Smart-Pikachu: Extending Interactivity of Stuffed Animals with Large Language Models

Toma Itagaki
toma@cs.columbia.edu
Columbia University
New York, NY, USA

Richard Li
lichard@cs.washington.edu
Paul G. Allen School of Computer Science & Engineering
Seattle, WA, USA

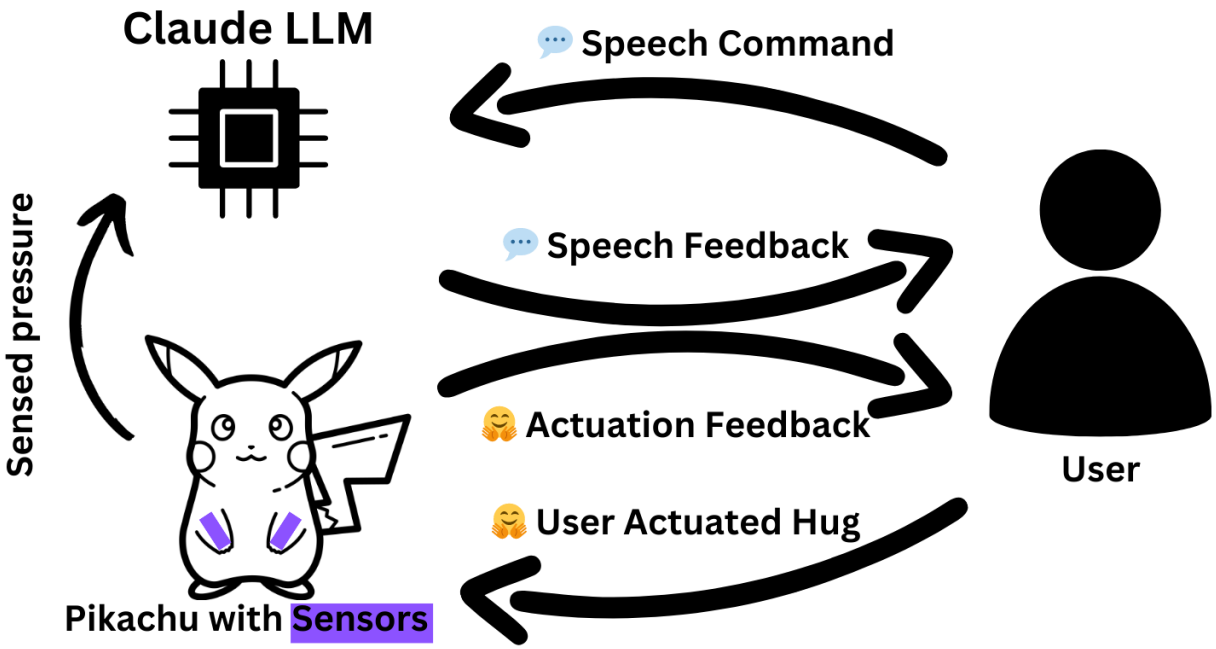

Figure 1: Overview of Pikachu's interaction and feedback loops.

ABSTRACT

We propose Smart-Pikachu, a stuffed animal equipped with sensing and actuation to explore the use of large language models (LLM's) with sensor data inputs. The augmentation of pressure sensing will allow for the LLM to interpret various interactions such as hugs and handshakes with the user. Furthermore, the actuation capabilities will extend our system's interactivity by providing physical feedback to the user. We will also incorporate text-to-speech output from the LLM to add another mode of interaction between the system and user. In this Student Innovation Challenge, we intend to explore applications at the intersection of sensing and interaction through LLM's and demonstrate an extension of LLMs' multimodal capabilities.

CCS CONCEPTS

• **Human-centered computing** → *Interaction devices*; • **LLM**; • **Sensing**; • **Interaction**;

KEYWORDS

Empathetic Computing, Prompt Design

UIST '23 Adjunct, October 29-November 01, 2023, San Francisco, CA, USA

ACM ISBN 979-8-4007-0096-5/23/10.
https://doi.org/10.1145/3586182.3625219

ACM Reference Format:
Toma Itagaki and Richard Li. 2023. Smart-Pikachu: Extending Interactivity of Stuffed Animals with Large Language Models. In *The 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23 Adjunct), October 29-November 01, 2023, San Francisco, CA, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3586182.3625219

1 INTRODUCTION

Large language models (LLM's) have gained significant traction in the last year for a wide variety of tasks. They are already being used for productivity and automation, research and learning, and creativity and play. For example, LLM's now power and augment search engine results, a previously plateaued domain. LLM's are also revitalizing education, enabling chatbots to become effective, personalized tutors. These models develop their capabilities through training on massive text datasets, up to and including the entire Internet.

