



THE MULTIVERSE ACCORDING TO BEN

SATURDAY, FEBRUARY 06, 2010

Siri, the new iPhone personal assistant: Some useful niche applications, no interesting AI

CONTRIBUTORS

BEN GOERTZEL
BEN

TWITTER UPDATES

New sport: throw small dog in the 3 ft deep snow and watch him swim his way out ;-D about 8 hours ago

follow me on Twitter

PREVIOUS POSTS

- China Ascendant? (a comment on Robin Hanson's comm...
- Dialoguing with the US Military on the Ethics of B...
- 100 neural net cycles to produce consciousness?
- Dream of the Multiversal Cylinder Causation as Storytelling
- AGI, Ethics, Cognitive Synergy and Ethical Synergy...
- Why Bother to Vote?: A Novel Multiversal Answer
- Will China Build AGI First?
- The First AGI Summer School in Xiamen, China (Retr...
- Quantum Logic Models of Classical

Today I tried out Siri, the new AI "personal assistant" app for the iPhone. It has some very smart people behind it, and is based on some code and ideas from the DARPA-funded CALO project. Siri's earlier prototype version impressed me with its integration of dialogue and maps, so I was eager to check it out.

Systems Give a N...



The Siri website says:

Just like a real assistant, Siri understands what you say, accomplishes tasks for you and adapts to your preferences over time.

Ahem....

You may want to scroll to the end of this post, and read my dialogue with Siri, before reading the rest of what I have to say about the app.

This is the first release, and one doesn't want to judge the whole Siri project too harshly based on a first impression. But all I can report on now is my reaction to the product I just downloaded on to my phone and chatted with: and ... well ... let's just say that

- Indeed, this version of Siri may be very useful for carrying out a very limited set of very specific functionalities
- It's not anything like a real assistant; and worse than that, it gives no appearance of being an AI system making any kind of effort to really understand anything
- The basic "chatbot" functionality seems unnecessarily crude and quirky

As an AI developer I'm well aware that sometimes you can make mediocre (or worse) products or demos based on deeply powerful technology. So I'm open to the possibility that there is some profound or at least interesting tech underlying Siri. But I was unable to find it via playing with the product for an hour or so.

What I found, rather, is

- a rather crude keyword based chatbot (i.e. crude even by the standards of keyword based chatbots), without much attempt

at dialogue management

- straightforward, rule-based integration with a very small set of knowledge bases (about restaurants and movies, for instance) and with a map engine
- straightforward integration with TrueKnowledge for answering of factual questions
- decent speech-to-text with a very nice interactive interface

What surprised me most was the crudity of the dialogue management, which you'll see in the transcript below, of my initial conversation with Siri. So often Siri's responses had nothing to do with the questions I asked.

And its persistence of information between questions is rudimentary and awkward. Once you ask one question about New York, it pretty much assumes all your subsequent questions are about New York ... but it doesn't understand linguistic references to previous queries, not even simple ones.

Is Siri useful as a virtual assistant? I suppose I *might* use it to find restaurants or movies, or to check flight status. And just the other day, in the midst of a conversation in the car with the kids, I wanted to know Hitler's birth year, and I asked Wikipedia on my iPhone -- it would have been nicer to ask Siri instead.

So, yeah, for a few specific functionalities, where Siri's language engine and database integration are well-tuned -- yeah, it may be genuinely useful.

But the useful functionality is really VERY narrow and brittle. If you go even slightly beyond what the application has been specifically tweaked for, the results seem to be useless and annoying.

As a single example, consider the following snippet from my first conversation with Siri, given in full at the end of this post:

Ben: What is Kate Braverman's latest book?

Siri: OK, here are some businesses named "Kate" a few miles from here

This is really an unnecessary gaffe', but it's not exceptional; Siri does that sort of thing quite frequently. It makes this mistake because the query is about books and authors, rather than about stuff it's tuned for: restaurants, movies, flights, TrueKnowledge facts. And even for some things it's tuned for, like flights, the results are often quite weird and confusing, as you'll see in the example dialogue below.

