to natural, complex instructions.

In this paper we introduce *Smart home Agent with Grounded Execution (SAGE)*, an LLM-based autonomous agent optimized for smart home applications. SAGE's long-term goal is to become an assistant that users can interact with as naturally as they would with a human companion. This requires the system to integrate three major categories of information: personal preferences, physcial grounding, and external grounding. Personalization is critical as it forms the foundation needed to understand natural and implicit requests by the user. For example, when asking a friend to "put the game on", a person assumes that their friend understands that their favorite sport is hockey, and that their favorite team is the Montreal Canadiens. Physical grounding informs the home assistant about the devices located in the user's home, their current states, and the possible actions that can be performed on them. Revisiting the "put the game on" example, a user would expect their friend to know which TV they are currently watching and put the game on that TV. Finally, *external grounding* gives the agent access to the external world outside the home such as weather information, TV schedules, emails, etc. In the "put the game on" example, the system requires access to a television schedule in order to understand if and how the request can be fulfilled.

The SAGE system architecture, shown in Figure 1, integrates a collection of tools, further detailed in Table 1, that allow it to exhibit personalization, physical grounding, and external grounding. A user command triggers a sequence of LLM queries, with the LLM's responses being used to determine which tools to employ, how to use them, and when to terminate execution. This approach to smart home automation shifts the control paradigm of home devices from a constrained setting, where users can choose among a predefined set of routines, to an unconstrained one, where users can successfully request novel behaviors that were not envisioned by the system's designers.

The main contributions of this paper are:

1. **An LLM-based agent for home automation (SAGE)** that can perform complex tasks within the home by autonomously using a sequence of tools which allow it to observe and act upon its environment.

2. **A collection of novel tools for home automation**. Our system develops a family of tools (see Table 1) that allow SAGE to be personalized and grounded.

3. **Evaluation on a real-world scenario**. We introduce a benchmark that consists of 43 highly challenging smart home tasks that test SAGE's personalization, intent resolution, device resolution, persistence, and command chaining capabilities. SAGE is shown to successfully perform 22 out of 43 of these tasks, significantly outperforming existing LLM-enabled baselines.

3 Related Work

3.1 Home Automation

A smart home system consists of connected IoT (Internet of Things) devices that enable the simultaneous monitoring, sensing, and control of the home environment. Automation of the control of these systems can lead to improvements in quality of life, comfort, and resource utilization [3]. Recent work has aimed to used machine learning to enhance the capabilities of smart home systems. For example, [4] proposed to perform home automation based on activity recognition, developing a deep learning algorithm to recognise users' activities from accelerometer data. Another focus is on voice-based home assistants, where the system is tasked with understanding users' voice utterances and executing the requested commands. For instance, [5] developed a voice controlled home automation system based on NLP (Natural Language Processing) and IoT to control four basic appliances. Leveraging advanced NLP techniques, current commercial solutions such as Bixby, Google Assistant, and Alexa offer a user-friendly interface capable of handling a variety of commands and questions from shopping and setting reminders to device control and home automation. However, these modern home assistants usually struggle with implicit and