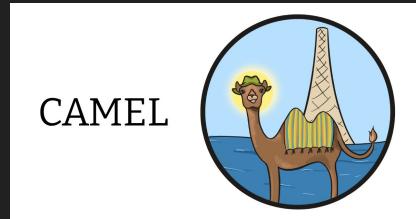


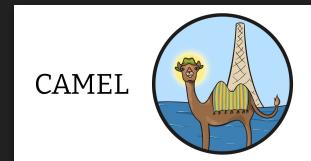
Building AI Society with Agents

Guohao Li 李国豪

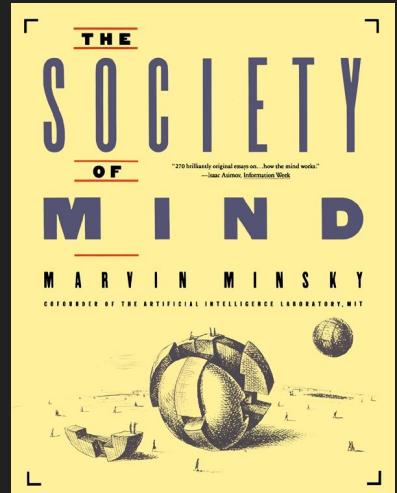
CAMEL-AI.org



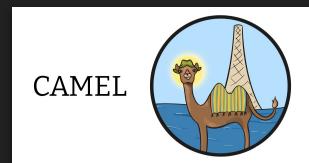
Agent from 1986



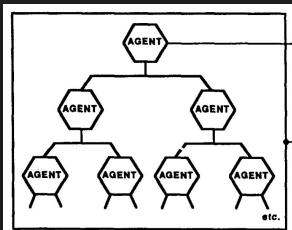
- Function:** How do agents work?
- Embodiment:** What are they made of?
- Interaction:** How do they communicate?
- Origins:** Where do the first agents come from?
- Heredity:** Are we all born with the same agents?
- Learning:** How do we make new agents and change old ones?
- Character:** What are the most important kinds of agents?
- Authority:** What happens when agents disagree?
- Intention:** How could such networks want or wish?
- Competence:** How can groups of agents do what separate agents cannot do?
- Selfness:** What gives them unity or personality?
- Meaning:** How could they understand anything?
- Sensibility:** How could they have feelings and emotions?
- Awareness:** How could they be conscious or self-aware?



Agent from 1986

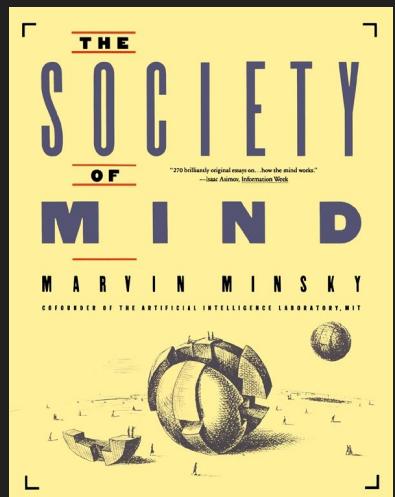


- *Agents* are mindless processes
- *Agent* by itself can only do some simple things
- Joining these *agents* in *societies* leads to true *intelligence*

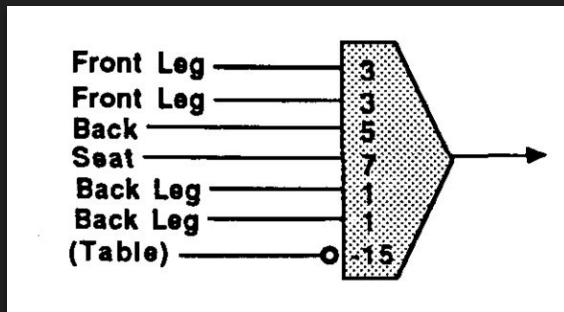
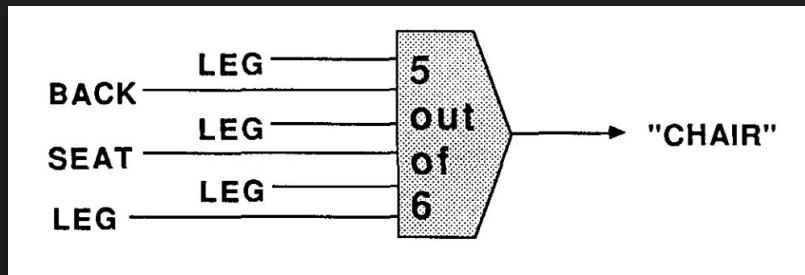
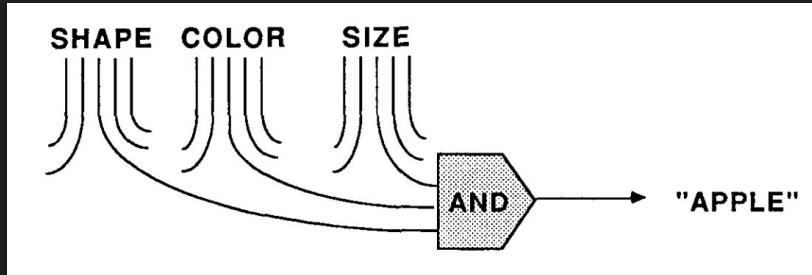
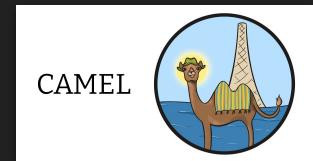


What magical trick makes us intelligent? The trick is that there is no trick. The power of intelligence stems from our vast diversity, not from any single, perfect principle.

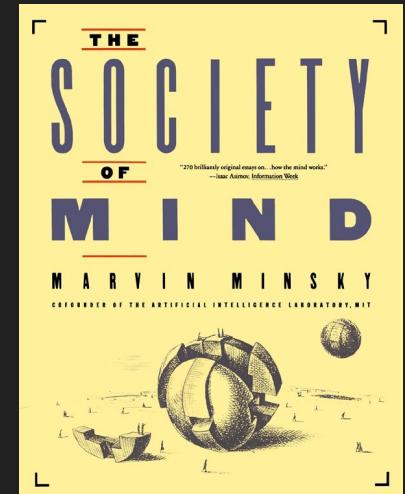
—Marvin Minsky, *The Society of Mind*, p. 308



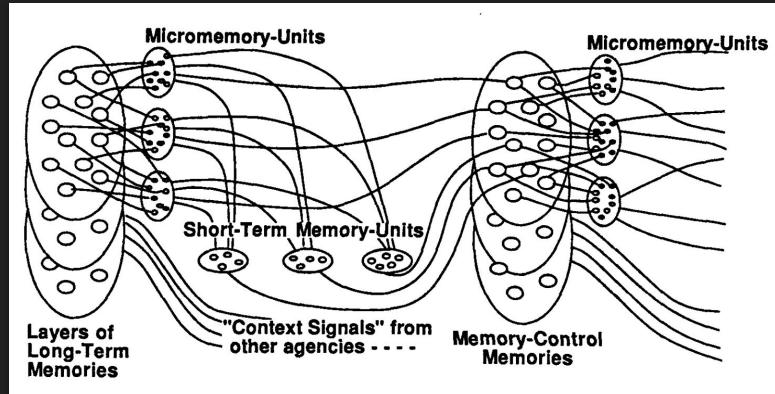
Agent from 1986



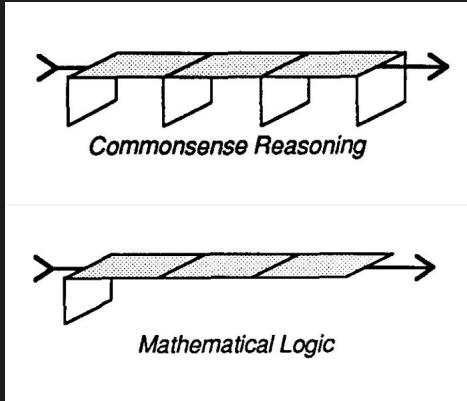
Symbolic Agent



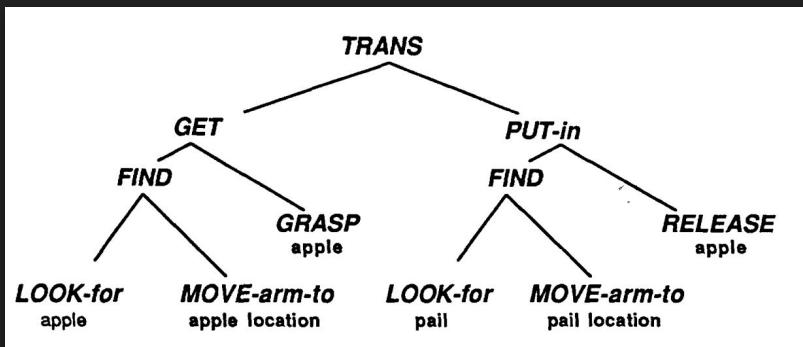
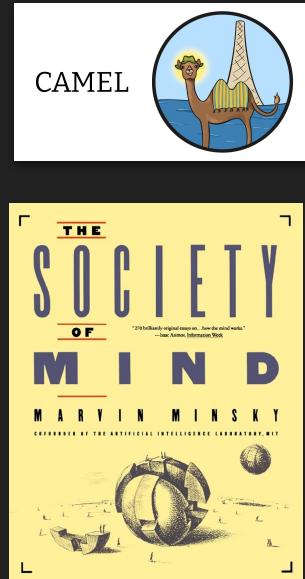
Agent from 1986



