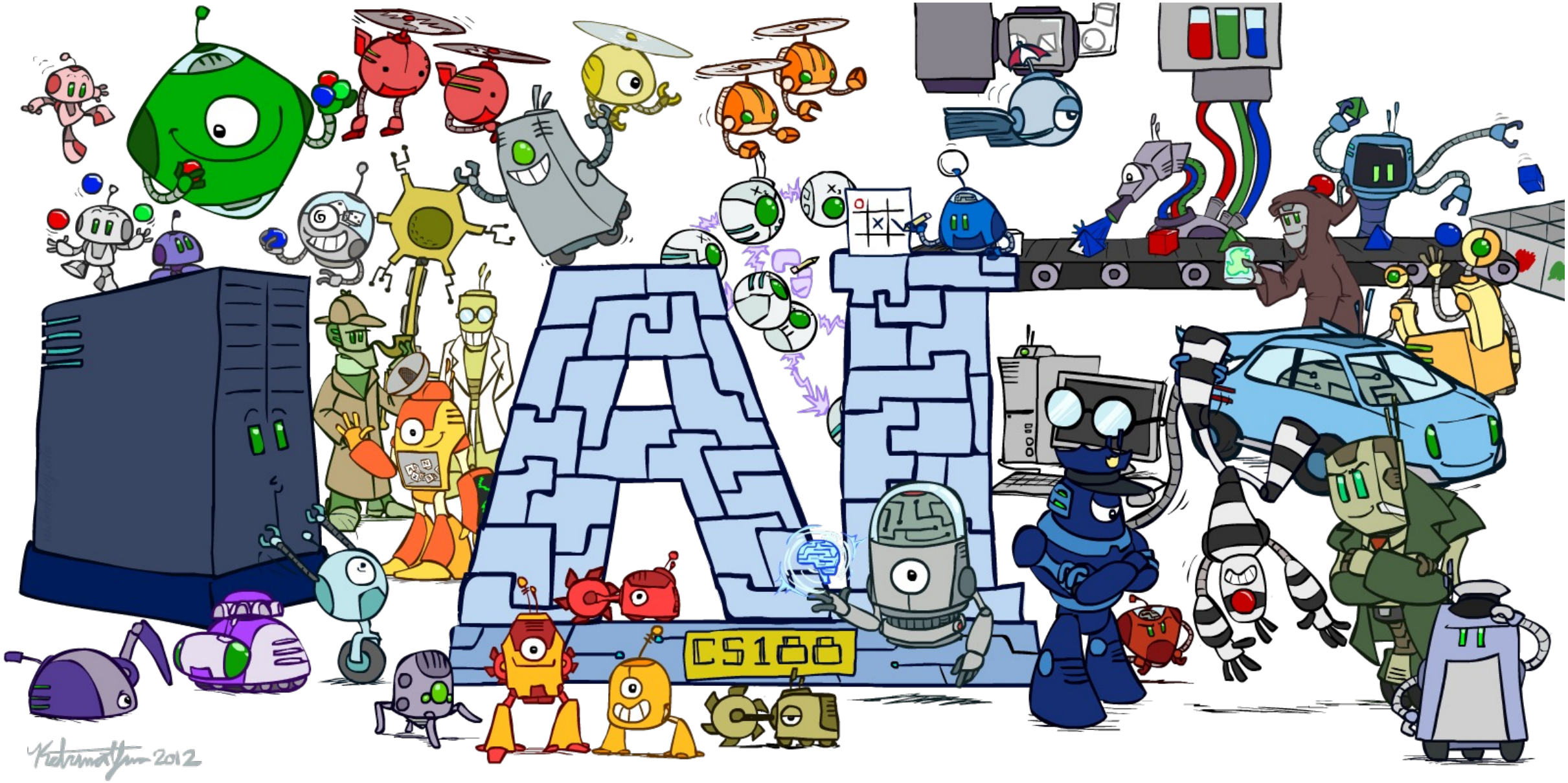
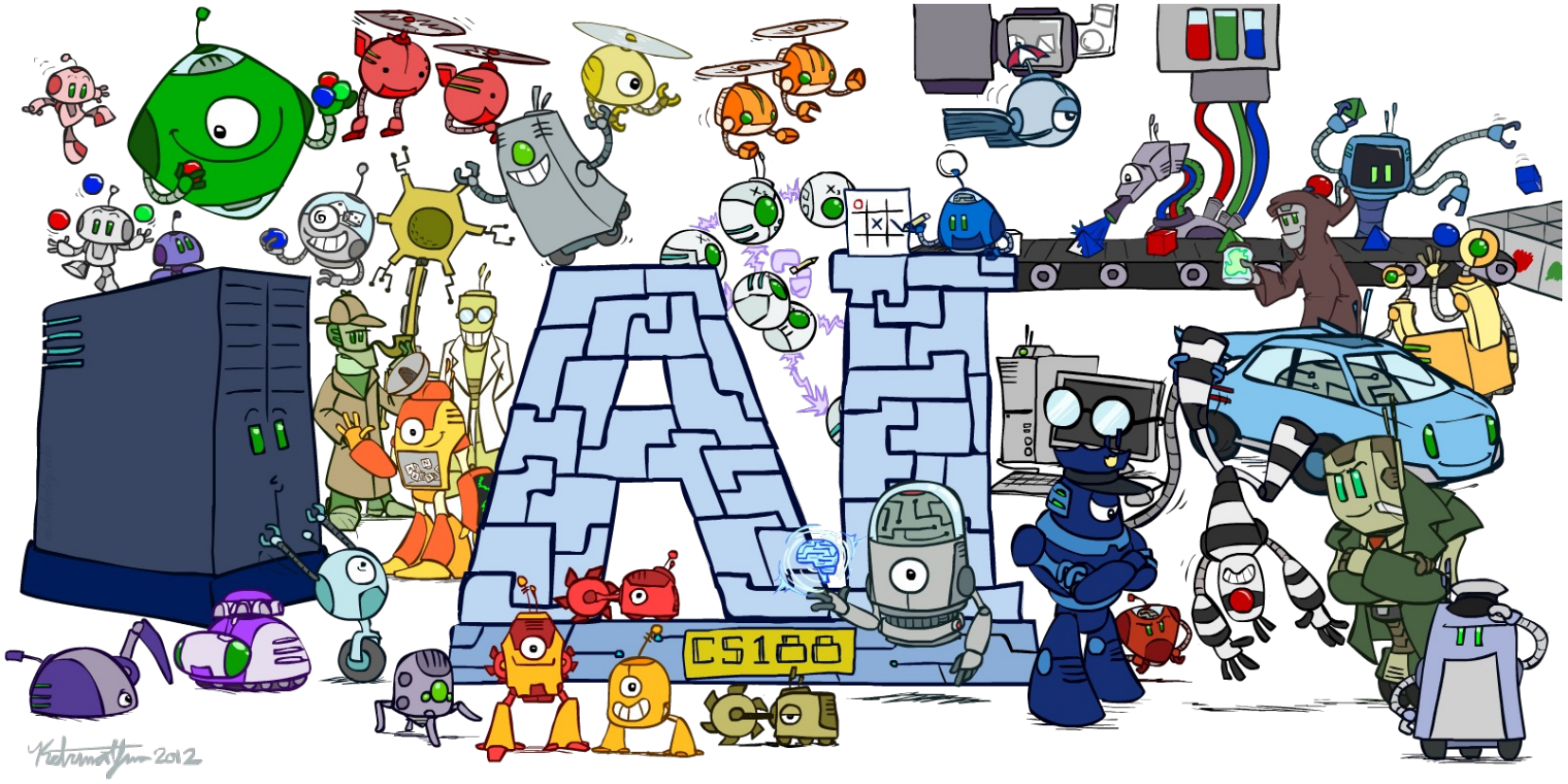
CS 188: Artificial Intelligence

