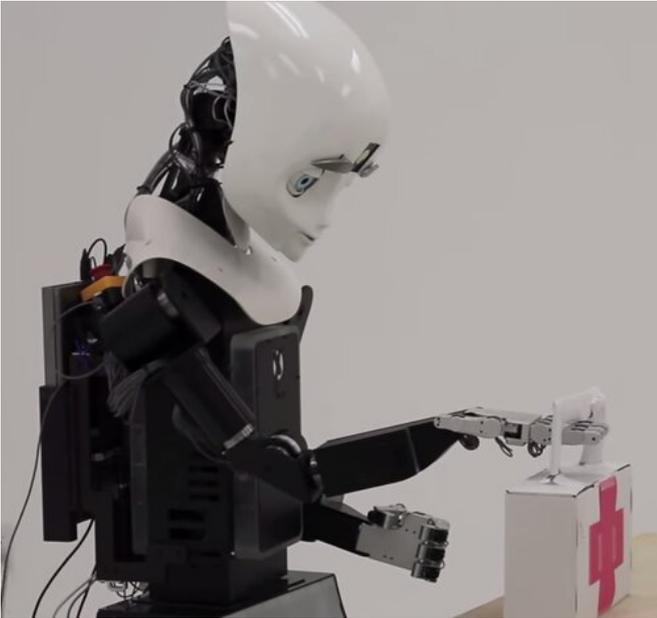


A statistical model of cognitive status for natural language generation

24 June 2020, by Ingrid Fadelli



Credit: MIRRORLab, mirrorlab.mines.edu/research/

In order for robots to be used a wide variety of settings, they need to be able to communicate seamlessly with humans. In recent years, researchers have thus been developing increasingly advanced computational models that could allow robots to process human language and formulate adequate responses.

An important aspect of [human language](#) that machines should acquire is the use of pronouns in sentences. According to an established linguistic theory known as the "Givenness Hierarchy" (GH), humans choose what pronouns to use based on their implicit assumptions about the "cognitive statuses" the objects have in the minds of their listeners. For example, if a speaker assumes that their target [object](#) is "in focus" (which is a cognitive [status](#)) within the current conversation, they may choose to use the pronoun "it."

Researchers at MIRRORLab at the Colorado School of Mines, have recently presented two models of cognitive status in a paper pre-published on arXiv. The first [model](#) is a theoretical rule-based Finite State Machine model directly informed by the GH literature, whereas the second model is a statistical probabilistic model (Cognitive Status Filter) that predicts the cognitive status of an object under uncertainty.

"My advisor, Dr. Tom Williams, and his peers had already started working on using the concept of cognitive status to aid robotic natural language understanding (NLU), where a listener has to identify the target object given their cognitive status/referring form information," Poulomi Pal, one of the researchers who carried out the study, told TechXplore. "The main idea/objective behind our recent paper was to create a computational model for cognitive status filtering based on the linguistic theory of the Givenness Hierarchy (GH) for the purpose of natural language generation (NLG), more specifically, to enhance machine use of pronouns (e.g., it, this, that, etc.)."

The first model presented by Pal and her colleagues is a Finite State Machine (FSM) model that generates the cognitive status of an object based on the rules laid out by the GH literature. The second model presented in the paper is a Cognitive Status Filter (CSF) that learns these rules automatically from textual data. The researchers then trained and evaluated their CSF model on data collected through the online Amazon Mechanical Turk platform.

During the experimental design of their CSF model, the researchers used a subset of the silver-standard English translation of [the OFAI multimodal task description corpus](#), which is a collection of human-human and human-robot multimodal interactions. They found that the CSF handled uncertainty better than the FSM model, as it did not follow pre-established rules, but instead acquired

rules directly from the data it was analyzing.

"Our results suggest that the CSF model is slightly better than the theoretical FSM model in terms of its accuracy in predicting the cognitive status of an object," Pal said. "The CSF model may thus be preferable when trying to assess the cognitive status of an object (especially when data is large), compared to a rule based theoretical model, as it can automatically learn the rules from the data."

The CSF model devised by Pal and her colleagues could ultimately help to enhance natural language interactions between humans and robots by improving upon the latter's ability to use pronouns in conversations. In the future, these findings could inspire other teams to develop similar models for robotics applications, as well as analogous techniques rooted in other fields of study, such as computational linguistics or cognitive psychology.

"We believe that developing a [computational model](#) like the CSF would help in the advancement of cognitively informed approaches toward both natural language generation and understanding," Pal said. "My plans for further research include developing and implementing a GH-informed anaphora generation model that accounts for the [cognitive status](#) of an object leveraging the CSF model during the selection of different referring forms for NLG."

More information: Givenness hierarchy theoretic cognitive status filtering. arXiv: 2005.11267 [cs.AI]. arxiv.org/abs/2005.11267

© 2020 Science X Network

APA citation: A statistical model of cognitive status for natural language generation (2020, June 24) retrieved 24 May 2022 from <https://techxplore.com/news/2020-06-statistical-cognitive-status-natural-language.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.