Latest AI Developments Explained (OpenAI SORA, World Models, Q*)

Pavel Kordík

Department of Applied Mathematics,
Faculty of Information Technology,
Czech Technical University in Prague

Machine Learning and
Computational Intelligence Group

SORA by Open AI

- Text to video LLM
  - Released on February 15, 2024
  - Simulator of the physical world
  - Natural language understanding
  - Story, video, sound generation
  - Transformers, World models, Latent diffusion, …

- Example text prompt:

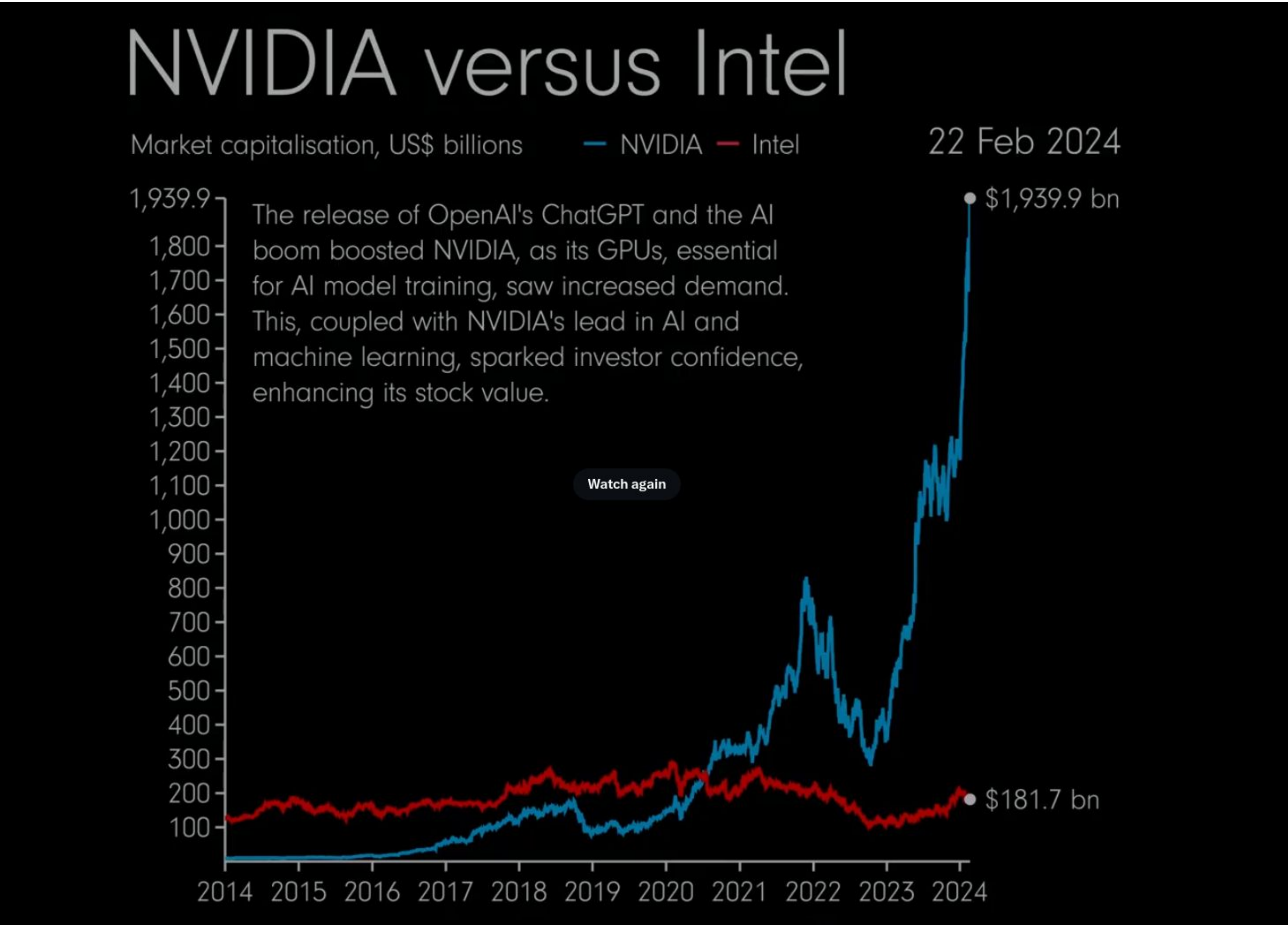
This close-up shot of a chameleon showcases its striking color changing capabilities. The background is blurred, drawing attention to the animal's striking appearance.

https://openai.com/research/video-generation-models-as-world-simulators

SORA by Open AI

What makes this possible?

- **Massive compute**
- Massive datasets
- Massive human effort
    - Researching new algorithms
    - Data preparation, benchmarks
    - Foundational model training
    - Model alignment
    - Model productization (in progress for SORA)

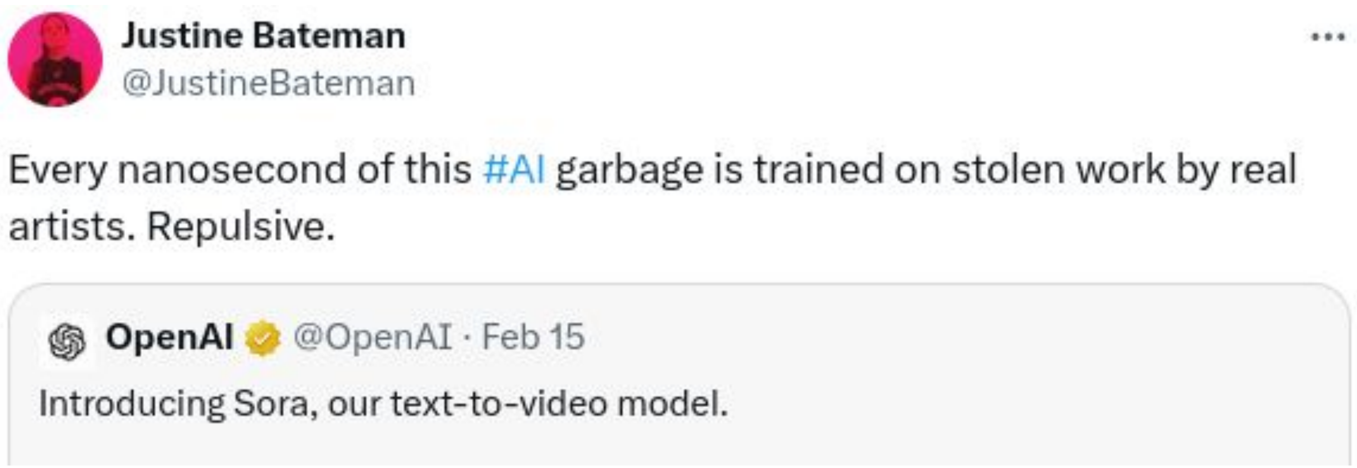

NVIDIA versus Intel
Market capitalisation, US$ billions
NVIDIA
Intel
22 Feb 2024
The release of OpenAI's ChatGPT and the AI boom boosted NVIDIA, as its GPUs, essential for AI model training, saw increased demand. This, coupled with NVIDIA's lead in AI and machine learning, sparked investor confidence, enhancing its stock value.
$1,939.9 bn
$181.7 bn
Watch again
1,939.9
1,800
1,700
1,600
1,500
1,400
1,300
1,200
1,100
1,000
900
800
700
600
500
400
300
200
100
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024

What makes this possible?

- Massive compute
- **Massive datasets**
- Massive human effort
    - Researching new algorithms
    - Data preparation, benchmarks
    - Foundational model training
    - Model alignment
    - Model productization (in progress for SORA)

What was Sora trained on? Creatives demand answers.

We think we know, but OpenAI refuses to tell us.

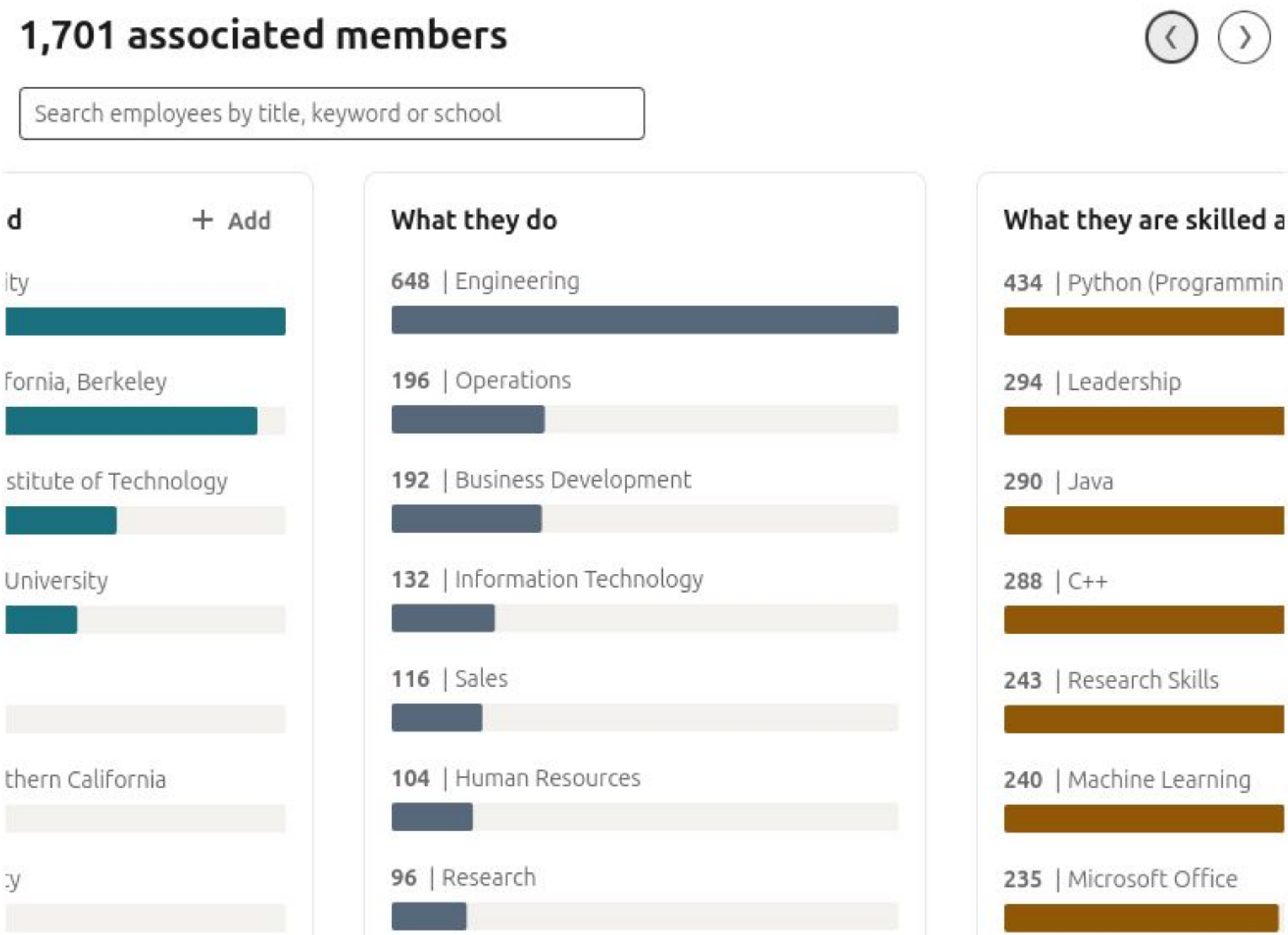


https://openai.com/blog/data-partnerships

What makes this possible?

- Massive compute
- Massive datasets
- **Massive human effort**
    - Researching new algorithms
    - Data acquisition, curation, benchmarks
    - Foundational model training
    - Model alignment
    - Model productization (in progress for SORA)

Just OpenAI - from 45 people in 2017 …



https://openai.com/blog/team-update-january

What made SORA possible?
Research perspective

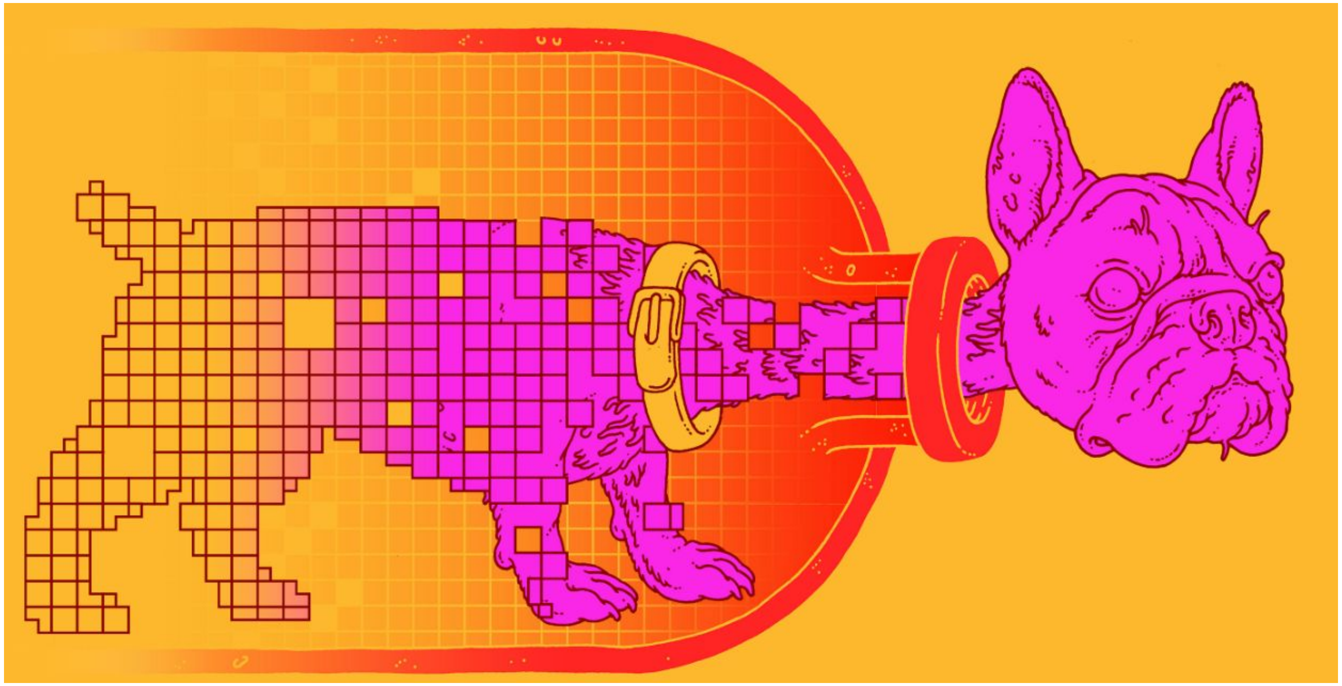
- Self-supervised training
- Understanding as compression
- Neural autoencoders
- Diffusion models
- Vision transformers
- Latent diffusion transformers
- World models

Self-supervised training

- Labels not needed
- Massive data available - text, images, sound, videos
- Learning tasks:
  - Masked modeling
  - Next sentence prediction
- Standard optimization methods available
  - SGD, ADAM, BackProp, …

Understanding as compression

- Preserve important information, Generalize, Predict
- Various forms of information bottlenecks