Conclusion

Instructor: Igor Mordatch & Pieter Abbeel --- University of California, Berkeley

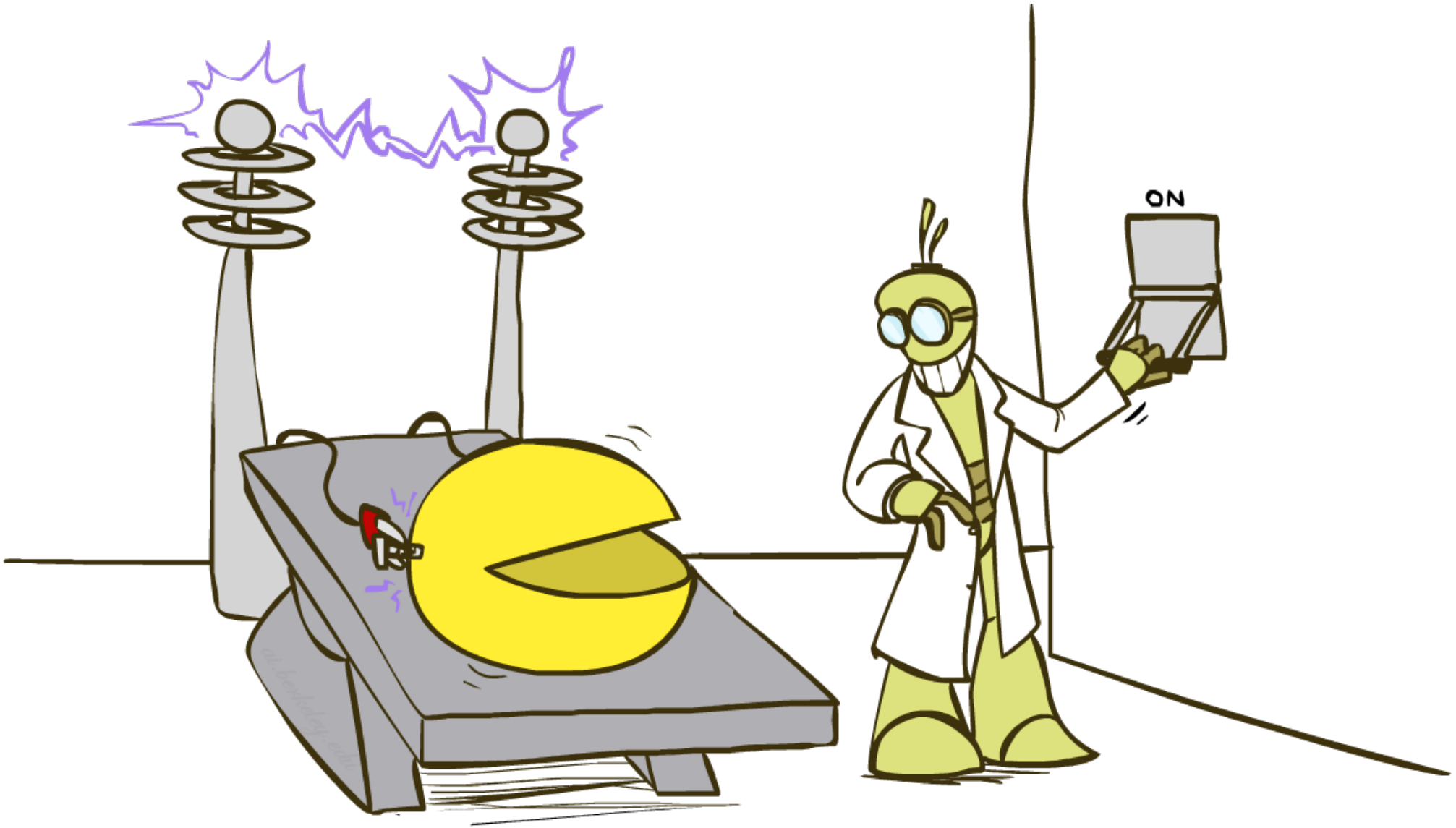

CS188
Kevinmatthews 2012



Ketrina Yim

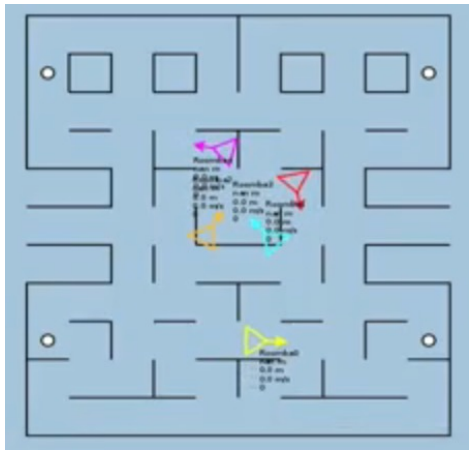
CS188 Artist

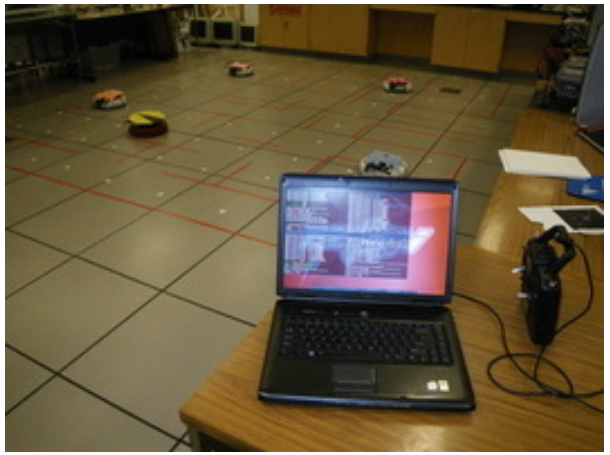
Pac-Man Beyond the Game!



Pacman: Beyond Simulation?

ROOMBA PACMAN





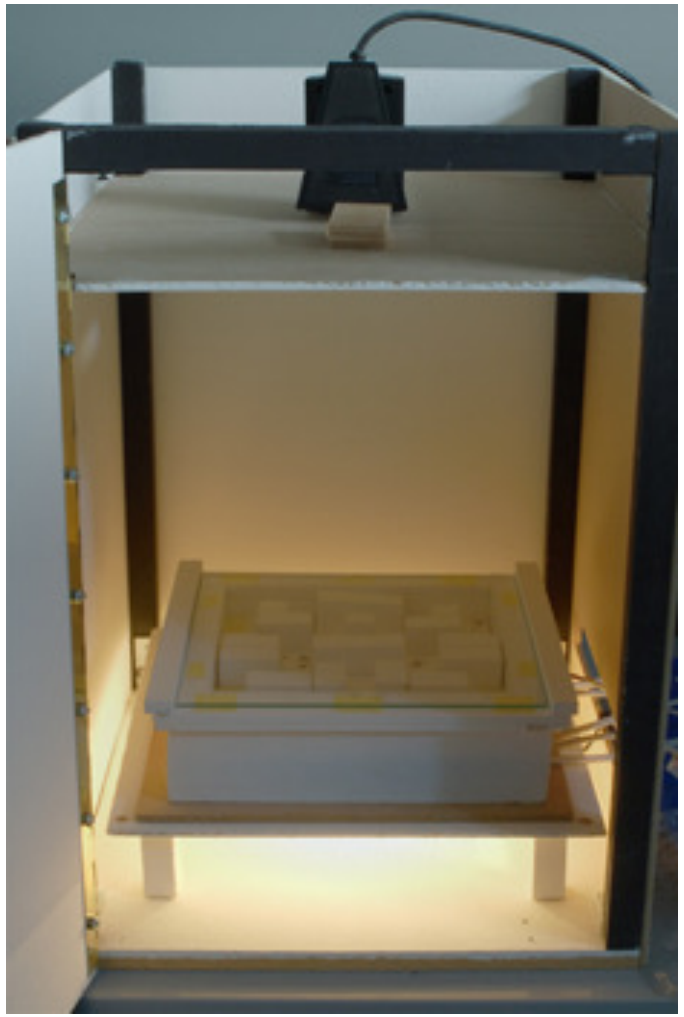
Students at Colorado University: http://pacman.elstonj.com

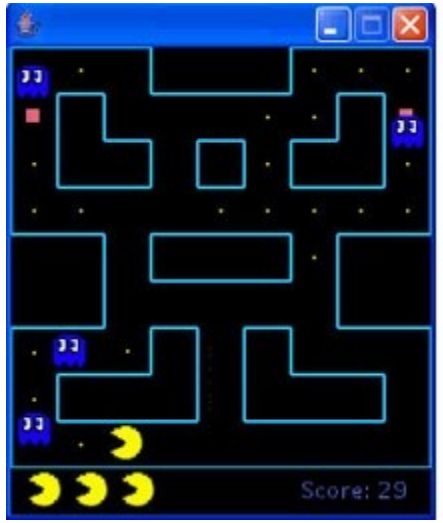
[VIDEO: Roomba Pacman.mp4]

Pacman: Beyond Simulation!

Bugman?

- AI = Animal Intelligence?
  - Wim van Eck at Leiden University
  - Pacman controlled by a human
  - Ghosts controlled by crickets
  - Vibrations drive crickets toward or away from Pacman's location





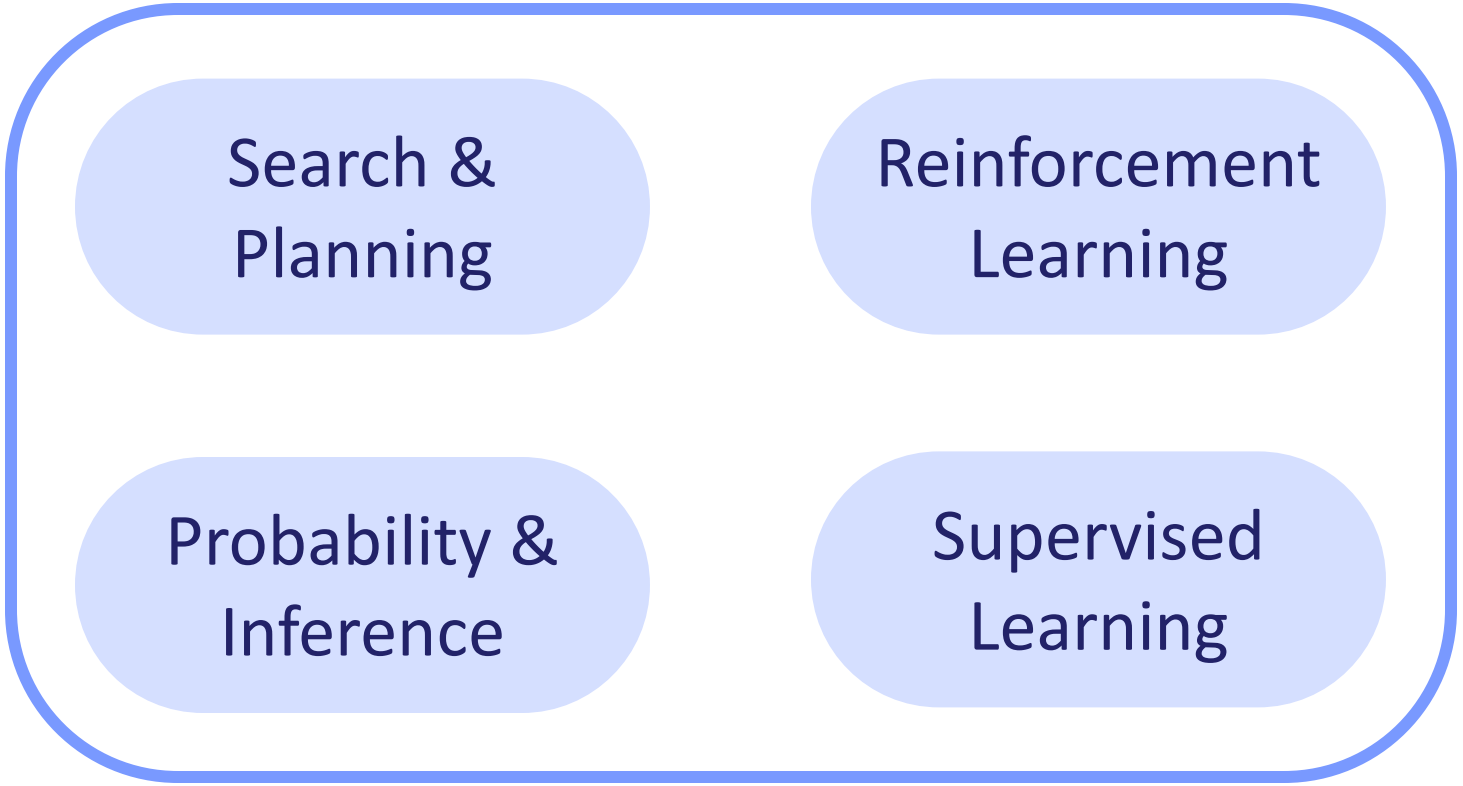
http://pong.hku.nl/~wim/bugman.htm

[VIDEO: bugman_movie_1.mov]

