fun-ai-talk

A Glimpse into AI for Games

SJ Zhu

hululu.zhu@gmail.com

08/20/2023

github.com/hululuzhu/mario-ai-2023

Today's Agenda

- About me and this talk
- Intro to Deep Learning AI
  - **Lab**: Tensorflow Playground
- Game AI is fun!
  - **Lab**: AI for Cart Pole
- Why is Super Mario AI hard?
  - **Lab**: AI for Mario, homework
- AI for even harder games?
- Beyond games

About me, ~11 years at Google & DeepMind

Intern as PhD Student	Software Engineer (and Manager)	Research Engineer

"Follow your passion", try things you like

Logos above are from images.google.com (fetched Jan 2023)

Disclaimer

This talk is my personal voluntary effort, prepared and conducted during my personal time outside of working hours.

All content is derived from publicly available sources, and the views expressed herein only represent my personal opinions, and do not reflect the positions of DeepMind®, Google®, or Alphabet®.

Images in this slides without references were generated by me using Midjourney® AI

Purpose of this talk

- Call for awareness in AI, especially for games
- Build your first AI programs with us together
- Make you a bit more interested and excited about AI

Intro to AI and Deep Neural Network (DNN)

AI = Artificial Intelligence

- Artificial: made by humans
- Intelligence: ability to be smart and figure out things

AI was brought up in 1950s, and has many ways of implementations

- Heuristics such as thousands of if/else logic
- Statistical or traditional (non-DNN) approaches
- Deep Neural Network (DNN, aka deep learning)

When we refer “AI” in this talk, it refers to “Deep Neural Network” only!

Unbox Deep Neural Network (DNN) a bit

Inspired by the structure of human brain neurons. DNN is often

- **Deep**: Neurons are grouped into "layers." A deep network is a stacked by many layers
- **Wide**: In each layer, it has many neurons that do complex signal/data processing and decision making together
- **Activation**: Similar to brain neurons, not all neurons need to be completely activated in every signal processing and decision.
    - Imagine this: Visual neurons probably don't need to react to sound or taste.

What are you talking about? Alright, let's play with a live example!

Lab time:

Tensorflow Playground

Quick recap of DNN AI

A powerful computer program that

- Human programmed architecture and allocated compute budget
- Take input of signals (measurements, text, visual, sound, and so on)
- Try to make decisions on these signals
  - Either by being taught by humans with “labels” (ground truth answers, this image is Mario not Elsa)
  - Or figure out “action” per step by humans define a “goal” (e.g. Win the Chess game)
  - Start with random guess
- “Learning”
  - Fill the gap of AI guess and outcome through “optimization”
- In the end: A powerful AI

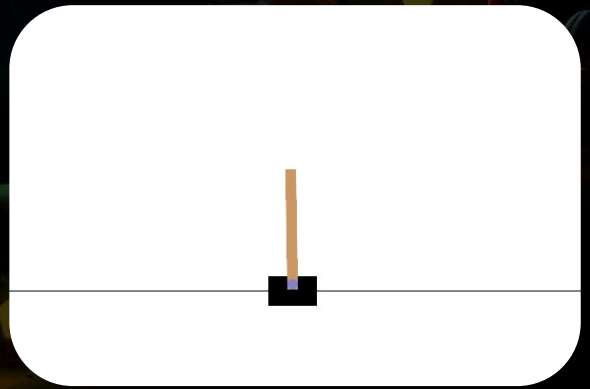
Game AI is fun!

Let's refer "Game" to "Video Games" for now. Because they are fun!

AI which can play the games will be fun!

To build and train a Game AI by yourself? Even more fun!