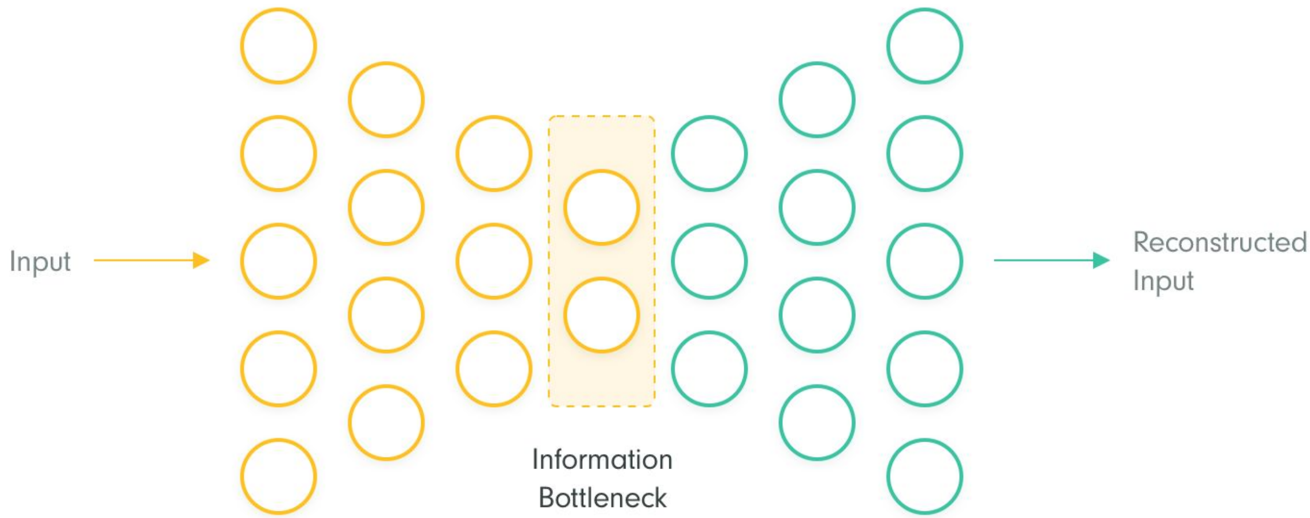


https://mural.maynoothuniversity.ie/10327/1/PM-Understanding-2016.pdf

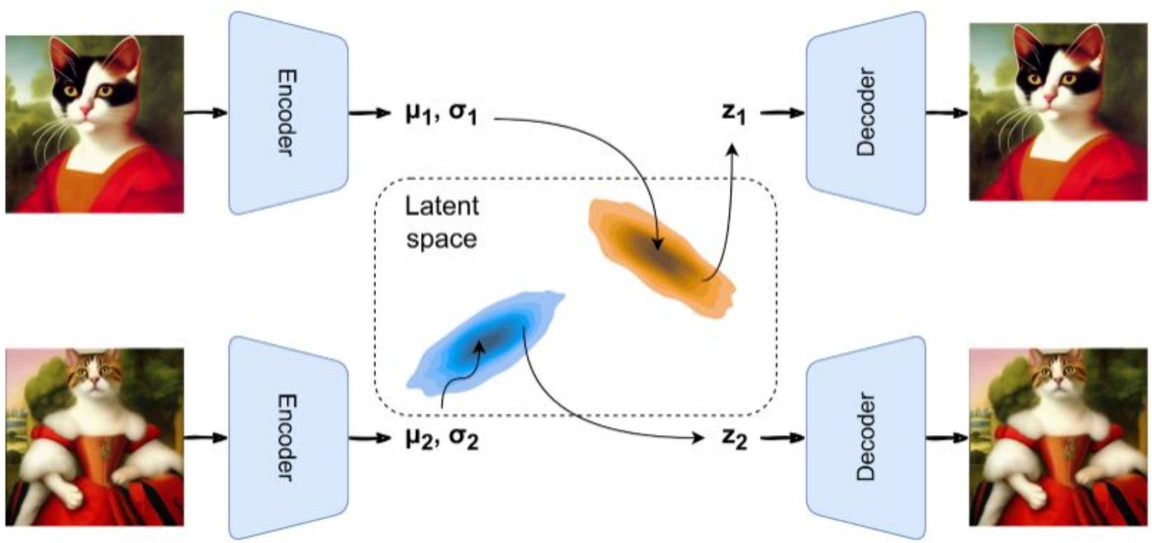
https://link.springer.com/article/10.1007/s11098-018-1152-1

Neural compression - AE, VAE

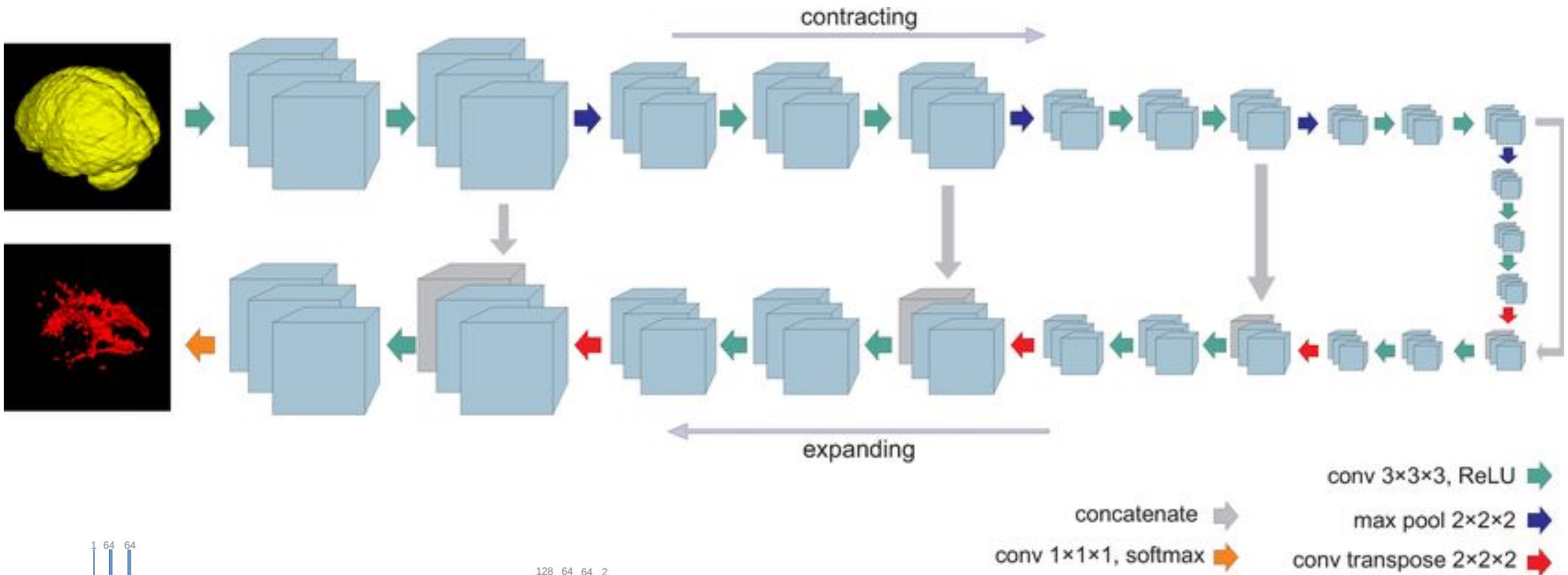
[Auto Encoder]() (AE)

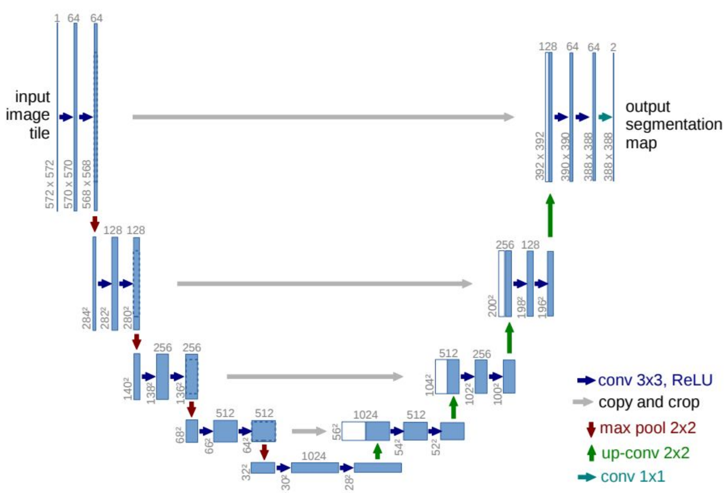


[Variational Auto Encoder]() (VAE)



Neural compression - UNet





U-Net: Convolutional Networks for Biomedical Image Segmentation:
https://arxiv.org/pdf/1505.04597.pdf

see also Hourglass networks:
https://arxiv.org/pdf/1603.06937.pdf

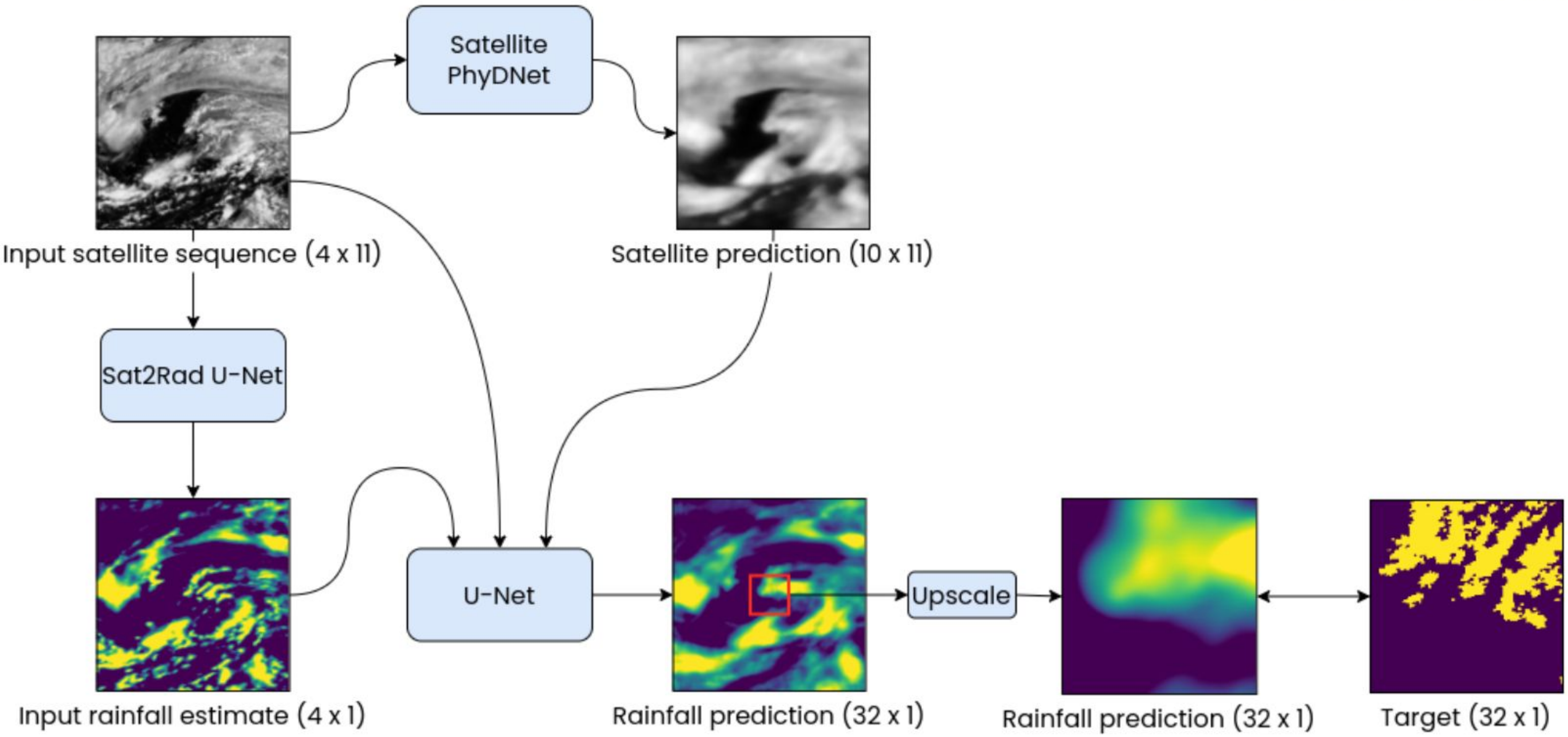
UNets are quite useful

WeatherFusionNet (WFN)

FACULTY OF INFORMATION TECHNOLOGY CTU IN PRAGUE

3 networks trained separately

METEOPRESS
WEATHER RADARS & AI NOWCASTING



https://github.com/Datalab-FIT-CTU/weather4cast-2022

Diffusion models

Adding gaussian noise

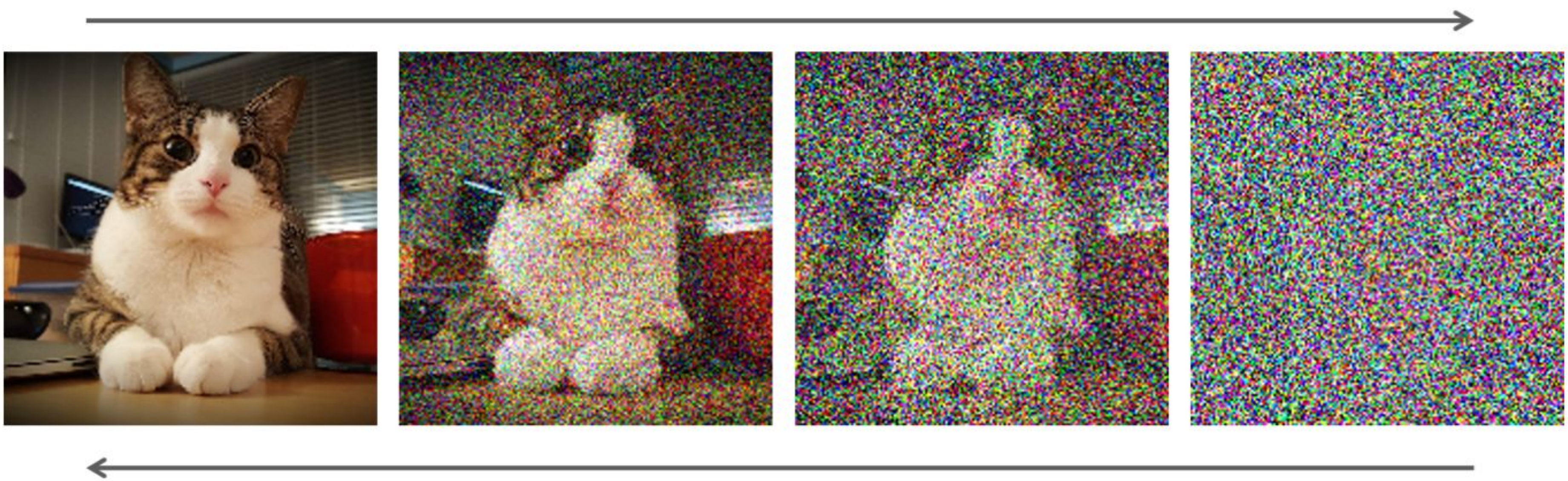
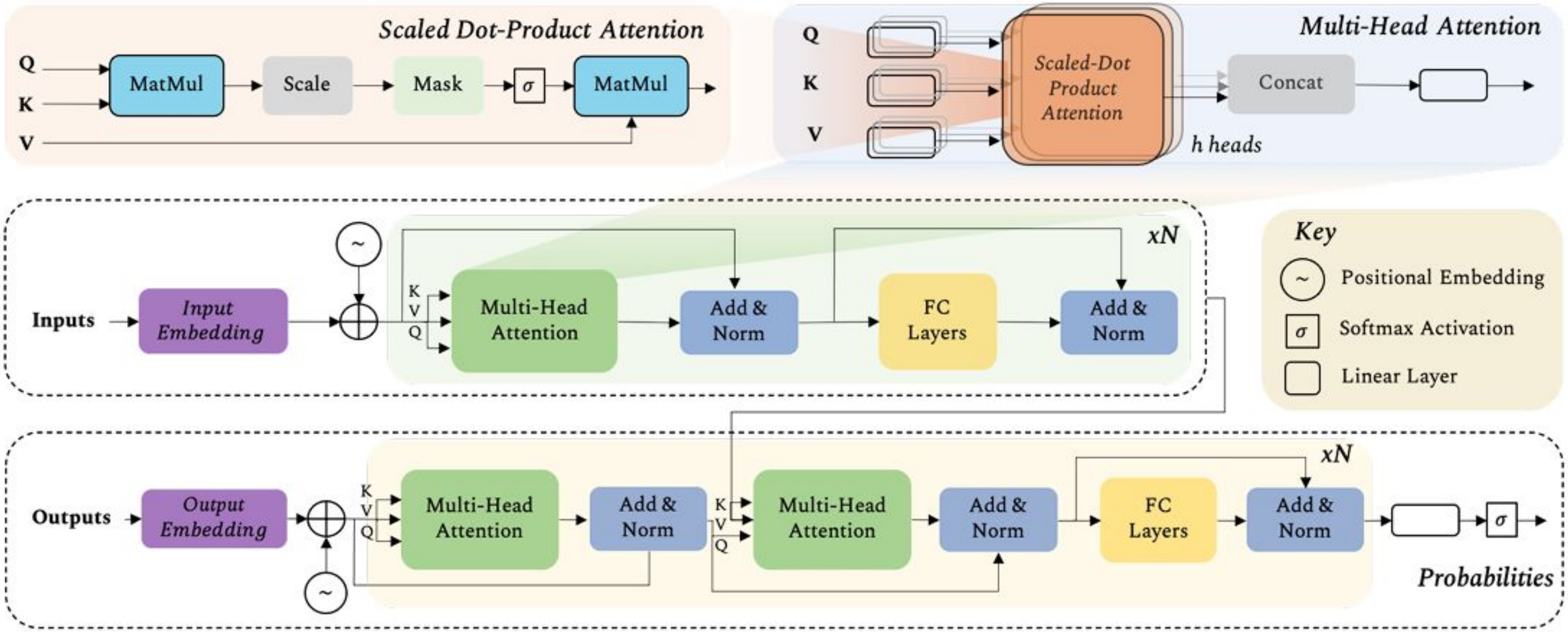