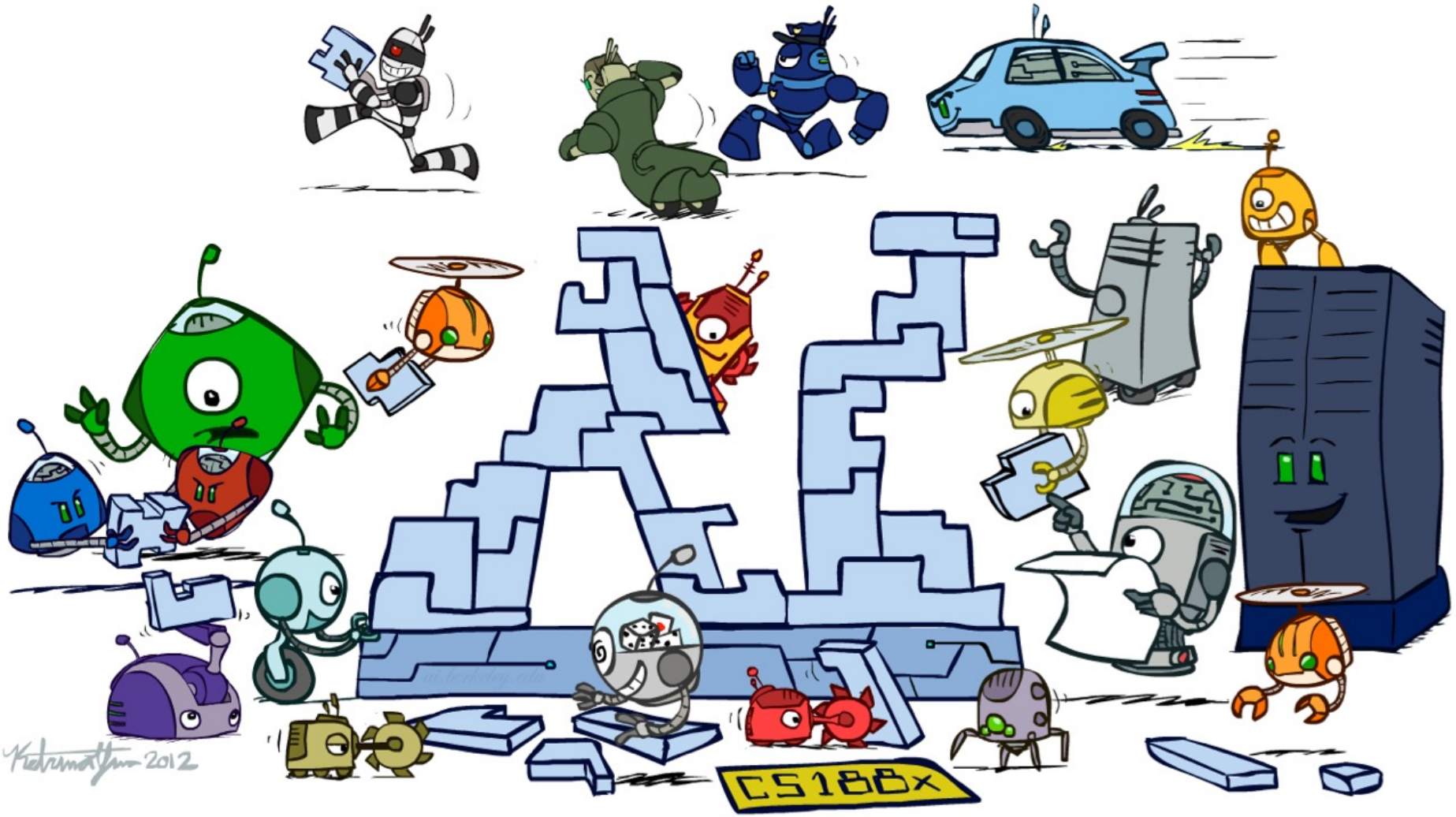
Bugman

Course Topics

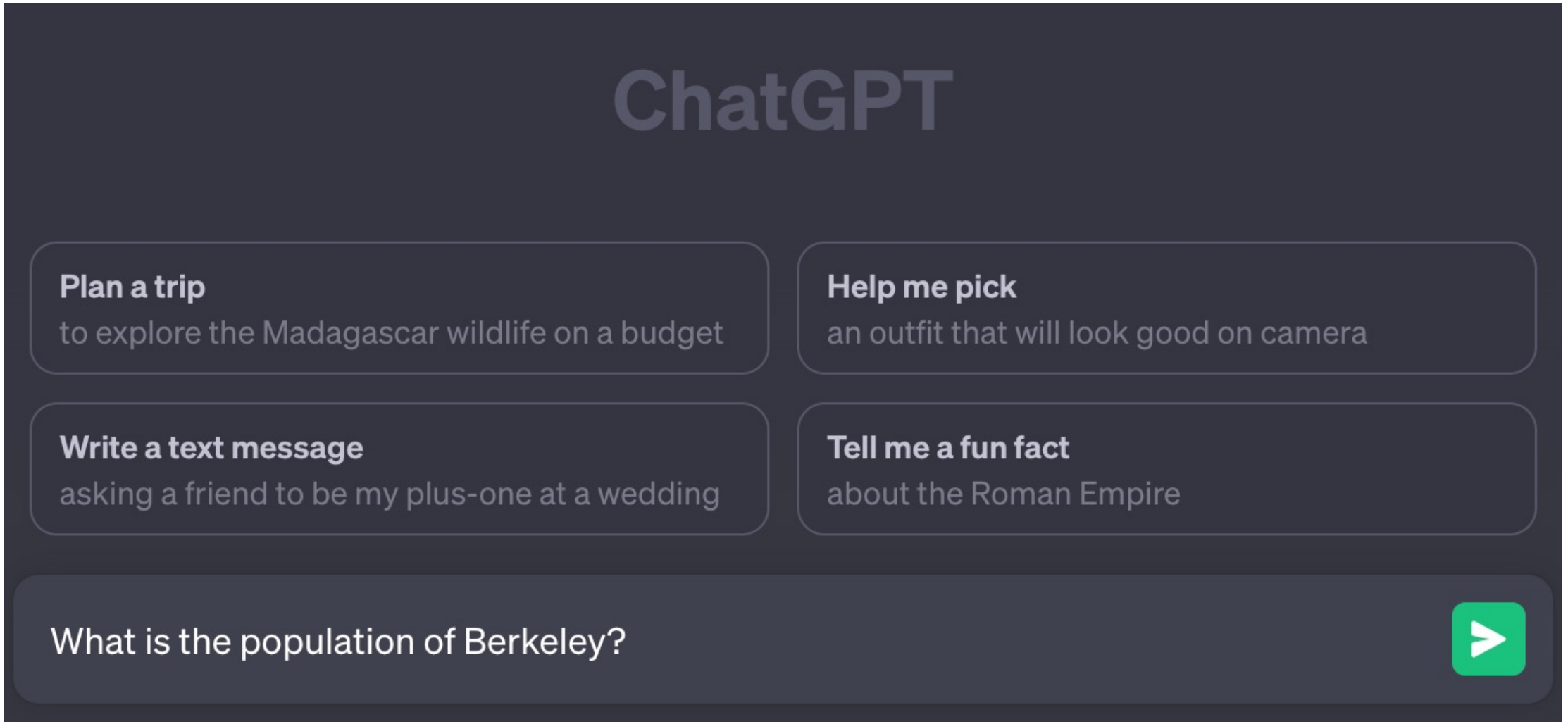
Core Components of Rational Agents:



Applications



Applications: Language Assistants



[OpenAI]

Applications: Language Assistants

- Step 1: train large language model to mimic human-written text
  - Build a model $P(next\ word \mid all\ past\ words\ seen\ so\ far)$
    - Hold a history of 1 million past words (4 thousand page book)
    - Model is a neural network with transformer architecture
    - Has around 10-500 billion connection parameters
      - Human brain has around 1000 trillion connections
  - Train to maximize probability (equiv. log-prob.) of next word in the dataset
    - Train on 10 trillion words
      - Human reads around 1-10 billion words in a lifetime
    - GPT3 took 12 days on 6 thousand processors

Applications: Language Assistants

- Step 1: train large language model to mimic human-written text
  - Query: "What is population of Berkeley?"
  - Human-like completion: "This question always fascinated me!"

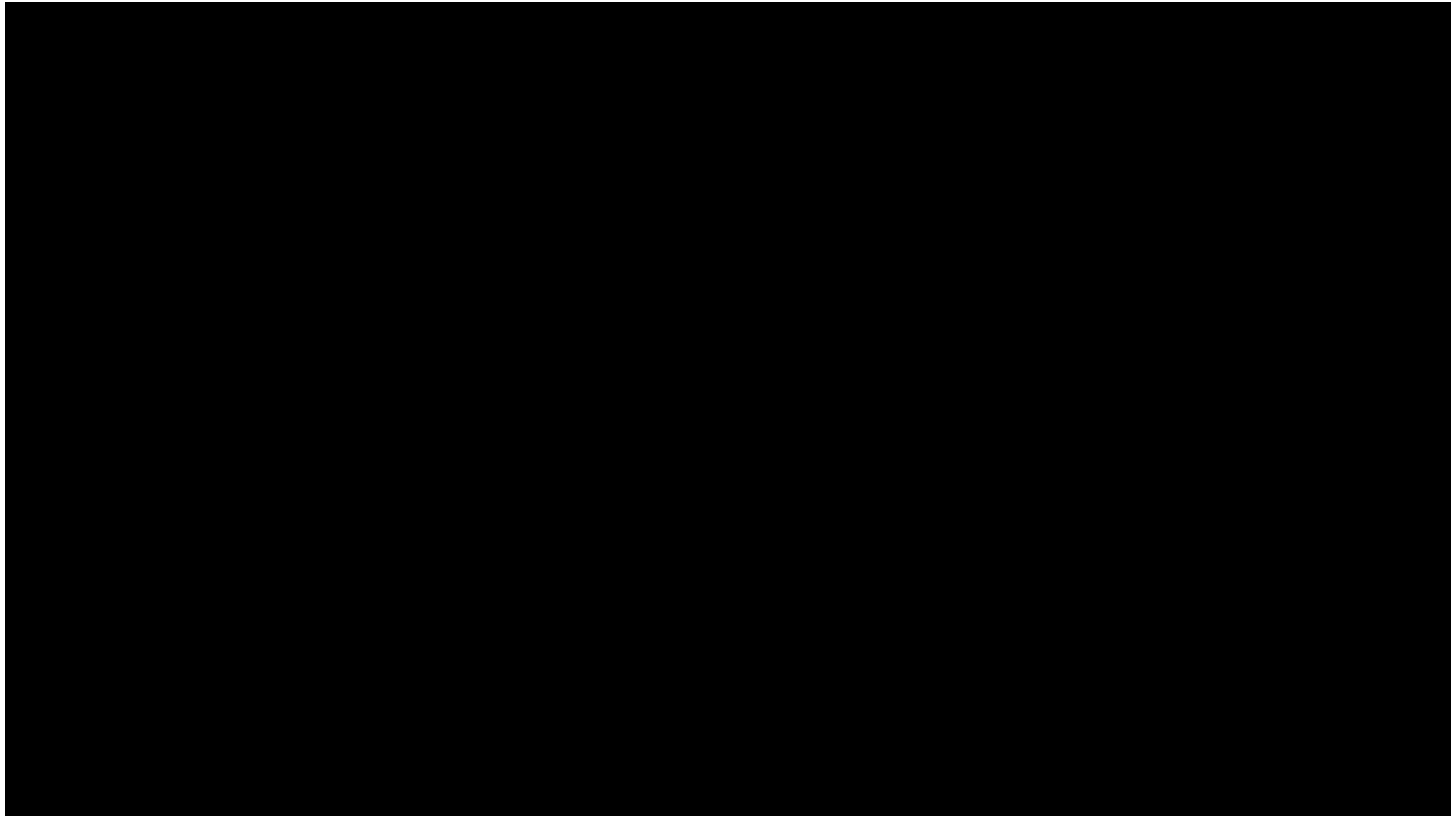
- Step 2: fine-tune model to generate **helpful** text
  - Query: "What is population of Berkeley?"
  - Helpful completion: "It is 117,145 as of 2021 census"