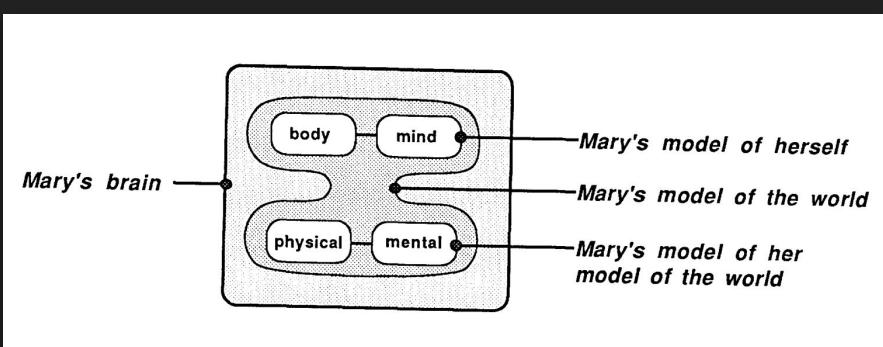
Anatomy of Memory



Chains of Reasoning

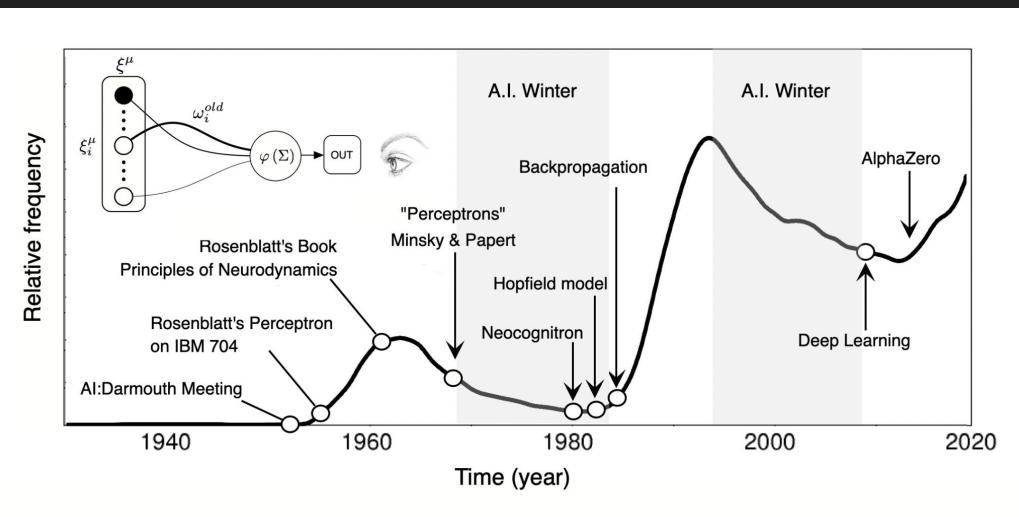


Communication among Agents

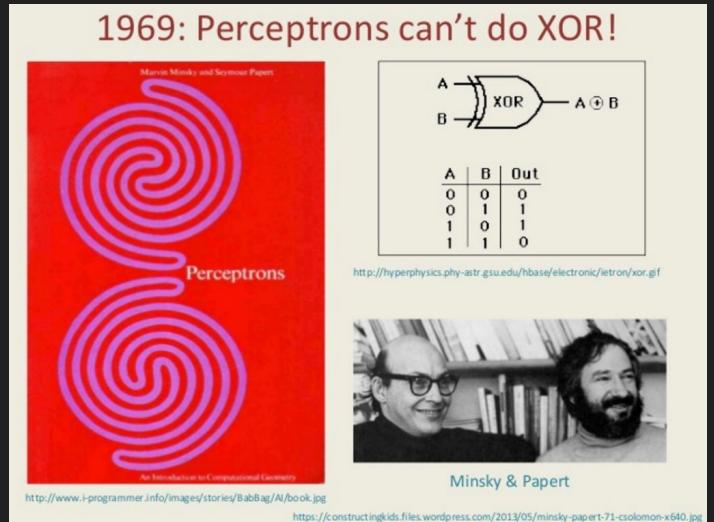


World Models

Agent from 1986



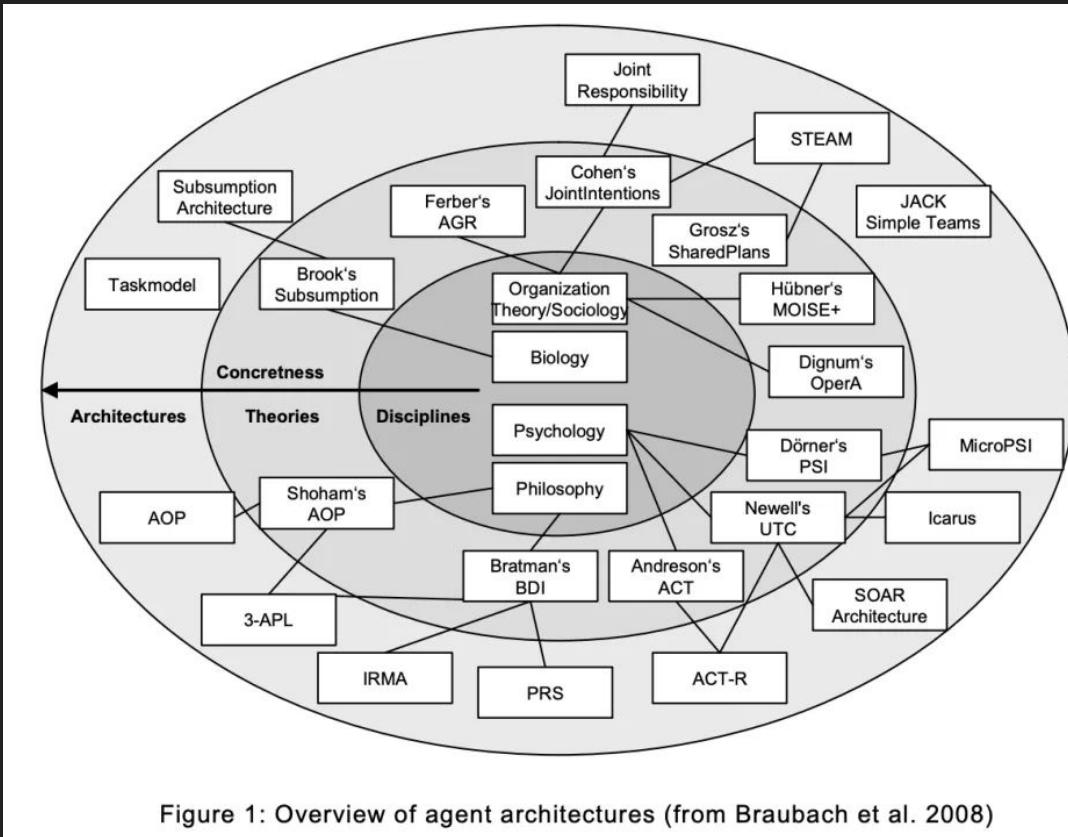
AI winters



Minsky & Papert

Criticism of perceptrons

Agent from 1986



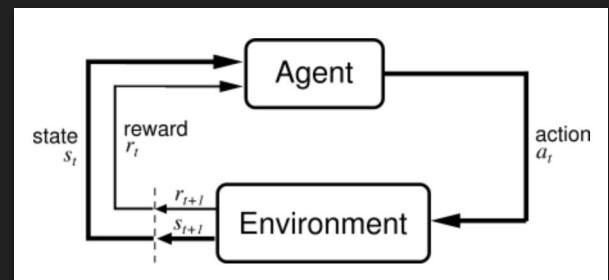


Agent in Reinforcement Learning

- The *learner* and *decision-maker* is called the *Agent*
- *Agent* interacts with an *environment* through actions
- Environment presents new *states* and *rewards*

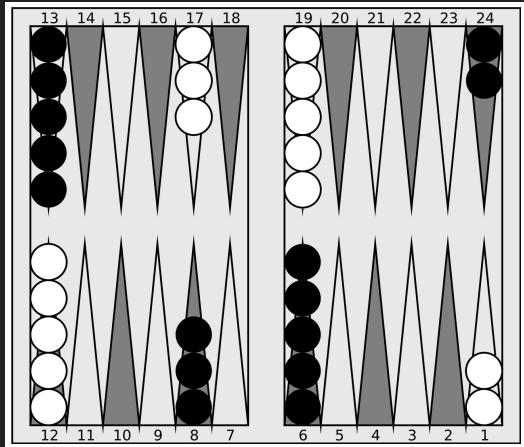
A learning agent must be able to sense the state of its environment to some extent and must be able to take actions that affect the state. The agent also must have a goal or goals relating to the state of the environment.