Image denoising process (remember masked modeling?)

https://developer.nvidia.com/blog/improving-diffusion-models-as-an-alternative-to-gans-part-1/

https://arxiv.org/pdf/2307.06304.pdf

Vision transformers



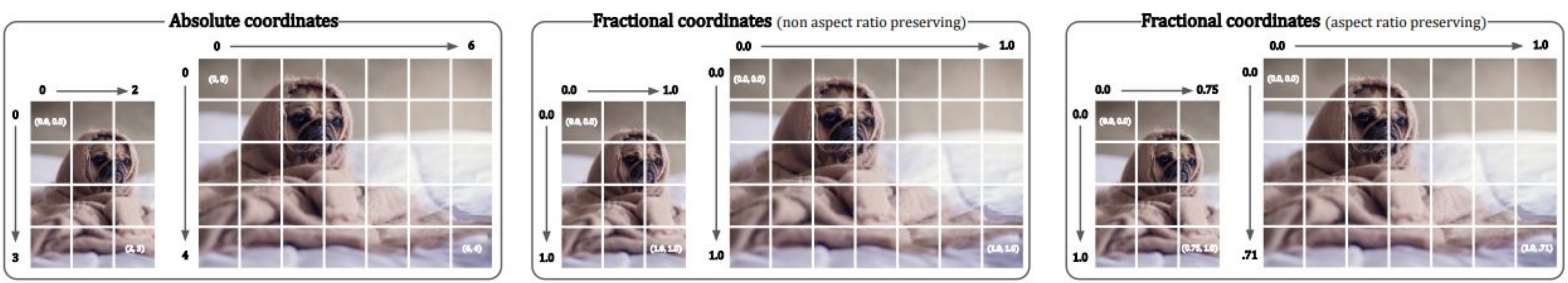
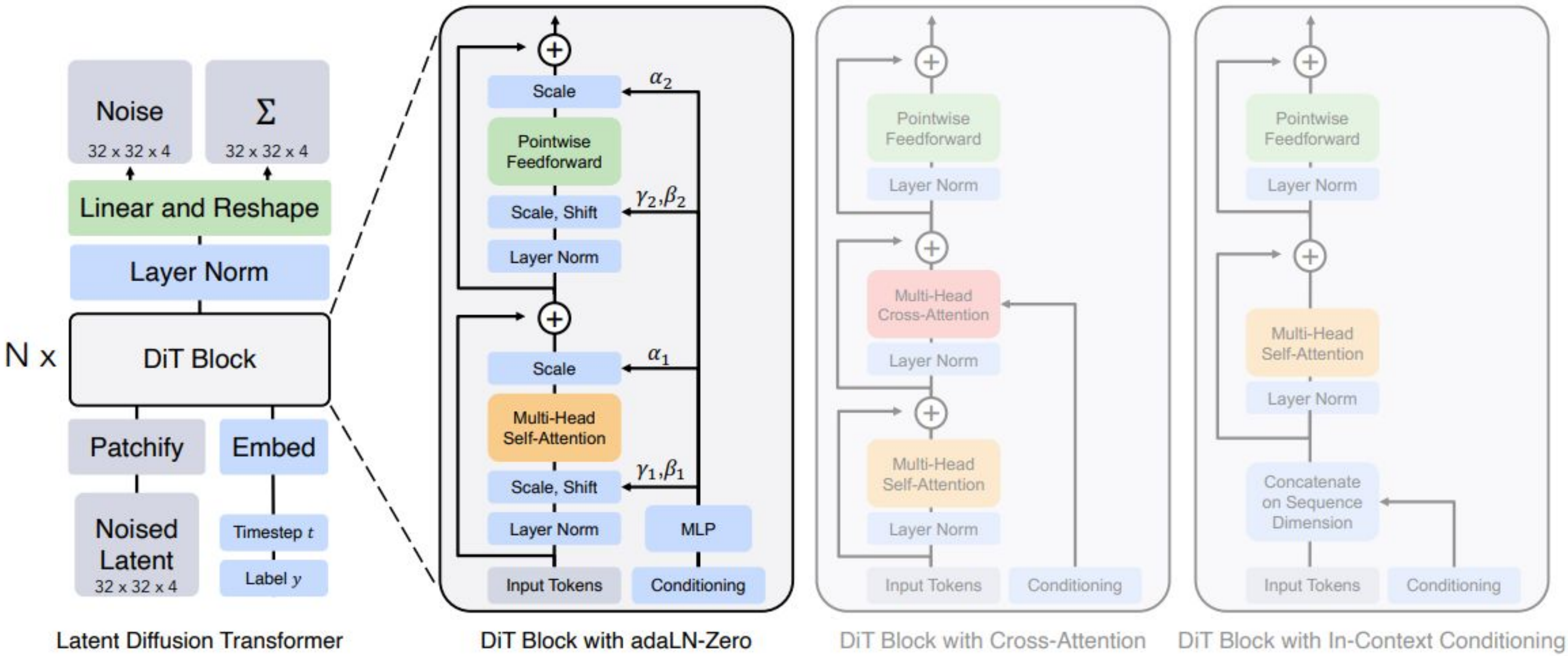


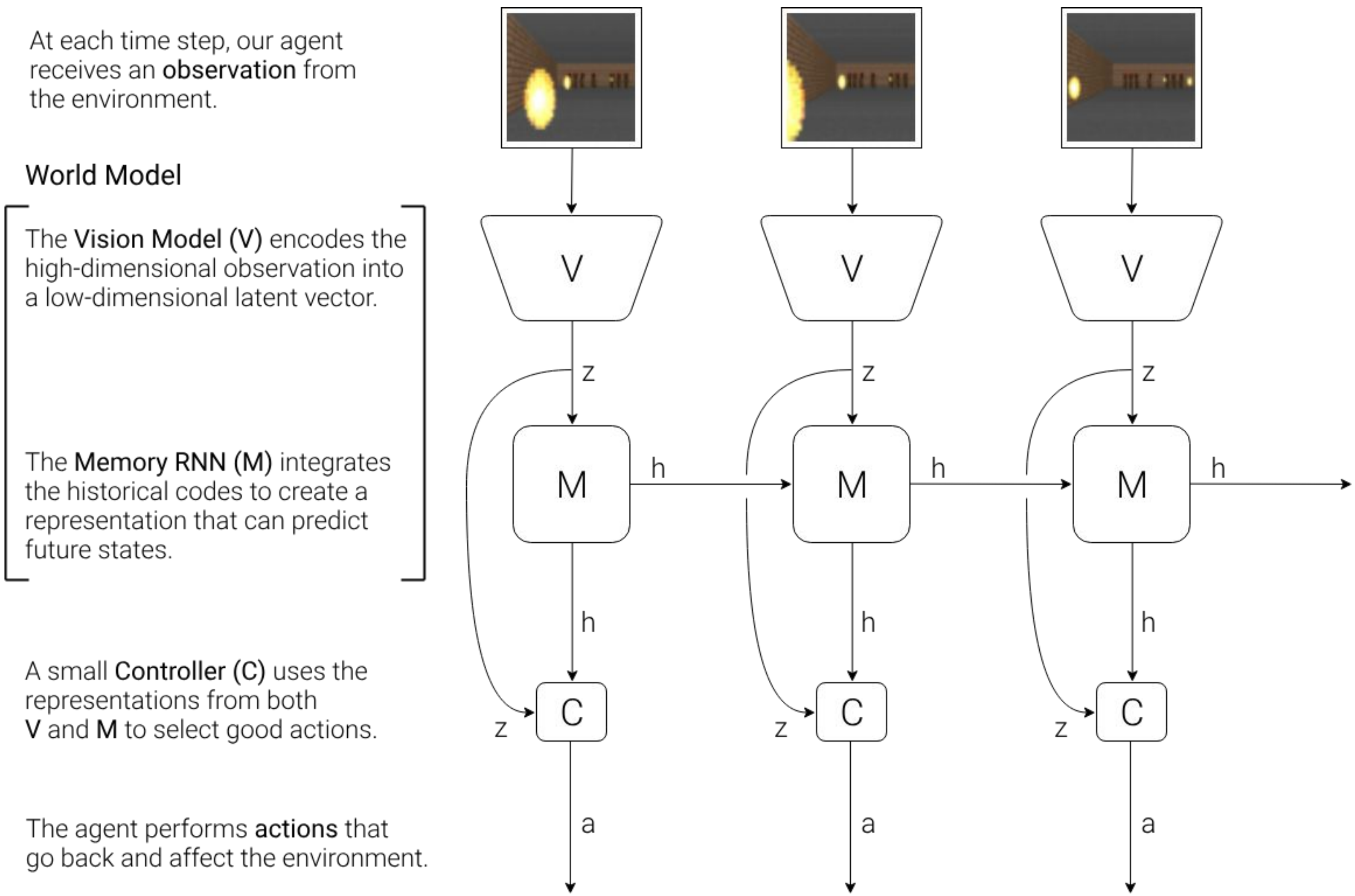
Image credit: *Matthew Henry* `burst.shopify.com/photos/dog-staying-warm`

Latent diffusion transformers



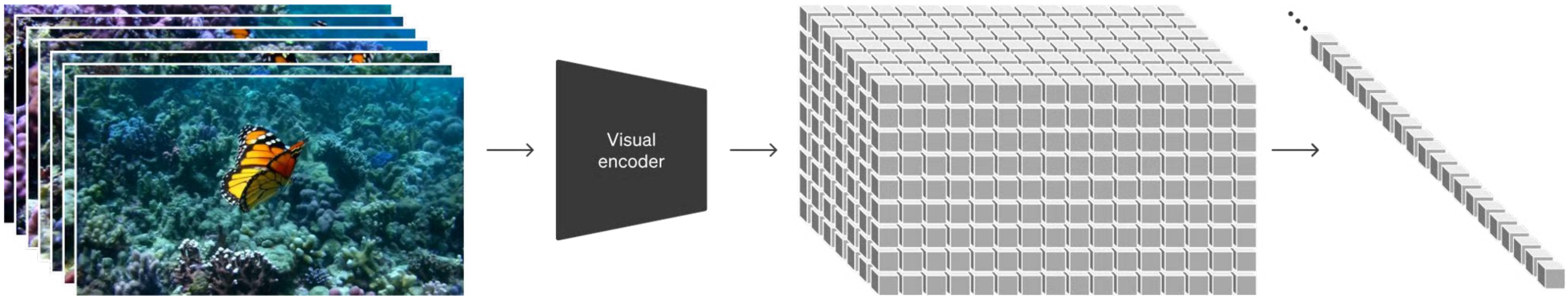
https://openaccess.thecvf.com/content/ICCV2023/papers/Peebles_Scalable_Diffusion_Models_with_Transformers_ICCV_2023_paper.pdf

World models



https://worldmodels.github.io/

SORA: Video generation models as world simulators



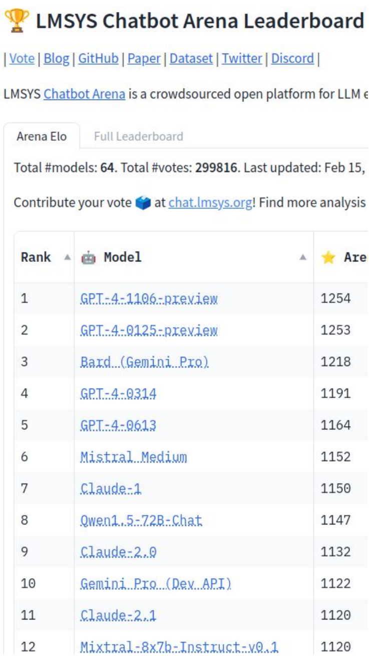
"...transformer architecture that operates on spacetime patches of video and image latent codes..."

EMO - Generating Expressive Portrait Videos with Audio2Video Diffusion

https://arxiv.org/pdf/2402.17485.pdf

SORA, EMO as Foundation models?

[Google Gemma, Gemini 1.5 Pro]



🏆 LMSYS Chatbot Arena Leaderboard

| Vote | Blog | GitHub | Paper | Dataset | Twitter | Discord |

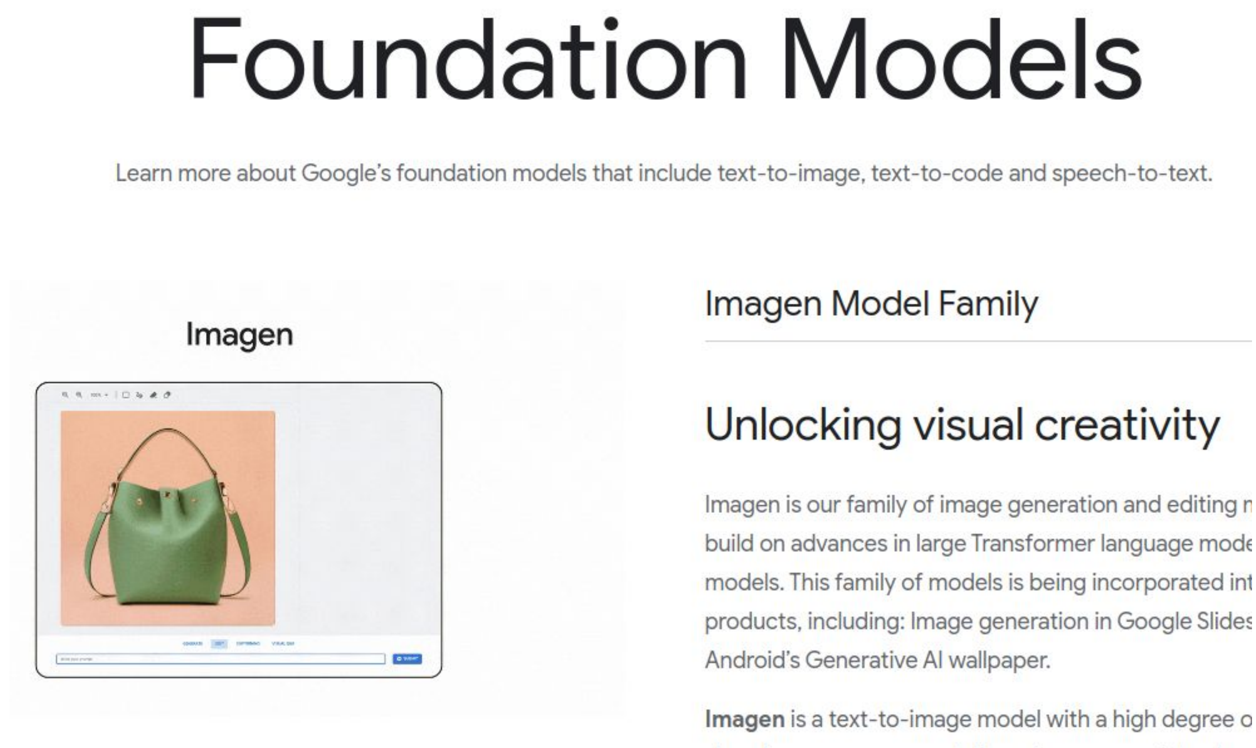
LMSYS Chatbot Arena is a crowdsourced open platform for LLM e

Arena Elo   Full Leaderboard

Total #models: **64**. Total #votes: **299816**. Last updated: Feb 15,

Contribute your vote 🗳️ at chat.lmsys.org! Find more analysis

Rank	🤖 Model	⭐ Are
1	GPT-4-1106-preview	1254
2	GPT-4-0125-preview	1253
3	Bard (Gemini Pro)	1218
4	GPT-4-0314	1191
5	GPT-4-0613	1164
6	Mistral Medium	1152
7	Claude-1	1150
8	Qwen1.5-72B-Chat	1147
9	Claude-2.0	1132
10	Gemini Pro (Dev API)	1122
11	Claude-2.1	1120
12	Mixtral-8x7b-Instruct-v0.1	1120



[Gov.uk initial report on foundational models]

[Governing open foundational models]

[CS 886: Recent Advances on Foundation Models]

What is next in AI research?