How about the speech-to-text? (Supplied by Nuance, and performed on a server not on the phone.) It's so-so.... Which may be a *great* achievement technically given the quality of the iPhone's mike -- but still, it's only so-so.

The iterative graphical interface for speech-to-text is GREAT -- being able to review Siri's interpretations of your speech and correct them on the phone before they're sent to the server is very nice. But it makes enough mistakes that, all in all, using its speech-to-text is many times slower for me than using the iPhone keyboard.

I can see some genuine niche applications for the current Siri version: restaurant and movie location, flight status checking, fact searching, and maybe a few other similar applications, *while driving*. Or while not driving, for users who aren't comfortable typing.

This is all very well, but it's a far far cry from being like a human assistant, right?

The website warns us that this is an early-stage product:

Siri is young and, like a child taking its first steps, may be awkward at times. Siri may occasionally misunderstand things you ask it to do even within its range of understanding.

but IMO, the comparison with a child is inappropriate. Most of the mistakes Siri makes are not mistakes of misunderstanding. They are mistakes of *not even trying to understand* -- mistakes of replying in the manner of a simplistic chatbot acting on keyword cues.

If I had an iPhone app that made mistakes of genuine misunderstanding, like a child, I'd devote time to teaching it regardless of whether it assisted me in any way. In the case of Siri, I don't get the feeling of any intelligence or learning going on.

The current version is, for better or worse, a simplistic tool with a nice interface and a very, very limited scope.

Perhaps later versions will add enough functionality to constitute a more generally useful "assistant." But without some fundamentally different (and more intelligent) approach to dialogue management, it's hard to see how the product could grow into anything but an assemblage of *a few dozen specialized information-gathering widgets glued together by a chatbot*.

I read Nova Spivack's very insightful discussion on Siri a number of months ago, and studied the Siri prototype fairly carefully, and based on that prior experience I actually expected more from the first release. I hoped for a little more sense of understanding, of there being an "assistant with a personality" there. Nope. Maybe the next version will have some fundamentally different technology inside it ... one can always hope.....

Apologies if this review is a bit harsh -- but I'm reviewing Siri not just as an iPhone app, but relative to the rhetoric associated with it about being "like a child taking its first steps" and "just like a human assistant." If Siri were merely marketed as an iPhone app with a few interesting niche uses, I probably wouldn't bother to write a blog post about it.... But I've devoted much of my life to the quest to make AI systems that actually learn like children, and ultimately will display intelligence similar to and then transcending that of adult humans. The quest to make humanlike AI is a serious thing. Siri just doesn't feel to me like any kind of step along the path to serious AI systems, and I don't really like it when somebody's marketing department uses "real AI" as a marketing slogan for a product (even if a nice one in some ways) that actually has nothing to do with humanlike general intelligence.

To finish off this post, here's the transcript of my first conversation with Siri, with some comments interspersed and prepended by "COMMENT." I talked to Siri some more after this conversation, but it didn't reveal any more intelligence or anything particularly surprising ... I think this transcript will be more than enough to give you the idea.

Siri: Hello, I'm Siri, I'm your new personal assistant

Ben: Where is New York?

Siri: [shows map of New York]

Ben: Do computers have dicks?

Siri translates to: Do computers have Dick's?

Ben OKs the translation (just for fun)

Siri: OK, here are some computer stores close to New York, NY: [gives list]

COMMENT: OK, the maps integration is nice. And it's not a bad answer as there are some computer stores named Dick's.

Ben: Do computers have penises?

Siri: OK, I found these computer stores close to New York, NY: [gives list]

COMMENT: Total lack of any attempt to understand the question

Ben: Do you have a brain?

[The speech-to-text took 4 tries to get this, but finally did]

Siri: OK, here are some local businesses close to New York, NY: [gives list]

COMMENT: What???

Ben: Where can I buy a new head?

Siri: Here are movies that are new to theaters playing near New York: [gives list]

COMMENT: OK, it understood "new" and used a default for the rest...

Ben: Will you call Laising for me?