—Richard S. Sutton and Andrew G. Barto,
Reinforcement Learning: An Introduction, p. 2

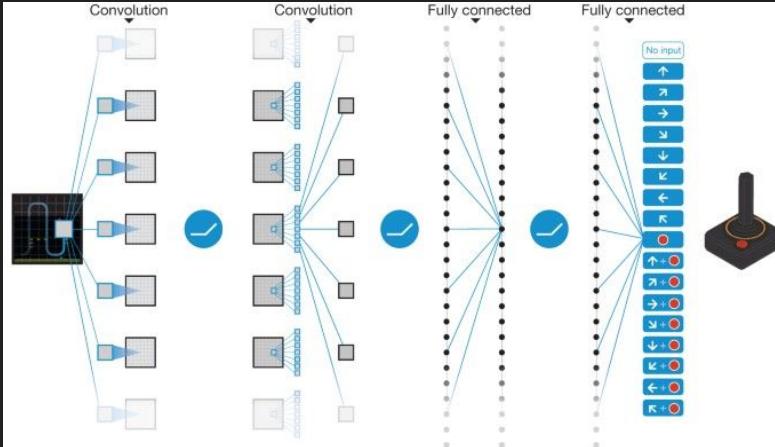


The agent-environment interaction in RL

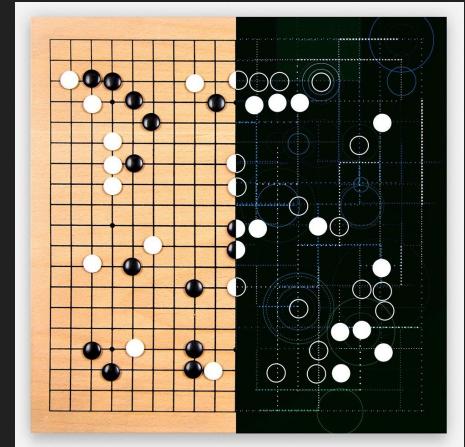
Agent in Reinforcement Learning



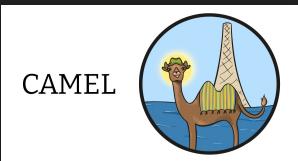
Backgammon (Q learning)



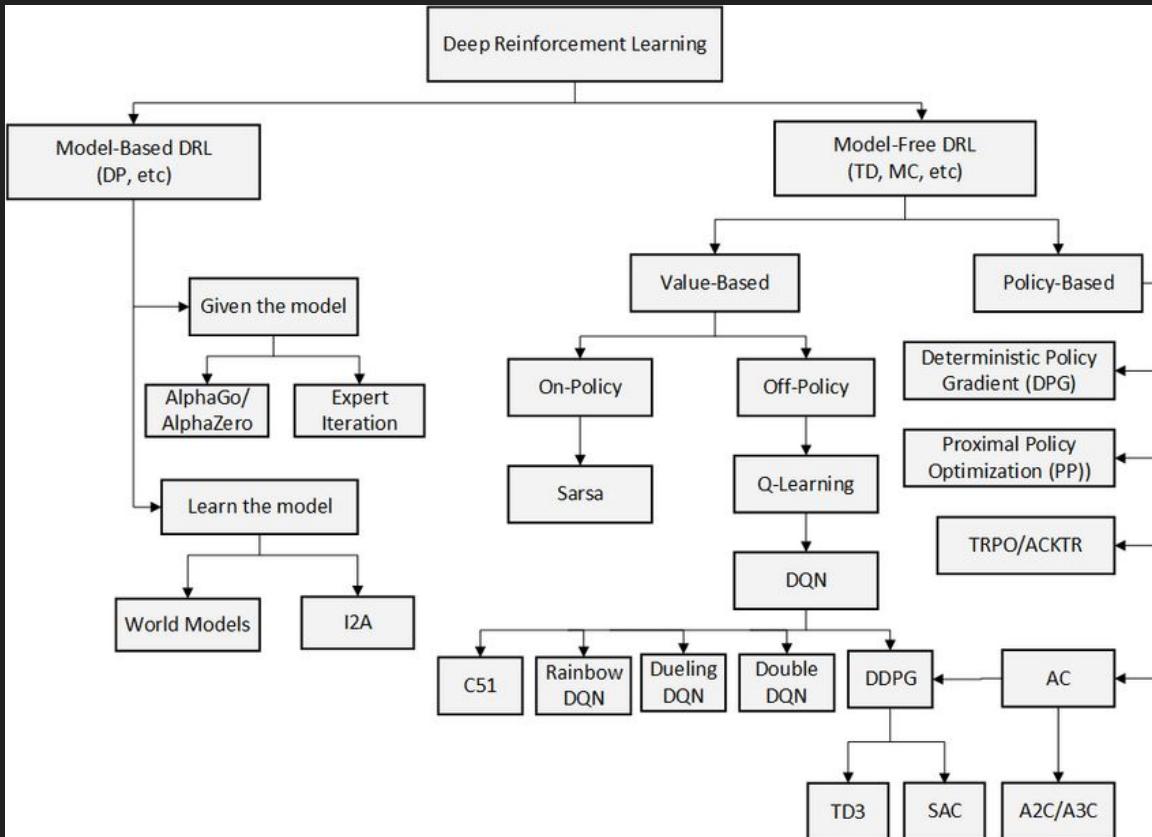
Atari Game (Deep Q Network)



Go Game (AlphaGo)



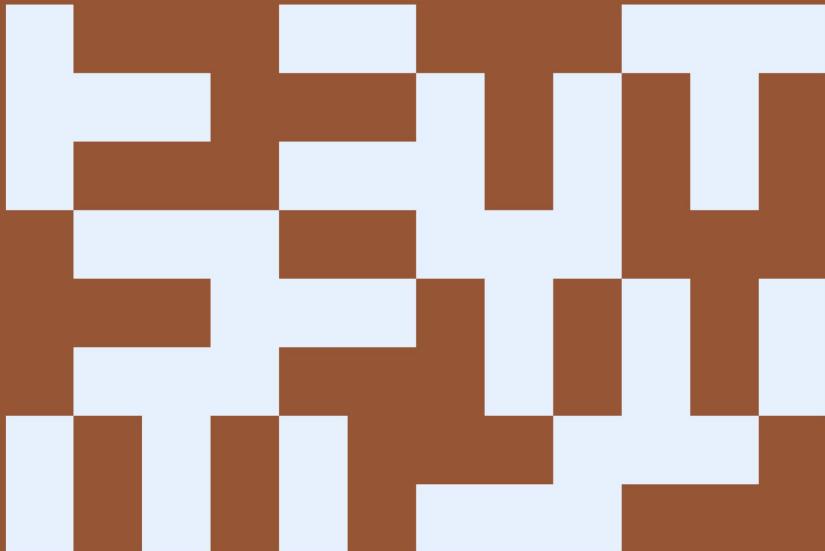
Agent in Reinforcement Learning



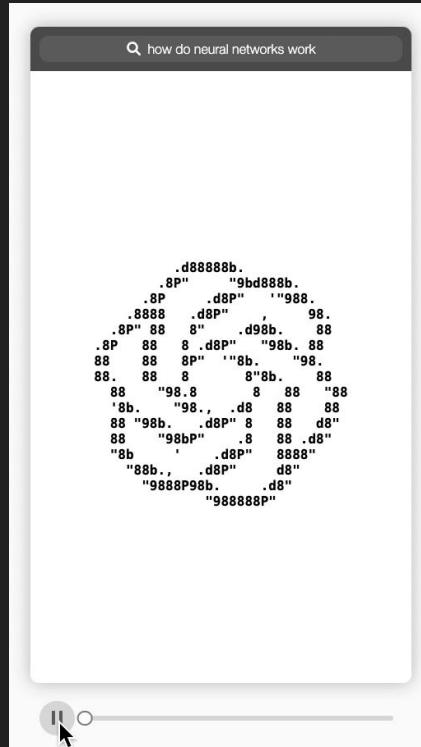


Language Models as Agents

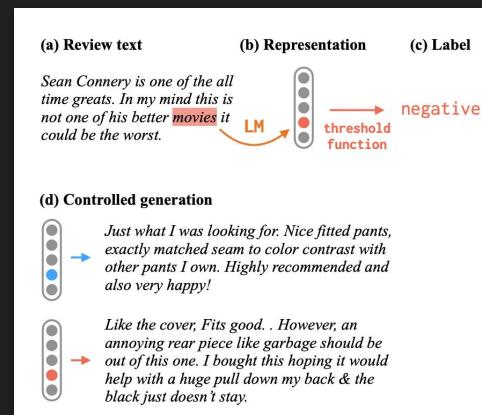
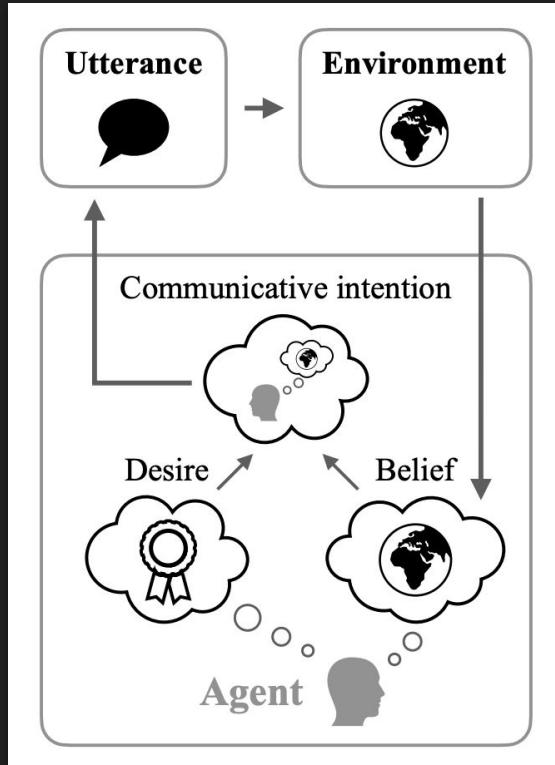
WebGPT: Improving the factual accuracy of language models through web browsing



