Natural Language Engineering
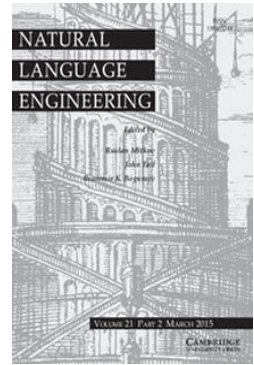
Additional services for *Natural Language Engineering:*

The limits of intelligent personal assistants

ROBERT DALE

Link to this article: http://journals.cambridge.org/abstract_S1351324915000042

How to cite this article:

Request Permissions : Click here

Industry Watch

The limits of intelligent personal assistants

R O B E R T   D A L E

Chief Technology Officer, Arria NLG plc
e-mail: robert.dale@arria.com

Abstract

In almost every science fiction movie you'll see people conversing with machines. Of course, the rise of intelligent personal assistants means you probably do this yourself already. This posting asks: what's the difference? Also, recent news on Facebook acquisitions, spoken language translation, and sentiment analysis.

Science fiction fascinates us because it paints a picture of a world that is different from our own. Different but not too different: we need to recognise the world of the future as belonging to the realm of the maybe-possible rather than the realm of the actual. It has to be just the other side of what you might call 'the reality horizon'. Stray too far on the other side of the horizon and you end up with science fantasy. Good science fiction is just at the boundaries of the believable.

The onward march of technology keeps science fiction writers in business. As technological capabilities move forward, the reality horizon, or what counts as believable, moves too. And so the science fiction writer has to dream up ever more ingenious devices – forms of transport, types of weapons, and ways of manipulating and extending our brains and bodies – that continue to stretch our imaginations.

But as future technologies go, portrayals of Artificial Intelligence are a bit different, because the reality horizon doesn't move. Human intelligence today is what it was 100 years ago. So, as time goes on, cinematic portrayals of machine intelligence, and particularly those that are manifested via the use of language, seem less and less like fiction. If you screw up your eyes a bit, and stop watching before the AI actually gets out of hand (I'm trying to avoid a spoiler here), the machine version of Johnny Depp in the 2014 movie *Transcendence*[1] does seem just a wee bit like IBM's Watson on steroids.

Ok, maybe you have to screw up your eyes quite a lot, but there's at least a family resemblance in there. Alternatively, consider Samantha, 'the first artificially intelligent operating system' in the 2013 romantic comedy *Her*.[2] Samantha has some amazing capabilities, a bit like a version of 2001's HAL[3] that has lightened up a

---

[1] http://www.imdb.com/title/tt2209764/
[2] http://www.imdb.com/title/tt1798709/
[3] http://en.wikipedia.org/wiki/HAL_9000

little. But her first tasks for the protagonist, Theodore, are within the bounds of the familiar: she cleans up his contacts and sorts out his mailbox. Again, with eyes screwed up and mufflers on your ears, you could mistake Samantha for a souped-up Siri.

So: each iteration of Hollywood's vision of AI seems a bit more plausible than the last, but not because the vision has changed. Rather, reality seems to be catching up.

Don't get me wrong. I'm in the camp that believes we're not going to see 'real AI' for quite some time, if ever. I suspect we'll wipe ourselves out as a species before we build machines that are smart enough to wipe us out. But I think there's indisputably a narrowing of the distance between what we have already achieved technologically and movie portrayals of machines-we-can-talk-to.

What's really interesting is the nature of the gap that remains. If you were to plot machine intelligence and human intelligence over time on a graph, human intelligence would be a horizontal line. Is the curve that represents machine intelligence going to cross that line? Or is the line corresponding to human intelligence an asymptote that the machine intelligence curve draws ever closer to, but never meets?

My money's on the latter, at least if the current trend of development in intelligent personal assistants (IPAs) is anything to go by. I'm referring here primarily to Siri,[4] Cortana[5] and Google Now,[6] although there are of course a host of minor players with similar characteristics.

By now, everyone is likely to have had exposure to one of these three. Just which one you're familiar with will be a secondary effect of your choice of mobile phone. I doubt there are many people who will buy a phone on the basis of the capabilities of the bundled personal assistant, although the amazing number of bake-offs between Siri, Cortana and Google Now that you'll find on the web might suggest otherwise.