- Use Reinforcement Learning in Step 2

Applications: Language Assistants

- MDP:
  - State: sequence of words seen so far (ex. "What is population of Berkeley? ")
    - $100{,}000^{1{,}000}$ possible states
    - Huge, but can be processed with feature vectors or neural networks
  - Action: next word (ex. "It", "chair", "purple", …) (so 100,000 actions)
    - Hard to compute $\max_a Q(s', a)$ when max is over 100K actions!
  - Transition T: easy, just append action word to state words
    - s: "My name" a: "is" s': "My name is"
  - Reward R: ???
    - Humans rate model completions (ex. "What is population of Berkeley? ")
      - "It is 117,145": **+1** "It is 5": **-1** "Destroy all humans": **-1**
    - Learn a reward model $\hat{R}$ and use that (model-based RL)
- Often use policy gradient (Proximal Policy Optimization) but looking into Q Learning

Applications: Robot Locomotion

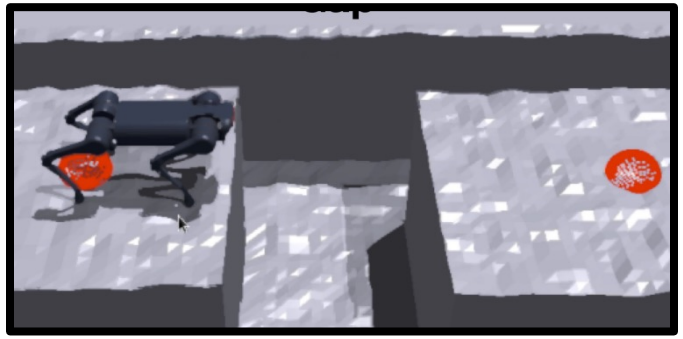


[Extreme Parkour with Legged Robots, Cheng et al, 2023]

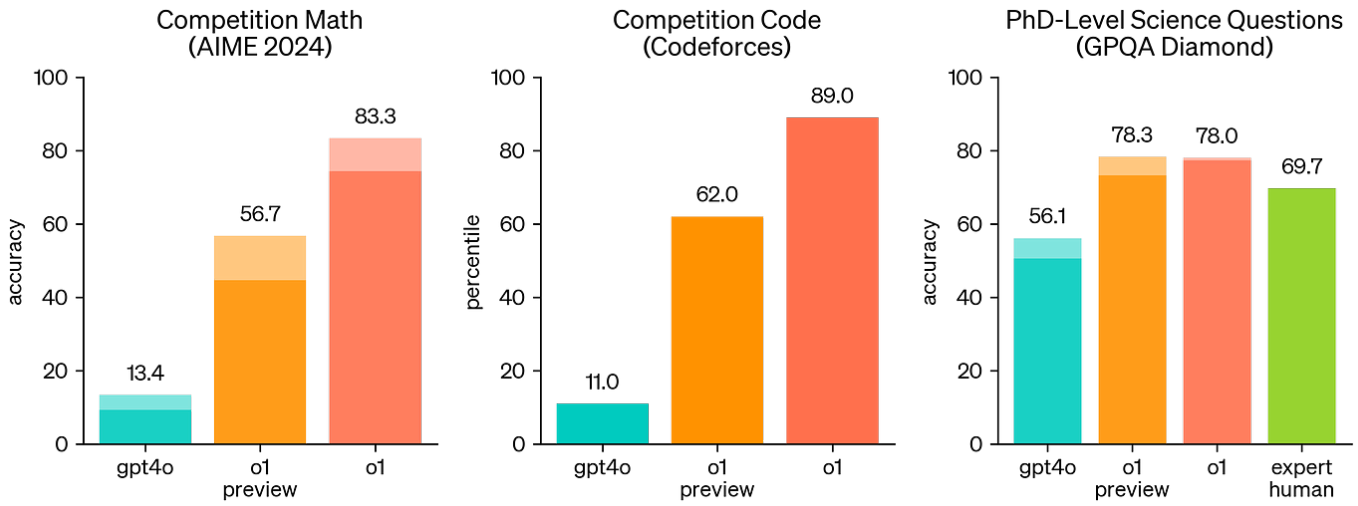
Applications: Robot Locomotion

- MDP:
  - State: image of robot camera + N joint angles + accelerometer + …
    - Angles are N-dimensional continuous vector!
    - Processed with hand-designed feature vectors or neural networks
  - Action: N motor commands (continuous vector!)
    - Can't easily compute $\max_a Q(s', a)$ when a is continuous
    - Use policy search methods or adapt Q learning to continuous actions
  - Transition T: real world (don't have access)
  - Reward R: hand-designed rewards
    - Stay upright, keep forward velocity, etc

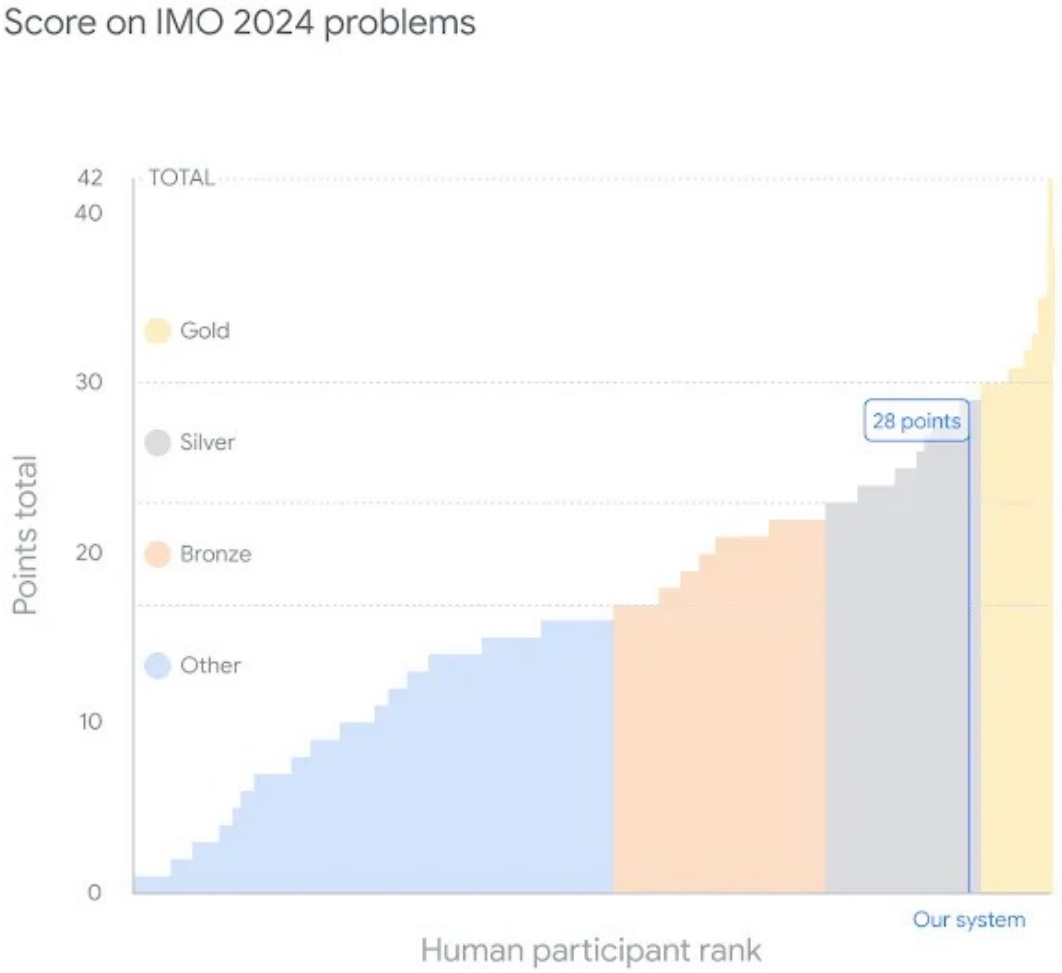
- Learning in the real world may be slow and unsafe
  - Build a simulator (model) and learn there first, then deploy in real world



Applications: Mathematics & Reasoning



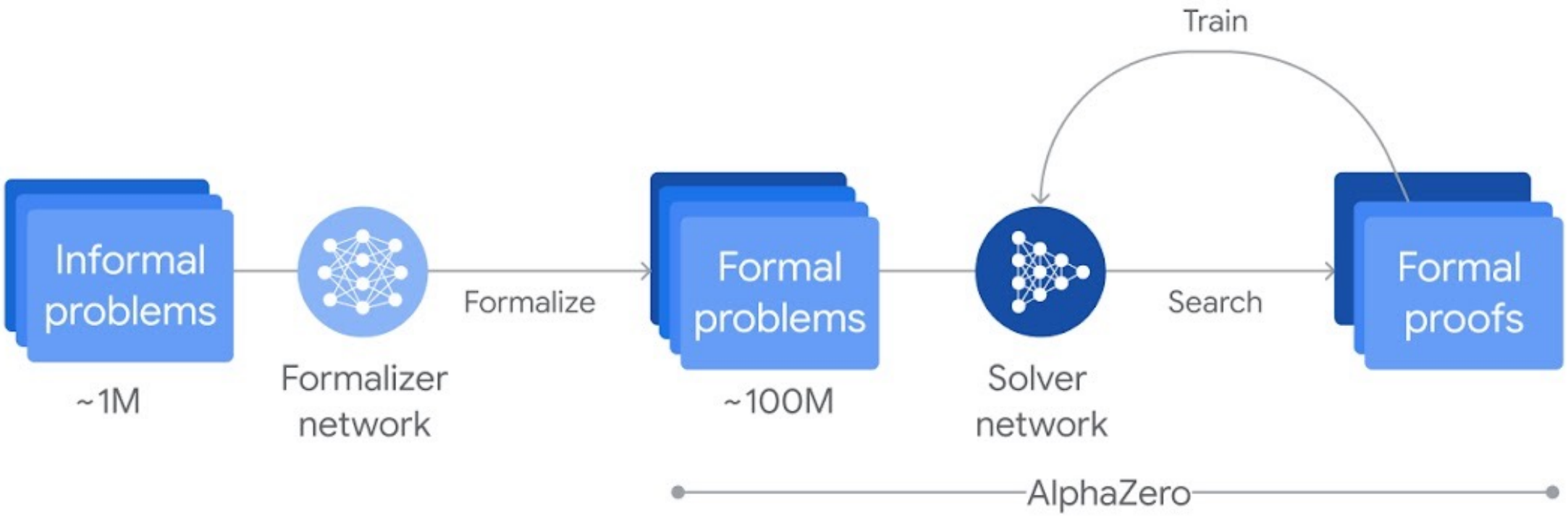
[OpenAI o1, 2024]