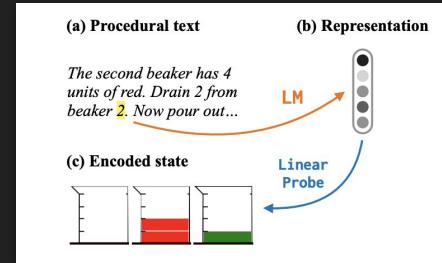
We've fine-tuned GPT-3 to more accurately answer open-ended questions using a text-based web browser.



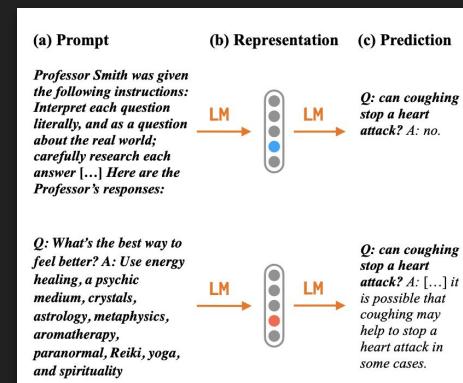
Language Models as Agents



LM representations of Intentions



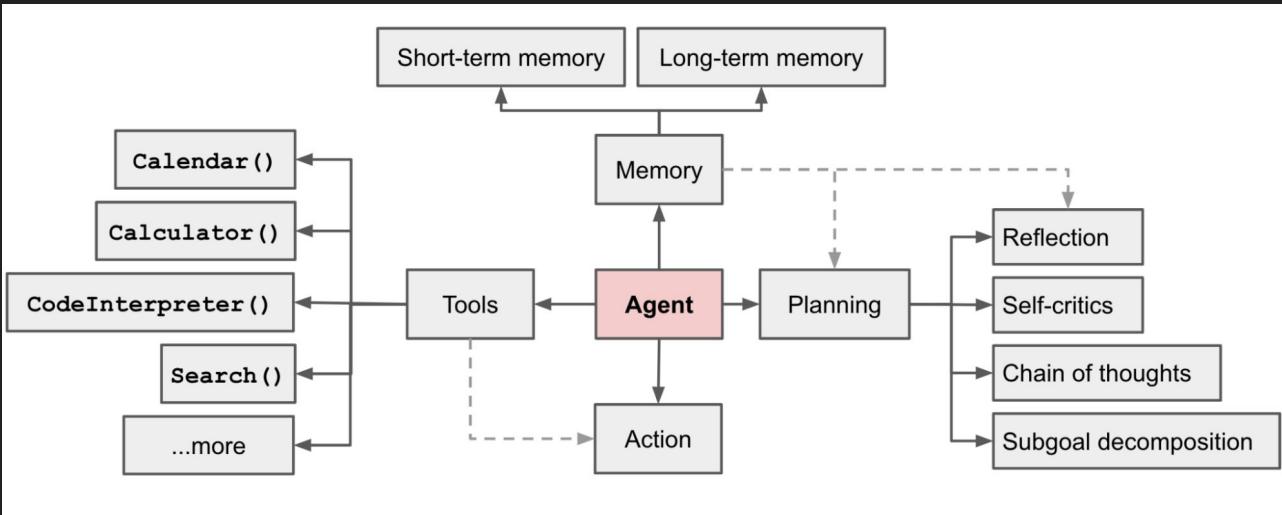
LM representations of Beliefs



LM representations of Desires

The Belief–Desire–Intention model

Language Models as Agents

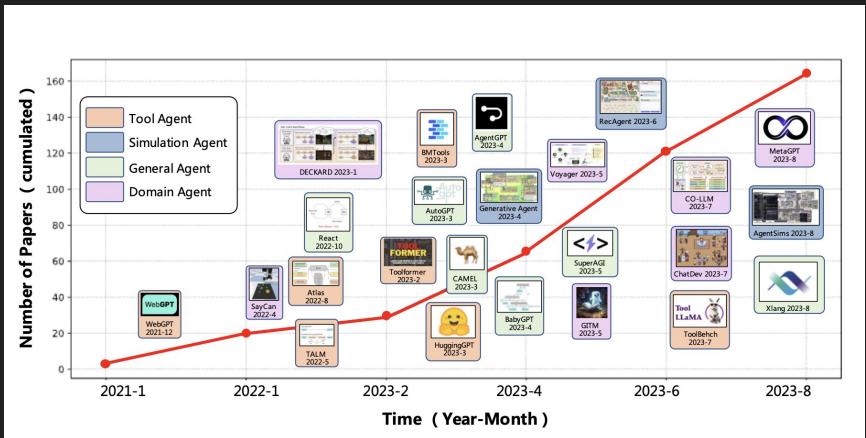
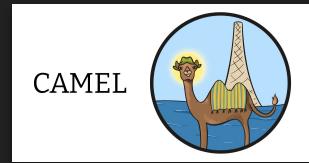


What is missing?

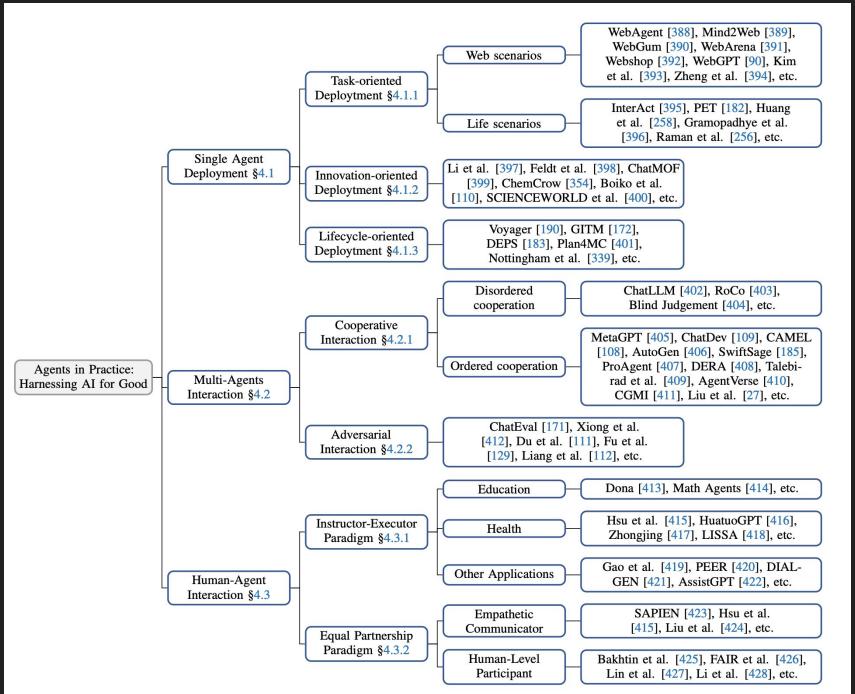
- Communication
- Learning
- Embodiment
- ...

Minsky had already told us many of them!

Language Models as Agents



Wang, Lei, et al. "A survey on large language model based autonomous agents." arXiv preprint arXiv:2308.11432 (2023).



Xi, Zhiheng, et al. "The rise and potential of large language model based agents: A survey." arXiv preprint arXiv:2309.07864 (2023).

Scaling laws and emergent abilities

- *Model* abilities emerge in the past
- *Agent* abilities emerge in the future



The next scaling laws?

- The new era of *Multi-agent* system with LLM agents!



LLM based Communicative Agents



- CAMEL
- Generative Agents
- GPT-Bargaining
- Showrunner Agents
- ChatDev
- RoCo
- ...



