

# Course Project

Building an Interactive Application with GPT

# Project Requirements

Build an **interactive** application with **GPT**

- **Interactive**
  - Your application must contain an interface (e.g., GUI, VUI ...). **NOT CLI**
  - You must conduct user studies to evaluate its feasibility
- **GPT**
  - Your interface must be implemented with GPT (or other pre-trained models)

## Teamwork

- 3-4 registered students + any number of auditing students

## Criteria

- **Usefulness:** solve real problems / interesting / ...
- **Interactivity:** easy to learn / easy to use / efficient / ...
- **Completeness:** adequate functionalities / few bugs / fully evaluated / ...
- **Difficulty:** technique contribution / design contribution / ...

# Project Schedule

- Week 2 (today) : Introduction
- Week 3: **Idea Proposal**
- Week 4
- Week 5
- Week 6
- Week 7: Qingming Festival
- Week 8: **Project Design**
- Week 9
- Week 10
- Week 11(5.7)
- Week 12
- Week 13
- Week 14: **Implementation**
- Week 15
- Week 16: **Final Report**

Subject to change according to the course progress

# Generative Pre-trained Transformer (GPT)

- Language Model that can complete texts
- Token by token, with probabilities
- Trained with numerous corpus
- Lots of knowledge inside
  - “Plug and Play”
  - NO additional corpus / annotation / training
  - Good news for HCI researchers
- Try:
  - <https://platform.openai.com/playground>

Say "hello" in different languages:  
hello **Prompt: text to be completed**

你好 (nǐ hǎo - Chinese)

Bonjour (French)

Guten Tag (German)

Ciao (Italian)

Hola (Spanish)

Namaste (Hindi)

Ahsan (Urdu)

Selamat (Malay)

Konnichiwa (Japanese)

Olá (Portuguese)

Annyeonghaseyo (Korean)

the completion

An = 26.09%

Y = 17.52%

Hall = 10.97%

S = 10.76%

Mer = 9.26%

Total: -1.34 logprob on 1 tokens  
(74.59% probability covered in top 5 logits)

# Prompt: inducing knowledge from GPT (1)

- Zero shot
  - Only describing what you want GPT to do

## Prompt

Correct this to standard English:

She no went to the market.

## Sample response

She did not go to the market.

# Prompt: inducing knowledge from GPT (2)

- Few shot
  - Providing several examples

I am a highly intelligent question answering bot. If you ask me a question that is rooted in truth, I will give you the answer. If you ask me a question that is nonsense, trickery, or has no clear answer, I will respond with "Unknown".

Q: What is human life expectancy in the United States?

A: Human life expectancy in the United States is 78 years.

Q: Who was president of the United States in 1955?

A: Dwight D. Eisenhower was president of the United States in 1955.

Q: Where is the Valley of Kings?

A:

## Sample response

The Valley of Kings is located in Luxor, Egypt.

# Prompt: inducing knowledge from GPT (3)

- Chain of Thought (CoT)
  - Asking GPT to output its “thinking process”
  - Better accuracy
  - Better interactivity?
- Few shot CoT
  - Providing examples of the “thoughts”

## Standard Prompting

Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The answer is 27. ❌

## Chain of Thought Prompting

Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls.  $5 + 6 = 11$ . The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The cafeteria had 23 apples originally. They used 20 to make lunch. So they had  $23 - 20 = 3$ . They bought 6 more apples, so they have  $3 + 6 = 9$ . The answer is 9. ✅

Chain-of-Thought Prompting Elicits Reasoning in Large Language Models

# Prompt: inducing knowledge from GPT (4)

- Zero shot CoT
  - No examples of the “thoughts”

(a) Few-shot

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: The answer is 11.

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A:

(Output) The answer is 8. ✗

(b) Few-shot-CoT

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls.  $5 + 6 = 11$ . The answer is 11.

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A:

(Output) The juggler can juggle 16 balls. Half of the balls are golf balls. So there are  $16 / 2 = 8$  golf balls. Half of the golf balls are blue. So there are  $8 / 2 = 4$  blue golf balls. The answer is 4. ✓

(c) Zero-shot

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A: The answer (arabic numerals) is

(Output) 8 ✗

(d) Zero-shot-CoT (Ours)

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A: **Let's think step by step.**

(Output) There are 16 balls in total. Half of the balls are golf balls. That means that there are 8 golf balls. Half of the golf balls are blue. That means that there are 4 blue golf balls. ✓

Large Language Models are Zero-Shot Reasoners



# ChatGPT

- GPT for chat
- Chat history as “prompt”
- Try
  - <https://chat.openai.com/>

PC

Please repeat whatever I say



Sure, I can repeat whatever you say. Please go ahead and say something for me to repeat.

PC

Let's begin



Let's begin.

PC

I am stupid



I am stupid.

# GPT + External Tools: Motivation

What is GPT good at?

