Digital Dullard

Billionaire Paul Allen's latest project to build electronic science tutors falls short

By Steven Cherry

Aristotle was one of the world’s greatest thinkers ever. Digital Aristotle, on the other hand, knows as much about chemistry as a reasonably bright high school student, and nothing else. Yet a Seattle boutique investment firm, Vulcan Inc., is spending millions of dollars over the next few years trying to turn Digital Aristotle into, well, a digital Aristotle.

And why not? That's just a fraction of a percent of Vulcan founder Paul G. Allen’s net worth. Even a noble failure would surely be at least as worthwhile as the Portland Trail Blazers professional basketball team, another of Allen's many interests. Allen, whose fortune derives from his standing as Bill Gates's original partner in Microsoft Corp., created Vulcan in 1986 to manage his investments.

Digital Aristotle began in 2003 as a contest, dubbed Project Halo. Three sets of high-powered researchers competed to create software that could do well on a high school advanced-placement exam in chemistry. They all succeeded. The winning program, written by a collaborative team from SRI International, in Menlo Park, Calif.; the University of Texas at Austin; and Boeing Phantom Works, in Seal Beach, Calif., scored a 3.00 on the exam out of a possible 5.00. That's better than the human student median grade of 2.82.

It's worth noting, though, that the program "learned" the knowledge contained in 71 printed pages of chemistry information at a cost of about US $10 000 per page. In fact, each of the three competing teams spent that, making the total outlay for knowledge acquisition a whopping $2.1 million. Much of the cost was for the salaries of artificial intelligence researchers, but one team also employed expert chemists.

Now, for a second stage of the contest, the same three teams are designing software tools that would allow Ph.D. graduate students to create collections of facts and...
inferences, so-called knowledge bases, much more cheaply. These tools would turn ordinary sentences of scientific knowledge—a definition of electrical resistance, for example, or the fact that all mammals are vertebrates—into what are called "knowledge constructs": well-defined concepts and quasi-mathematical relationships among them. Once the constructs have been collected and stored, tried-and-true problem-solving methods and other AI technologies could be brought to bear on the task of answering questions on the chemistry test, for example.

Of course, it will take millions of dollars more to build those tools. But for Allen, the third richest man in the world, according to Forbes magazine's most recent list, that's just the change that drops down behind the sofa cushions. And if Allen is the wealthiest knowledge suitor to be smitten by the charms of AI, he's hardly the first. That honor might go to Aristotle himself, the human one. Some 2200 years ago, he dreamed of an "instrument" that "could accomplish its own work, obeying or anticipating the will of others." At the dawn of the computer age, Alan Turing dreamed of a machine so humanlike that a panel of judges wouldn't be able to distinguish it from a real person. Since then, AI has had more flashes in the pan than a French restaurant.

TO BE SURE, AI HAS ITS SUCCESSES. Factory robots use machine vision to track parts. Automotive suspension systems and camcorders use fuzzy logic to smooth out jarring motions. Hospitals use large knowledge bases of drug effects and interactions to ensure that prescribed drugs don't conflict with one another. Computer programs now repeatedly beat the world's chess champions. Part of AI's image problem stems from the fact that whenever a development moves from lab to market, it's no longer artificial intelligence; it's just software.

Still, these successes have something in common: each mimics a human capability in an artificially and greatly narrowed domain. Robots "see" stages of a predefined assembly process; they couldn't tell a rabbit from a fox running through the woods. The camcorder system isn't a general way of holding steady a variety of electronic devices, let alone a martini glass at a crowded cocktail party. And chess-playing computers still can't hold an intelligent conversation about the game, or anything else for that matter.

Similarly, in Project Halo's first contest, Noah Friedland, the program manager responsible for the project at Vulcan, stuck to a very well-defined problem, the chemistry advanced-placement exam. Even Halo's second phase is limited to the hard sciences—physics, chemistry, and biology. Indeed, the chosen task is almost exactly the same, answering advanced-placement-type questions in a narrow domain. The goal now is to slash the per-page cost.

DIGITAL ARISTOTLE

GOAL: Use artificial intelligence techniques to create electronic tutors for physics, chemistry, and biology students

WHY IT'S A LOSER: Despite a nominal goal of solving a real-world problem, Digital Aristotle is a narrowly defined proof-of-concept project that won't create a commercially viable product or enterprise

ORGANIZATION: Vulcan Inc.

