"Get ready for a party": Exploring smarter smart spaces with help from large language models

Evan King, Haoxiang Yu, Sangsu Lee, Christine Julien
The University of Texas at Austin
{e.king, hxyu, sethlee, c.julien}@utexas.edu

"Set up for a party"

You are an Al that controls a smart home. Here is the state of the

Here's the updated state of the devices in the home, in JSON format:

devices in the home, in JSON format:

}

The user issues the command: set up for a party. Change the device

Explanation: In response to the command "set up for a party", the Al

state as appropriate. Provide your response in JSON format.

has updated the device state as follows:

GPT models can infer meaning behind ambiguous user commands and control smart home devices in response. When told to "set up for a party", GPT-3 produces valid JSON that sets the lights to a color loop and turns on the stereo for music playback.

Abstract-The right response to someone who says "get ready

I. INTRODUCTION

for a party" is deeply influenced by meaning and context. For a smart home assistant (e.g., Google Home), the ideal response

An exciting prospect of smart homes at their advent was might be to survey the available devices in the home and change the potential to reduce user burden by providing seamless, their state to create a festive atmosphere. Current practical

systems cannot service such requests since they require the
unobtrusive, and "smart" interfaces to everyday devices. While
ability to (1) infer meaning behind an abstract statement and
smart assistants have improved significantly over the years
(2) map that inference to a concrete course of action appropriate
with respect to speech recognition 25 24 and user satis-
for the context (e.g., changing the settings of specific devices).
faction 20 16], a central challenge remains: how can these
In this paper, we leverage the observation that recent task-
agnostic large language models (LLMs) like GPT-3 embody a
assistants be made to respond appropriately to ambiguous user
vast amount of cross-domain, sometimes unpredictable contextual
commands that may be influenced by context or are otherwise
knowledge that existing rule-based home assistant systems lack,
impossible for system developers to anticipate beforehand? An
which can make them powerful tools for inferring user intent
example of such a command might be a user preparing their
and generating appropriate context-dependent responses during
home to entertain for guests, who asks their smart assistant to
smart home interactions. We first explore the feasibility of a
"get ready for a party". The hope is that the assistant-if it is
system that places an LLM at the center of command inference
and action planning, showing that LLMs have the capacity to
truly smart-might be able to help by inferring the meaning
infer intent behind vague, context-dependent commands like "get
of the statement and determining how to change the state of

ready for a party" and respond with concrete, machine-parseable
available devices in response: perhaps to start up the user's
instructions that can be used to control smart devices. We
party playlist on a smart speaker and change their smart lights
furthermore demonstrate a proof-of-concept implementation that
to a festive color scheme. In practice, however, such a request
puts an LLM in control of real devices, showing its ability to infer
intent and change device state appropriately with no fine-tuning
is beyond the capacity of current smart home systems. Google
or task-specific training. Our work hints at the promise of LLM-
Home will sadly admit: "I'm sorry, I didn't understand."
driven systems for context-awareness in smart environments,
In this paper, we are motivated by the observation that large
motivating future research in this area.
language models (LLMs) like OpenAI's GPT-3 3 have shown
an impressive ability to generalize to new tasks with high zero-
shot performance, as well as the capacity to infer meaning
Index Terms-human-centered computing, artificial intelli-
behind semantically complex or abstract statements [15]. We
gence, internet of things
thus ask the question: can this powerful capacity for cross-