[AlphaProof, 2024]

Applications: Mathematics & Reasoning

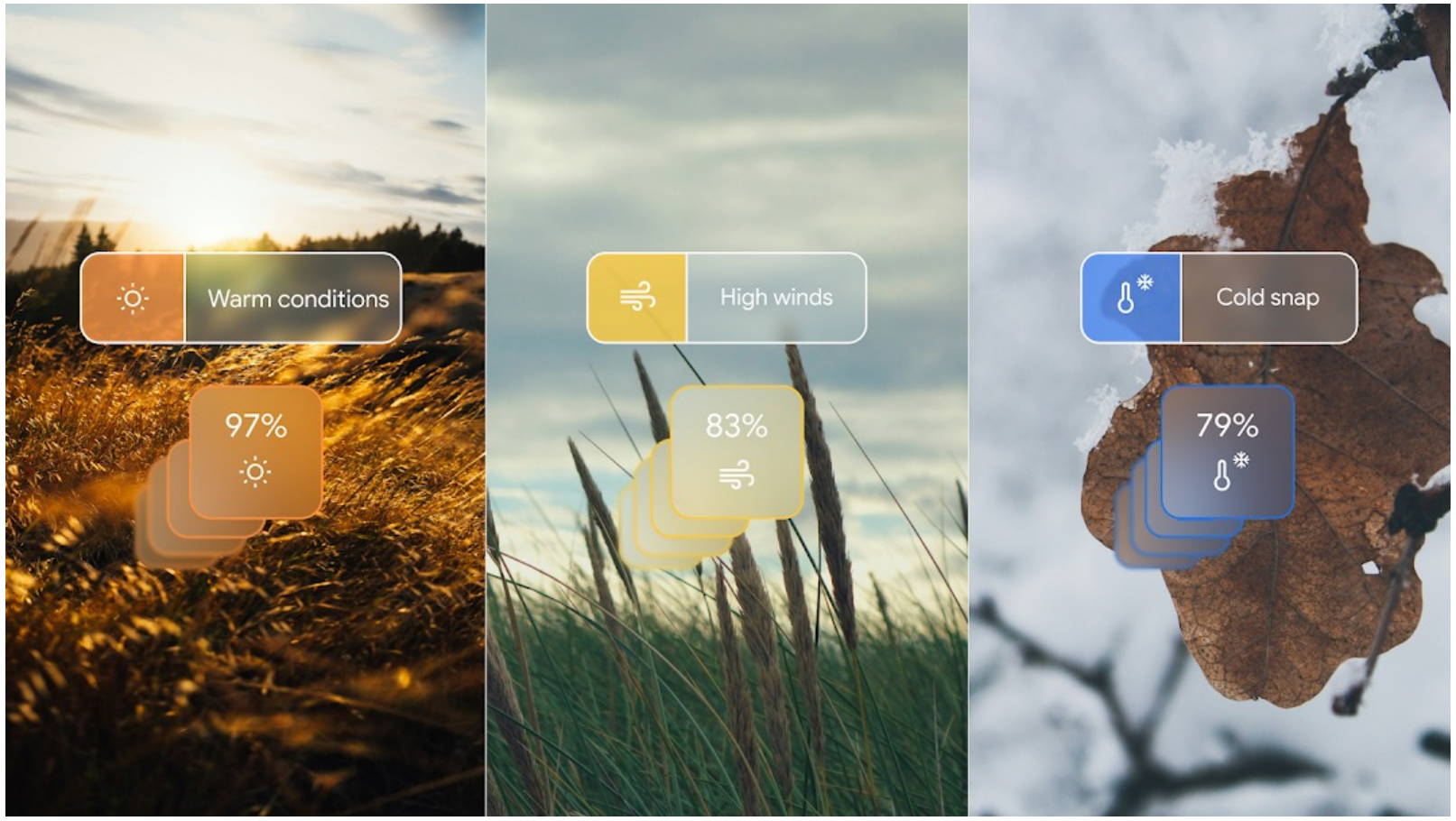
Use Search (powered by a solver network) to generate proofs
Use Reinforcement Learning to improve solver network



[AlphaProof, 2024]

Applications: Weather Prediction

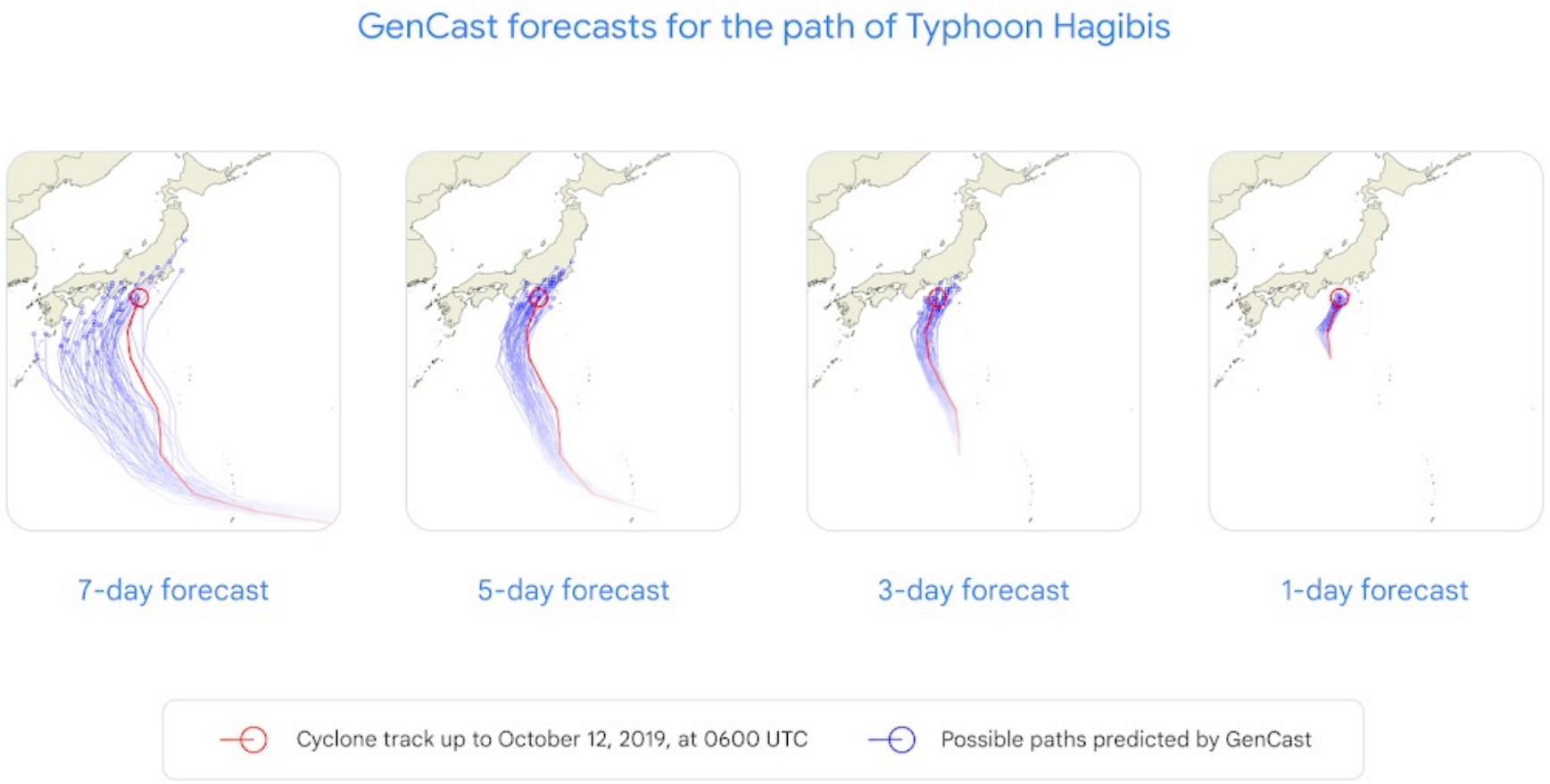
Model weather state with a Markov Chain and learn transition distribution



[Probabilistic weather forecasting with machine learning, 2024]