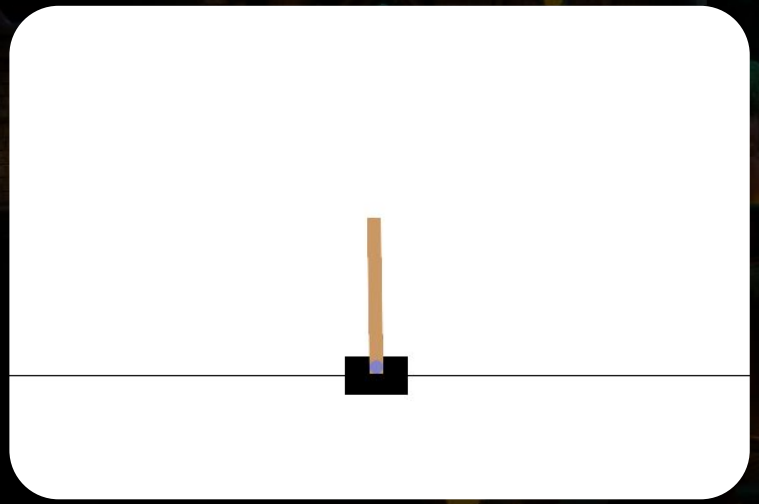
With limited class time (and compute budget), let's use this cart pole game as our first Game AI example!



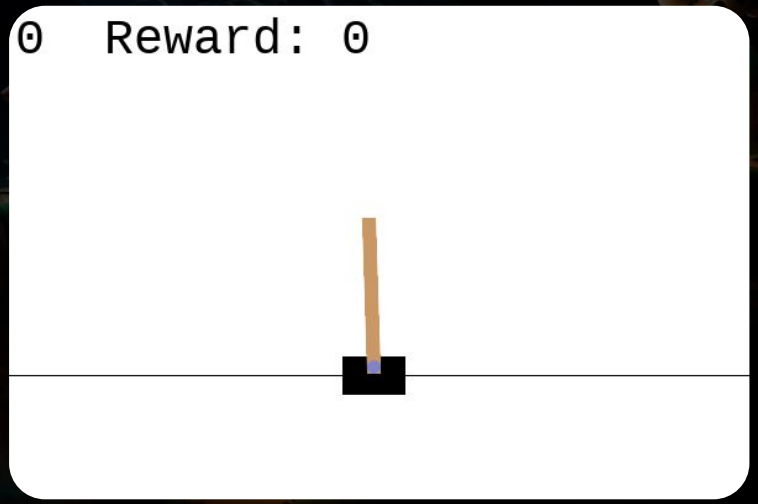
Cart Pole Zero to Hero is not Easy!

- We will play with it soon!

Hard to survive at first



Pro to keep alive after



AI for cart pole game is still challenging, some physics

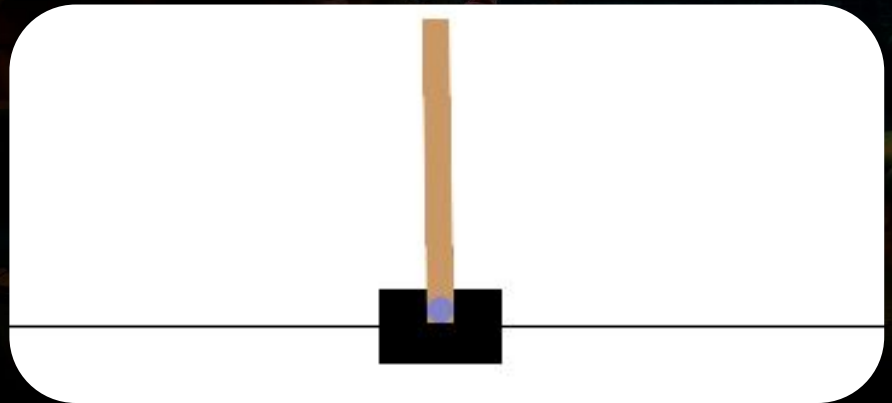
```
import gym

env = gym.make('CartPole-v1')
env.reset().tolist()
```

```
[0.048271920531988144,
 -0.04601958766579628,
 0.04441657289862633,
 0.045917950570583344]
```

What are the 4 numbers?

Why can the 4 numbers enough to represent the game state?



