February 16, 2011

Computer Wins on 'Jeopardy!': Trivial, It's Not

By **JOHN MARKOFF**

YORKTOWN HEIGHTS, N.Y. — In the end, the humans on "Jeopardy!" surrendered meekly.

Facing certain defeat at the hands of a room-size I.B.M. computer on Wednesday evening, Ken Jennings, famous for winning 74 games in a row on the TV quiz show, acknowledged the obvious. "I, for one, welcome our new computer overlords," he wrote on his video screen, borrowing a line from a "Simpsons" episode.

From now on, if the answer is "the computer champion on "Jeopardy!," the question will be, "What is Watson?"

For I.B.M., the showdown was not merely a well-publicized stunt and a $1 million prize, but proof that the company has taken a big step toward a world in which intelligent machines will understand and respond to humans, and perhaps inevitably, replace some of them.

Watson, specifically, is a "question answering machine" of a type that artificial intelligence researchers have struggled with for decades — a computer akin to the one on "Star Trek" that can understand questions posed in natural language and answer them.

Watson showed itself to be imperfect, but researchers at I.B.M. and other companies are already developing uses for Watson's technologies that could have a significant impact on the way doctors practice and consumers buy products.

"Cast your mind back 20 years and who would have thought this was possible?" said Edward Feigenbaum, a Stanford University computer scientist and a pioneer in the field.

In its "Jeopardy!" project, I.B.M. researchers were tackling a game that requires not only encyclopedic recall, but also the ability to untangle convoluted and often opaque statements, a modicum of luck, and quick, strategic button pressing.

The contest, which was taped in January here at the company's T. J. Watson Research Laboratory before an audience of I.B.M. executives and company clients, played out in three

televised episodes concluding Wednesday. At the end of the first day, Watson was in a tie with Brad Rutter, another ace human player, at $5,000 each, with Mr. Jennings trailing with $2,000.

But on the second day, Watson went on a tear. By night's end, Watson had a commanding lead with a total of $35,734, compared with Mr. Rutter's $10,400 and Mr. Jennings's $4,800.

Victory was not cemented until late in the third match, when Watson was in Nonfiction. "Same category for $1,200," it said in a manufactured tenor, and lucked into a Daily Double. Mr. Jennings grimaced.

Even later in the match, however, had Mr. Jennings won another key Daily Double it might have come down to Final Jeopardy, I.B.M. researchers acknowledged.

The final tally was $77,147 to Mr. Jennings's $24,000 and Mr. Rutter's $21,600.

More than anything, the contest was a vindication for the academic field of artificial intelligence, which began with great promise in the 1960s with the vision of creating a thinking machine and which became the laughingstock of Silicon Valley in the 1980s, when a series of heavily financed start-up companies went bankrupt.

Despite its intellectual prowess, Watson was by no means omniscient. On Tuesday evening during Final Jeopardy, the category was U.S. Cities and the clue was: "Its largest airport is named for a World War II hero; its second largest for a World War II battle."

Watson drew guffaws from many in the television audience when it responded "What is Toronto?????"

The string of question marks indicated that the system had very low confidence in its response, I.B.M. researchers said, but because it was Final Jeopardy, it was forced to give a response. The machine did not suffer much damage. It had wagered just $947 on its result. (The correct answer is, "What is Chicago?")

"We failed to deeply understand what was going on there," said David Ferrucci, an I.B.M. researcher who led the development of Watson. "The reality is that there's lots of data where the title is U.S. cities and the answers are countries, European cities, people, mayors. Even though it says U.S. cities, we had very little confidence that that's the distinguishing feature."

The researchers also acknowledged that the machine had benefited from the "buzzer factor."

Both Mr. Jennings and Mr. Rutter are accomplished at anticipating the light that signals it is possible to "buzz in," and can sometimes get in with virtually zero lag time. The danger is to

buzz too early, in which case the contestant is penalized and "locked out" for roughly a quarter of a second.

Watson, on the other hand, does not anticipate the light, but has a weighted scheme that allows it, when it is highly confident, to hit the buzzer in as little as 10 milliseconds, making it very hard for humans to beat. When it was less confident, it took longer to buzz in. In the second round, Watson beat the others to the buzzer in 24 out of 30 Double Jeopardy questions.

"It sort of wants to get beaten when it doesn't have high confidence," Dr. Ferrucci said. "It doesn't want to look stupid."

Both human players said that Watson's button pushing skill was not necessarily an unfair advantage. "I beat Watson a couple of times," Mr. Rutter said.

When Watson did buzz in, it made the most of it. Showing the ability to parse language, it responded to, "A recent best seller by Muriel Barbery is called 'This of the Hedgehog,' " with "What is Elegance?"

It showed its facility with medical diagnosis. With the answer: "You just need a nap. You don't have this sleep disorder that can make sufferers nod off while standing up," Watson replied, "What is narcolepsy?"

The coup de grâce came with the answer, "William Wilkenson's 'An Account of the Principalities of Wallachia and Moldavia' inspired this author's most famous novel." Mr. Jennings wrote, correctly, Bram Stoker, but realized that he could not catch up with Watson's winnings and wrote out his surrender.

Both players took the contest and its outcome philosophically.

"I had a great time and I would do it again in a heartbeat," said Mr. Jennings. "It's not about the results; this is about being part of the future."

For I.B.M., the future will happen very quickly, company executives said. On Thursday it plans to announce that it will collaborate with Columbia University and the University of Maryland to create a physician's assistant service that will allow doctors to query a cybernetic assistant. The company also plans to work with Nuance Communications Inc. to add voice recognition to the physician's assistant, possibly making the service available in as little as 18 months.

"I have been in medical education for 40 years and we're still a very memory-based curriculum," said Dr. Herbert Chase, a professor of clinical medicine at Columbia University

who is working with I.B.M. on the physician's assistant. "The power of Watson- like tools will cause us to reconsider what it is we want students to do."

I.B.M. executives also said they are in discussions with a major consumer electronics retailer to develop a version of Watson, named after I.B.M.'s founder, Thomas J. Watson, that would be able to interact with consumers on a variety of subjects like buying decisions and technical support.

Dr. Ferrucci sees none of the fears that have been expressed by theorists and science fiction writers about the potential of computers to usurp humans.

"People ask me if this is HAL," he said, referring to the computer in "2001: A Space Odyssey." "HAL's not the focus; the focus is on the computer on 'Star Trek,' where you have this intelligent information seek dialogue, where you can ask follow-up questions and the computer can look at all the evidence and tries to ask follow-up questions. That's very cool."

This article has been revised to reflect the following correction:

Correction: February 24, 2011

An article last Thursday about the I.B.M. computer Watson misidentified the academic field vindicated by Watson's besting of two human opponents on "Jeopardy!" It is artificial intelligence — not computer science, a broader field that includes artificial intelligence.