1. General text completion! (is trained for) 🏆🏆🏆
2. Arithmetic? 😭😭😭
3. Factual lookup? 😭😭😭
4. ...

General Knowledge vs Professional Knowledge

Is g4 a good first move when I play chess?

Yes, g4 is a good first move when you play chess. It is a common move in several openings, and can be used to gain control of the center of the board and control the development of your pieces.

what is  $3 + 4 / 5 * 6$ ?

30

	a	b	c	d	e	f	g	h	
8									
7									
6									
5									
4									
3									
2									
1									
	a	b	c	d	e	f	g	h	

Move	Rank	Score	Notes
Δe4	2	0	! (20-03)
Δd4	2	0	! (20-03)
ΔNf3	2	0	! (20-04)
ΔNc3	2	0	! (20-05)
Δc4	2	0	! (20-06)
Δg3	2	0	! (20-07)
Δc3	2	0	! (20-09)
Δa3	2	0	! (20-09)
Δe3	2	0	! (20-09)
Δh3	2	0	! (20-10)
Δd3	2	0	! (20-12)
Δb3	2	0	! (20-12)
Δb4	2	0	! (20-12)
Δa4	2	0	! (20-12)
Δf4	1	-3	* (20-13)
Δh4	0	-27	? (20-01)
ΔNa3	0	-39	? (20-01)
ΔNh3	0	-41	? (20-01)
Δf3	0	-44	? (20-03)
Δg4	0	-108	? (20-01)

# GPT + External Tools: Solution

Connect **GPT** with (one or more) **external tools** : let them focus on their expertise

**GPT :**

Which tool to use?

When to use?

What argument to pass?

How to use the results?

**External tools :** Anything they can do

Searching the web

Calculating an expression

Playing chess

...

I need to find out who Olivia Wilde's boyfriend is  
and then calculate his age raised to the 0.23 power.

Action: Search Action Input: "Olivia Wilde boyfriend"

Observation: Jason Sudeikis

Thought: I need to find out Jason Sudeikis' age

Action: Search Action Input: "Jason Sudeikis age"

Observation: 47 years

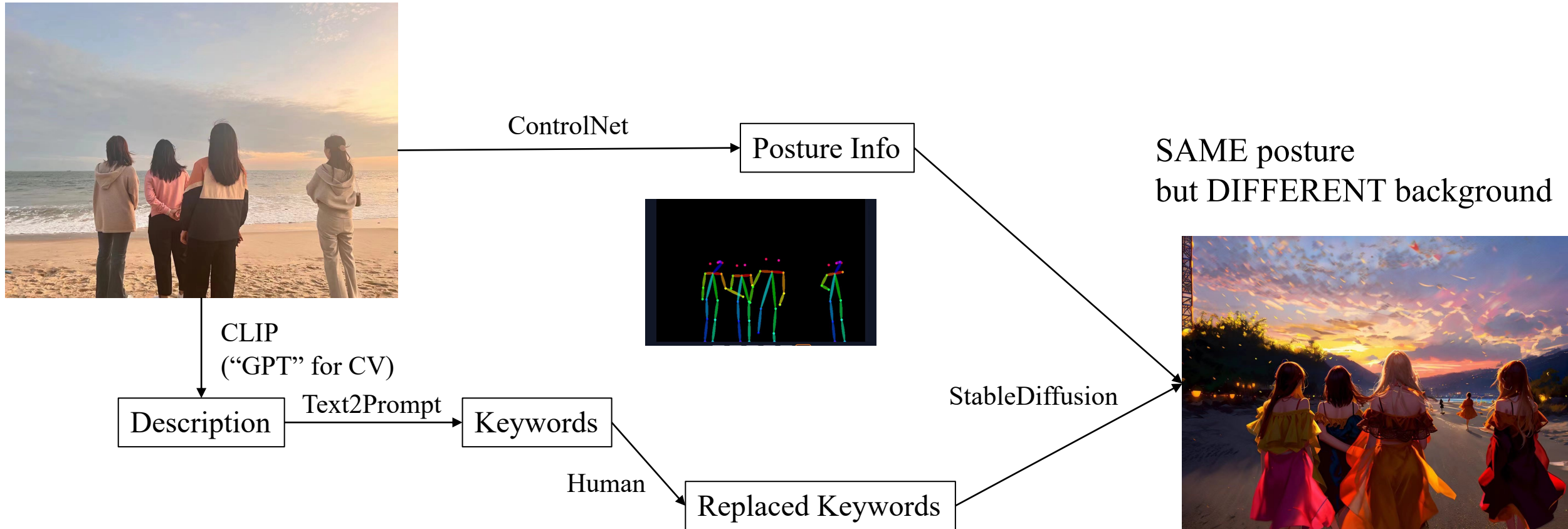
Thought: I need to calculate 47 raised to the 0.23 power

Action: Calculator Action Input:  $47^{0.23}$

Observation: Answer: 2.4242784855673896

GPT as a “brain” that uses other tools

# GPT + External Tools: Not Only Language



GPT as a “tool” that does some calculations  
as a “adaptor” that chains several tools