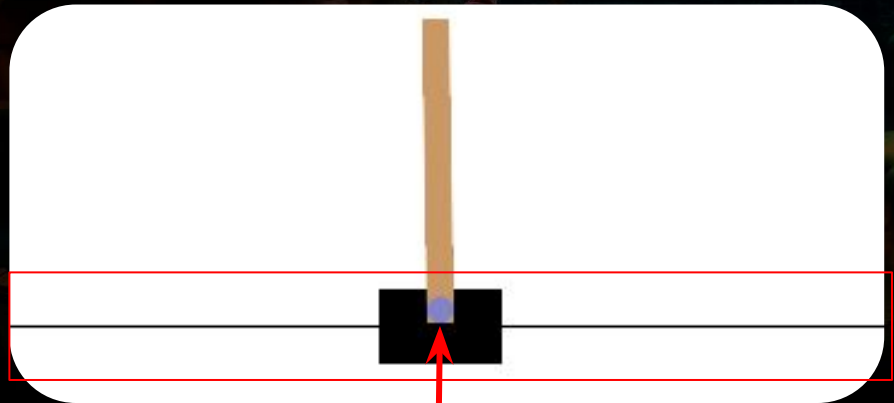
AI for cart pole game is still challenging, some physics

```
import gym

env = gym.make('CartPole-v1')
env.reset().tolist()
```

```
[0.048271920531988144,
 -0.04601958766579628,
 0.04441657289862633,
 0.045917950570583344]
```

Num	Observation	Min	Max
0	Cart Position	-4.8	4.8
1	Cart Velocity	-Inf	Inf
2	Pole Angle	~ -0.418 rad (-24°)	~ 0.418 rad (24°)
3	Pole Angular Velocity	-Inf	Inf



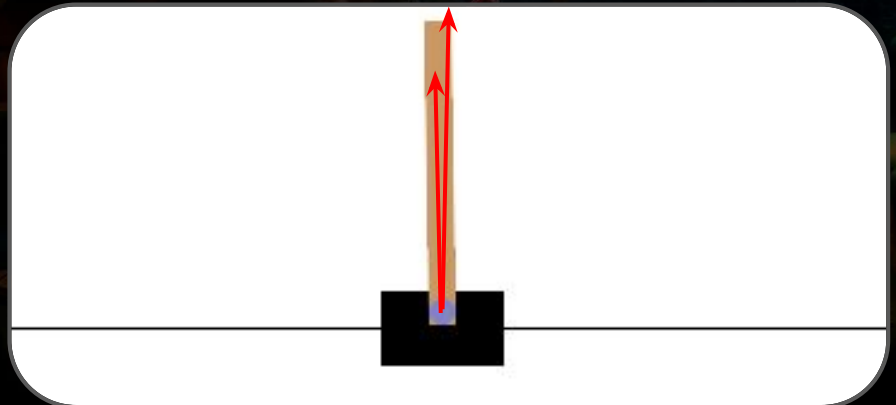
AI for cart pole game is still challenging, some physics

```
import gym

env = gym.make('CartPole-v1')
env.reset().tolist()
```

```
[0.048271920531988144,
 -0.04601958766579628,
 0.04441657289862633,
 0.045917950570583344]
```

Num	Observation	Min	Max
0	Cart Position	-4.8	4.8
1	Cart Velocity	-Inf	Inf
2	Pole Angle	~ -0.418 rad (-24°)	~ 0.418 rad (24°)
3	Pole Angular Velocity	-Inf	Inf



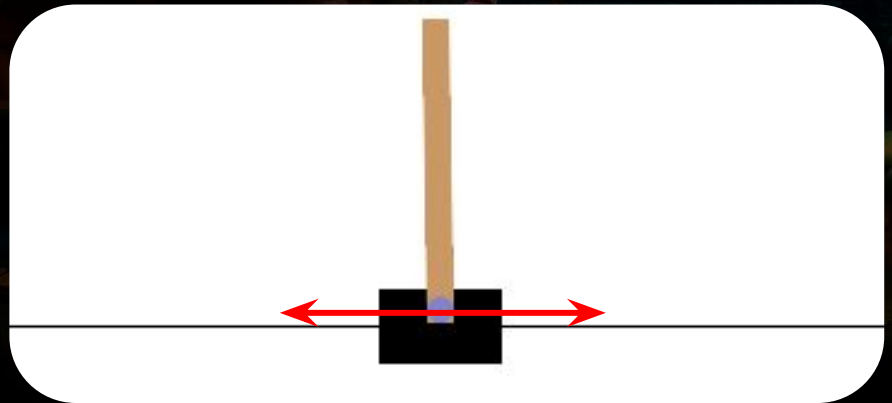
AI for cart pole game is still challenging, some physics

```
import gym

env = gym.make('CartPole-v1')
env.reset().tolist()
```

```
[0.048271920531988144,
 -0.04601958766579628,
  0.04441657289862633,
  0.045917950570583344]
```