There are of course differences between the three. Most obvious at first glance is the difference in terms of personification: while Siri and Cortana very much play the 'virtual human' card, the only pronoun that seems appropriate as a means of reference to Google Now is 'it'. Perhaps related to this, there's also a difference in terms of the bounds of embodiment. Siri is just in my phone, for the moment at least.[7] There's something comforting about that; when the phone is in the other room, I know Siri's not listening (or at least, she can't hear me). Google Now is another story. I've come to appreciate its reminders about where I parked the car when I'm out and about, but the first time a map popped up unprompted on my desktop with the expected duration of my commute, just around the time I normally leave, made me feel like I was being watched. Amazon's Echo[8] will take that omnipresence a step further, and it will be interesting to see whether people will consider this too invasive in a manner similar to the reactions we've seen to

---

[4] https://www.apple.com/ios/siri
[5] http://www.microsoft.com/en-us/mobile/campaign-cortana/
[6] https://www.google.com/landing/now/
[7] http://www.macworld.co.uk/news/mac-software/siri-for-mac-launch-date-reports-suggests-siri-will-feature-in-yosemite-3536158/
[8] http://www.amazon.com/oc/echo/ref_=ods_dp_ae

Google Glass. At the time of writing, there is a flurry of media coverage on the topic of Samsung's televisions' alleged eavesdropping on lounge room conversations,[9] suggesting that people might not welcome these new members of the household.

But I digress. Despite the differences, there's something all of today's IPAs have in common, and it's this characteristic that makes me think that the current approach to their development is going to leave us with an unbridgeable canyon between human and machine capabilities.

Put simply, the architectural philosophy underlying today's IPAs is fundamentally reductionist. The assistant is really an interface to a collection of discrete and effectively independent underlying functionalities. The task of the interface is to channel the request you make to a service that deals only with requests of that particular kind, whether that be modifying your phone settings, finding movie times or making a restaurant booking. If it's not the case that 'there's an app for that', then the IPA falls back on a search engine query.

Of course this is, in its own way, a great model. A divide-and-conquer strategy makes it possible to focus on providing functionalities that address high value or high frequency tasks, leaving the long tail of minority interests for later. And it makes it possible to spread the burden of development in a way that encourages innovation: with an appropriate API, third parties can add new functionalities to the base platform, providing an ever-growing range of capabilities.

But at the end of the day, what you have is a collection of essentially distinct capabilities. While there's no doubt that these capabilities can be extremely useful, the whole is no more than the sum of the parts. A collection of diverse domain-specific functionalities doesn't add up to intelligence.

In the medium term, there are two ways this might play out. On the one hand, the number of distinct functionalities supported by the platform might increase beyond what can sensibly be managed by this kind of architecture, perhaps resulting in the user frequently being routed to a functionality that wasn't the one they were actually looking for: the more categories you have, the greater the risk that categorisation will drop you in the wrong bin. Or, we might acknowledge that there's a threshold to the number of distinct capabilities that can be navigated via such an interface without frustration, in the same way as having multiple screens of app icons on your phone just makes it harder to find things, so that you end up removing apps to make navigation more manageable.

Either way, the point remains: assembling a collection of specific capabilities doesn't amount to a broad-coverage intelligence. Sorry, ScarJo's Samantha isn't coming to a phone near you any time soon.

Of course, we've seen this phenomenon repeatedly in the history of AI. Early AI research explored what a general problem-solving machine would look like, but in practical terms it was domain-specific expert systems that actually demonstrated value. As distinct from Star Trek's Universal Translator,[10] machine translation technology still works best when tuned to a particular domain. Similarly, a true

---

[9] http://edition.cnn.com/2015/02/11/opinion/schneier-samsung-tv-listening/
[10] http://www.startrek.com/database_article/universal-translator

open-domain question-answering system would be more than a collection of domain-specific question-answering systems, but current activity around IBM's Watson very sensibly aims to deliver value by specific application targeting. (Disclosure: Arria NLG has a partnership agreement with IBM.)