Despite their all of their sophistication, LLM agents are limited in their expressivity. As the name suggests, LLM's are fundamentally designed to process and generate textual language. More recent work has explored techniques for adapting LLM's to work with other types of data, such as GPT-4's ability to accept image inputs [2]. Other recent multi-modal LLM work has proposed the use of LLM's for training a student model for embodied agents [6]. Building on this related work, we are interested in investigating how these advances might enhance a physical agents' ability to interact with humans.

We propose Smart-Pikachu, a stuffed animal augmented with fabric pressure sensors and mechanical actuators. Pikachu's pressure sensors will enable it to recognize different kinds of interaction, such as hand shakes and hugs. In turn, it will also be able to respond with its mechanical actuators. For example, they will enable Pikachu to hug the user back, or hold the user's hand with its paw.

In this UIST SIC proposal, we will explore the use of LLM's for creating a Specifically, we propose the following plan:

(1) Augment our Pikachu stuffed animal with pressure sensors and compliant actuators
(2) Investigate how to design prompts that help Claude process the pressure sensor signals and generate corresponding responses through its mechanical actuators
(3) Illustrate a design space in which we explore applications for a multi-modal interactive Pikachu

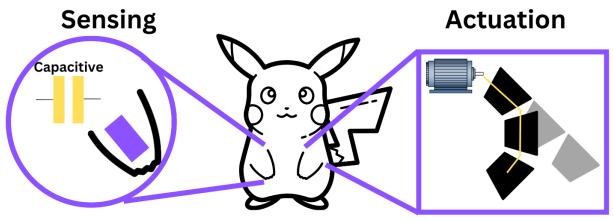

Figure 2: Sensing and actuation in Pikachu system.

2 PIKACHU SYSTEM

Our proposal manifests itself in the form of a Pikachu, a stuffed animal. We will augment an off-the-shelf Pikachu toy with pressure sensing and actuation capabilities.

We will employ pressure sensing capabilities similar to the fabric-based capacitive sensors introduced in FabToy [4]. These sensors will enable the system to recognize different types of interactions such as hand-shakes and hugs. The sensor signals will be processed by Claude to recognize interactions similar to the techniques introduced in [5].

Moreover, we will adopt techniques previously used for actuating robotic grippers [1, 3] to power Pikachu's arms. These actuators will allow our system to further provide physical interactions such as reciprocating to hugs and handshakes. These actuators will allow our system to further provide physical interactions, such as reciprocating hugs and handshakes. We will devise techniques for processing Claude's output and mapping it to actuation patterns.

In addition to Pikachu's pressure sensing and actuation, we will also include a speaker and microphone system for speech input and output. We hope that Pikachu's multi-modality creates a richer, more engaging interaction experience.

3 ETHICS AND BROADER IMPACT

In order to ensure that no users are harmed with the mechanical actuators, the finger mechanism will be built with a limited range of motion. The firmware to control the servo motor attached to the actuating finger will also be constraint.

During the demo, users will be given the option to interact with Pikachu with or without actuation.

We believe this work will introduce an approach to improve the variety of input modalities into LLMs. Furthermore, we hope that our demonstration can bring more awareness towards the utility of timeseries sensing in the realm of LLMs.

REFERENCES

[1] Mike Binnard and Mark R Cutkosky. 2000. Design by composition for layered manufacturing. *J. Mech. Des.* 122, 1 (2000), 91–101.
[2] Sébastien Bubeck, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, Peter Lee, Yin Tat Lee, Yuanzhi Li, Scott Lundberg, et al. 2023. Sparks of artificial general intelligence: Early experiments with gpt-4. *arXiv preprint arXiv:2303.12712* (2023).
[3] Aaron M Dollar and Robert D Howe. 2006. A robust compliant grasper via shape deposition manufacturing. *IEEE/ASME transactions on mechatronics* 11, 2 (2006), 154–161.
[4] Ali Kiaghadi, Jin Huang, Seyedeh Zohreh Homayounfar, Trisha Andrew, and Deepak Ganesan. 2022. FabToys: plush toys with large arrays of fabric-based pressure sensors to enable fine-grained interaction detection. In *Proceedings of the 20th Annual International Conference on Mobile Systems, Applications and Services*. 1–13.
[5] Xin Liu, Daniel McDuff, Geza Kovacs, Isaac Galatzer-Levy, Jacob Sunshine, Jiening Zhan, Ming-Zher Poh, Shun Liao, Paolo Di Achille, and Shwetak Patel. 2023. Large Language Models are Few-Shot Health Learners. *arXiv preprint arXiv:2305.15525* (2023).
[6] Yue Wu, So Yeon Min, Yonatan Bisk, Ruslan Salakhutdinov, Amos Azaria, Yuanzhi Li, Tom Mitchell, and Shrimai Prabhumoye. 2023. Plan, Eliminate, and Track–Language Models are Good Teachers for Embodied Agents. *arXiv preprint arXiv:2305.02412* (2023).