Num	Observation	Min	Max
0	Cart Position	-4.8	4.8
1	Cart Velocity	-Inf	Inf
2	Pole Angle	~ -0.418 rad (-24°)	~ 0.418 rad (24°)
3	Pole Angular Velocity	-Inf	Inf



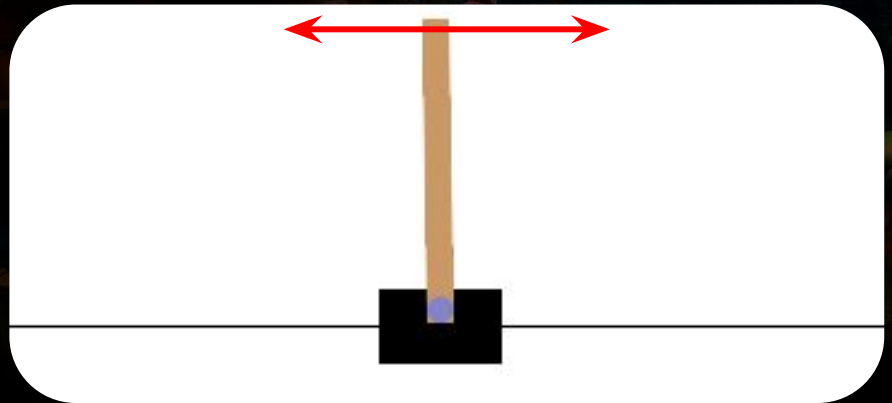
AI for cart pole game is still challenging, some physics

```
import gym

env = gym.make('CartPole-v1')
env.reset().tolist()
```

```
[0.048271920531988144,
 -0.04601958766579628,
 0.04441657289862633,
 0.045917950570583344]
```

Num	Observation	Min	Max
0	Cart Position	-4.8	4.8
1	Cart Velocity	-Inf	Inf
2	Pole Angle	~ -0.418 rad (-24°)	~ 0.418 rad (24°)
3	Pole Angular Velocity	-Inf	Inf



What can we control for cart pole?

Num	Action
0	Push cart to the left
1	Push cart to the right

Note: The velocity that is reduced or increased by the applied force is not fixed and it depends on the angle the pole is pointing. The center of gravity of the pole varies the amount of energy needed to move the cart underneath it

https://www.gymlibrary.dev/environments/classic_control/cart_pole/

Lab time:
Build our Cart Pole AI!
Notebook link

Super Mario AI is Harder!

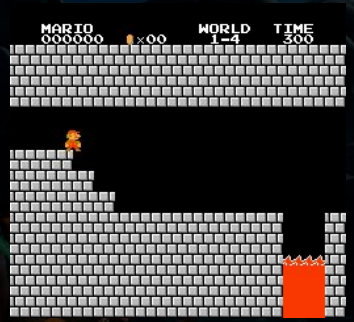
Instead of 4 numbers to represent game state like CartPole

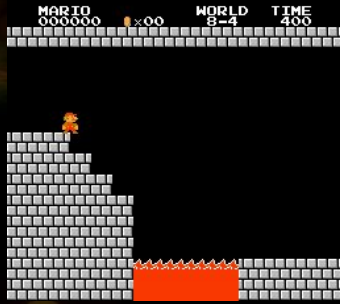
Mario has pictures (256*256*3 numbers!!!) for each frame, 30 frames per second!

8 Worlds * 4 Stages = 32 Stages!

More actions to take!

All gifs are generated using [my colab notebook]





Our strategy to train Super Mario AI

Simplify the visual inputs to AI

- Thus AI only needs a simpler view and less computation

Simplify the actions to take

- To get started, instead of 10+ possible actions, we ask Mario to only move Right, or Up+Right

But, we cannot simplify one thing, sequence of data

- We have to pass seq of images (or similar), otherwise, AI cannot figure out the trajectory (e.g. enemy moving towards or away from Mario)

Hard to understand? Let's take a look at the code and demo!