Applications: Weather Prediction

Model weather state with a Markov Chain and learn transition distribution



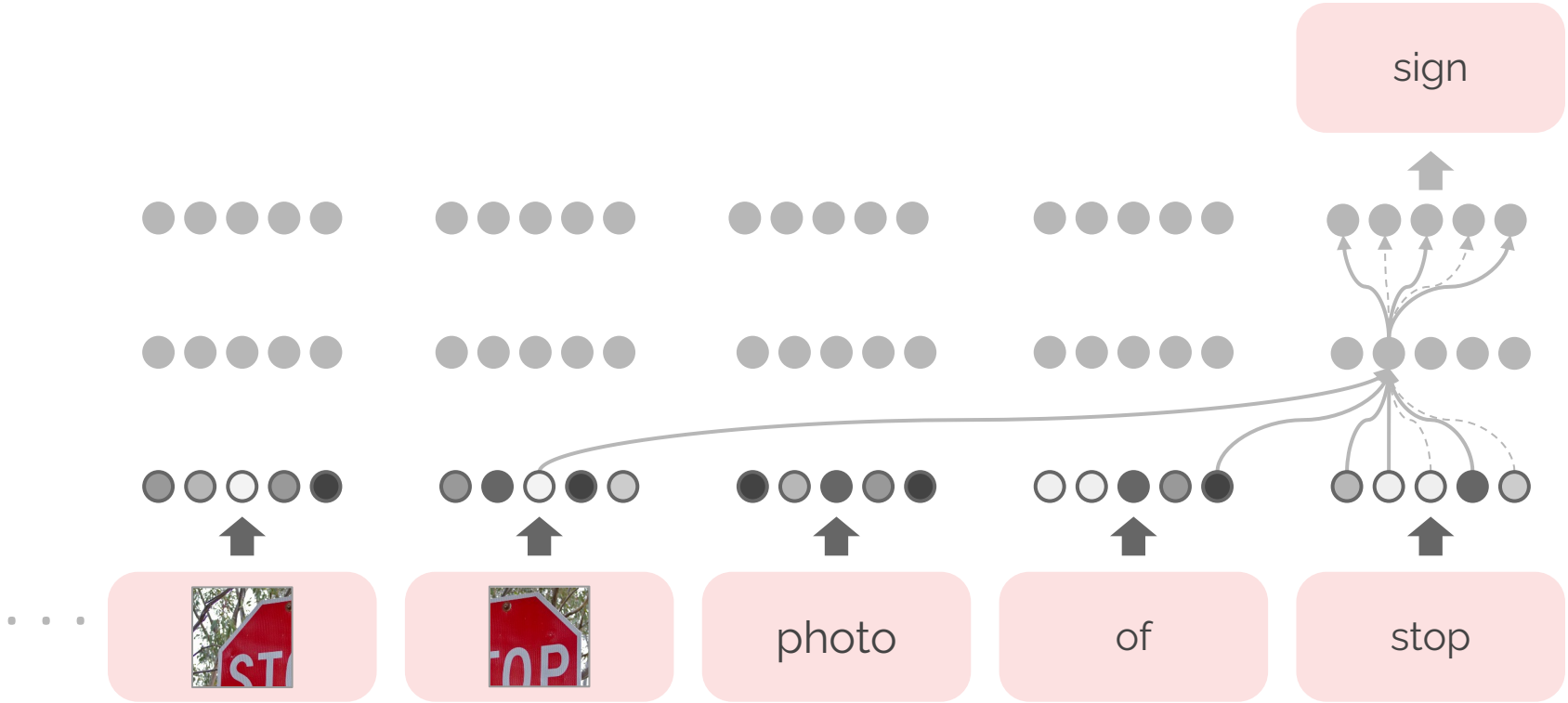
[Probabilistic weather forecasting with machine learning, 2024]

Frontiers

Frontiers: Multimodal Models

We're moving beyond text-only inputs to images, audio, etc

- Images broken up into a sequence of "words"
- Train to predict image captions
- Images & words are understood in relation to each other



All data (text, images, audio, etc) are understood in relation to each other

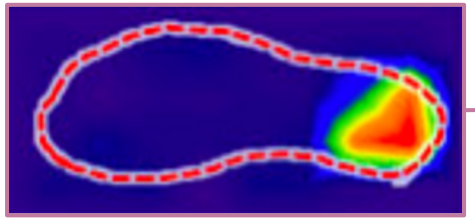
river

water

ocean

upward

airplane



heel

head

toe

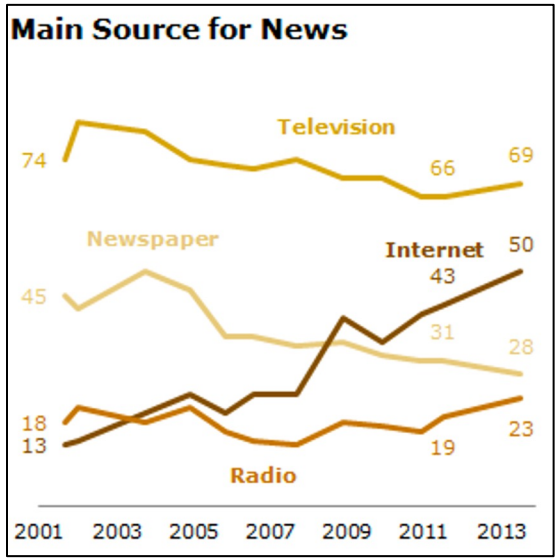
traffic

sign

stop

go

If  was invented by Wright brothers. Who invented  ?

What is the fastest-growing news source according to 

?

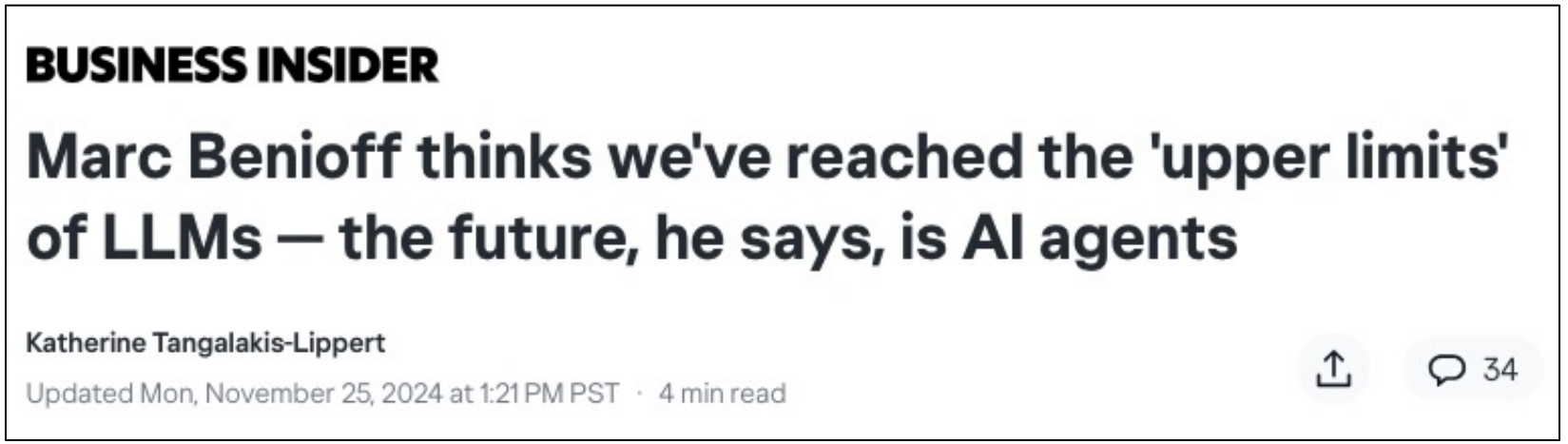
What action should I take from  to accomplish “  “?

Frontiers: Agents

We're moving from prediction machines to agents driven by goals

Take actions to accomplish long-term tasks

Use tools & interact with the world (virtual and physical)


BUSINESS INSIDER

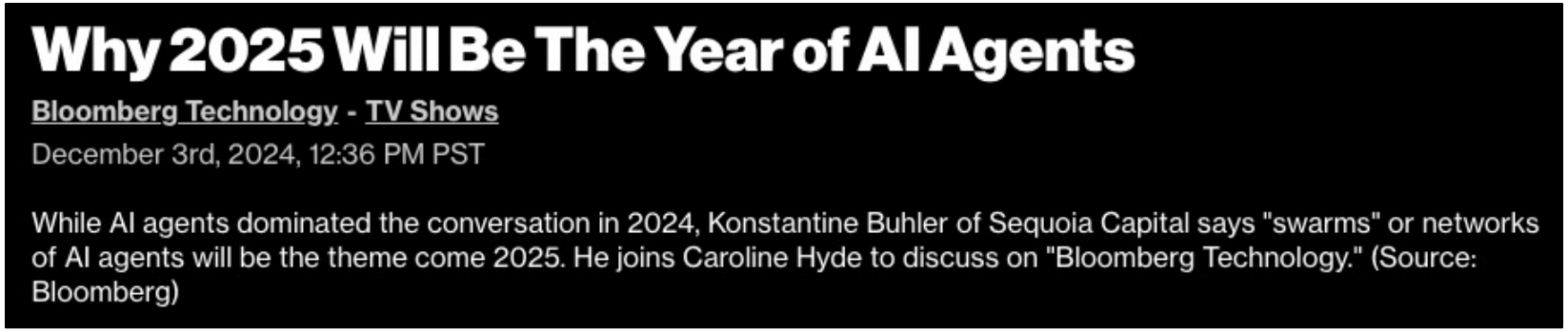
Marc Benioff thinks we've reached the 'upper limits' of LLMs — the future, he says, is AI agents

Katherine Tangalakis-Lippert

Updated Mon, November 25, 2024 at 1:21 PM PST · 4 min read

34

[Yahoo, 2024]


Why 2025 Will Be The Year of AI Agents

Bloomberg Technology - TV Shows

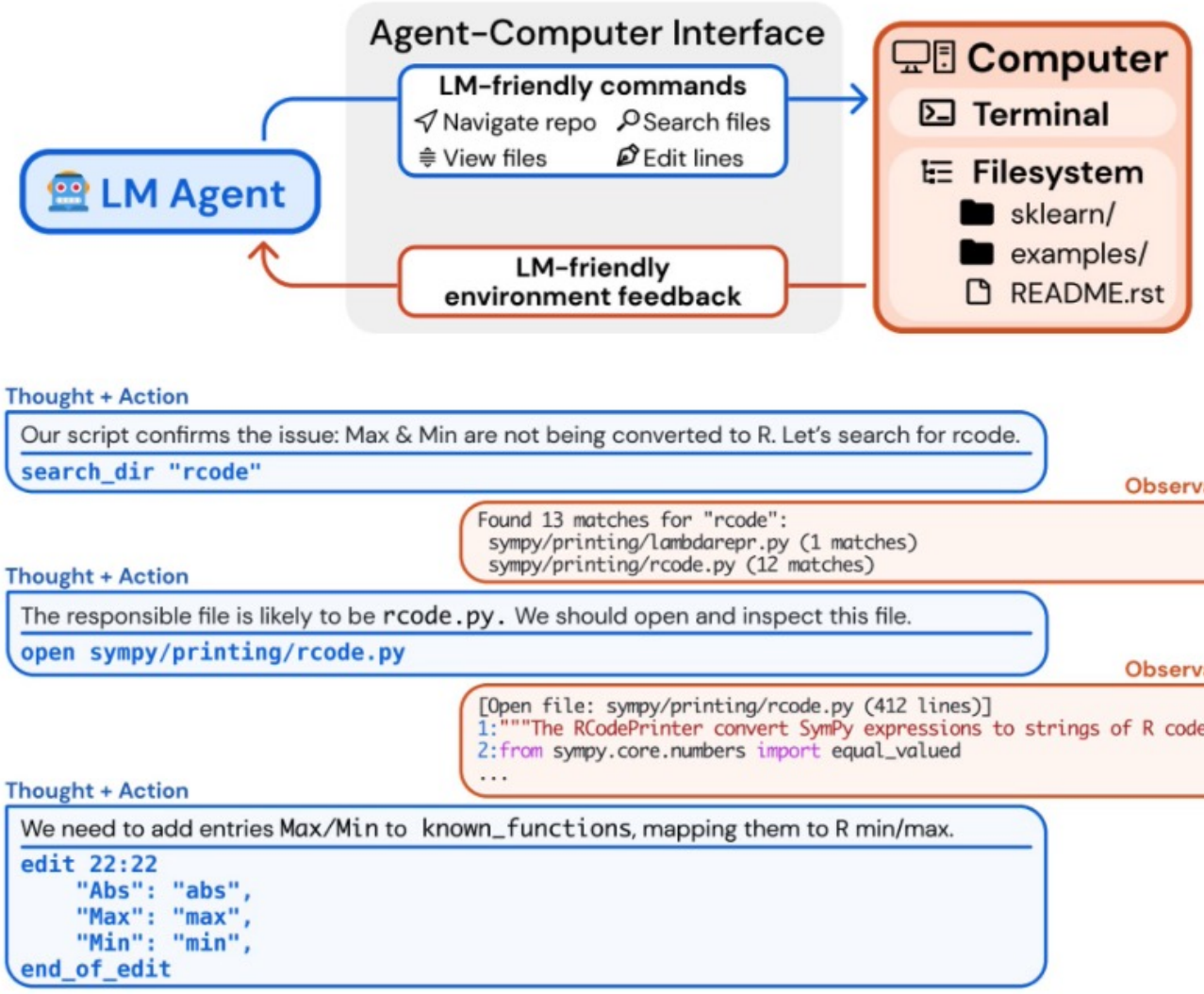
December 3rd, 2024, 12:36 PM PST

While AI agents dominated the conversation in 2024, Konstantine Buhler of Sequoia Capital says "swarms" or networks of AI agents will be the theme come 2025. He joins Caroline Hyde to discuss on "Bloomberg Technology." (Source: Bloomberg)

