What's next

Enterprise

"How many ounces in a cup?" or "Show me pictures of a pup."?

In order for Conversational AI to work, Natural Language Understanding needs to see both context and the human condition. Ken Arakelian helps us understand what happens when speech technology only gets part of the picture.

Ken Arakelian
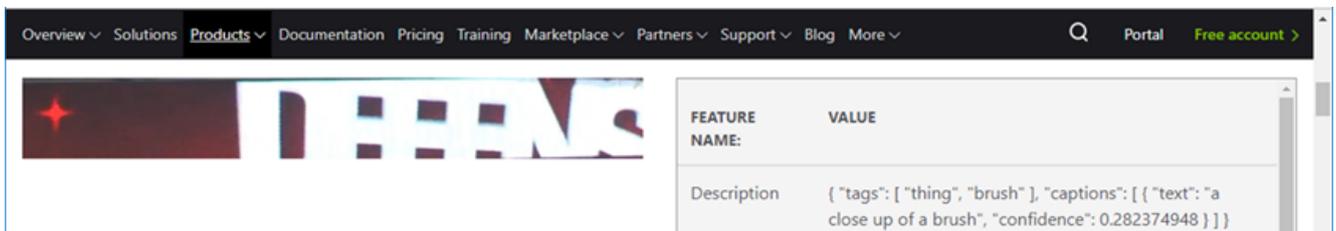Posted December 6, 2018

Quantum Leap

There was a TV show called "Quantum Leap" where Sam Beckett, a scientist, is trapped in a time travel experiment gone wrong and "leaps" into a different person's body each week. Every episode starts with the moments after he leapt into the next body and he found himself in a funny (or dangerous) situation. Sam had to make a quick decision on how to get out of it; this is basically what happens when you say, "Alexa, how many ounces in a cup?" Alexa wakes up with little context and makes a split-second decision about the question you're asking. What makes Alexa's job harder than Sam Beckett's is that Sam was a person and always leapt into a human body and Alexa doesn't understand what it means to be human.

Out of context

Let's leap into a situation and try to understand what we're seeing when we only get a snippet out of context. What do you see below?
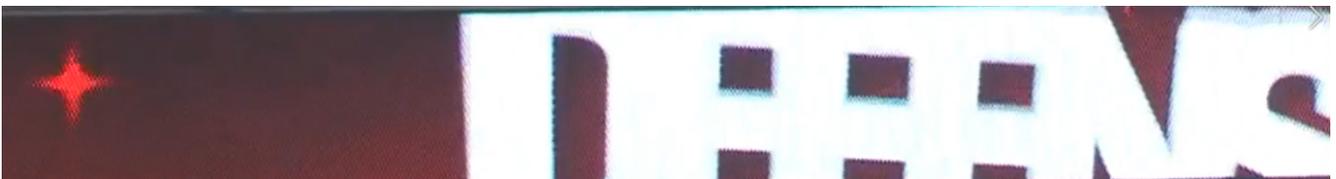
You're probably saying, "Not much," or "Not sure," and that is what the latest AI vision algorithms are saying, too. The results from MS AI Vision API is below, and it's worse than "not sure" because it's 28.2% sure this is a "close up of a brush". It's definitely **not** a closeup of a brush and 99% of people know that.

What's missing from being able to identify this picture is <u>context of the situation</u> and the <u>knowledge that comes with being human</u>. Without these two things our AI solutions have hit a wall. The dumb and funny things that smart speakers and IVRs (Interactive Voice Response – aka "Press 1 for billing") recognize out of context have been a source of amusement for years. I can think of dozens of movie scenes where the computer mis-recognizes and hilarity ensues (think Idiocracy | Not Sure – YouTube). So how do we make our robots smarter? Give them context by letting them always watch and listen, and teach them about what it means to be human, with all of our fears, desires, etc.

The human condition (or – I can't make back-to-back meetings at opposite ends of the campus)

Sam Beckett had the benefit of being human when he leapt into a new body and had to work

his way through a situation. Our robots today do not have this understanding and make ridiculous "errors" because of that. For example, the problem of finding a meeting time for a large group at a company is an age-old problem. Smart people have tried to solve this problem with automated meeting scheduler bots. They gave their bots context about what rooms are available and what everyone's schedule looks like. These meeting scheduling programs still fail because they miss the subtleties of being human-like: just because 18 of 20 people are available 1-2 pm on Friday doesn't mean it's a good meeting time because the two people that aren't available are the CEO and the Customer and their scheduling needs take precedence. Also, most people have a town hall meeting across campus from 12-1 pm so they probably can't make it back in time for the 1 pm meeting. These are things outside the computer's understanding and why automated schedulers still miss the mark.