Lab time:
Super Mario

1. Visualization Notebook, take a glance of all stages!
2. AI Training Notebook (skeleton), homework to build a more powerful AI

Homework

Based on our AI Training Notebook

1. Can you train an AI to finish at least 1 stage?
2. Can you think of a plan to train an AI that could finish multiple stages?
3. How about doing some research on PPO, MlpPolicy and CnnPolicy?

A more Complex AI: AlphaStar 2019 by DeepMind

- Starcraft is more complex
  - Super Mario has ~20 actions
  - Starcraft has huge number of actions to take!
    - E.g. image pixel level
    - Or "what, who, where, when" action sequence!
- To make things more challenging!
  - Decision on imperfect info (e.g. no direct visibility to opponent)
  - AI has to respond in ~100ms

A story of AlphaStar First Author: Oriol Vinyals

VP of DeepMind Research

Used to be Pro StarCraft Player in Spain

Tried to convince parents StarCraft is meaningful, but failed

Worked hard at Berkeley on StarCraft AI, but failed

Continued his dream at DeepMind and succeeded

Bought Nature magazine, flied back to Spain, and told his parents “I told you 15 years ago!”

Image from https://www.youtube.com/watch?v=Kedt2or9xlo

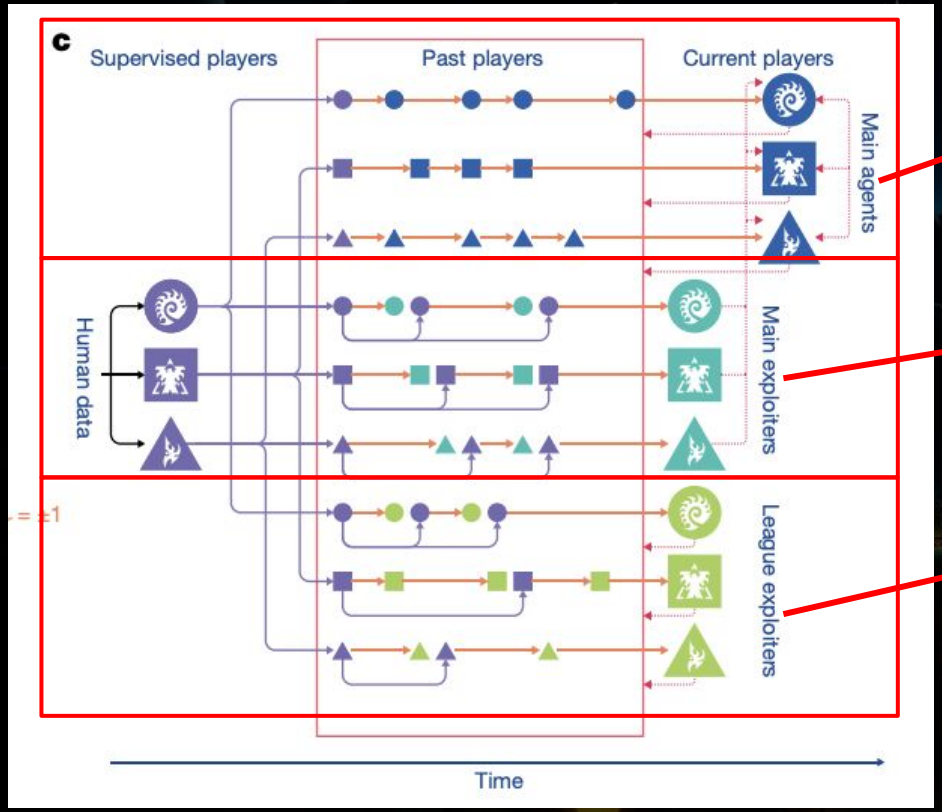
Let's take a look at the StarCraft game

Basic operations

Team Fights

Reference:
https://techcrunch.com/2017/08/09/blizzard-and-deepmind-turn-starcraft-ii-into-an-ai-research-lab/
https://www.reddit.com/r/starcraft/comments/s4qr1/what_is_your_favorite_starcraft_gif/