Can this silo-based approach to building up intelligence incrementally ever get us to 'real' AI? Is there an alternative approach that is more likely to succeed, and what would that look like? Others are asking the same question: Siri's inventors, now at Viv Labs,[11] seem to recognise the issue here, and claim that the solution they are working on 'radically simplifies the world by providing an intelligent interface to everything', but it remains to be seen whether this will turn out to be anything other than a refactored version of the reductionist approach.

We'll know that consumers have cottoned on to the limitations of this approach when we see a science fiction movie where the machine intelligences that attempt to overthrow the world are ultimately defeated because of their siloed and fragmented patchwork intelligence.

What's new

Here are some interesting language-technology related news items that have popped up over the last month or so.

2015 kicked off with Facebook acquiring voice recognition start-up Wit.AI,[12] whose tagline is/was 'Natural Language for the Internet of Things'. So far there's only speculation about how Facebook will use the technology, but late last year they also acquired Mobile Technologies, the Pittsburgh-based developer behind speech translation app Jibbigo.[13] Facebook also went on to release some of its AI tools as open source.[14]

Notwithstanding the comments above about open-domain translation, things are clearly moving forwards in this space, with the long-inevitable combination of speech technology and machine translation now made real in Google's latest Translate app[15] and Microsoft's Skype Translator.[16] The integration of Word Lens[17] into the Translate app makes this a really useful Swiss Army Knife for travellers. And Google Translate is also responsible for 2015's first contender for language-technology headline of the year: Google Translate saves baby in Irish roadside birth![18]

Sentiment analysis is currently the main game in town in terms of commercialising text-based language technologies, so much so that the technology now seems to be

---

[11] http://viv.ai/
[12] https://wit.ai/
[13] http://en.wikipedia.org/wiki/Jibbigo
[14] http://www.wired.com/2015/01/facebook-open-sources-trove-ai-tools/
[15] http://thenextweb.com/google/2015/01/12/googles-real-time-voice-translation-service-mobile-devices-way/
[16] http://www.skype.com/en/translator-preview/
[17] http://mashable.com/2015/01/14/google-translate-word-lens/
[18] http://www.theregister.co.uk/2015/02/13/google_translate_baby_delivery_swahili/

achieving commodity status. Every week it seems that there's a new start-up or application in the area, each providing a slightly different spin on the same core competency. Recently active players worth a look here are BirdEye,[19] who have announced a range of enterprise-level reputation management tools; Securly,[20] who announced sentiment analysis geared towards detecting when children are at risk; Sysomos,[21] who call what they do social intelligence; and Luminoso,[22] who raised $ 6.5M in the middle of last year for an approach that claims to have 'cracked the code on text mining'. We'll dig into this space in more detail in a subsequent posting.

There has been a fair bit of traffic recently around people complaining about how hopeless the voice recognition in their cars is. I can't work out whether self-driving cars will increase the need for good technology in this space, or whether they will make the technology redundant, turning us all into back-seat drivers who just get ignored. Either way, undeterred by these reports, Honeywell is setting its sights higher, testing voice recognition technology on aircraft flight decks.[23] Scary? Check out what they have in mind on their YouTube video.[24] I think we're safe for a while.

Finally, this month's QULT (Quirky Use of Language Technology): take a look at hitchBOT,[25] 'a chatty internet-connected robot that hitchhiked across Canada last summer', and which at the time of writing has just started its German tour. This is just great. Child-sized, and built from bits and pieces like a beer cooler bucket, pool noodles and rubber boots and gloves, it uses Wikipedia as a source of conversation topics that are 'worthy of its trivia-loving persona'. The perfect passenger for a self-driving car?

[19] `http://birdeye.com/`
[20] `http://www.securly.com/`
[21] `http://www.sysomos.com/`
[22] `http://www.luminoso.com/`
[23] `https://aerospace.honeywell.com/news/honeywell-testing-in-flight-voice-recognition-for-future-cockpits`
[24] `https://www.youtube.com/watch?v=RcvCaTlnYaI`
[25] `http://www.cbc.ca/news/technology/hitchbot-heads-on-adventure-in-germany-1.2918345`