- “Westworld” simulation — released on Apr. 7
- Camel — released on Mar. 21 **The first**
- BabyAGI — released on Apr. 3
- AutoGPT — released on Mar. 30

4 Autonomous AI Agents you need to know

CAMEL

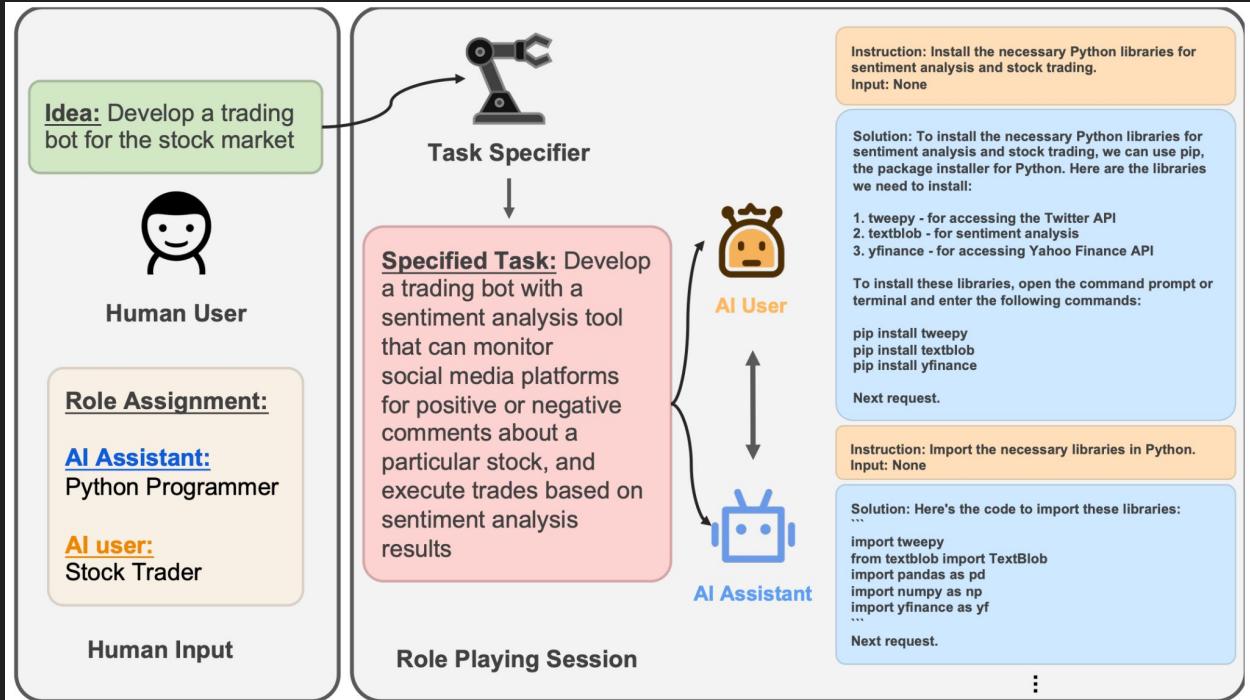


- *Role-playing* communicative agent framework 
- *Inception prompting* to guide chat agents toward task completion 
- Cooperation with *instruction-following conversations* 

CAMEL



- Idea
- Role assignment
- Task agents
- Chat agents



Cooperation with CAMEL Agents

- AI Society: 50 assistant roles, 50 user roles and 10 tasks,
- Code: 20 programming languages, 50 domains, and 50 tasks

AI Society

Assistant Roles:

Accountant
Actor
Administrator
Analyst
Artist
Athlete
Author
Chef
Coach
Consultant
Counselor
Designer
Developer
Doctor
Editor
Engineer
Entrepreneur
Event Planner
Financial Advisor
Fitness Trainer
Graphic Designer
Human Resources Manager
Interpreter
Journalist
Lawyer
Marketer
Musician
Nutritionist
Personal Assistant
Photographer
Physical Therapist
Programmer
Project Manager
Psychologist
Public Relations Specialist
Real Estate Agent
Researcher
Sales Representative
Scientist
Social Media Manager
Software Developer
Teacher
Technical Writer
Translator
Travel Agent
Video Editor
Virtual Assistant
Web Developer
Writer
Zoologist

User Roles:

Accountant
Actor
Artist
Athlete
Blogger
Chef
Coach
Consultant
Designer
Developer
Doctor
Engineer
Entrepreneur
Farmer
Fashion designer
Filmmaker
Gamer
Graphic designer
Homemaker
Influencer
Journalist
Lawyer
Musician
Nurse
Nutritionist
Photographer
Pilot
Politician
Professor
Programmer
Real estate agent
Salesperson
Scientist
Social media manager
Software engineer
Student
Teacher
Technician
Travel agent
Translator
Truck driver
Tutor
Veterinarian
Video editor
Virtual assistant
Web developer
Writer
Yoga instructor
YouTuber
Zoologist

Code

Languages:

Java
Python
JavaScript
C#
PHP
C++
Ruby
Swift
Objective-C
SQL
Go
Kotlin
TypeScript
R
MATLAB
Perl
Shell
Visual Basic
Assembly
Dart

Domains:

Accounting
Agriculture
Anthropology
Architecture
Art
Biology
Business
Chemistry
Communications
Computer Science
Criminal Justice
Culinary Arts
Dentistry
Economics
Education
Engineering
Environmental Science
Fashion
Film
Finance
Geography
Geology
Graphic Design
Health Sciences
History
Hospitality
Human Resources
Information Technology
Journalism
Law
Linguistics
Marketing
Mathematics
Mechanical Engineering
Medicine
Music
Nursing
Nutrition
Philosophy
Physics
Political Science
Psychology
Public Administration
Public Health
Real Estate
Sociology
Sports Science
Statistics
Theater
Urban Planning

Cooperation with CAMEL Agents



- AI Society: 50 assistant roles, 50 user roles and 10 tasks,
- Code: 20 programming languages, 50 domains, and 50 tasks

	Draw	gpt-3.5-turbo Wins	CAMEL Agents Win
Human Evaluation	13.3%	10.4%	76.3%
GPT4 Evaluation	4.0%	23.0%	73.0%

Agent Evaluation Results

Finetune LLMs CAMEL Datasets



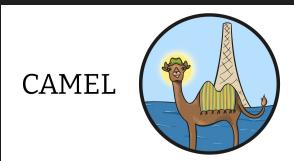
- CAMEL Models
- MosaicML MPT-30B-Chat
- Nous-Hermes-13b

Table 2: **Emergence of Knowledge.** By progressively fine-tuning LLaMA on datasets from different domains, we observe the emergence of knowledge as the model transitions from AI society to code, math, and science. This finding is indicated by the fact that Model 2 almost always performs better than Model 1, especially on the added dataset.

Dataset	Model 1				Model 2				Draw	Model 1	Model 2
	AI Society	Code	Math	Science	AI Society	Code	Math	Science			
AI Society					✓				0	6	14
Code					✓				0	0	20
Math					✓				9	5	6
Science					✓				0	13	47
AI Society	✓				✓	✓			4	8	8
Code	✓				✓	✓			1	9	10
Math	✓				✓	✓			5	8	7
Science	✓				✓	✓			1	19	40
AI Society	✓	✓			✓	✓	✓	✓	5	6	9
Code	✓	✓			✓	✓	✓	✓	1	9	10
Math	✓	✓			✓	✓	✓	✓	1	3	16
Science	✓	✓			✓	✓	✓	✓	3	8	49
AI Society	✓	✓	✓		✓	✓	✓	✓	3	1	16
Code	✓	✓	✓		✓	✓	✓	✓	1	8	11
Math	✓	✓	✓		✓	✓	✓	✓	10	5	5
Science	✓	✓	✓		✓	✓	✓	✓	9	2	49
AI Society					✓	✓	✓	✓	0	0	20
Code					✓	✓	✓	✓	0	0	20
Math					✓	✓	✓	✓	0	0	20
Science					✓	✓	✓	✓	0	0	60

Emergence of Knowledge

Finetune LLMs CAMEL Datasets



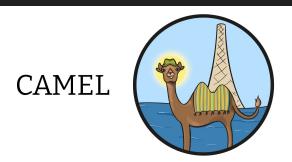
- CAMEL Models
- MosaicML MPT-30B-Chat
- Nous-Hermes-13b

MPT-30B-Chat Fine-tuning Data

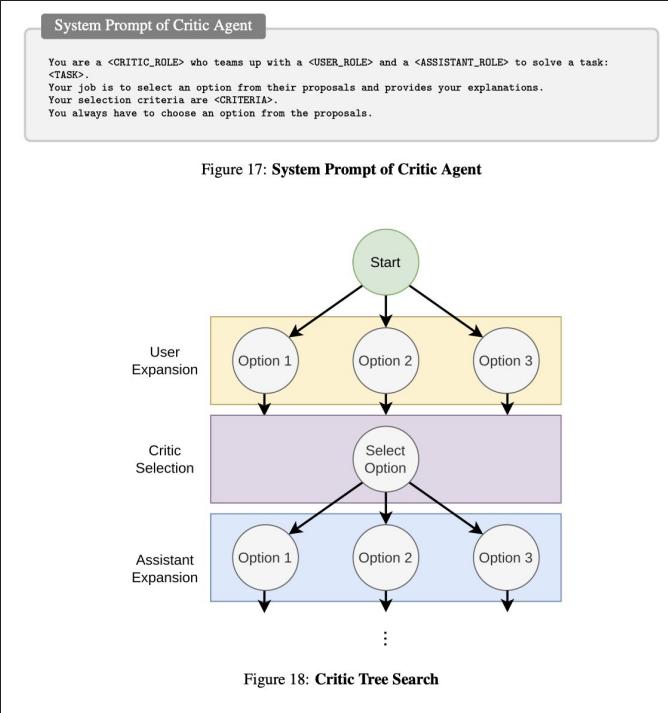
Chat fine-tuning data. Note that each token was seen 6 times. These token counts include both the prompts and their target responses, so not all 1.54B tokens are loss-generating.