Siri translates this to: Will you call great thing for me?

After many tries Siri cannot get the Chinese name "Laising" so Ben OKs the wrong translation just for fun

Siri: Here are the most popular movies playing near New York: [gives list]

COMMENT: Where the heck did that come from???

Ben: I want to buy a new phone

Siri: Here are the most popular movies that are new

to theaters playing near New York: [gives list]

Ben: I do not want to see a movie. I want to buy a new phone.

Siri: Here are the most popular movies that are new to theaters playing near New York: [gives list]

COMMENT: Again, a total lack of understanding ... no attempt to find the arguments of the verb "want" or understand the meaning of "not." Just statistical word-matching (on "new") used in a crude way.

Ben: Where can I find Mongolian food?

Siri: OK, I found these Mongolian restaurants a short drive from New York, NY

COMMENT: Useful. Again, the maps integration is nice, and it knows that restaurants are associated with food

Ben: Where can I find Mongolian women?

Siri: OK, here are some Mongolian restaurants a short drive from New York, NY

COMMENT: Oops, maybe it didn't know restaurants were associated with food. Of course, Mongolian restaurants ARE a good place to find Mongolian women, but it didn't know that. It just associated "Mongolian" with "restaurants" by default because it doesn't know anything else Mongolian in New York, and it's biased toward location queries.

Ben: What is the capital of Mongolia?

Siri: I think I learned that in grad school ... I found answer for you: [answer in box: Ulaanbaatur, the city in Ulaaanbaatar, Mongolia (source:TrueKnowledge.com)]

COMMENT: Good, TrueKnowledge is integrated in the obvious way

Ben: How many people live there?

Siri: I think I learned that in grad school.... I'm not sure I can help you with that yet, but here's what a search engine has to say: [points to results of a Web search, which are useless, e.g. "Math Lesson: How many people...", "Is there bad people in Los Angeles...", etc.]

COMMENT: It can't deal with "there." Very weak understanding of dialogue.

Ben: How many people live in Ulaanbaator, Mongolia?

Siri: I found an answer for you: [in a box, 1067472]

COMMENT: But it actually knew the answer, via TrueKnowledge again...

Ben: How many of those people are men?

Siri: I'm not sure of the answer, but here's what a search engine has to say: [points to results of a

Web search, which are useless...]

Ben: What is Kate Braverman's latest book?

Siri: OK, here are some businesses named "Kate" a few miles from here

COMMENT: Unnecessarily confused

Ben: I want to buy a book written by Kate Braverman

Siri: Sorry, I can't get you information on books yet, but can help you look on the Web [points to a search engine, and to Amazon's website]

Ben: What can you tell me about Kate Braverman?

Siri: I can check on flights for you, but I need a little more information. I need to know at least one airport code, or an airline and a flight number.

COMMENT: Unnecessarily confused again.

Ben: Are airplanes flying from Reagan National Airport now, or have they been delayed because of the snow?

Siri: Hmm. I couldn't find any direct flights from RLA for February 8. For now, I can only find the status for direct flights. Soon I'll be able to book flights, and show connecting flights. [This is confusing since it is not Feb 8 now, but rather Feb

6]

COMMENT: That was exactly the kind of question I would like to be able to ask a virtual assistant and get an answer. But Siri badly botched it.

Ben: What is RLA?

[this took 3 tries for the speech to text]

Siri: I couldn't find any direct flights from RLA for February 8. For now, I can only find the status for direct flights. Soon I'll be able to book flights, and show connecting flights.

COMMENT: More unnecessary confusion, as my question was very simple and direct and didn't merit such a complex answer

Ben: I thought Reagan National Airport was DCA, not RLA

[this took 3 tries for the speech to text]

Siri: I'm not sure what you are asking, Ben. One of these, perhaps [points to empty list]

Ben: Who programmed you?

[after 6 tries the speech-to-text couldn't get it, and I gave up]

POSTED BY BEN GOERTZEL AT 11:23 PM



0 COMMENTS:

POST A COMMENT

<< Home