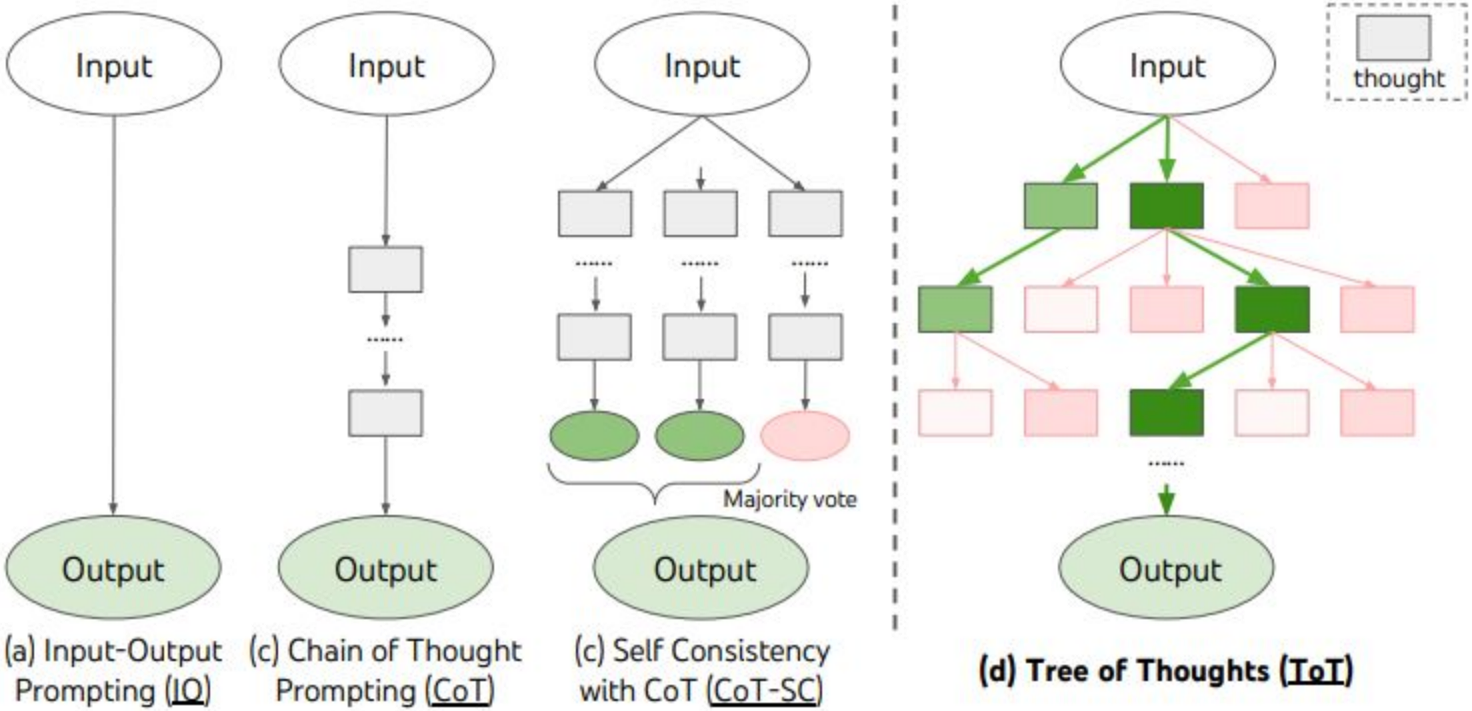
OpenAI board warned Project Q* could 'threaten humanity' prior to Sam Altman's sacking

Newest algorithm is said to have solved simple maths problems – a major step forward for AI

Matthew Field

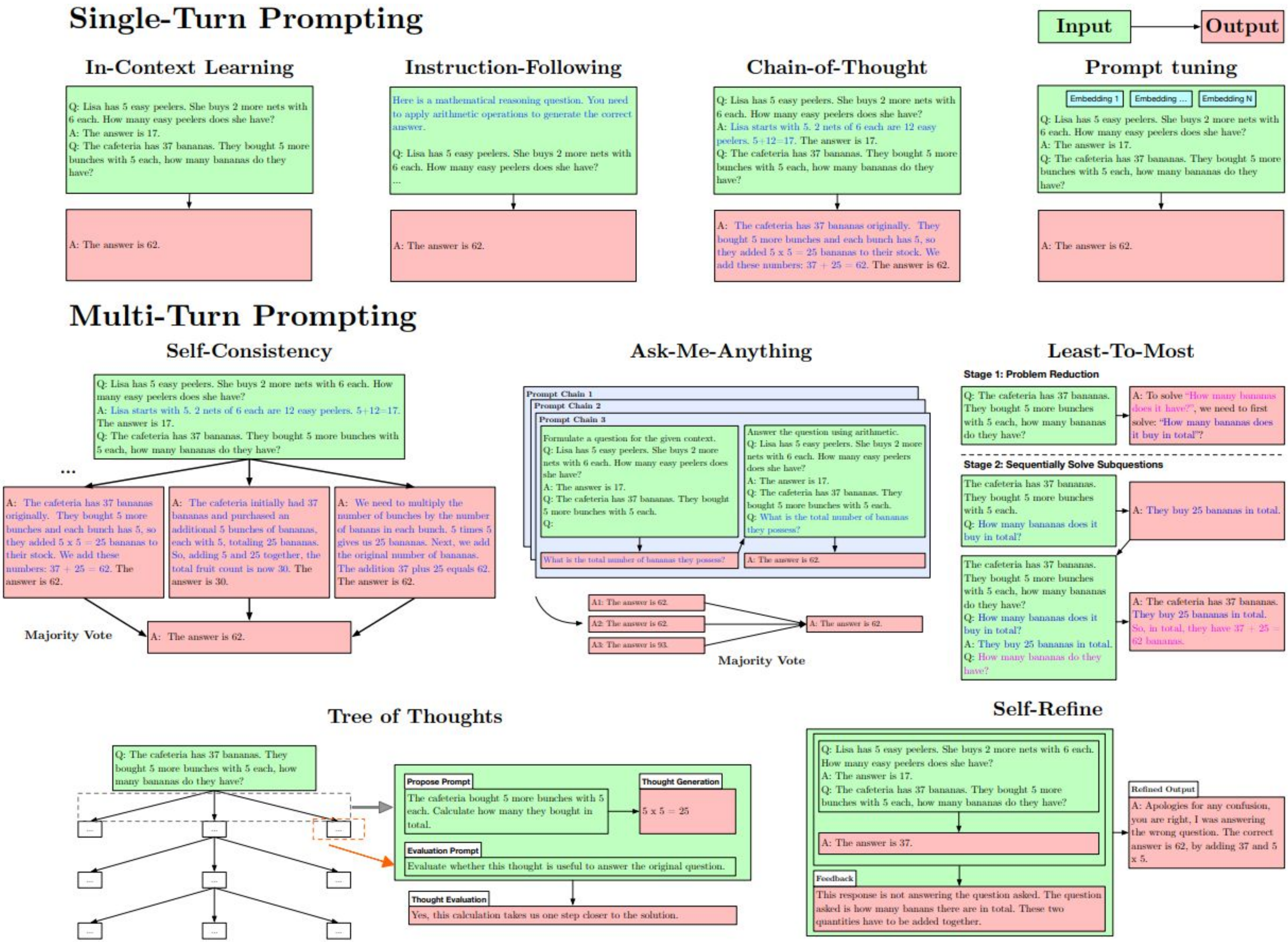
23 November 2023 • 12:36pm

Some speculations …



[Tree of Thoughts: Deliberate Problem Solving with Large Language Models]

4. Search algorithm. Finally, within the ToT framework, one can plug and play different search algorithms depending on the tree structure. We explore two relatively simple search algorithms and leave more advanced ones (e.g. A* [11], MCTS [2]) for future work:



Single-Turn Prompting
Input
Output
In-Context Learning
Q: Lisa has 5 easy peelers. She buys 2 more nets with 6 each. How many easy peelers does she have?
A: The answer is 17.
Q: The cafeteria has 37 bananas. They bought 5 more bunches with 5 each, how many bananas do they have?
A: The answer is 62.
Instruction-Following
Here is a mathematical reasoning question. You need to apply arithmetic operations to generate the correct answer.
Q: Lisa has 5 easy peelers. She buys 2 more nets with 6 each. How many easy peelers does she have?
...
A: The answer is 62.
Chain-of-Thought
Q: Lisa has 5 easy peelers. She buys 2 more nets with 6 each. How many easy peelers does she have?
A: Lisa starts with 5. 2 nets of 6 each are 12 easy peelers. 5+12=17. The answer is 17.
Q: The cafeteria has 37 bananas. They bought 5 more bunches with 5 each, how many bananas do they have?
A: The cafeteria has 37 bananas originally. They bought 5 more bunches and each bunch has 5, so they added 5 x 5 = 25 bananas to their stock. We add these numbers: 37 + 25 = 62. The answer is 62.
Prompt tuning
Embedding 1
Embedding ...
Embedding N
Q: Lisa has 5 easy peelers. She buys 2 more nets with 6 each. How many easy peelers does she have?
A: The answer is 17.
Q: The cafeteria has 37 bananas. They bought 5 more bunches with 5 each, how many bananas do they have?
A: The answer is 62.
Multi-Turn Prompting
Self-Consistency
Q: Lisa has 5 easy peelers. She buys 2 more nets with 6 each. How many easy peelers does she have?
A: Lisa starts with 5. 2 nets of 6 each are 12 easy peelers. 5+12=17. The answer is 17.
Q: The cafeteria has 37 bananas. They bought 5 more bunches with 5 each, how many bananas do they have?
...
A: The cafeteria has 37 bananas originally. They bought 5 more bunches and each bunch has 5, so they added 5 x 5 = 25 bananas to their stock. We add these numbers: 37 + 25 = 62. The answer is 62.
A: The cafeteria initially had 37 bananas and purchased an additional 5 bunches of bananas, each with 5, totaling 25 bananas. So, adding 5 and 25 together, the total fruit count is now 30. The answer is 30.
A: We need to multiply the number of bunches by the number of banans in each bunch. 5 times 5 gives us 25 bananas. Next, we add the original number of bananas. The addition 37 plus 25 equals 62. The answer is 62.
Majority Vote
A: The answer is 62.
Ask-Me-Anything
Prompt Chain 1
Prompt Chain 2
Prompt Chain 3
Formulate a question for the given context.
Q: Lisa has 5 easy peelers. She buys 2 more nets with 6 each. How many easy peelers does she have?
A: The answer is 17.
Q: The cafeteria has 37 bananas. They bought 5 more bunches with 5 each.
Q:
What is the total number of bananas they possess?
Answer the question using arithmetic.
Q: Lisa has 5 easy peelers. She buys 2 more nets with 6 each. How many easy peelers does she have?
A: The answer is 17.
Q: The cafeteria has 37 bananas. They bought 5 more bunches with 5 each.
Q: What is the total number of bananas they possess?
A: The answer is 62.
A1: The answer is 62.
A2: The answer is 62.
A3: The answer is 93.
A: The answer is 62.
Majority Vote
Least-To-Most
Stage 1: Problem Reduction
Q: The cafeteria has 37 bananas. They bought 5 more bunches with 5 each, how many bananas do they have?
A: To solve "How many bananas does it have?", we need to first solve: "How many bananas does it buy in total"?
Stage 2: Sequentially Solve Subquestions
The cafeteria has 37 bananas. They bought 5 more bunches with 5 each.
Q: How many bananas does it buy in total?
A: They buy 25 bananas in total.
The cafeteria has 37 bananas. They bought 5 more bunches with 5 each, how many bananas do they have?
Q: How many bananas does it buy in total?
A: They buy 25 bananas in total.
Q: How many bananas do they have?
A: The cafeteria has 37 bananas. They buy 25 bananas in total. So, in total, they have 37 + 25 = 62 bananas.
Tree of Thoughts
Q: The cafeteria has 37 bananas. They bought 5 more bunches with 5 each, how many bananas do they have?
Propose Prompt
The cafeteria bought 5 more bunches with 5 each. Calculate how many they bought in total.
Thought Generation
5 x 5 = 25
Evaluation Prompt
Evaluate whether this thought is useful to answer the original question.
Thought Evaluation
Yes, this calculation takes us one step closer to the solution.
Self-Refine
Q: Lisa has 5 easy peelers. She buys 2 more nets with 6 each. How many easy peelers does she have?
A: The answer is 17.
Q: The cafeteria has 37 bananas. They bought 5 more bunches with 5 each, how many bananas do they have?
A: The answer is 37.
Feedback
This response is not answering the question asked. The question asked is how many banans there are in total. These two quantities have to be added together.
Refined Output
A: Apologies for any confusion, you are right, I was answering the wrong question. The correct answer is 62, by adding 37 and 5 x 5.

Path towards Super-intelligence and AGI

What is needed?

- Memory augmented models
- Planning and reasoning - blend with reinforcement learning?
- Better training datasets
- Better learning from human feedback
- Better benchmarks
- Open ended learning?

https://arxiv.org/pdf/2301.04589.pdf

Memory augmented LLMs are computationally universal

We show that transformer-based large language models are computationally universal when augmented with an external memory. Any deterministic language model that conditions on strings of bounded length is equivalent to a finite automaton, hence computationally limited. However, augmenting such models with a read-write memory creates the possibility of processing arbitrarily large inputs and, potentially, simulating any algorithm. We establish that an existing large language model, Flan-U-PaLM 540B, can be combined with an associative read-write memory to exactly simulate the execution of a universal Turing machine, $U_{15,2}$. A key aspect of the finding is that it does not require any modification of the language model weights. Instead, the construction relies solely on designing a form of stored instruction computer that can subsequently be programmed with a specific set of prompts.

	A	B	C	D	E	F	G	H
0	$0, +, B$	$1, +, C$	$0, -, G$	$0, -, F$	$1, +, A$	$1, -, D$	$0, +, H$	$1, -, I$
1	$1, +, A$	$1, +, A$	$0, -, E$	$1, -, E$	$1, -, D$	$1, -, D$	$1, -, G$	$1, -, G$

	I	J	K	L	M	N	O
0	$0, +, A$	$1, -, K$	$0, +, L$	$0, +, M$	$0, -, B$	$0, -, C$	$0, +, N$
1	$1, -, J$	halt	$1, +, N$	$1, +, L$	$1, +, L$	$0, +, O$	$1, +, N$

Table 1: *Transition table for the universal Turing machine $U_{15,2}$. Rows are indexed by the read symbol σ, columns are indexed by the state q, and each table entry (σ', m, q') specifies the write symbol σ', the tape head move $m \in \{-1, +1\}$, and the next state q'.*

AlphaGO Zero (2017)

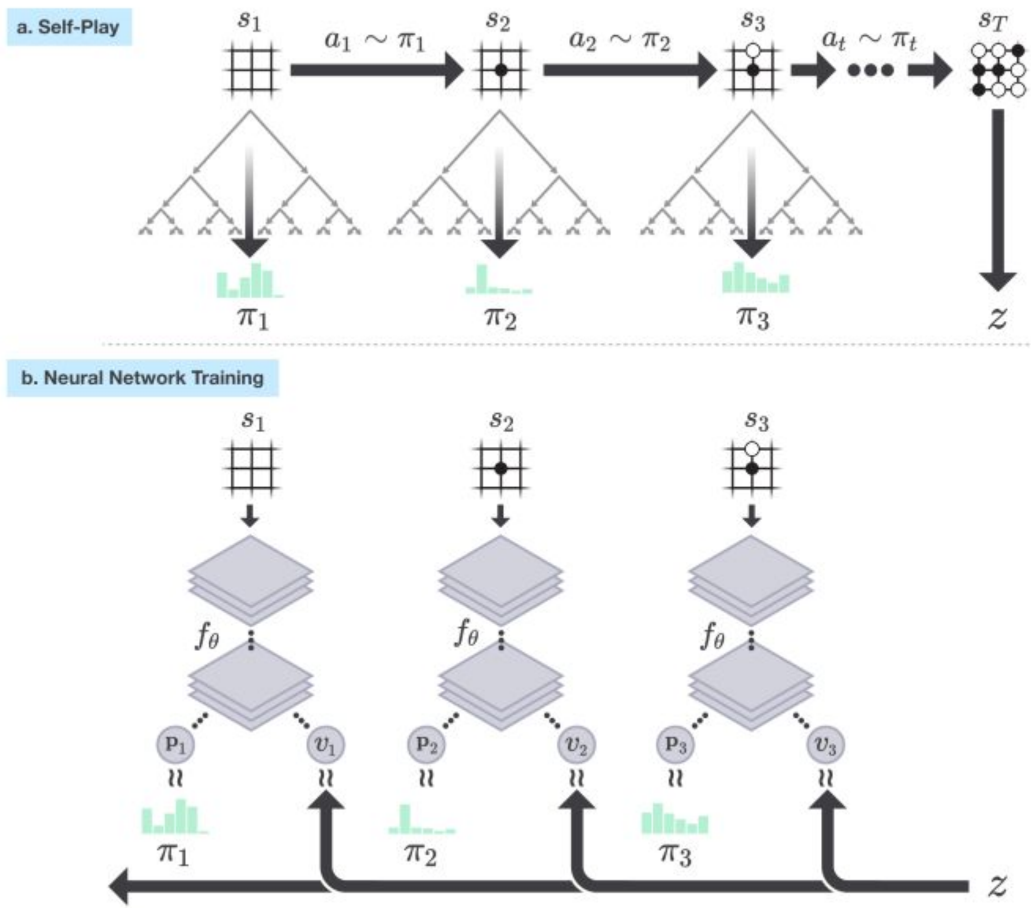
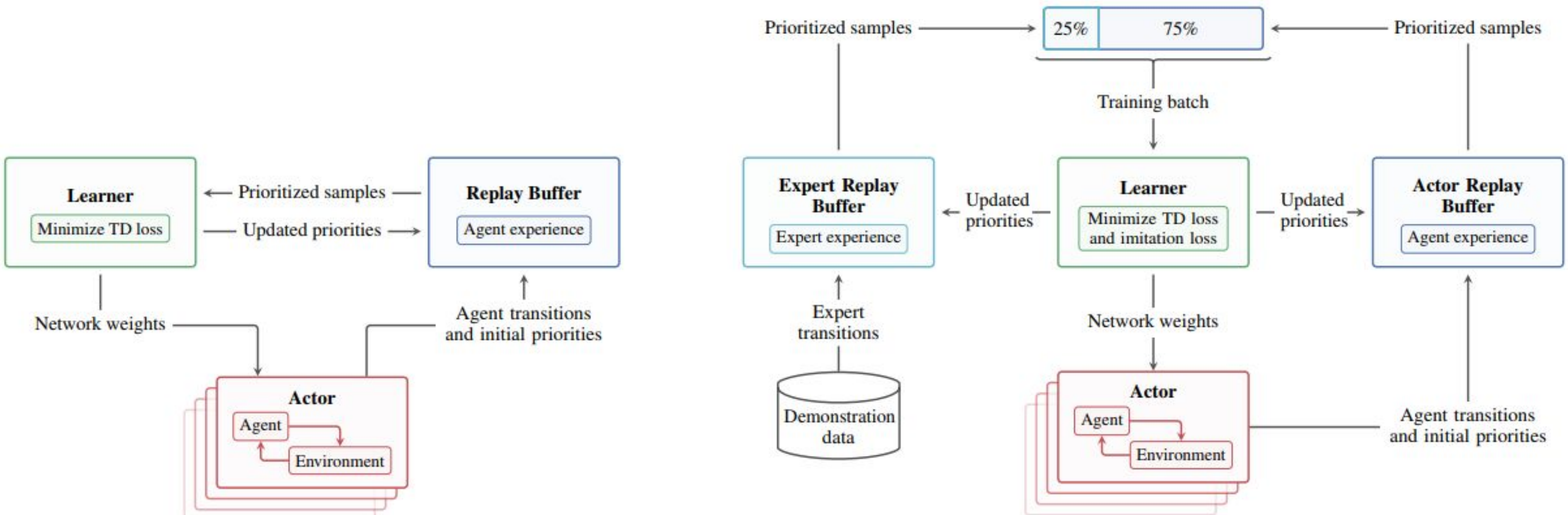