Another example of why understanding the human condition is so important is in healthcare and healthy living programs. Why does the "eat healthy and exercise" message have such a low take-rate (or any healthy living program for that matter – "smoking cessation" etc.)? Kerry Evers, CEO of www.prochange.com, explains that the message needs to be tailored to whichever state of change the person is in on the Transtheoretical Model. Put simply, does the person even understand that they have a problem? Understanding whether a person acknowledges their health problem and is ready to make a behavior change, or if the person is still in denial of this, is important in deciding how to craft the message and conversation. Our Conversational AI doesn't do this today.

Situation context + human condition = human-level understanding

We're getting better at context, but having our automated systems understand the human condition is only just being thought about now. For **context** we've used Enterprise CRM and other data sources to understand what's happening in the account so that we don't ask dumb questions, at worst, and, at best, that we're able to predict the customer's needs before they have to ask.

Table stakes for context today are the use account histories from CRM systems (Are you calling about the higher than usual bill you just received?), weather reports from the interwebs (Are you chatting about the power outage due to a storm?), or news reports (Are you texting about the stock market correction?), etc. State of the art context is a combination of aggregate context (What is the most asked question during tax season?) and individual context (This person is an independent contractor and wants copies of her 1099s). At Nuance we build machine learning models to use context to help personalize and predict our bots' conversations, and this is why we're the best in the industry at building expert conversational bots.
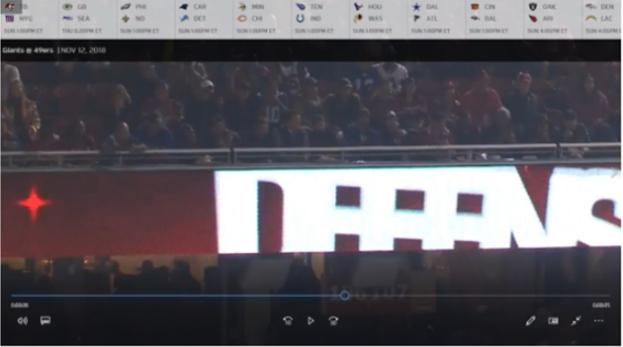
Giving computers/bots an **understanding of the human condition** is a much harder problem, and we're only scratching the surface today. There are efforts underway to give computers general knowledge about the world through ontologies. Simply put – ontologies organize knowledge in a giant, connected graph – one example would be this snippet of a movie ontology.
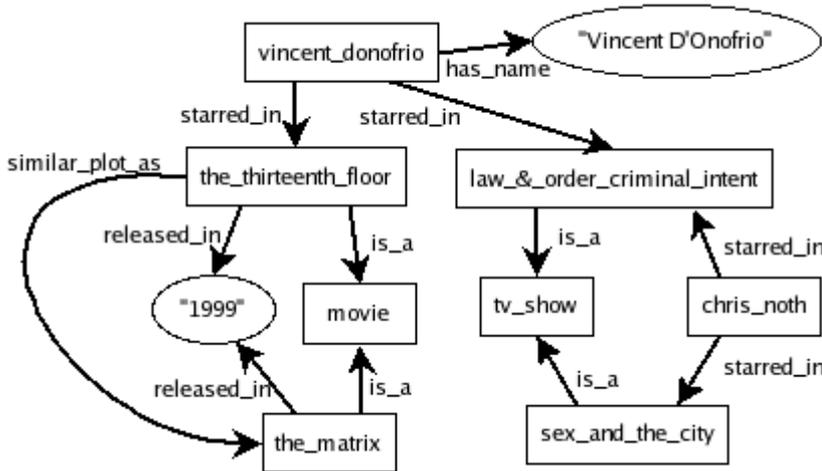
What hasn't been attempted yet is to give computers an understanding of what it means to be hungry, lost, hesitant, etc. and how a computer should change its conversational strategy to accommodate these things. There are no conversational bots today that understand that I wouldn't want a meeting in my open 9-10 am slot because I will have taken the red eye home the night before.

"How many ounces in a cup?" or "Show me pictures of a pup."?

Back to our "leap" experiment – let's zoom out a bit and give more context. Do you know what the picture below is now? I bet most of you do, but MS Vision API still doesn't.

Watch the video below for full context. You still can't see the image on the screen that well, but you know what it says because you have context.

The next leap forward

The next leap forward comes from context and understanding the human condition/perspective. At Nuance our deep industry experience allows us to maximize the

value of data sources, as well as to preemptively account for unexpected situations, allowing users to get back on track. Anyone can design solutions that follow happy paths, but not everyone knows how to build robust solutions for when users act, well, human.

Tags: Artificial Intelligence, human-machine interaction, natural language understanding

More Information

	Understand your customers, naturally
	Find out how NLU with Nuance can deliver better business results.
	Learn more

About Ken Arakelian

Ken Arakelian is the director of product management for Nuance's On Demand business. With more than 15 years of experience, Ken has worked in the contact center industry as a consultant, account manager and product manager.

View all posts by Ken Arakelian