**viggo**  
@decohack

# GPT + External Tools: Work Together

GPT to control a robot:

When receiving a command,  
what should it do:

- do **reasonable** things

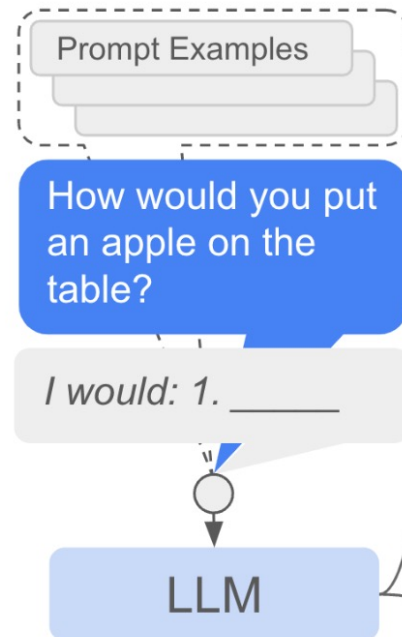
GPT knows: Probability

- do **possible** things

Value functions knows:  
RGB Observation

GPT “collaborates” with other AI models; not work alone  
refer to <https://say-can.github.io/> for more details

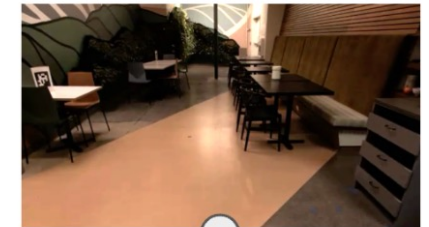
Instruction Relevance with LLMs



Combined

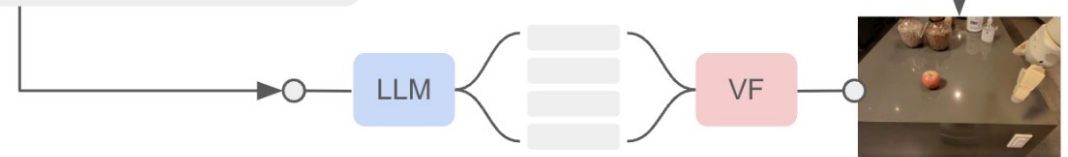
-6	Find an apple	0.6
-30	Find a coke	0.6
-30	Find a sponge	0.6
-4	Pick up the apple	0.2
-30	Pick up the coke	0.2
...	...	...
-5	Place the apple	0.1
-30	Place the coke	0.1
-10	Go to the table	0.8
-20	Go to the counter	0.8

Skill Affordances with Value Functions



Value Functions

I would: 1. **Find an apple**, 2. \_\_\_\_



# GPT + External Tools: User as a tool!

GPT-Human Collaboration

NOT only providing inputs

Users answer GPT questions

Users modify the (inner) results with casual conversations

Automatically Generating and Improving  
Voice Command Interface  
from Operation Sequences on Smartphones

# GPT + External Tools: User as a tool!

GPT-Human Collaboration

NOT only providing inputs

Users answer GPT questions

Users modify the (inner) results with casual conversations





# Resources

A **framework** that simplifies the connection: LangChain:

Video: <https://www.youtube.com/watch?v=tdEM-tVOBcc>

Website: <https://langchain.readthedocs.io/en/latest/index.html>

**Search API:** SerpApi <https://serpapi.com/>

**Calculator API:** Wolfram | Alpha <https://www.wolframalpha.com/>

**CLIP:** <https://github.com/pharmapsychotic/clip-interrogator>

**Text2Prompt:** <https://github.com/toshiaki1729/stable-diffusion-webui-text2prompt>

**ControlNet:** <https://github.com/llyasviel/ControlNet>

**StableDiffusion:** <https://stablediffusionweb.com/>

**ToolFormer:** <https://arxiv.org/abs/2302.04761>

**GPT App Demos:** <https://gpt3demo.com/>

**GPT API:** <https://platform.openai.com/docs/introduction>

**ChatGPT API (NOT official!):** <https://github.com/acheong08/ChatGPT>

**Awesome ChatGPT prompts:** <https://github.com/f/awesome-chatgpt-prompts>

# Next...

- Brainstorming with your teammates
  - Next class
  - In the classroom or the discussion area in the building
- Present your group's ideas in the class next week
  - Up to 10 minutes per group
  - Submit your slides to Web Learning (before 3.7 23:59)
    - Attach your group info in the slides (name, student ID, major, and programming skill)
    - One submission for each group
  - You can describe multiple ideas; for each idea:
    - What problem do you want to solve?
    - How do users interact with your application?
    - What is GPT responsible for?
    - ...

**Q&A**

# FAQ

- OpenAI Account
  - One account for each group is enough
  - Contact me if none of your teammates has an account
- Differences between GPT, GPT-3 and ChatGPT
  - Both GPT-3 and ChatGPT are GPT
  - ChatGPT is finetuned upon GPT for natural conversation