Figure 1: **Self-play reinforcement learning in *AlphaGo Zero*. a** The program plays a game $s_1, ..., s_T$ against itself.

https://discovery.ucl.ac.uk/id/eprint/10045895/1/agz_unformatted_nature.pdf

https://arxiv.org/pdf/1805.11593.pdf

Observe and Look Further (2018)



(a) Ape-X DQN (b) Ape-X DQfD (ours)

… we ease the exploration problem by using human demonstrations that guide the agent towards rewarding states …

https://arxiv.org/pdf/1802.07740.pdf

Machine Theory of Mind (2018)

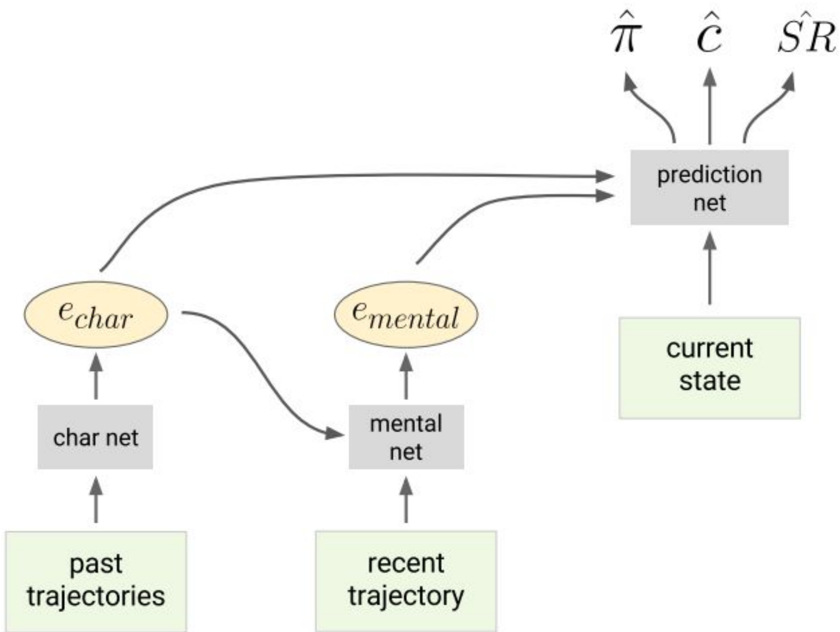


Figure 1. **ToMnet architecture.** The *character net* parses an agent's past trajectories from a set of POMDPs to form a character embedding, e_{char}. The *mental state net* parses the agent's trajectory on the current episode, to form an embedding of its mental state, e_{mental}. These embeddings are fed into the *prediction net*, which is then queried with a current state. This outputs predictions about future behaviour, such as next-step action probabilities ($\hat{\pi}$), probabilities of whether certain objects will be consumed ($\hat{c}$), and predicted successor representations ($\widehat{SR}$; Dayan, 1993).

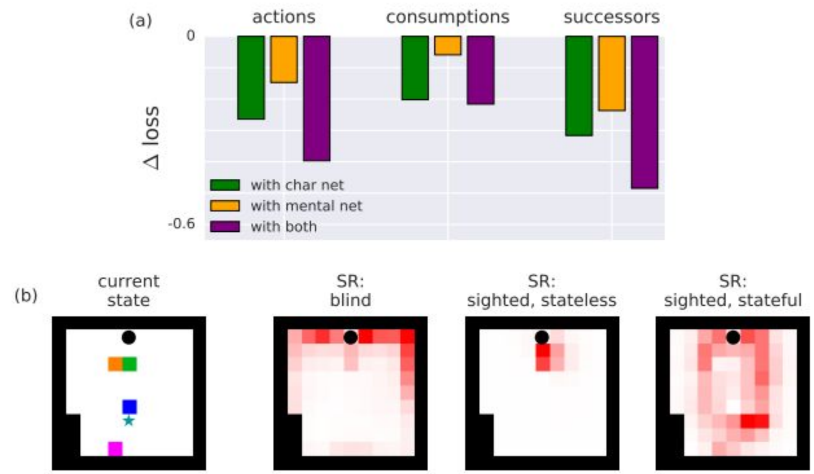
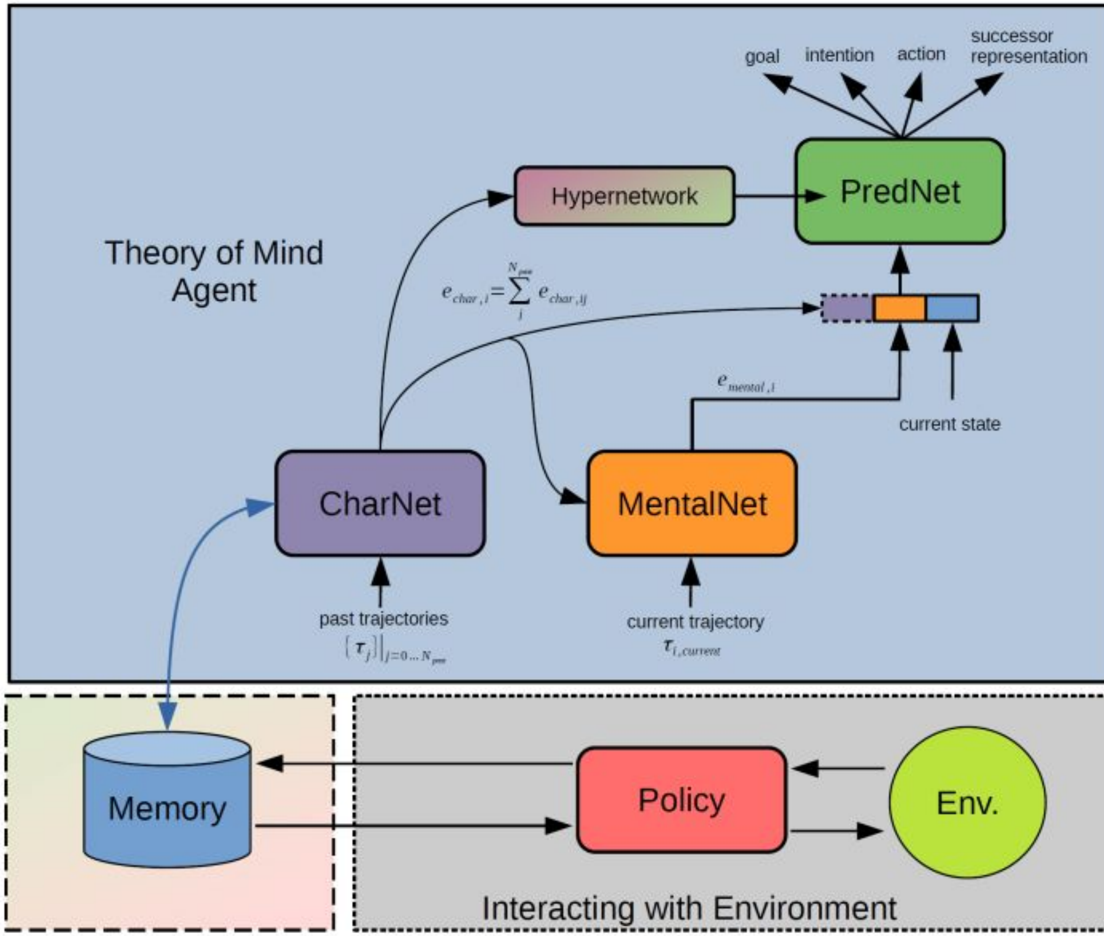


Figure 7. **Using the ToMnet to characterise trained neural-net agents. (a)** Usefulness of ToMnet components for the three behavioural prediction targets, compared with a simple ToMnet with no character nor mental net. Longer bars are better; including both character and mental nets is best. More details are given in Table A1. **(b)** A ToMnet's prediction of agents' future state occupancy given a query POMDP state at time $t = 0$ (left), as per Fig 4d. Star denotes the subgoal. The maps on the right are produced after observing behaviour on $N_{\mathrm{past}} = 5$ past POMDPs from a sampled agent of each subspecies (always preferring the pink object). The ToMnet does not know a priori which subspecies each agent belongs to, but infers it from past behaviour.

A ToM architecture

- Observer maintains memory of previous episodes of the agent.
- It theorizes the “traits” of the agent.
    - Implemented as Hyper Networks.
- Given the current episode, the observer tries to infer goal, intention, action, etc of the agent.
    - Implemented as memory retrieval through attention mechanisms.



https://proceedings.mlr.press/v119/badia20a/badia20a.pdf

Agent57: Outperforming the Atari Human Benchmark

Adrià Puigdomènech Badia [*1] Bilal Piot [*1] Steven Kapturowski [*1] Pablo Sprechmann [*1]
Alex Vitvitskyi [1] Daniel Guo [1] Charles Blundell [1]

Abstract

Atari games have been a long-standing benchmark in the reinforcement learning (RL) community for the past decade. This benchmark was proposed to test general competency of RL algorithms. Previous work has achieved good average performance by doing outstandingly well on many games of the set, but very poorly in several of the most challenging games. We propose Agent57, the first deep RL agent that outperforms the standard human benchmark on all 57 Atari games. To achieve this result, we train a neural network which parameterizes a family of policies ranging from very exploratory to purely exploitative. We propose an adaptive mechanism to choose which policy to prioritize throughout

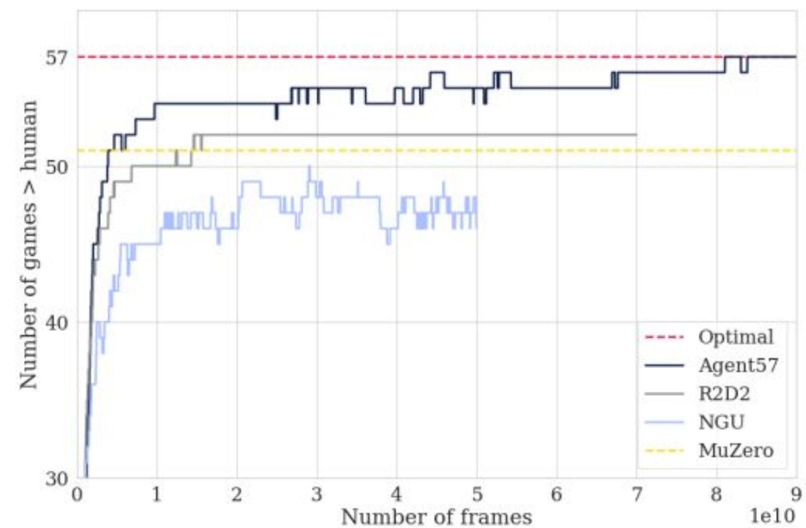
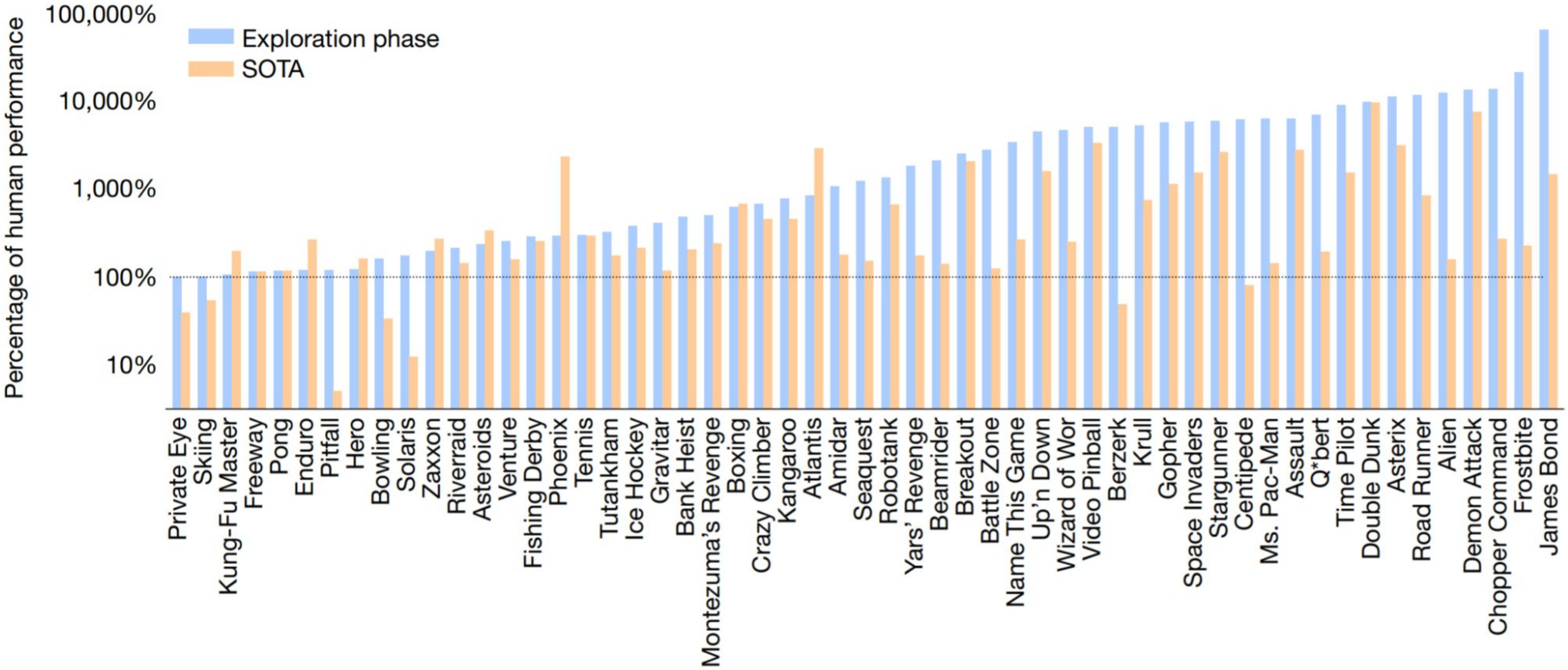


Figure 1. Number of games where algorithms are better than the human benchmark throughout training for Agent57 and state-of-the-art baselines on the 57 Atari games.