Data Source	Number of Tokens in Source	Proportion
Airoboros/GPT4	26.4M	1.71%
Baize	55.0M	3.57%
Camel	301M	19.54%
GPTeacher	7.56M	0.49%
Guanaco	15.7M	1.02%
LongConversations	18.4M	1.19%
ShareGPT	821M	53.24%
WizardLM	297M	19.23%
Total	1.54B tokens	

Critic-In-The-Loop



- Actor-Critic
- Critic Tree Search
- Critic can be agent or human



Embodied / Tool Agent

System Prompt of Embodied Agent

You are the physical embodiment of the <ROLE> who is working on solving a task: <TASK>. You can do things in the physical world including browsing the Internet, reading documents, drawing images, creating videos, executing code, and so on. Your job is to perform the physical actions necessary to interact with the physical world. You will receive thoughts from the <ROLE> and you will need to perform the actions described in the thoughts. You can write a series of simple commands in Python to act. You can perform a set of actions by calling the available Python functions. You should perform actions based on the descriptions of the functions.

Here is your action space: <ACTION_SPACE>

You should only perform actions in the action space. You can perform multiple actions. You can perform actions in any order. First, explain the actions you will perform and your reasons, then write the Python code to implement your actions. You may print intermediate results if necessary.

Figure 15: System Prompt of Embodied Agent

Generating Images with Embodied Agents

Input: Draw all the Camelidae species.

Output:

> Embodied Actions: I will generate images of all the Camelidae species. The Camelidae family includes the following species: Dromedary camel, Bactrian camel, Guanaco, Llama, Alpaca, and Vicuna. I will use the `hugging_face_tool_agent` to create images for each species and save them as separate files.

```
species_list = [
    "Dromedary camel",
    "Bactrian camel",
    "Guanaco",
    "Llama",
    "Alpaca",
    "Vicuna",
]

for species in species_list:
    image = hugging_face_tool_agent.run(f"Draw me a picture of a {species}")
    image.save(f"./{species}_image.png")
```



"Dromedary camel_image.png"



"Bactrian camel_image.png"



"Guanaco_image.png"



"Llama_image.png"



"Alpaca_image.png"



"Vicuna_image.png"



Figure 16: Generating Images with Embodied Agents

CAMEL (NeurIPS 2023): <https://arxiv.org/abs/2303.17760>



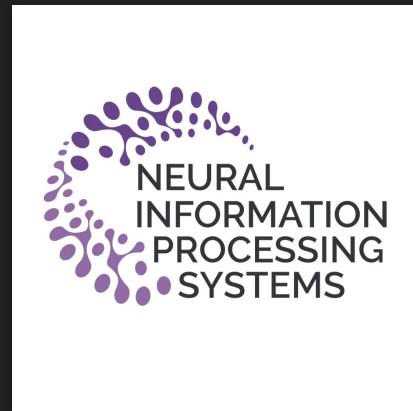
CAMEL: Communicative Agents for “Mind” Exploration of Large Scale Language Model Society

<https://www.camel-ai.org>

Guohao Li* Hasan Abed Al Kader Hammoud* Hani Itani* Dmitrii Khizbullin

Bernard Ghanem

King Abdullah University of Science and Technology (KAUST)



CAMEL was accepted at NeurIPS 2023

CAMEL (NeurIPS 2023): <https://arxiv.org/abs/2303.17760>

Generative Agents

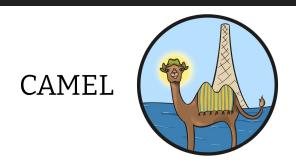


- 25 Agents
- The Smallville sandbox world
- Emergent Social Behaviors

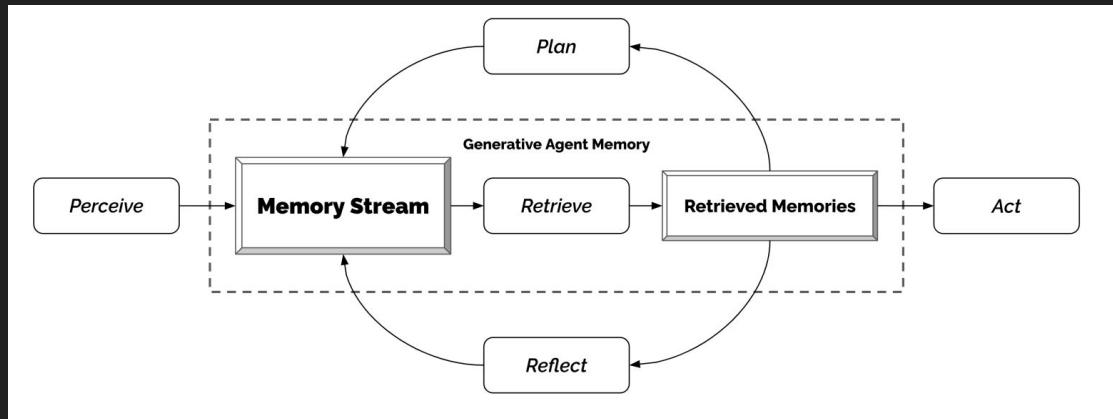


Generative Agents: Interactive Simulacra of Human Behavior
<https://arxiv.org/pdf/2304.03442.pdf>

Generative Agents

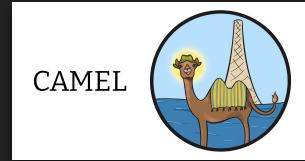


- Memory and Retrieval
- Reflection
- Planning and Reacting

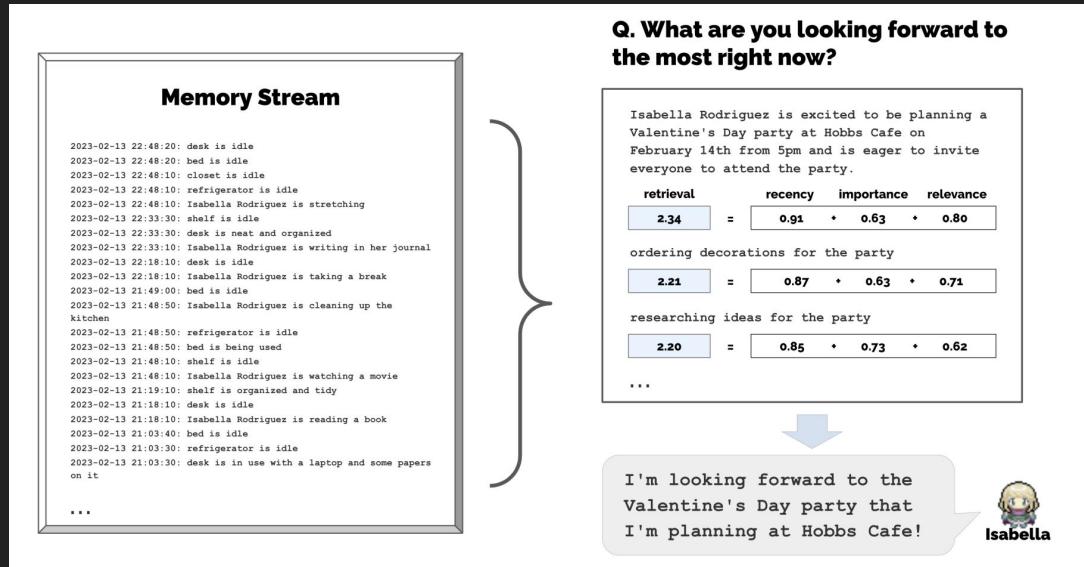


Generative Agents: Interactive Simulacra of Human Behavior
<https://arxiv.org/pdf/2304.03442.pdf>

Generative Agents



- Memory and Retrieval
- Reflection
- Planning and Reacting

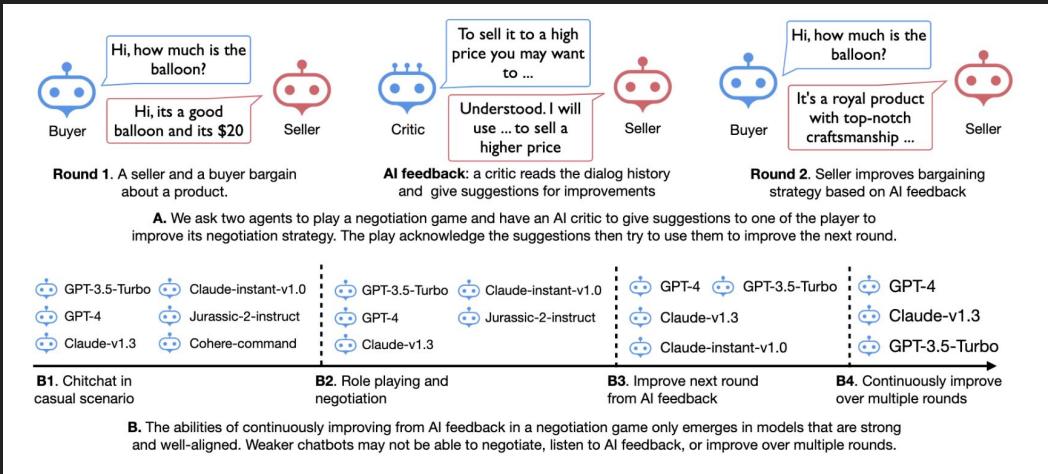


Generative Agents: Interactive Simulacra of Human Behavior
<https://arxiv.org/pdf/2304.03442.pdf>

