough to help scientists discover unknown rules governing other physical phenomena, an area with many potential applications.

“Essentially this sort of algorithm would be useful anywhere where there is a theoretical gap despite abundance of data,” Lipson said. “There are many such areas, from cosmology to particle physics to biology. We can even start looking for quantitative laws in areas like social behavior.”

The researchers also foresee that using machines for mathematical details could give scientists more freedom to consider creative and conceptual work; scientists could form conceptual frameworks to find predictive explanations for observed phenomena, and the AI program could work within that framework to find a solution.

Lipson and Schmidt have already moved ahead to fulfill that promise, and say that the program has found a new predictive biological law as part of yet-to-be-published research.

Their currently published work features algorithms that used bootstrap-
genes in baker’s yeast are the engine for enzymes that catalyze biochemical reactions.

The system followed through on its hypothesis from start to finish in a process called active learning. Using robotics, Adam formulated its own experiments to test the hypothesis, performed the necessary steps, and used a plate reader to analyze the results, repeating itself when needed. Researchers checked the final results once the experiments were done to confirm that Adam was correct.

The robot’s results don’t represent a major genomics breakthrough—scientists say it’s roughly on par with a graduate student’s work—but the results mean big things for automation. Laboratory robotics has increased productivity in recent years but the concurrent increase in results has created what Aberystwyth researchers say is an interpretation bottleneck. With scientists struggling to analyze an overabundance of experiments, Adam’s goal is automated understanding.

“Because biological organisms are so complex, it is important that the details of biological experiments are recorded in great detail,” said professor Ross King, who headed the research. “This is difficult and irksome for human scientists, but easy for robot scientists.”

Adam’s capabilities show just how easy that could be. The system can perform more than 1,000 new experiments each day, with experiments lasting up to four days, using more than 50 yeast strains.

The Aberystwyth team has already taken what it learned from building Adam and created a new machine, Eve, that will search for new types of drugs to combat diseases such as malaria and schistosomiasis. It’s the sort of work that King says would prove ideal for future robot scientists.

“If science was more efficient, it would be better placed to help solve society’s problems,” King said. “One way to make science more efficient is through automation. Automation was the driving force behind much of the 19th and 20th century progress, and this is likely to continue.”

**AI’s Future Role**

Although the projects represent important breakthroughs in computation, scientists don’t believe that robots will take over their jobs anytime soon. Adam and similar systems were designed as complementary tools for scientists to consider questions they couldn’t attempt to answer before.

“One of the main objectives of the research is to make science more efficient and therefore to speed up innovation,” King said. “I hope that the robot scientist idea will be widely taken up and that it achieves this objective.” Lipson’s vision for his natural-laws algorithm seems to share King’s idea regarding scientific impact.

“I think we are looking at a new age of discovery,” he said. “Just like design automation allows engineers to delegate some of the more mundane design tasks to computers and focus on higher-level creative work, so can algorithms of this sort allow scientists to focus on developing new conceptual frameworks, and use computers to see if these frameworks help explain data.”

Waltz sees automated science as a complement for scientific work that will require greater understanding of computation as well as the ability to find patterns with AI.

“I think that perhaps these papers will inspire others to try to do something similar,” he said. “There is work in other areas that I think could count as belonging to the same space: in particular, the astronomical databases that are truly enormous, and I think people are mining that data for some kind of understanding of structure. Ultimately, I think Earth and planetary sciences, measurements of Earth itself, and trying to model climate or weather could be another area where such methods could be used profitably.”
Supercomputer to Answer Jeopardy Challenge

James Figueroa

IBM, which made history in 1997 when its Deep Blue supercomputer defeated chess champion Gary Kasparov, has its sights set on another mind-bending goal. The company announced in late April that it’s putting the finishing touches on a machine that will compete against human contestants on the game show Jeopardy, a project that originated from IBM’s involvement in the Open Advancement of Question Answering (OAQA) initiative.

The new system is called Watson, in honor of IBM founder Thomas J. Watson Sr. (see Figure 2). It represents a potentially major step in computer intelligence and interaction with humans: a system that can understand complex questions expressed in human terms and parse language nuances such as puns and wordplay. It’s an ambitious step forward in the question-answering (QA) field, one that will require technological achievements in natural language processing, information retrieval, knowledge representation and reasoning, and machine learning. IBM engineers are anxious to find out how well the computer will play.