AlphaStar 2019, multi-agent league training



3 groups, 12 agents in total

- Main agents (1 x 3)
  - Self-play with all other 3 races of agents and try to improve
- Main exploiters (1 x 3)
  - Exploit weakness in "Main" group by playing with Main agents only
- League exploiters (2 x 3)
  - Try to find systematic weakness of the "entire league"

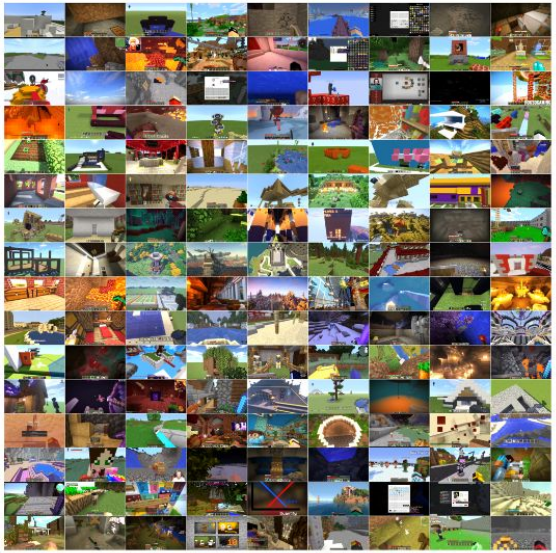
Exploiters periodically re-initialized (purple shapes) to increase diversity and robustness

MineDojo 2022 for MineCraft

Video from https://minedojo.org/ =

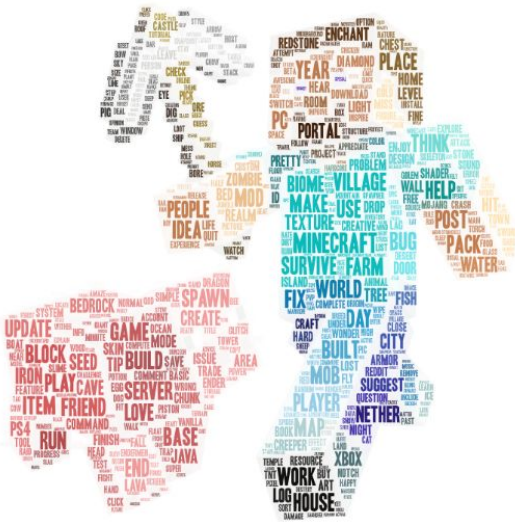
How is the awesome MineDojo trained?

- Look at how similarly this AI learns as compared to humans like us!