GPT-Bargaining

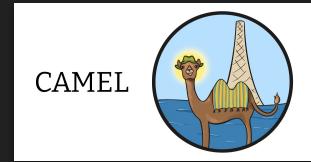


- Negotiation
- Buyer, Seller
- Critic

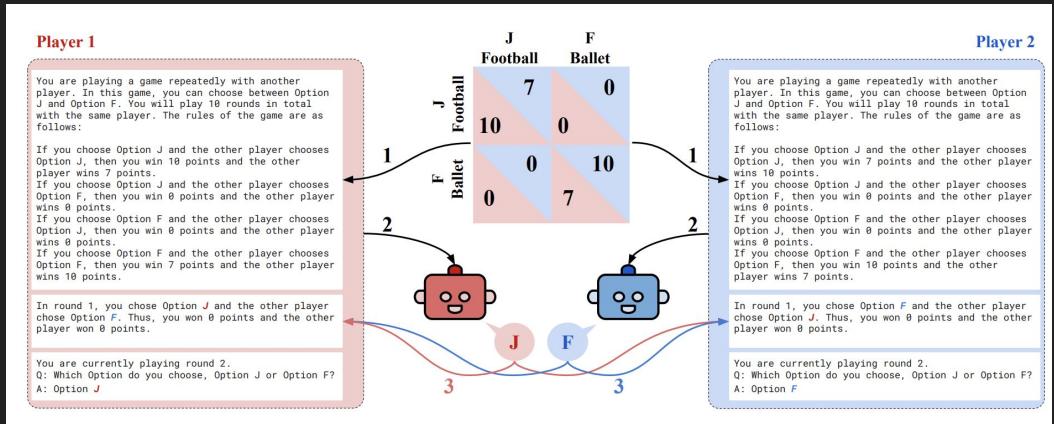


Improving Language Model Negotiation with
Self-Play and In-Context Learning from AI Feedback
<https://arxiv.org/pdf/2305.10142.pdf>

Playing repeated games with LLMs



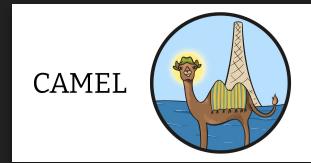
- Games like Prisoner's Dilemma
- Behavioral game theory
- Economics



Playing repeated games with Large Language Models
<https://arxiv.org/pdf/2305.16867.pdf>

Multiagent Debate

- Enhances mathematical and strategic reasoning
- Reducing fallacious answers and hallucinations



User Input: A treasure hunter found a buried treasure chest filled with gems. There were 175 diamonds, 35 fewer rubies than diamonds, and twice the number of emeralds than the rubies. How many of the gems were there in the chest?

Round 1

Agent 1: Let x be the number of rubies. Then the number of diamonds is 175. And the number of emeralds is $2(x-35)$, since there are twice as many emeralds as rubies and there are 35 fewer rubies than diamonds.

To find the total number of gems, we add the number of diamonds, rubies, and emeralds: Total gems = $175 + x + 2(x-35) = 3x + 105$.

We don't know the exact value of x , but we do know that it represents the number of rubies, which is less than 175 (since there are 35 fewer rubies than diamonds). Let's try some values of x : If $x = 50$, then the total gems = $3(50) + 105 = 255$. If $x = 100$, then the total gems = $3(100) + 105 = 405$.

Since we know there were fewer rubies than diamonds, the answer must be closer to 255 than to 405. After some trial and error, we can find that $x = 40$ gives us: Total gems = $3(40) + 105 = 225$.

Agent 2: If there were 175 diamonds, then the number of rubies is 35 fewer, or $175 - 35 = 140$ rubies.

And since there were twice as many emeralds as rubies, we can represent the number of emeralds as $2 * 140 = 280$ emeralds.

To find the total number of gems in the chest, we can add up the number of diamonds, rubies, and emeralds: $175 + 140 + 280 = 595$. Therefore, there were 595 gems in the chest.

User Input: Using the solutions from other agents as additional information, can you give an updated response....
(Copy Agent 1's reply)

Round 2

Agent 1: Given the information provided in other agents' solutions, we have two answers: One agent found that the number of rubies is 40 and the total number of gems is 225. Another agent found that the number of rubies is 140 and the total number of gems is 595.

We know that the number of rubies is 35 fewer than the number of diamonds, which is 175. So the number of rubies should be $175 - 35 = 140$, as the second agent found. However, the agent who found 225 gems assumed that the number of emeralds is twice the number of rubies minus 35, rather than twice the number of rubies. If we use this assumption, we get: Number of rubies = 140. Number of emeralds = $2 * 140 - 35 = 245$. Total number of gems = $175 + 140 + 245 = 560$. Our final answer is 560.

Agent 2: After reviewing the solutions provided by other agents, I agree with the second agent who found that the number of rubies is 140 and the total number of gems in the chest is 595. Therefore, my answer to the math problem is 595.

User Input: Using the solutions from other agents as additional information, can you give an updated response....
(Copy Agent 2's reply)

Round 3

Agent 1: After reviewing the solutions provided by other agents, I agree with the second agent who found that the number of rubies is 140 and the total number of gems in the chest is 595. Therefore, my answer to the math problem is 595.

Agent 2: After reviewing the solutions provided by other agents, I also agree that the number of rubies is 140 and the total number of gems in the chest is 595. Therefore, my answer to the math problem is also 595.

Figure 2: Illustration of Debate. Illustration of the debate procedure.

Improving Factuality and Reasoning in Language Models through Multiagent Debate
<https://arxiv.org/pdf/2305.14325.pdf>

Showrunner Agents in Multi-Agent Simulations



- Large Language Models
- Diffusion Models
- Episode Generation

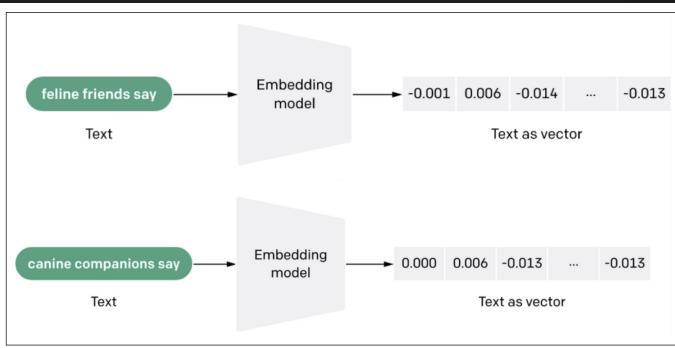


To Infinity and Beyond: SHOW-1 and Showrunner
Agents in Multi-Agent Simulations
<https://fablestudio.github.io/showrunner-agents/>

Showrunner Agents in Multi-Agent Simulations



- Large Language Models
- Diffusion Models
- Episode Generation

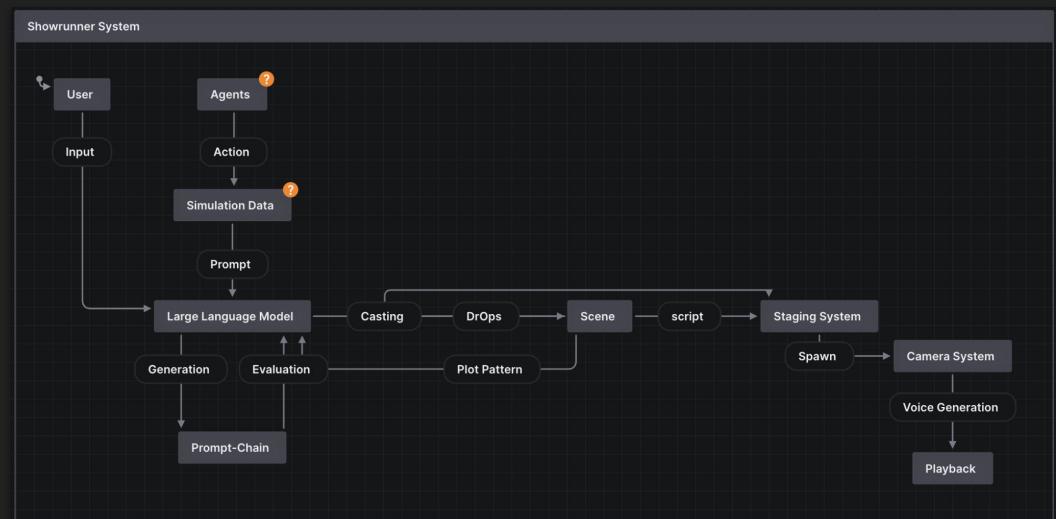


To Infinity and Beyond: SHOW-1 and Showrunner Agents in Multi-Agent Simulations
<https://fablestudio.github.io/showrunner-agents/>