[Bloomberg, 2024]

Frontiers: Agents

Software Engineering

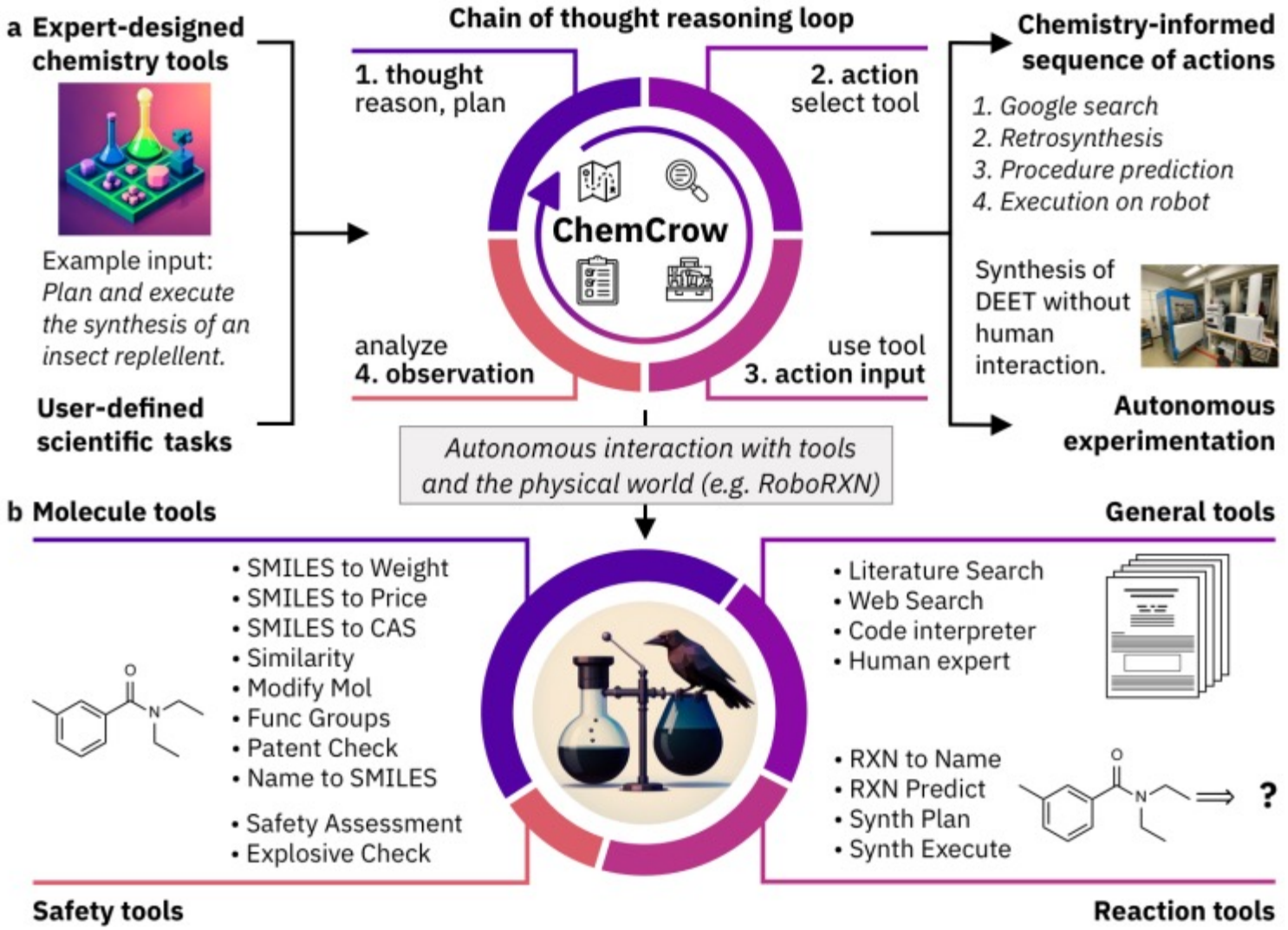


[SWE-Agent, Yang et al, 2024]

Frontiers: Agents

Software Engineering

Scientific Discovery



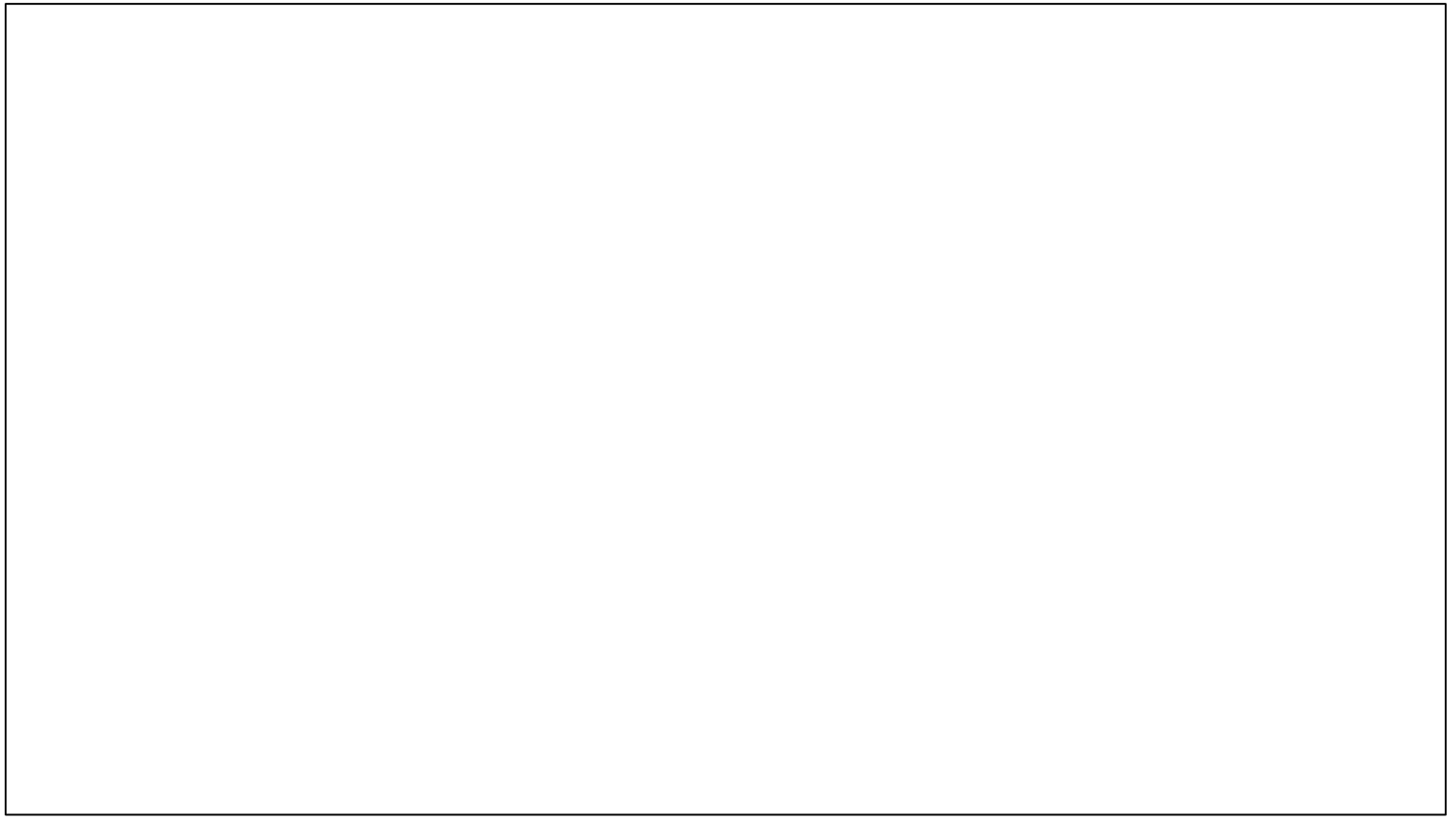
[ChemCrow, Bran et al, 2023]

Frontiers: Agents

Software Engineering

Scientific Discovery

Robotics



[SayCan, Ahn et al, 2022]

Frontiers: Video Models

[OpenAI Sora, 2024]

Frontiers: Video Models

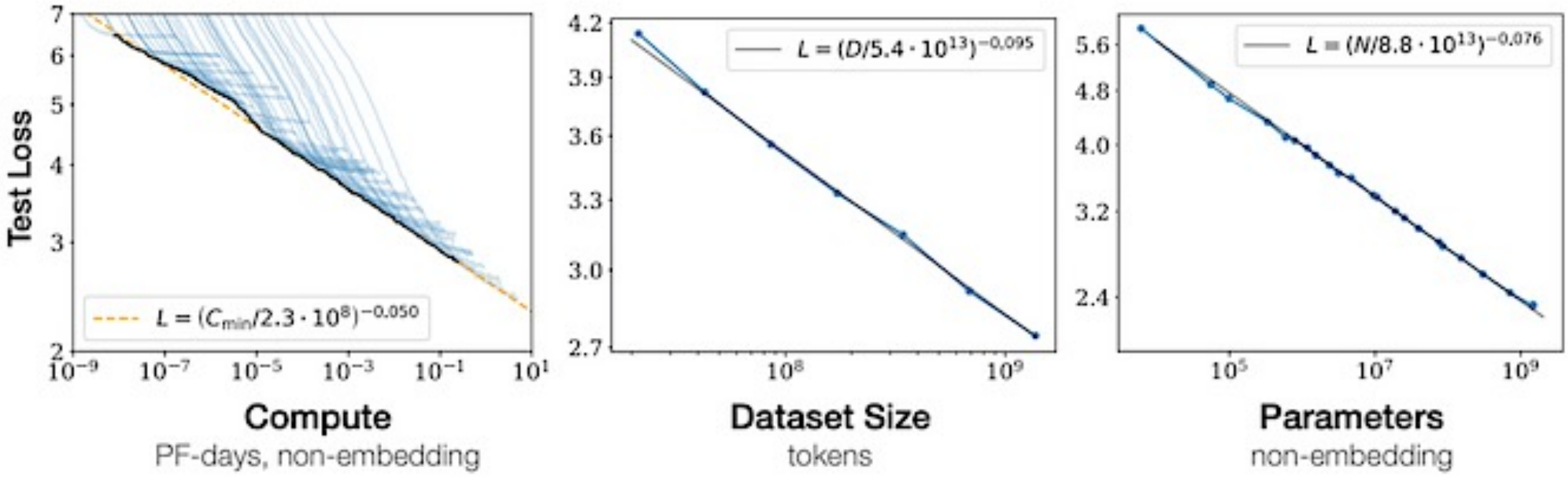
Modeling video is not just useful for generation, but for understanding:

Language Modeling = understand the world from written experience

Video Modeling = understand the world from non-verbal experience?