CENTER OF ACTIVITY: Seattle

NUMBER OF PEOPLE ON THE PROJECT: One full-time researcher and three small teams of contract researchers

BUDGET: Confidential

It seems impossible, though, to escape from the need for ever more general knowledge. Referring to the test scores of the first-phase Halo systems, Friedland admitted that "one of the main reasons for point loss was that the project team engineers didn't know enough chemistry—for example, classifying something as an acid irrespective of the
solution it's in."

Significantly, the SRI-led winning team used professional chemists, while the two others did not. "With SRI, a team of chemists in Menlo Park turned the textbook into 35 pages of rules," Friedland says. "They then shipped it to a knowledge engineering team in Austin, which rendered it into knowledge constructs."

"Even that had problems," says Friedland. "The chemists made one set of assumptions about rule formation, not articulating all the assumptions. SRI found these assumptions caused many failures."

Despite such problems, some view the Vulcan project as impressive. Chris Welty, a researcher in the IBM Thomas J. Watson Research Center's Knowledge Structures Group who has been following the progress of Digital Aristotle, says, "It may be that one of the grand challenges of AI is to make systems that can perform tasks that they were not designed for, but I don't believe Project Halo has such an ambition at this stage. With that qualification, I think the initial project was hugely successful."

IN THE NARROW WORLD of AI research, success is often measured in research papers that advance the field. There, at least, Project Halo does some new and interesting things. "Our methodology was well received," Friedland crows. "We had a peer-reviewed paper at KR2004"—the Ninth International Conference on Principles of Knowledge Representation and Reasoning. It's "a very competitive conference. Only one of every six papers was accepted. Ours was one of the top five papers."

One way Halo exemplified an advance in methodology, Friedland notes, was that in the chemistry test, the AI systems had to explain their answers. "The exams were graded not just on the correctness of the answer but on the quality of the explanation," Friedland says. "That's very important. To build trust with users, you have to have an explanation mechanism." Another novel element was a "taxonomy of failures." Halo researchers looked to see just what caused the programs to lose exam points.

It's something the tight-knit AI community has taken notice of. "My impression was that the Halo team did an extremely thorough job evaluating the results of the first phase," says IBM's Welty.

In Project Halo's second phase, the three teams from the first round will build knowledge-acquisition tools that will let Ph.D. students in physics, chemistry, and biology pass on knowledge of their subject areas.

The software tools they construct will have to be able to turn classroom learning into the knowledge rules that an AI machine has to have in order to, say, pass an advanced-placement exam. If the software can answer ordinary-language science questions, it can serve as a digital tutor for students taking science classes.

Friedland says, "If the systems can produce comparable results [to] Halo I, that would be a major achievement. There will be no knowledge engineers, only the Ph.D. students in the field. We will have brought the cost down to that of building a textbook, which is in the range of $500 to $1000 per page."

"We are on the verge of democratizing knowledge formulation," he declares.

But the knowledge engineering will be eliminated only by building it into the front end—the knowledge-acquisition system into which the Ph.D. students will feed what they know. "Imagine you have a textbook or scientific paper and a Halo II application," Friedland says. "The tool would guide the educator through the process of capturing the knowledge in that document." How many millions of dollars—exactly what minuscule fraction of Allen's wealth—will be spent building the front end? Friedland declined to say.

Paul Allen's successful investments—Charter Communications, The Sporting News, and
Ticketmaster, to name a few—have something in common: none started out as a research project. One that did, a Palo Alto, Calif., technology and marketing research enterprise called Interval Research Corp., provided gainful employment to some of the brightest minds in Silicon Valley.

Interval lasted eight years before running through its allocation of $100 million. In that time, it tried to spin off three separate companies, each of which was a short-lived commercial failure. Interval itself exists today as a line on people’s résumés, and as a 42-meter-long collection of corporate documents in the bowels of the department of special collections at Stanford University in California.

If Allen's experience with Interval Research is any guide, Digital Aristotle will result in no successful commercial products or spinoffs. Whether its research does more than gather library dust is anyone’s guess.*

ILLUSTRATION: DAVID PLUNKERT