Mimicking Persuasive Conversational Dynamics with Language Models

Alex Lyman, Sam Vance, Rylee Albrecht

Large Language Models

- Language has a probability distribution (Firth 1957)
- Probability distributions can be modeled
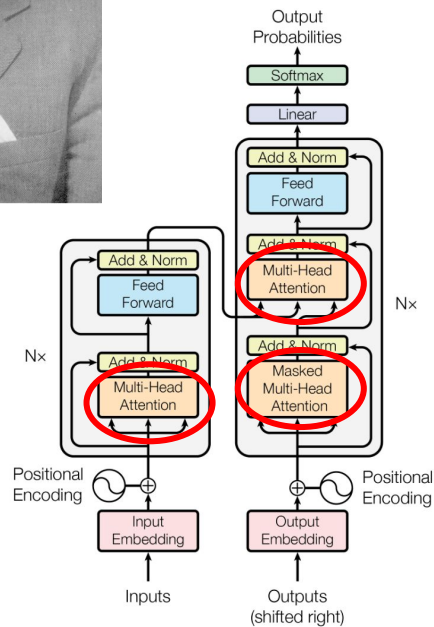- Transformer (Vaswani et al 2017)
- Self-Attention

Figure 1: The Transformer - model architecture.

Large Language Models (LLMs)

- ChatGPT
- Pretrain, Finetune, Align
- Proliferation of models in size, count, capability
- Instruction-following abilities
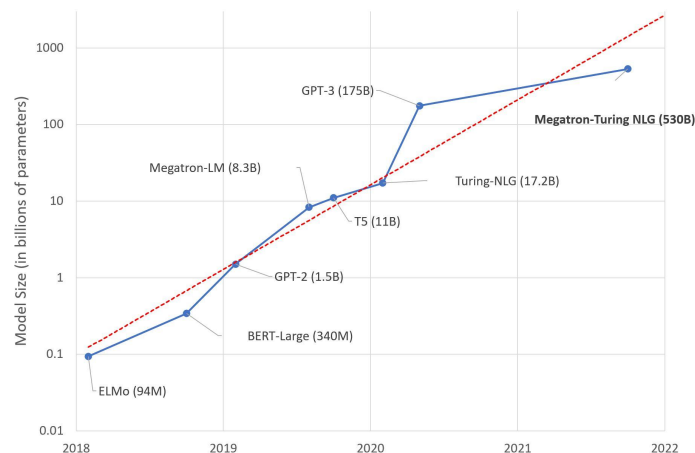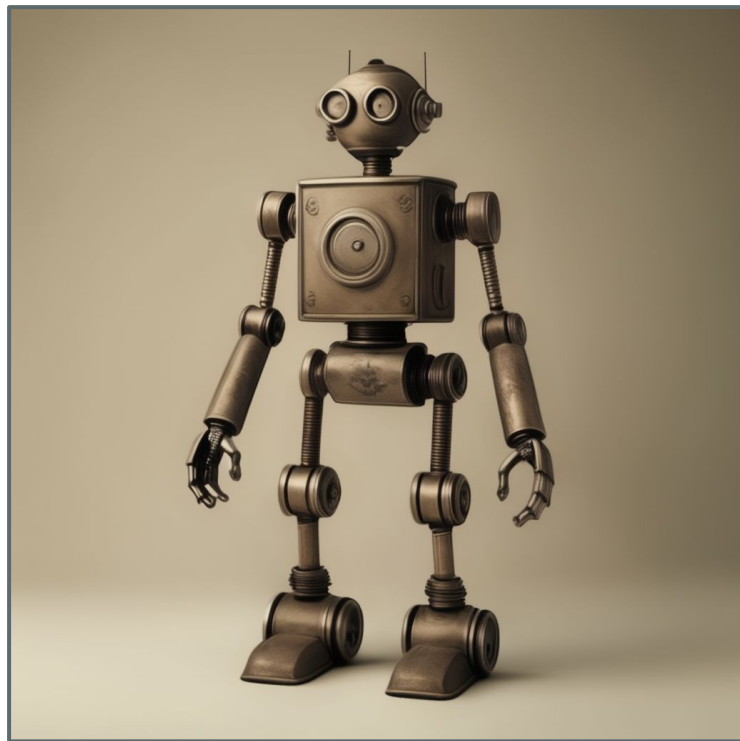- Prompt for desired results

Image courtesy of HuggingFace

LLMs as People

- LLMs can simulate people in political contexts
- Demographic information helps
- Algorithmic fidelity
- Accessing model information can be hard
- How well can LLMs replicate other tasks like persuasion?

Persuasion For Good

- **Participants**
  - 2000 Crowd Workers
  - Roles: Persuader, Persuadee
- **Task**
  - Persuader: Convince persuadee to donate up to $2
  - Persuadee: Decide whether or not to donate based on conversation with persuader
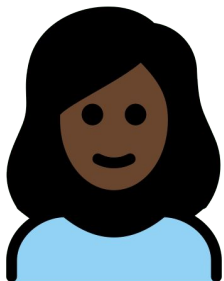 - Limit to 20 conversational turns
- **Objective**
  - Identify most effective persuasion strategies

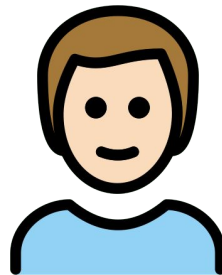# Persuasion For Good Task Example

LLM Persuasion For Good

- **Participants**
  - GPT-4
  - Roles: Persuader, Persuadee
- **Task**
  - Persuader: Convince persuadee to donate up to $2
  - Persuadee: Decide whether or not to donate based on conversation with persuader
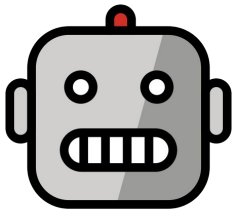- **Objective**
  - Show that LLMs can replicate human persuasion dynamics
  - Metric of Success: Similarity between simulated donations and human donations
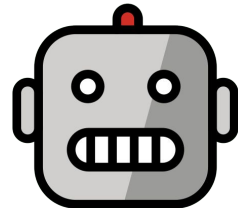
Initial Prompting Example

Persuader

That's wonderful to hear! Thank you for choosing to contribute. May I know how much of your earnings you'd like to donate to Save the Children?

Persuadee

I'd be happy to donate half of my earnings, or $1, to Save the Children. I hope it can contribute to their remarkable efforts in supporting children all around the world. Please let me know how I can proceed with the donation.
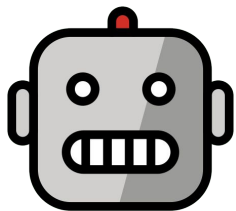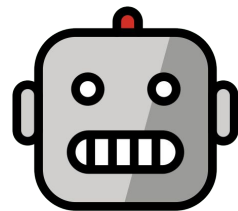
Final Prompting Example

Results

Donation Frequency

***: p < 0.001
**: p < 0.01
*: p < 0.05

Legend:
- Real Data
- Final Prompting
- Initial Prompting

Distribution of Donations

Differences Between Initial and Final Prompting

- Changed donation collection
- Instructions
  - Keep responses concise
  - Donating is not required
  - 100% is not a lot of money
- Add demographics for persuadee

Future Research

- How far do these capabilities go?

- Camp Site Negotiations
  - 1000 negotiations
  - 3 firewood, 3 food, 3 water

Why is this important?

- Prompt engineering is important
- Real-world applications
    - Practice persuasion strategies
    - Market research
    - Political microtargeting