“Watson is a computer system that is going to advance the state of the art in automatic question answering,” Watson project leader David Ferrucci said in a video to promote the supercomputer (www.youtube.com/watch?v=3e22ufcqfTs). “Under the hood in Watson is a natural language processing technology that’s going to advance the field. Jeopardy is a great showcase for that kind of technology, because what Jeopardy requires is that the computer competes with some of the best humans in the world, minds that can very rapidly access a huge breadth of knowledge, deliver precise answers—upwards of 85 or 90 percent precision—and deliver that with really great confidence.”

IBM is keeping many of the technical details about the supercomputer under wraps, but did reveal that it would be built through its Blue Gene architecture project and use the Unstructured Information Management Architecture (UIMA) framework for its analytic components.

Details of the competition have yet to be ironed out. Before staging a taped show, IBM and Jeopardy producers are planning a series of test matches this year to determine how well the human-versus-machine setup works in production. Developers expect Watson to be self-contained on the Jeopardy set, relying on its own knowledge base and natural language text without any help from an Internet connection. During the game, clues will be submitted to the machine as electronic text at the same time they’re revealed to human contestants. To beat its opponents, Watson must determine the correct response and submit its answer (via a voice synthesizer) within five seconds.

“While computers have demonstrated that they can quickly recall documents based on pre-indexed keywords, knowing that a term from the potentially thousands of returned results correctly answers the question is a whole other ball game,” IBM explained on its project Web site (www.research.ibm.com/deepqa/index.shtml). “It requires on-the-fly deep analysis of large volumes of language and the production of accurate probabilities that a term or combination of terms is the right answer—all in time to buzz.”

Jeopardy producers are already considering human contestants to pit against the machine, including a possible match against Ken Jennings, who won a record 74 consecutive times on the show in 2004.

Up to the Challenge
IBM researchers have been planning the system for nearly two years as part of the DeepQA project, developing a massively parallel computing platform that would have business applications beyond the Jeopardy appearance.

“The challenge is to build a system that, unlike systems before it, can rival the human mind’s ability to determine precise answers to natural language questions and to compute accurate confidences in the answers,” Ferrucci said. “This confidence-processing ability is key. It greatly distinguishes the IBM approach from conventional search, and is critical to implementing useful business applications of question answering.”

The confidence factor was identified early in Watson’s preparation, going back to the project’s origins as part of the OAQA initiative. In a December 2008 report drawing from IBM’s work with Carnegie Mellon and other universities, researchers identified speed, accuracy, and confidence as the most critical metrics to build into a Jeopardy-playing machine. Those requirements meant that the DeepQA team needed the ability to completely rethink many of its algorithms and engineering approaches, including techniques such as deep parsing and information extraction.

The Jeopardy system was one of five challenge problems that the OAQA team considered during its workshop, meant to spur advances that had only
been hinted at in question-answering research programs such as the Advanced Question Answering for Intelligence and evaluations such as the Text Retrieval Conference (TREC). The other challenges included a TREC task to answer 500 natural language questions derived from one million news articles (a week’s worth) and a sustained investigation involving a series of questions to arrive at a complete intelligence report.

The Jeopardy challenge, however, stood out because of the game show’s broad appeal and popularity in the US and other countries, making it a natural fit as IBM’s next challenge problem after Deep Blue.

**Answering Machines**

The general population is sure to compare Watson to popular imaginings of AI such as HAL 9000 in 2001: A Space Odyssey, but IBM is quick to say that its system will have more in common with the unobtrusive and reliable QA computer featured on Star Trek. The sort of human-computer dialog envisioned on the television show is along the same line as IBM’s business intelligence goals.

According to Ferrucci, Watson’s performance on Jeopardy will be a measuring stick for applying innovative QA techniques to business applications, possibly changing how people find and use information. The supercomputer’s technology could eventually find its way into help desks, Web self-service applications, and regulatory-compliance systems.

“Watson is a compelling example of how the planet—companies, industries, cities—is becoming smarter,” IBM CEO officer Samuel Palmisano said. “With advanced computing power and deep analytics, we can infuse business and societal systems with intelligence.”

![Figure 2. IBM’s Watson. IBM scientist and project director David Ferrucci presents Watson, a question-answering supercomputer in development to compete on the game show Jeopardy. (photo courtesy of IBM)](image-url)