April 27, 2009

Computer Program to Take On ‘Jeopardy!’

By JOHN MARKOFF

YORKTOWN HEIGHTS, N.Y. — This highly successful television quiz show is the latest challenge for artificial intelligence.

What is “Jeopardy”?

That is correct.

I.B.M. plans to announce Monday that it is in the final stages of completing a computer program to compete against human “Jeopardy!” contestants. If the program beats the humans, the field of artificial intelligence will have made a leap forward.

I.B.M. scientists previously devised a chess-playing program to run on a supercomputer called Deep Blue. That program beat the world champion Garry Kasparov in a controversial 1997 match (Mr. Kasparov called the match unfair and secured a draw in a later one against another version of the program).

But chess is a game of limits, with pieces that have clearly defined powers. “Jeopardy!” requires a program with the suppleness to weigh an almost infinite range of relationships and to make subtle comparisons and interpretations. The software must interact with humans on their own terms, and fast.

Indeed, the creators of the system — which the company refers to as Watson, after the I.B.M. founder, Thomas J. Watson Sr. — said they were not yet confident their system would be able to compete successfully on the show, on which human champions typically provide correct responses 85 percent of the time.

“The big goal is to get computers to be able to converse in human terms,” said the team leader, David A. Ferrucci, an I.B.M. artificial intelligence researcher. “And we’re not there yet.”

The team is aiming not at a true thinking machine but at a new class of software that can “understand” human questions and respond to them correctly. Such a program would have enormous economic implications.

Despite more than four decades of experimentation in artificial intelligence, scientists have made only modest progress until now toward building machines that can understand language and interact with humans.

The proposed contest is an effort by I.B.M. to prove that its researchers can make significant technical progress by picking “grand challenges” like its early chess foray. The new bid is based on three years of work by a team that has grown to 20 experts in fields like natural language processing, machine learning and information retrieval.

Under the rules of the match that the company has negotiated with the “Jeopardy!” producers, the computer will not have to emulate all human qualities. It will receive questions as electronic text. The human contestants will both see the text of each question and hear it spoken by the show’s host, Alex Trebek.

The computer will respond with a synthesized voice to answer questions and to choose follow-up categories. I.B.M. researchers said they planned to move a Blue Gene supercomputer to Los Angeles for the contest. To approximate the dimensions of the challenge faced by the human contestants, the computer will not be connected to the Internet, but will
make its answers based on text that it has “read,” or processed and indexed, before the show.

There is some skepticism among researchers in the field about the effort. “To me it seems more like a demonstration than a grand challenge,” said Peter Norvig, a computer scientist who is director of research at Google. “This will explore lots of different capabilities, but it won’t change the way the field works.”

The I.B.M. researchers and “Jeopardy!” producers said they were considering what form their cybercontestant would take and what gender it would assume. One possibility would be to use an animated avatar that would appear on a computer display.

“We’ve only begun to talk about it,” said Harry Friedman, the executive producer of “Jeopardy!” “We all agree that it shouldn’t look like Robby the Robot.”

Mr. Friedman added that they were also thinking about whom the human contestants should be and were considering inviting Ken Jennings, the “Jeopardy!” contestant who won 74 consecutive times and collected $2.52 million in 2004.

I.B.M. will not reveal precisely how large the system’s internal database would be. The actual amount of information could be a significant fraction of the Web now indexed by Google, but artificial intelligence researchers said that having access to more information would not be the most significant key to improving the system’s performance.

Eric Nyberg, a computer scientist at Carnegie Mellon University, is collaborating with I.B.M. on research to devise computing systems capable of answering questions that are not limited to specific topics. The real difficulty, Dr. Nyberg said, is not searching a database but getting the computer to understand what it should be searching for.

The system must be able to deal with analogies, puns, double entendres and relationships like size and location, all at lightning speed.

In a demonstration match here at the I.B.M. laboratory against two researchers recently, Watson appeared to be both aggressive and competent, but also made the occasional puzzling blunder.

For example, given the statement, “Bordered by Syria and Israel, this small country is only 135 miles long and 35 miles wide,” Watson beat its human competitors by quickly answering, “What is Lebanon?”

Moments later, however, the program stumbled when it decided it had high confidence that a “sheet” was a fruit.

The way to deal with such problems, Dr. Ferrucci said, is to improve the program’s ability to understand the way “Jeopardy!” clues are offered. The complexity of the challenge is underscored by the subtlety involved in capturing the exact meaning of a spoken sentence. For example, the sentence “I never said she stole my money” can have seven different meanings depending on which word is stressed.

“We love those sentences,” Dr. Nyberg said. “Those are the ones we talk about when we’re sitting around having beers after work.”