730K YouTube videos with **2.2B** words in English transcripts

~7K Wiki pages with interleaved text, images, tables, and diagrams



340K Reddit posts with **6.6M** comments in r/Minecraft subreddit

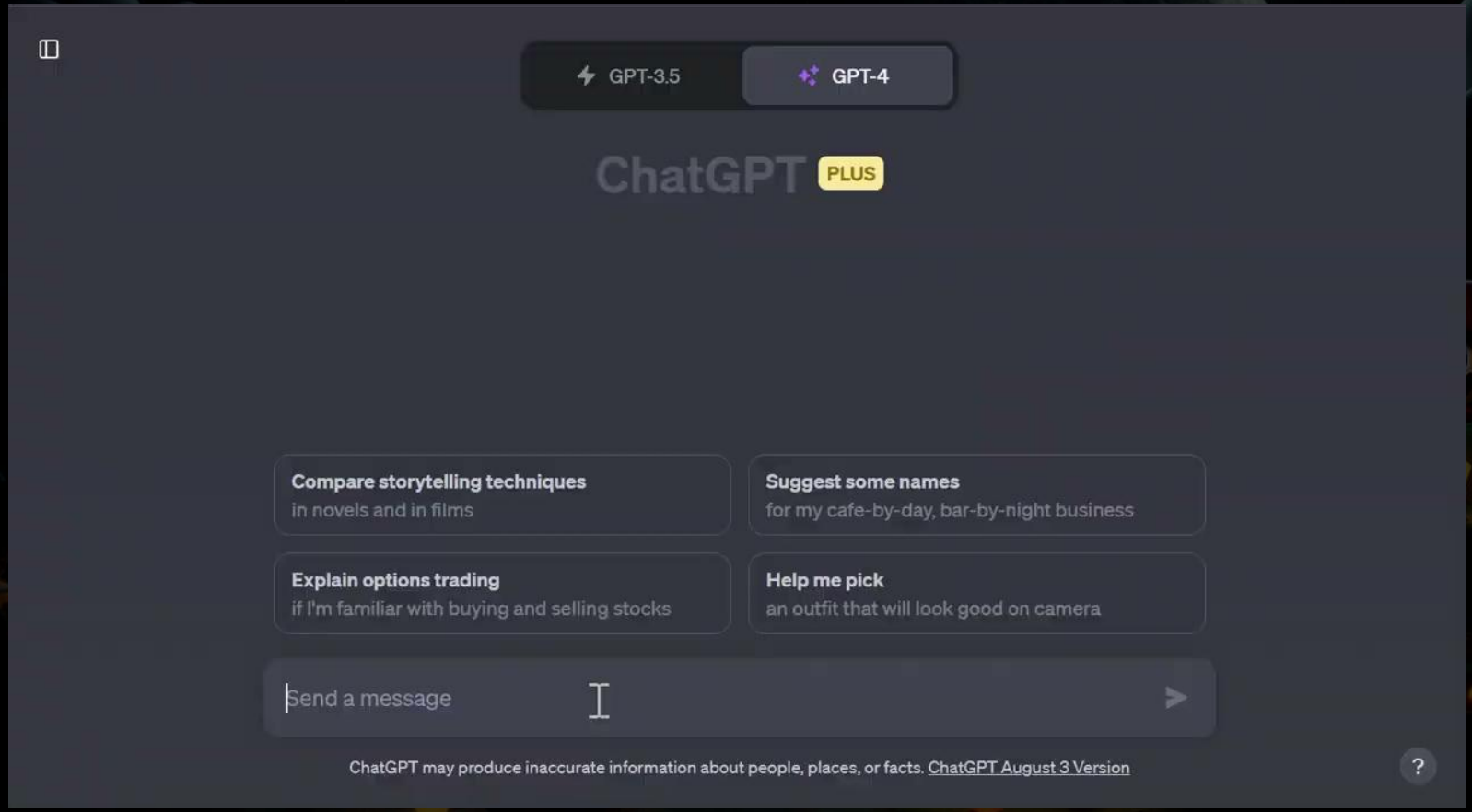
More Highlights of MineDojo 2022 from minedojo.org

- Jeff Bezos follows the 1st author of this Paper, Jim Fan's twitter!
- I asked Jim what inspired him on this research: "I just love MineCraft" !

Text commands to AI Actions!

The algorithm highlights!

Lastly, let's briefly talk about ChatGPT



Recorded 08/20/2023

How is ChatGPT related to Game AI we talked about

- Assume we want to train ChatGPT to play a game called "SAT"

1. Pretrain	*2. SFT (supervised finetune)*	*3. RM (Reward Model)*	*4. RLHF (Reinforcement learning from Human Feedback)*
Read textbooks, gain "knowledge"	Mimic from teachers and peers how to answer SAT questions	Think hard: How does SAT scoring work?	Self Improve (as Game AI learns) to expect higher SAT scores!

Guess what algorithm commonly used in RLHF?

How is ChatGPT related to Game AI we talked about

- Assume we want to train ChatGPT to play a game called "SAT"

1. Pretrain	*2. SFT (supervised finetune)*	*3. RM (Reward Model)*	*4. RLHF (Reinforcement learning from Human Feedback)*
Read textbooks, gain "knowledge"	Mimic from teachers and peers how to answer SAT questions	Think hard: How does SAT scoring work?	Self Improve (as Game AI learns) to expect higher SAT scores!

Guess what algorithm commonly used in RLHF?

PPO, which we used to train cartpole and Mario!

Q&A + Discussion now!

A Glimpse into AI for Games

- Turn off recording
- Anything related to games and AI plz!

Before we finish

Why do we still need to study hard in the AI Era?

Think about it and discuss with your friends!


Thank you!
May games & AI bring
more joy to you!