The Best Way to Get Your Coffee: An Exploration of Natural Language Understanding

9月 3, 2021

By Dr. Barbertje Streefkerk, User Interface Design Manager

Let me tell you about a situation that happens from time to time in my house: my husband is sitting outside on the terrace, while I'm inside in the kitchen. It's around 4 pm. He asks, with a big smile, "Is the coffee machine still on?"

It may seem simple: linguistically, this is a yes or no question. But perhaps you need to be married to my husband to understand the deeper meaning. You see, he's asking me to bring him a freshly brewed coffee. And of course, it will be with milk, because from years of drinking coffee together, I know that's how he takes it.

In this case, to correctly interpret the question, context is important. I was standing in the kitchen next to the coffee machine; he was sitting in a chair outside. It was much easier for me to grab a cup and push the button on the coffee machine. Also, local and even cultural behavior plays a role - we often drink coffee in the late afternoon, rather than tea as people in England would do at that time of the day. Knowledge from the past also plays a role, giving me the foresight to put some milk in.

In the world of speech science, the act of requesting something by saying something else is known as an indirect speech act. These speech acts are only possible because humans are not restricted to a literal interpretation but can also consider contextual clues.

As natural language understanding (NLU) becomes more commonplace, we are also seeing our speech assistants gain the ability to handle indirect speech acts, where comments like "I am cold" result in the in-car assistant turning on the heater.

But where are the guardrails, and which contextual interpretation do you need to understand how certain questions and commands should be executed? Comments like, "it is hot inside" could trigger the assistant to lower the heater, but if a driver is just entering his car after it's been parked in the sun on a 35-degree Celsius day, the system should first open all the windows for some minutes.

Other indirect speech acts like "I am bored" are open to all kinds of different interpretations of what the system should do. This is exactly the point where our speech assistants go beyond speech recognition and NLU and enter a new level of AI – the kind that can leverage context, behavior, past knowledge, and more to support drivers in ways they never thought possible.

For more on how we're advancing conversational AI for mobility, visit www.cerence.com, and keep up with us on LinkedIn.