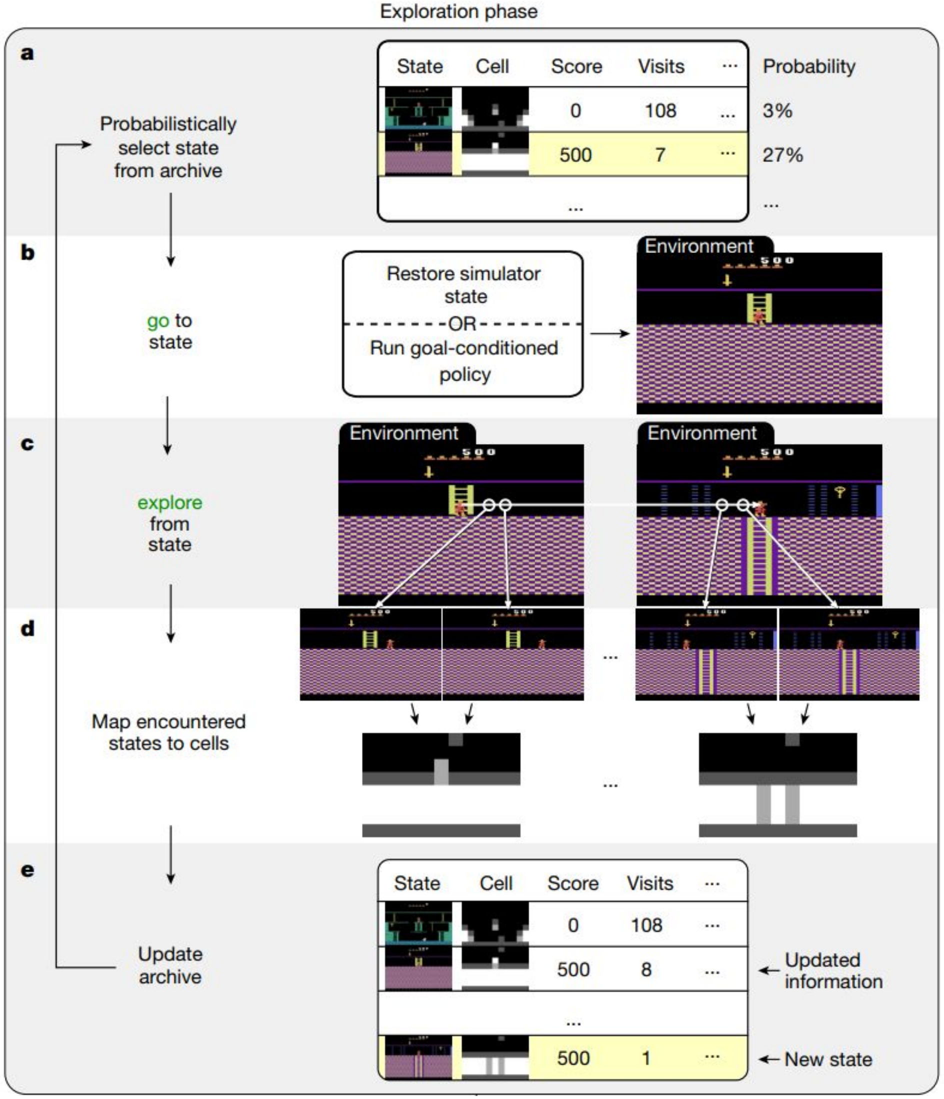
… we train a neural network which parameterizes a family of policies ranging from very exploratory to purely exploitative. We propose an adaptive mechanism to choose which policy to prioritize throughout the training process. …

Human competitiveness of results



https://www.nature.com/articles/s41586-020-03157-9.pdf

First return then explore (2021)



- What if rewards provide sparse and deceptive feedback?
- We can explicitly ‘remember’ promising states and returning to such states before intentionally exploring

https://arxiv.org/pdf/2305.14992.pdf

Reasoning with Language Model is Planning with World Model

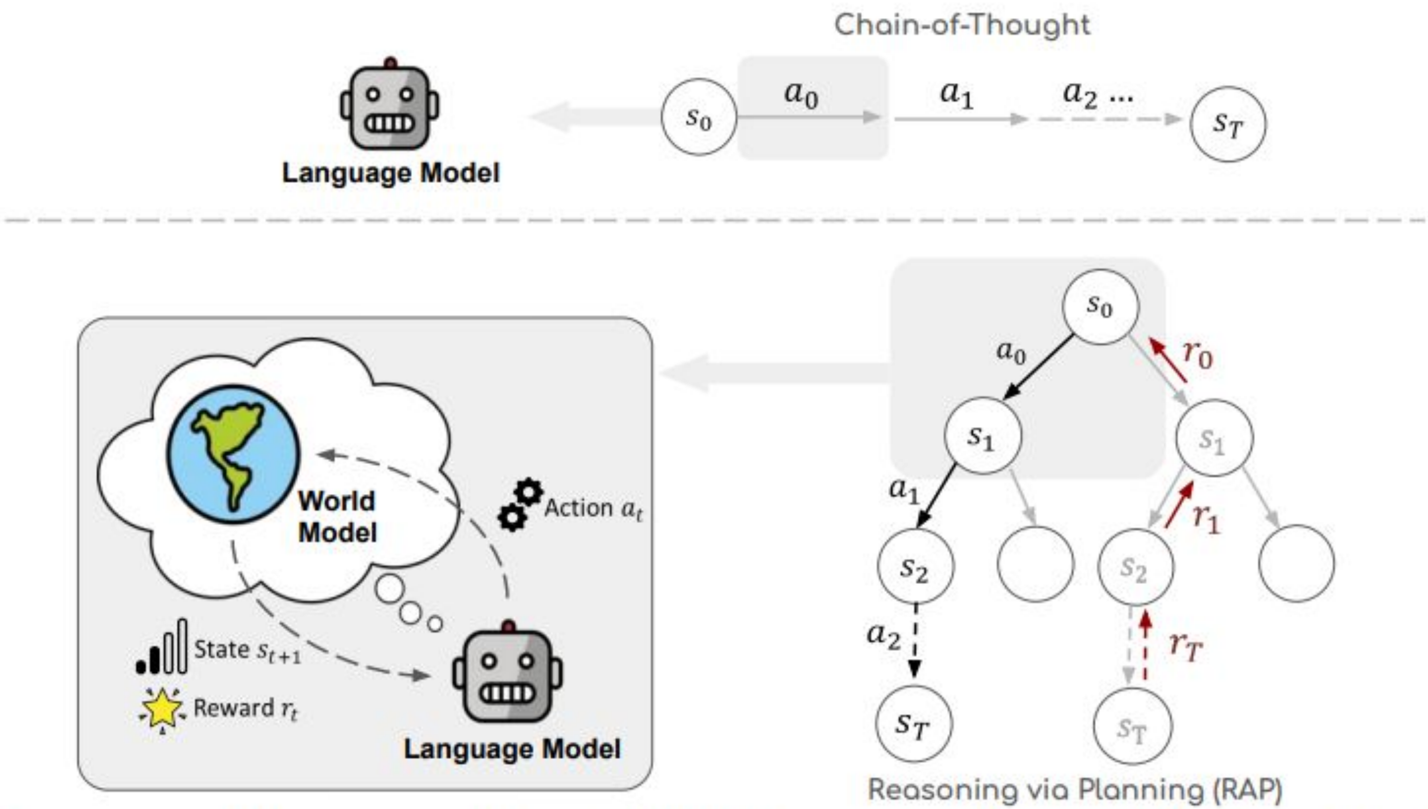
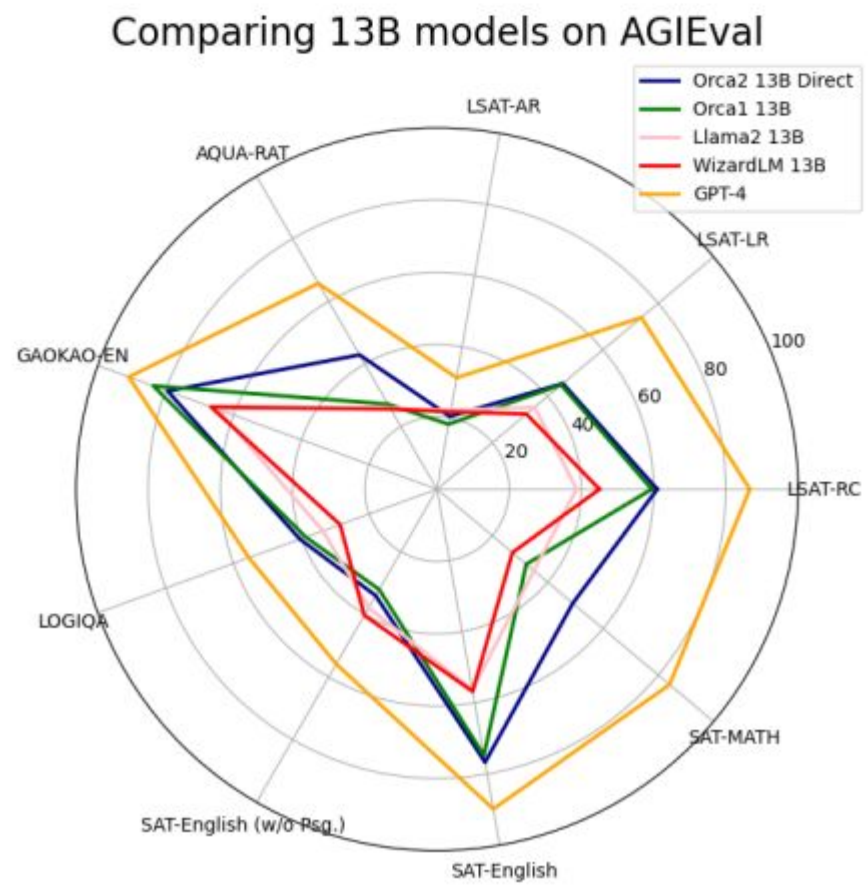


Figure 1: An overview of Reasoning via Planning (RAP). Compared with previous LLM reasoning methods like Chain-of-Thought (Wei et al., 2022), we explicitly model the world state from a world model (repurposed from the language model), and leverage advanced planning algorithms to solve the reasoning problems.

AGIeval - evaluating reasoning capabilities

Orca 2: Teaching Small Language Models How to Reason: https://arxiv.org/pdf/2311.11045.pdf

AGIEval: A Human-Centric Benchmark for Evaluating Foundation Models: https://arxiv.org/pdf/2304.06364.pdf



Fast open-ended environment for training reinforcement learning agents

Mining

Farming

Dungeon Crawling

Archery

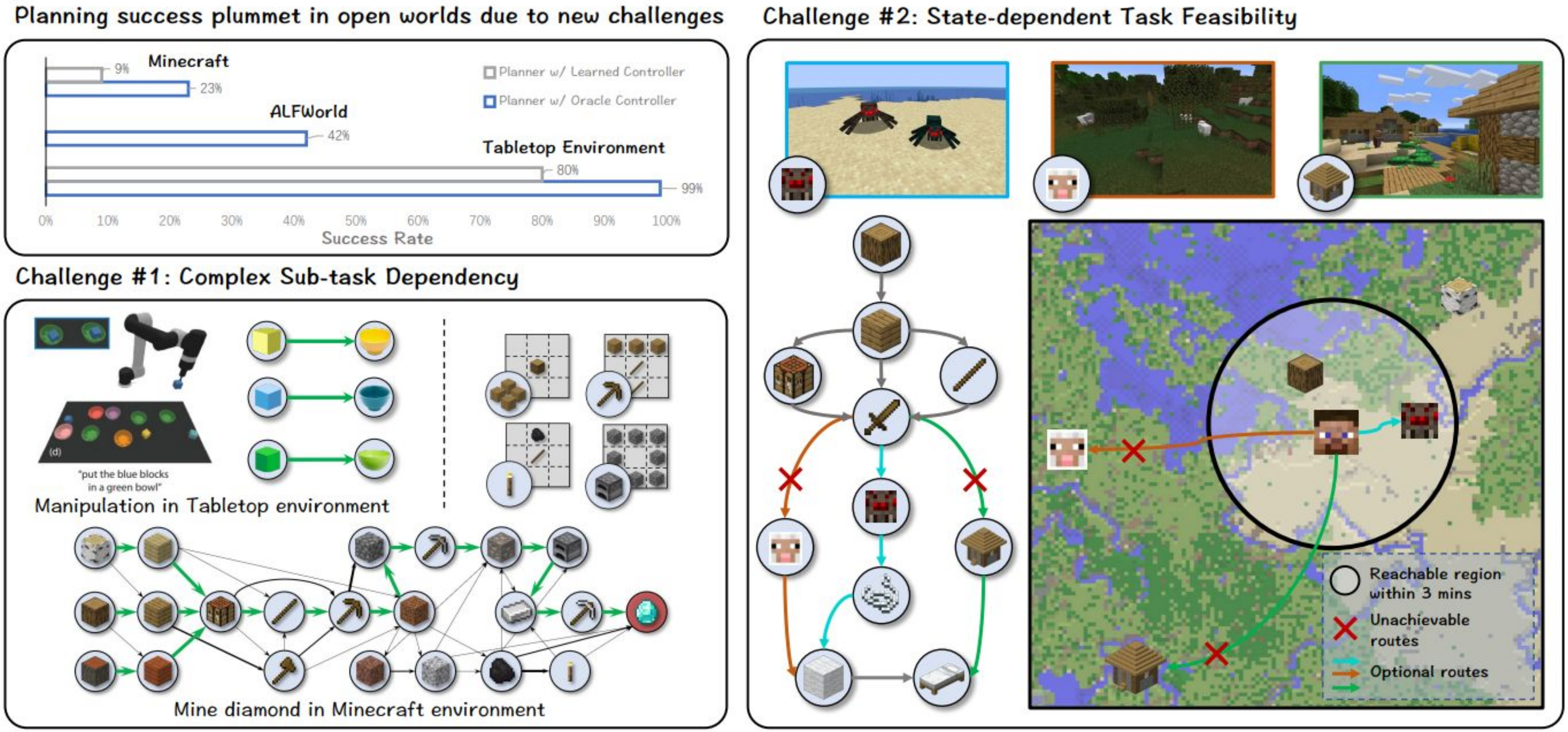
Building

Magic

https://craftaxenv.github.io/

https://arxiv.org/pdf/2302.01560.pdf

Describe, Explain, Plan and Select: LLMs Enables Open-World Multi-Task Agents



DreamerV3: universal world models

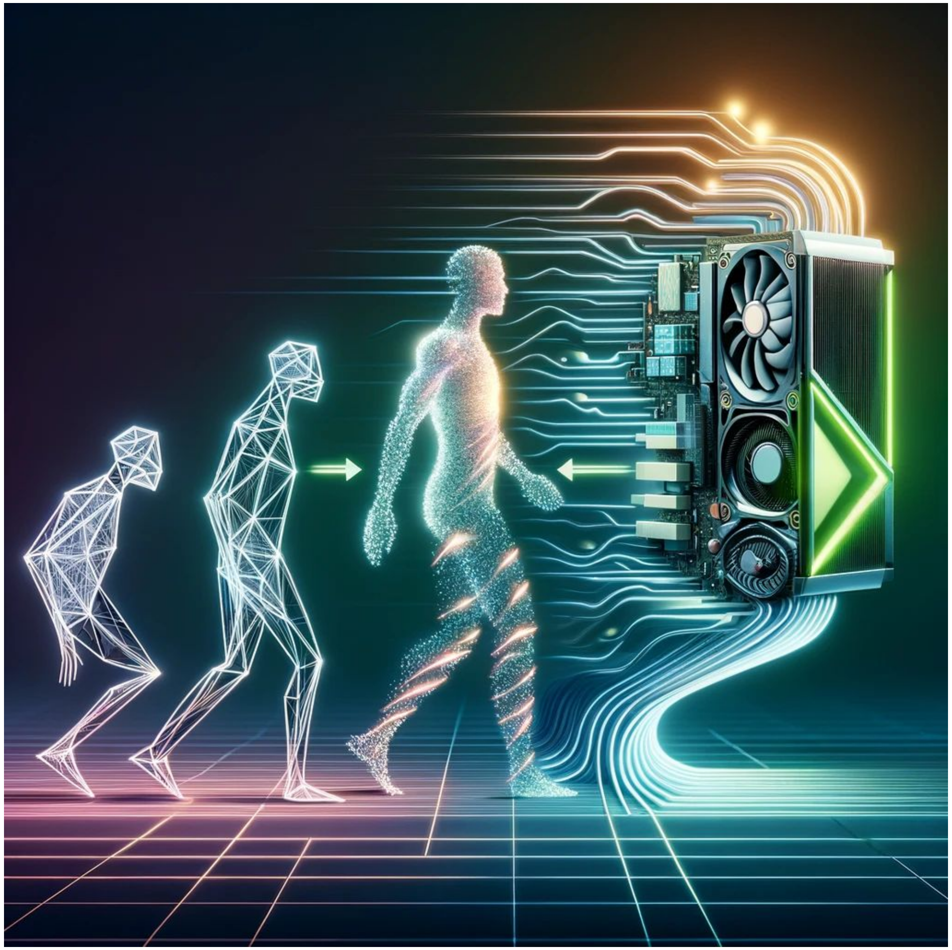
https://arxiv.org/pdf/2301.04104.pdf

Google Genie: generating playable environments from single image

Zero shot world models of diverse environments: https://arxiv.org/pdf/2402.15391v1.pdf