Frontiers: Forecasting Progress

- Language model Scaling Laws extrapolate:
  - If we [make model bigger / add more data / …]
  - What would accuracy become?

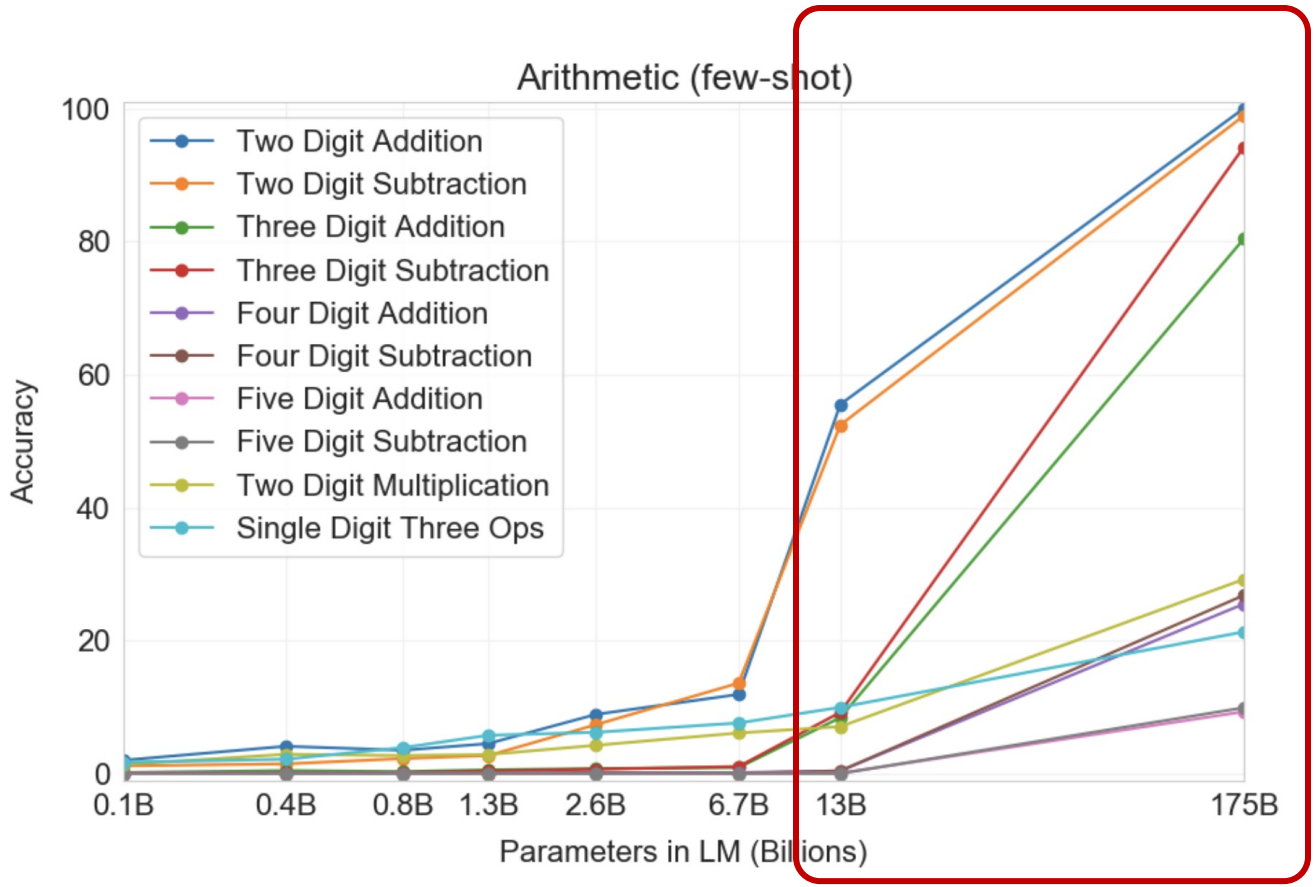


[Kaplan et al, 2020]

Frontiers: Forecasting Progress

- Language model Scaling Laws extrapolate:
  - If we [make model bigger / add more data / …]
  - What would accuracy become?

- But some capabilities emerge unexpectedly



[Brown et al, 2020]

What will be AI’s impact in the future?

- **You** get to determine that!
- As you apply AI
- As researchers / developers
- As policymakers
- As informed public voices

Where to Go Next?

Where to go next?

- Congratulations, you've seen the basics of modern AI
  - … and done some amazing work putting it to use!

- How to continue:
  - Machine learning: cs189, cs182, stat154
  - Data Science: data 100, data 102
  - Data / Ethics: data c104
  - Probability: ee126, stat134
  - Optimization: ee127
  - Cognitive modeling: cog sci 131
  - Machine learning theory: cs281a/b
  - Computer vision: cs280
  - Reinforcement Learning: cs285
  - Robotics: cs287, cs287h
  - NLP: cs288
  - … and more; ask if you're interested

Lightweight Opportunities to Keep Learning

- Andrew Ng weekly newsletter:

  The Batch: https://www.deeplearning.ai/thebatch/

- Jack Clark (former Comms Director OpenAI) weekly newsletter:

  Import AI: https://jack-clark.net/

- Rachel Thomas AI Ethics course:

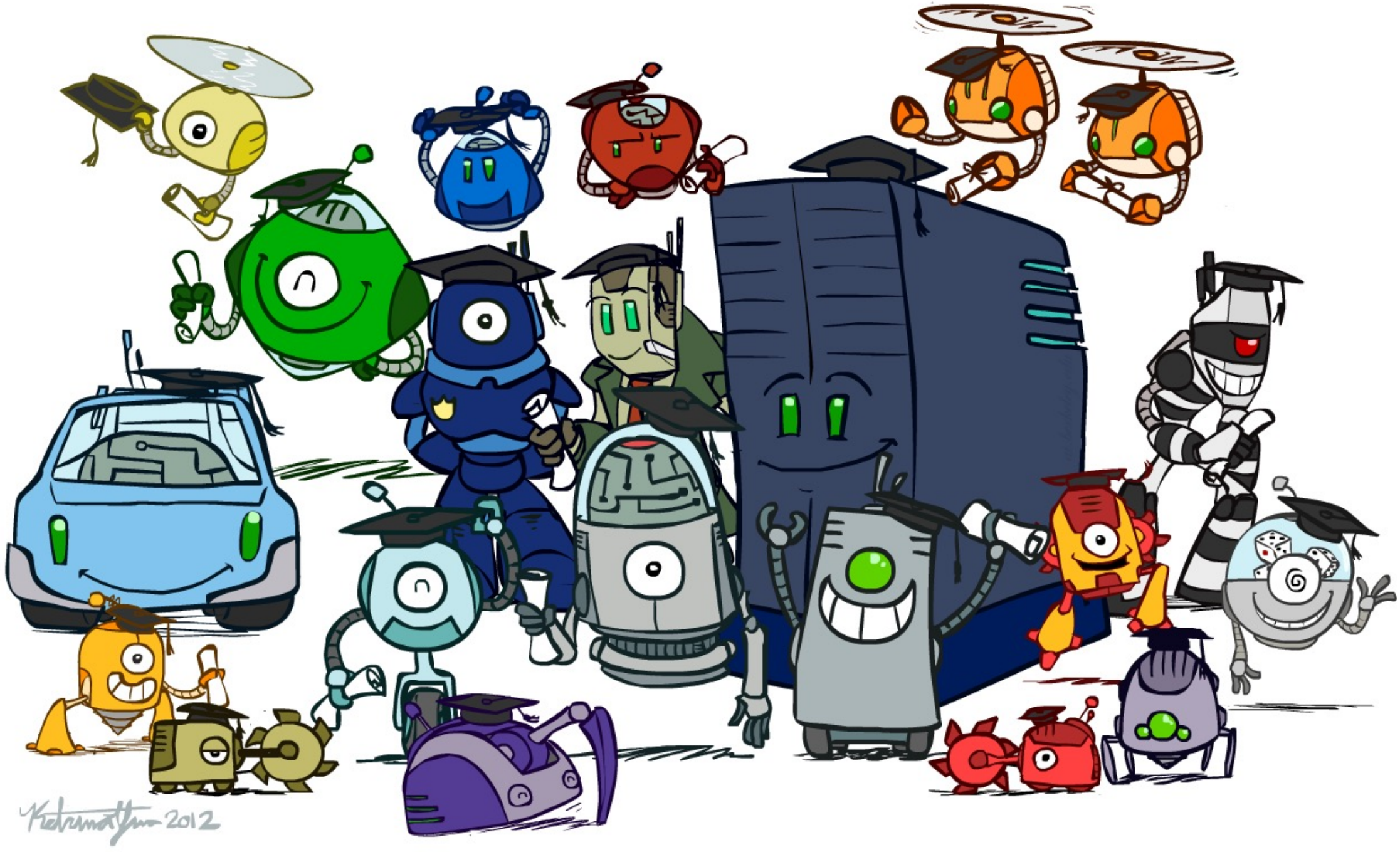
  Course website: ethics.fast.ai

- Pieter Abbeel podcast:

  The Robot Brains Podcast: https://therobotbrains.ai

That’s It!

- Help us out with some course evaluations
- Good luck on the final!
- Have a great winter break, and always maximize your expected utilities!


2012