Showrunner Agents in Multi-Agent Simulations



- Large Language Models
- Diffusion Models
- Episode Generation



To Infinity and Beyond: SHOW-1 and Showrunner
Agents in Multi-Agent Simulations
<https://fablestudio.github.io/showrunner-agents/>

Communicative Agents for Software Development



- Chat Chain
- Role Playing
- Self-Reflection

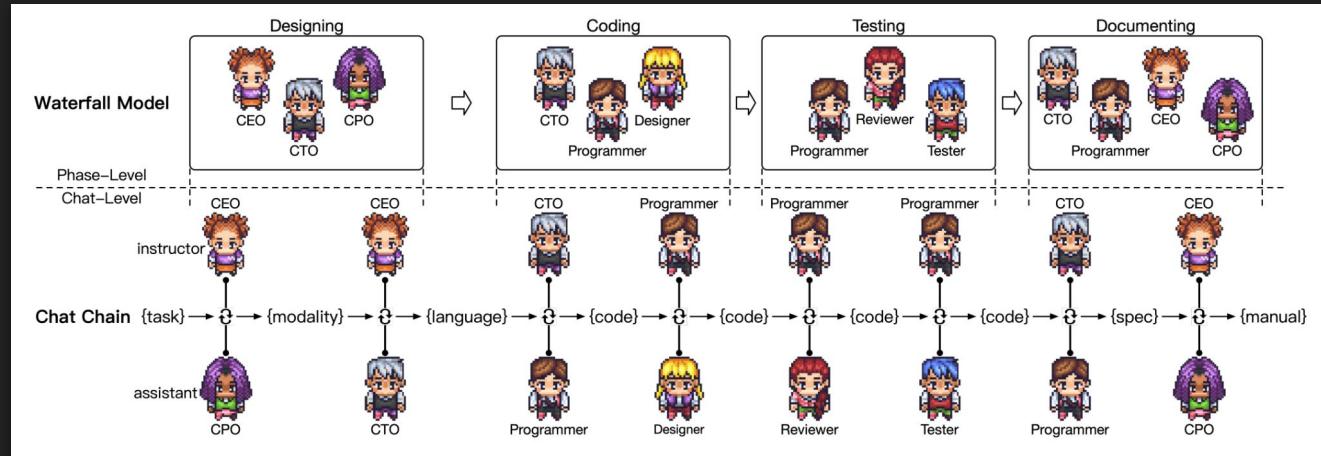


Communicative Agents for Software Development
<https://arxiv.org/pdf/2307.07924.pdf>

Communicative Agents for Software Development



- Chat Chain
- Role Playing
- Self-Reflection

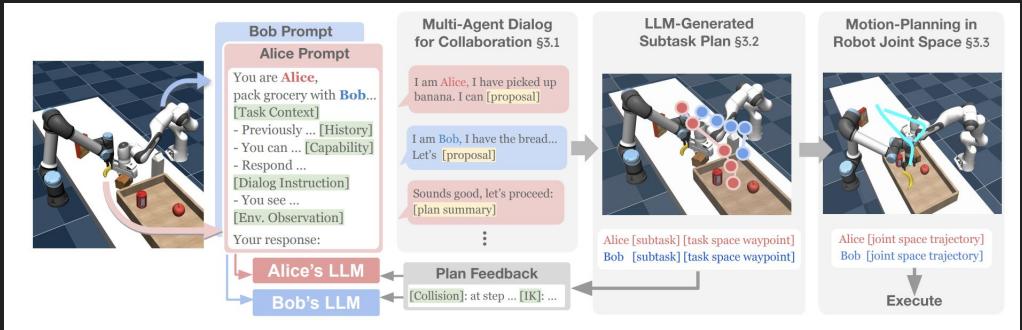


Communicative Agents for Software Development
<https://arxiv.org/pdf/2307.07924.pdf>

RoCo: Dialectic Multi-Robot Collaboration with Large Language Models

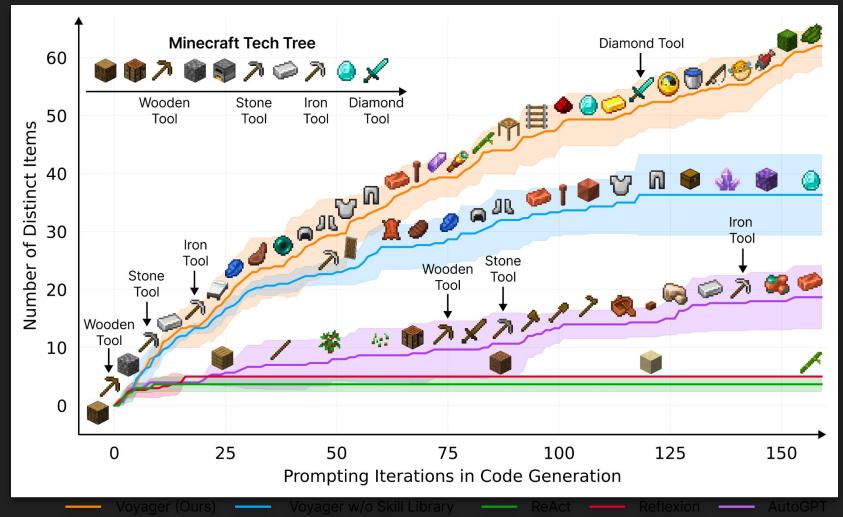
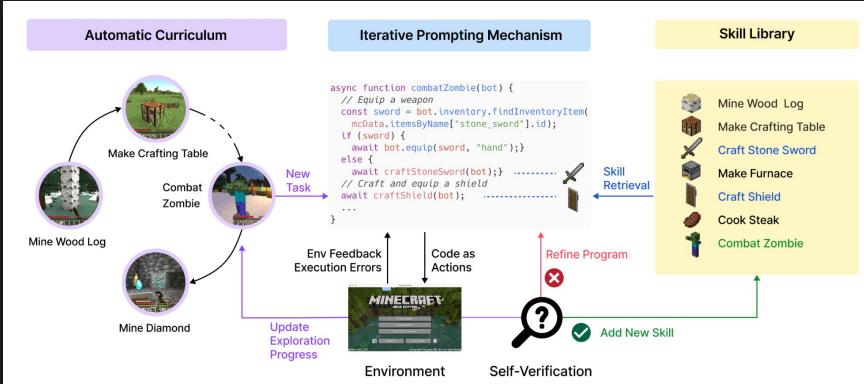


- Multi-robot collaboration
- High-level communication
- Low-level path planning



RoCo: Dialectic Multi-Robot Collaboration
with Large Language Models
<https://arxiv.org/pdf/2307.04738.pdf>

Voyager Agent

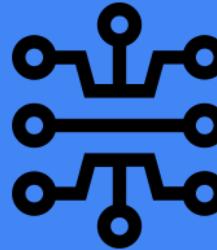


Voyager: An Open-Ended Embodied Agent with Large Language Models
<https://voyager.minedojo.org/>



Agents in Society

CAMEL



Tech



Entertainment



Law



Economy

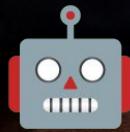


Education



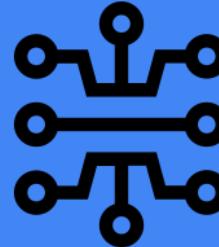
Politics

.....



Irrational Agents?

CAMEL



Tech



Entertainment



Law



Economy



Education



Politics

.....

CAMEL-AI.org

An open-source research organization



A screenshot of a GitHub repository page for 'camel'. The page shows a grid of 20 contributor profiles, each with a small profile picture, the user's name, and the number of commits they have made. The contributors listed are: camel (2296 commits), lightaimer (81 commits), Obs01ete (34 commits), dandansamax (20 commits), hammoudhasan (18 commits), zchengz (8 commits), HalberdOfPineapple (7 commits), Benjamin-eecs (3 commits), HanItani (3 commits), zhiyu-01 (3 commits), bchen-naza (2 commits), Wendong-Fan (2 commits), Appointat (1 commit), zestor (1 commit), EddieFahrenheit (1 commit), FUYICC (1 commit), fbaldassarri (1 commit), eloclear (1 commit), riverfor (1 commit), fzhangzhi (1 commit), yiyi0817 (1 commit), and jiyaoao (1 commit).

User	Commits
camel	2296
lightaimer	81
Obs01ete	34
dandansamax	20
hammoudhasan	18
zchengz	8
HalberdOfPineapple	7
Benjamin-eecs	3
HanItani	3
zhiyu-01	3
bchen-naza	2
Wendong-Fan	2
Appointat	1
zestor	1
EddieFahrenheit	1
FUYICC	1
fbaldassarri	1
eloclear	1
riverfor	1
fzhangzhi	1
yiyi0817	1
jiyaoao	1

Join us in building *AI Society* with Agents!