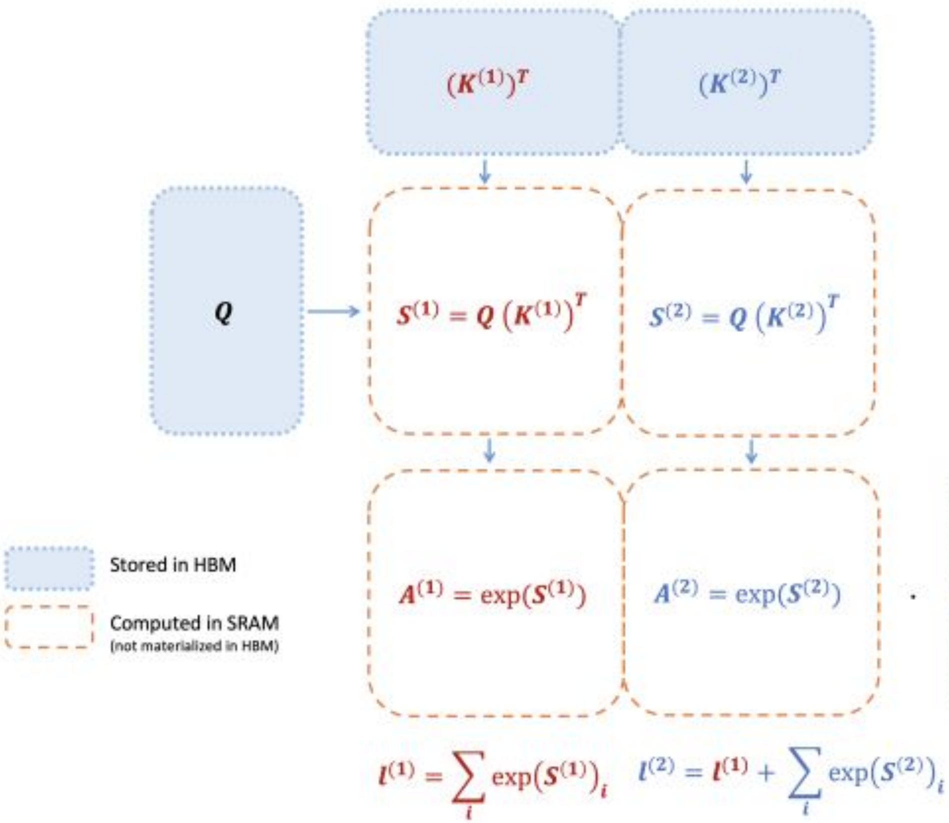
Making AI systems affordable

Transforming research prototypes into production AI systems that are fast and can handle millions of concurrent users …



https://github.com/Dao-AILab/flash-attention

Flash attention - accelerating transformers



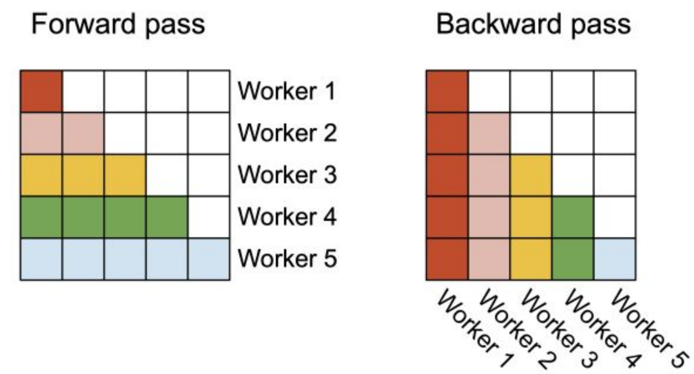


Figure 1: Diagram of how FlashAttention forward pass is performed, when the key $\mathbf{K}$ is partitioned into two blocks and the value $\mathbf{V}$ is also partitioned into two blocks. By computing attention with respect to each block and rescaling the output, we get the right answer at the end, while avoiding expensive memory reads/writes of the intermediate matrices $\mathbf{S}$ and $\mathbf{P}$. We simplify the diagram, omitting the step in softmax that subtracts each element by the row-wise max.

Ring attention

	Max context size supported (×1e3)				
	Vanilla	Memory Efficient Attn	Memory Efficient Attn and FFN	Ring Attention (Ours)	Ours vs SOTA
8x A100 NVLink					
3B	4	32	64	**512**	8x
7B	2	16	32	**256**	8x
13B	2	4	16	**128**	8x
32x A100 InfiniBand					
7B	4	64	128	**4096**	32x
13B	4	32	64	**2048**	32x
TPUv3-512					
7B	1	4	8	**2048**	256x
13B	1	2	8	**1024**	128x
TPUv4-1024					
3B	8	16	32	**16384**	512x
7B	4	8	16	**8192**	512x
13B	4	8	16	**4096**	256x
30B	2	4	8	**2048**	256x

https://arxiv.org/pdf/2310.01889.pdf

https://arxiv.org/pdf/2106.09685.pdf

LORA: LOW-RANK ADAPTATION OF LARGE LANGUAGE MODELS

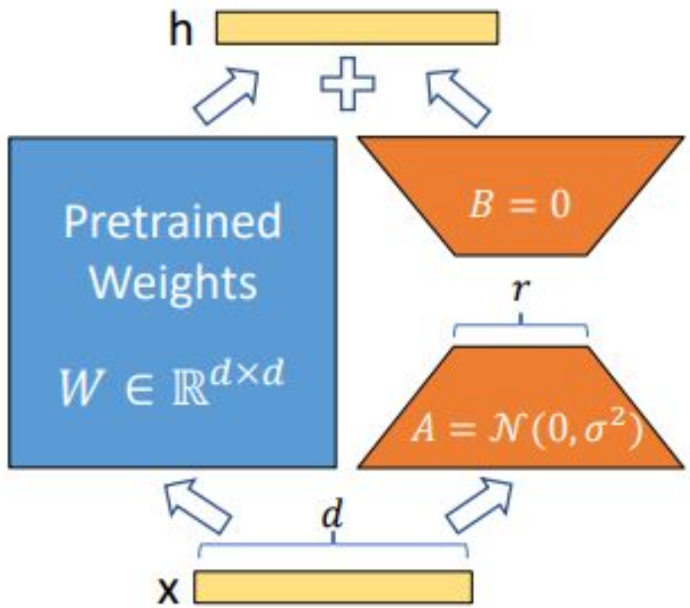


Figure 1: Our reparametrization. We only train A and B.

https://arxiv.org/pdf/2305.14314.pdf

QLoRA for transformers

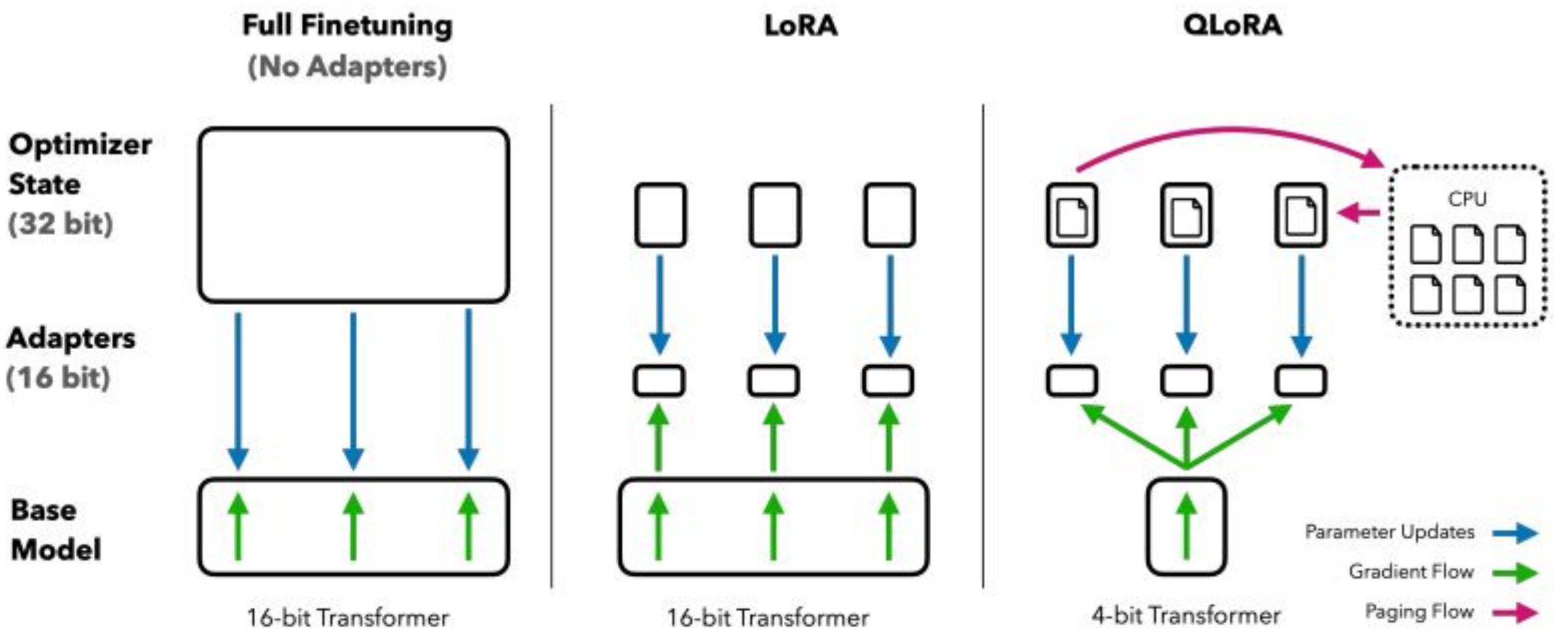
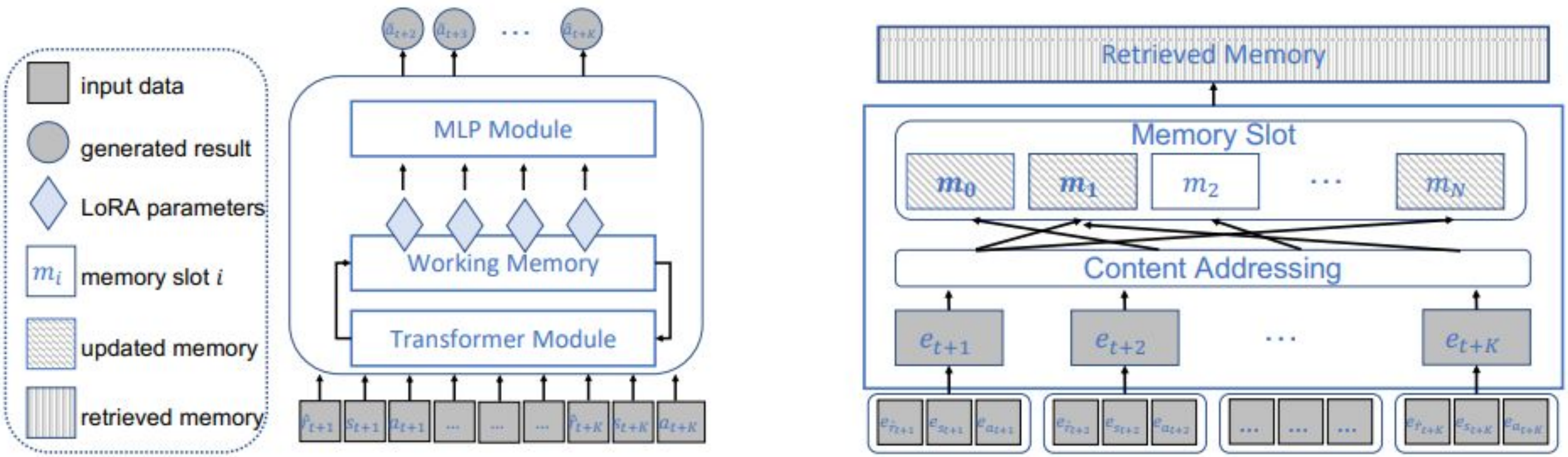


Figure 1: Different finetuning methods and their memory requirements. QLoRA improves over LoRA by quantizing the transformer model to 4-bit precision and using paged optimizers to handle memory spikes.

https://arxiv.org/pdf/2305.16338.pdf

Decision Transformers with Internal Working Memory



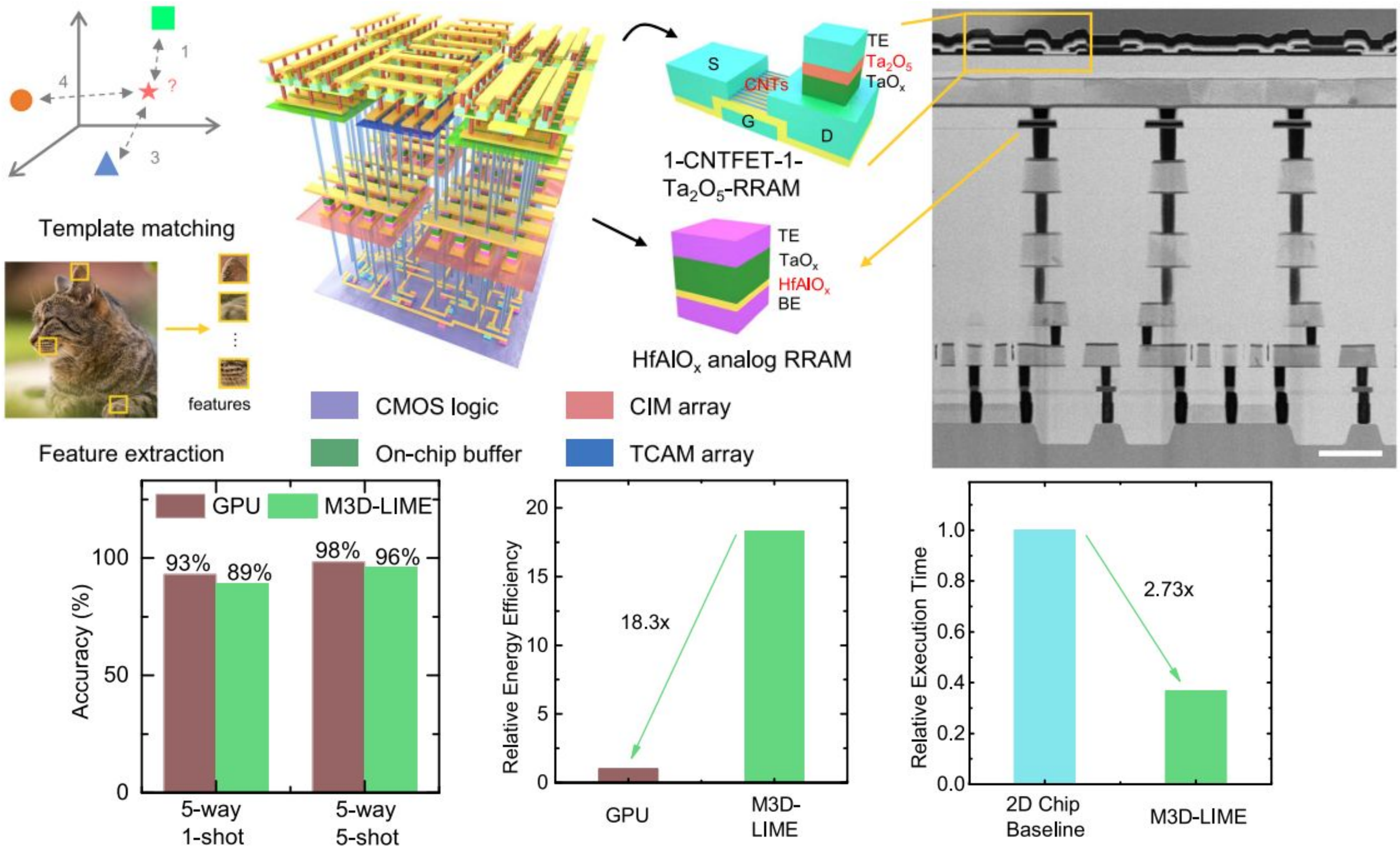
(a) DT-Mem architecture (b) Memory module architecture

Figure 2: An overview of the proposed DT-Mem architecture. In 2a, Transformer module interact with working memory multiple times.

DT-Mem achieves better generalization on Atari games with only 10% of the model parameters compared to the state-of-the-art method.

https://www.nature.com/articles/s41467-023-42981-1

Hardware acceleration in 3D



The classification accuracy of one-shot/few-shot learning on the Omniglot dataset using GPU and the M3D-LIME. The accuracy is the average of 5 randomly selected classes (5-way) in the dataset. Benchmark of the energy efficiency of the M3D-LIME chip and GPU. Benchmark of the execution time on the M3D-LIME and 2D chip baseline

Real world AI use cases

Autonomous driving

Autopilot Review: Tesla and Elon Musk have placed a big target on delivering autonomous Full Self-Driving and even a Tesla RoboTaxi service. By most accounts, that's extremely ambitious and it doesn't seem close-at-hand… yet.

That said, the latest Full Self-Driving Beta is showing some impressive improvements. Now with Hardware 4 having launched in 2023 we should continue to see a steady pace of improvements to the Autopilot and Full Self-Driving experience. However, it's important to remember, it's likely still years away from being truly autonomous.

—

Police investigate after Waymo driverless car vandalized, set ablaze in SF's Chinatown

Planning-oriented Autonomous Driving

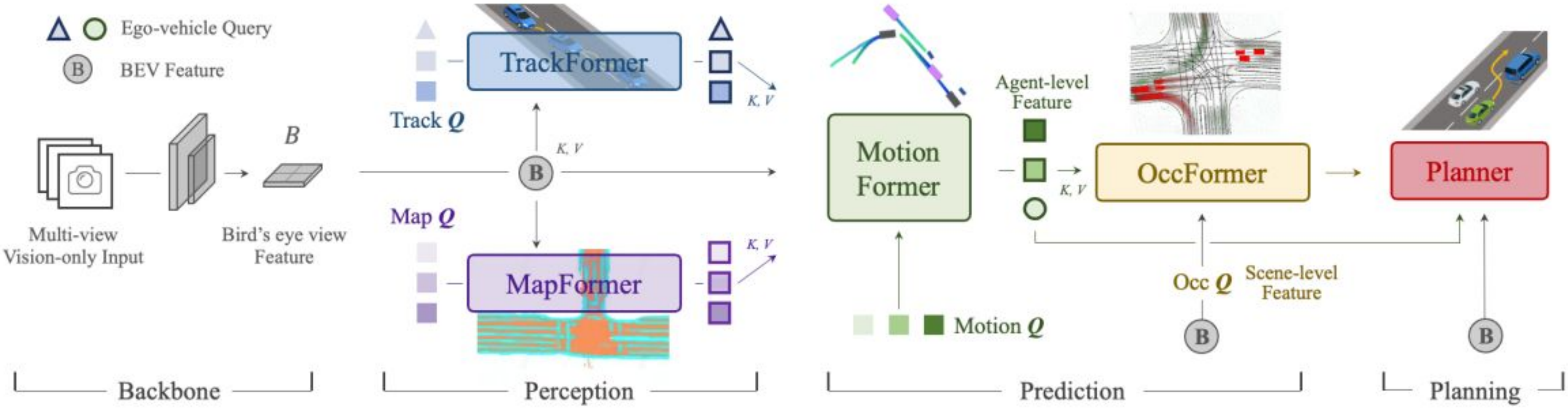


Figure 2. **Pipeline** of Unified Autonomous Driving (UniAD). It is exquisitely devised following planning-oriented philosophy. Instead of a simple stack of tasks, we investigate the effect of each module in perception and prediction, leveraging the benefits of joint optimization from preceding nodes to final planning in the driving scene. All perception and prediction modules are designed in a transformer decoder structure, with task queries as interfaces connecting each node. A simple attention-based planner is in the end to predict future waypoints of the ego-vehicle considering the knowledge extracted from preceding nodes. The map over occupancy is for visual purpose only.

https://openaccess.thecvf.com/content/CVPR2023/papers/Hu_Planning-Oriented_Autonomous_Driving_CVPR_2023_paper.pdf

https://arxiv.org/pdf/2309.17080.pdf

A Generative World Model for Autonomous Driving

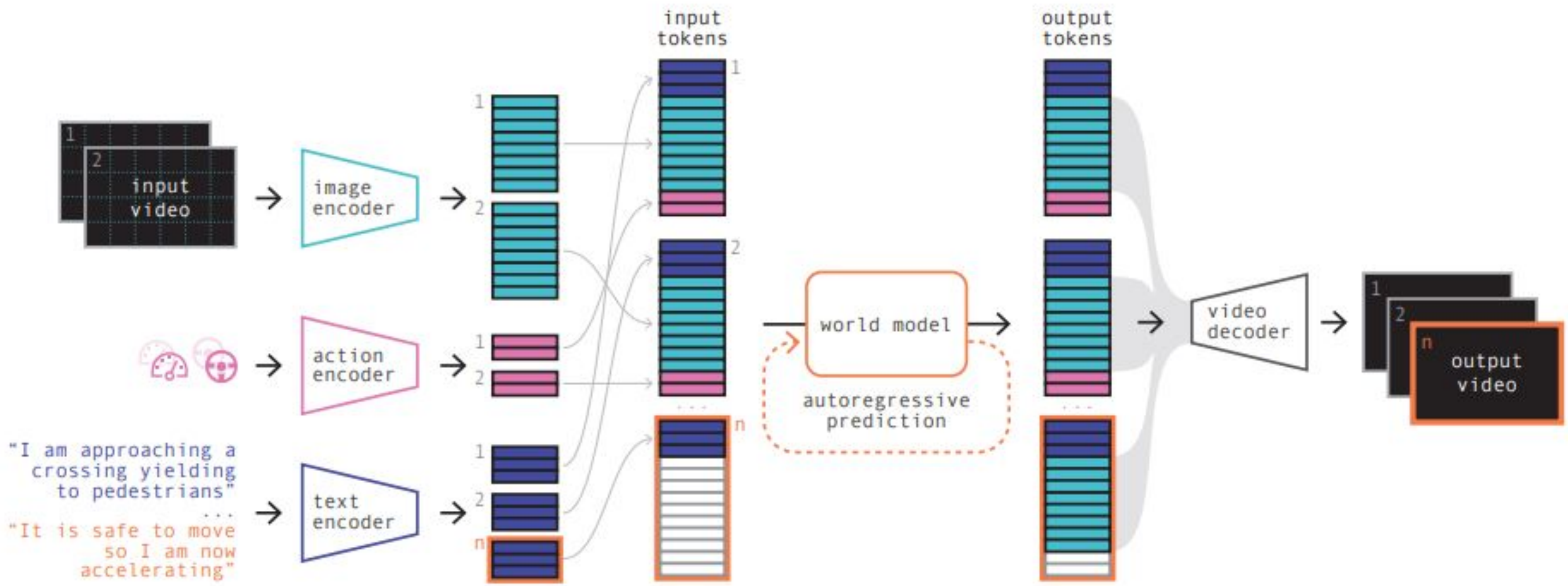


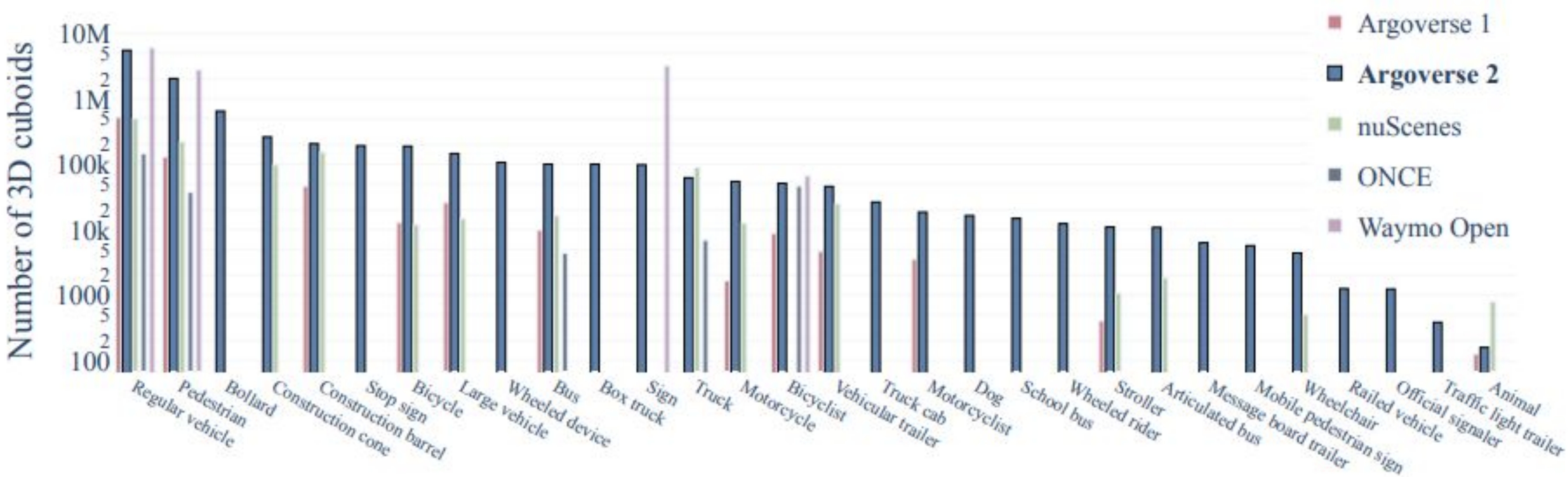
Figure 2: Architecture of GAIA-1. First, we encode information from all input modalities (video, text, action) into a common representation: images, text and actions are encoded as a sequence of tokens. The world model is an autoregressive transformer that predicts the next image token conditioned on past image, text, and action tokens. Finally, the video decoder maps the predicted image tokens back to the pixel space, at a higher temporal resolution.

https://arxiv.org/pdf/2301.00493.pdf

Argoverse 2: Next Generation Datasets for Self-Driving Perception and Forecasting

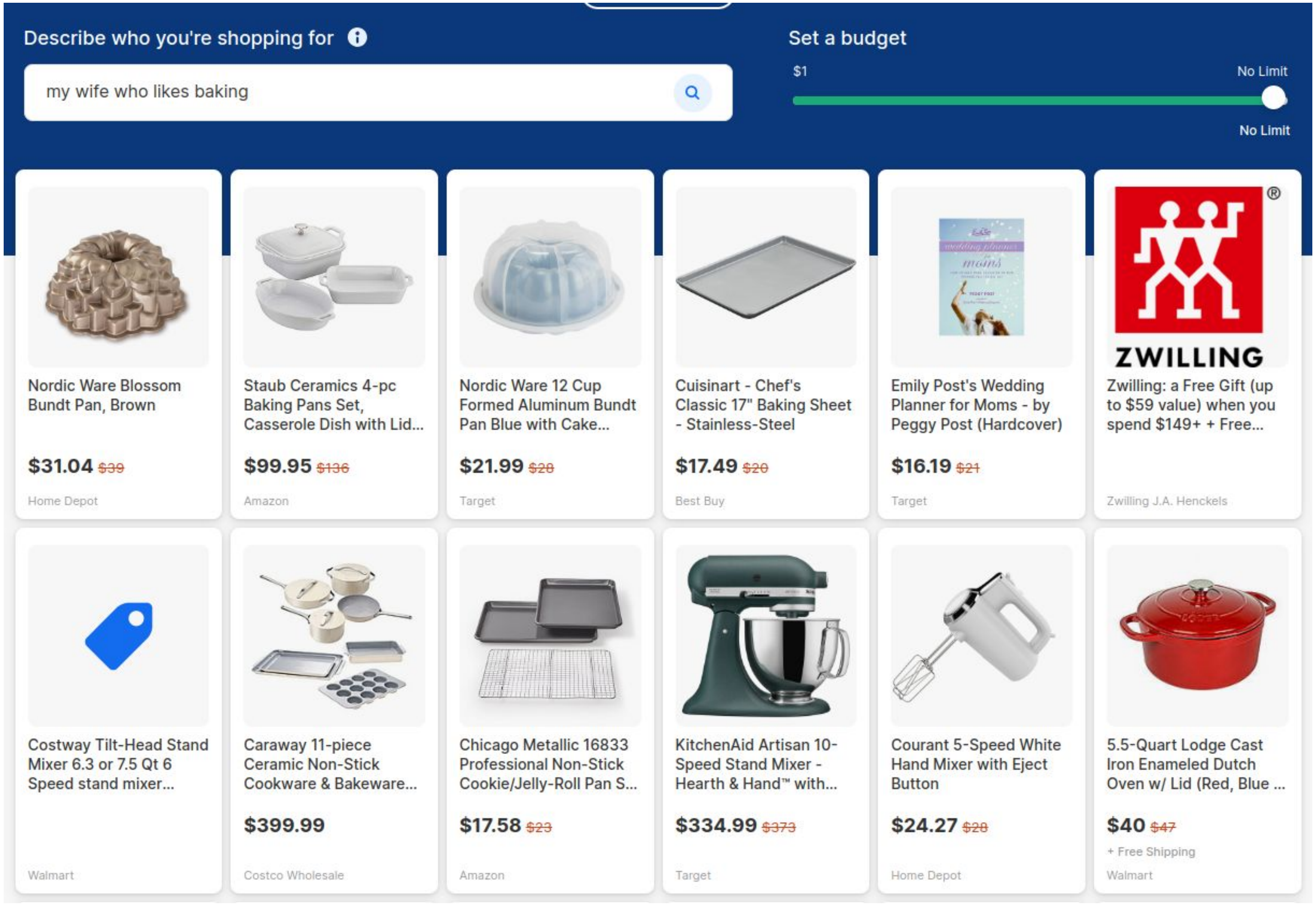
Table 1: Comparison of the Argoverse 2 *Sensor* and *Lidar* datasets with other sensor datasets.

Name	# Scenes	Cities	Lidar?	# Cameras	Stereo	HD Maps?	# Classes	# Evaluated Classes
Argoverse 1 [6]	113	2	✓	7	✓	✓	15	3
KITTI [17]	22	1	✓	2	✓		3	3
nuScenes [4]	1,000	2	✓	6		✓	23	10
ONCE [36]	581	–	✓	7			5	3
Waymo Open [45]	1,150	3	✓	5			4	4
Argoverse 2 Sensor	1,000	6	✓	9	✓	✓	30	26
Argoverse 2 Lidar	20,000	6	✓	-		✓	-	-



LLM search and recommendations

Given a natural query, return items from catalog relevant for given user …



Refining SBERT-based Semantic Search: Ethical Controls and Content Precision

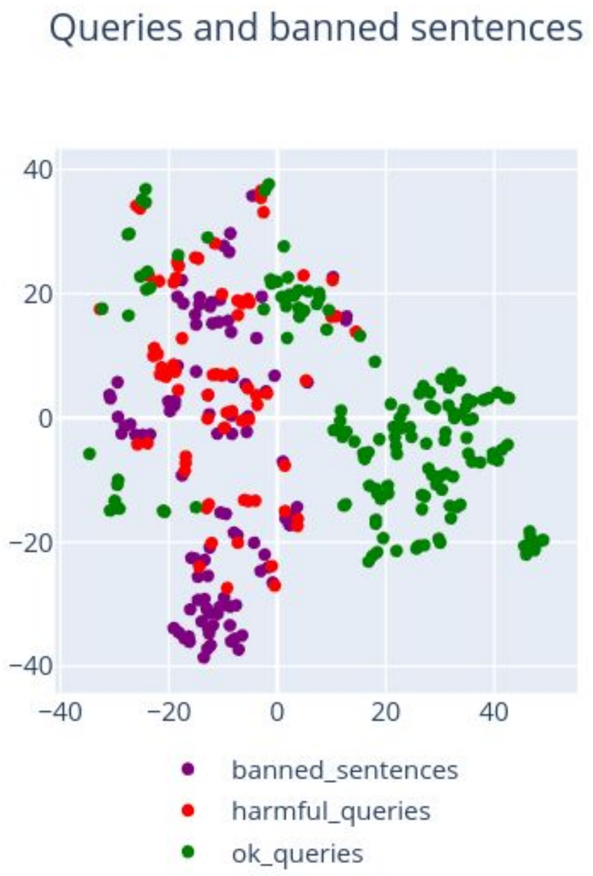
3 Data and model

Data for training the SBERT model were obtained from the client. The various queries used for testing the ethical filters were designed by the author, using several sources such as [2] or [1].

The SBERT model is derived from a pre-trained general MPNetModel from the SentenceTransformers library. It was then trained using the data provided by the customer to better perform on this specific task. It outputs 512-dimensional embedding vectors.

4 Ethical controls and content precision

The idea of ethical control is the follows: Create a list of banned sentences, that would cover harmful topics. If a query similarity to some banned sentence is higher than a given threshold, it should be ignored. To measure the similarity between two queries, it is first necessary to get the embedding vectors. These can be obtained using the trained



Personal assistants

Memory-augmented Dialogue Management for Task-oriented Dialogue Systems

https://arxiv.org/pdf/1805.00150.pdf

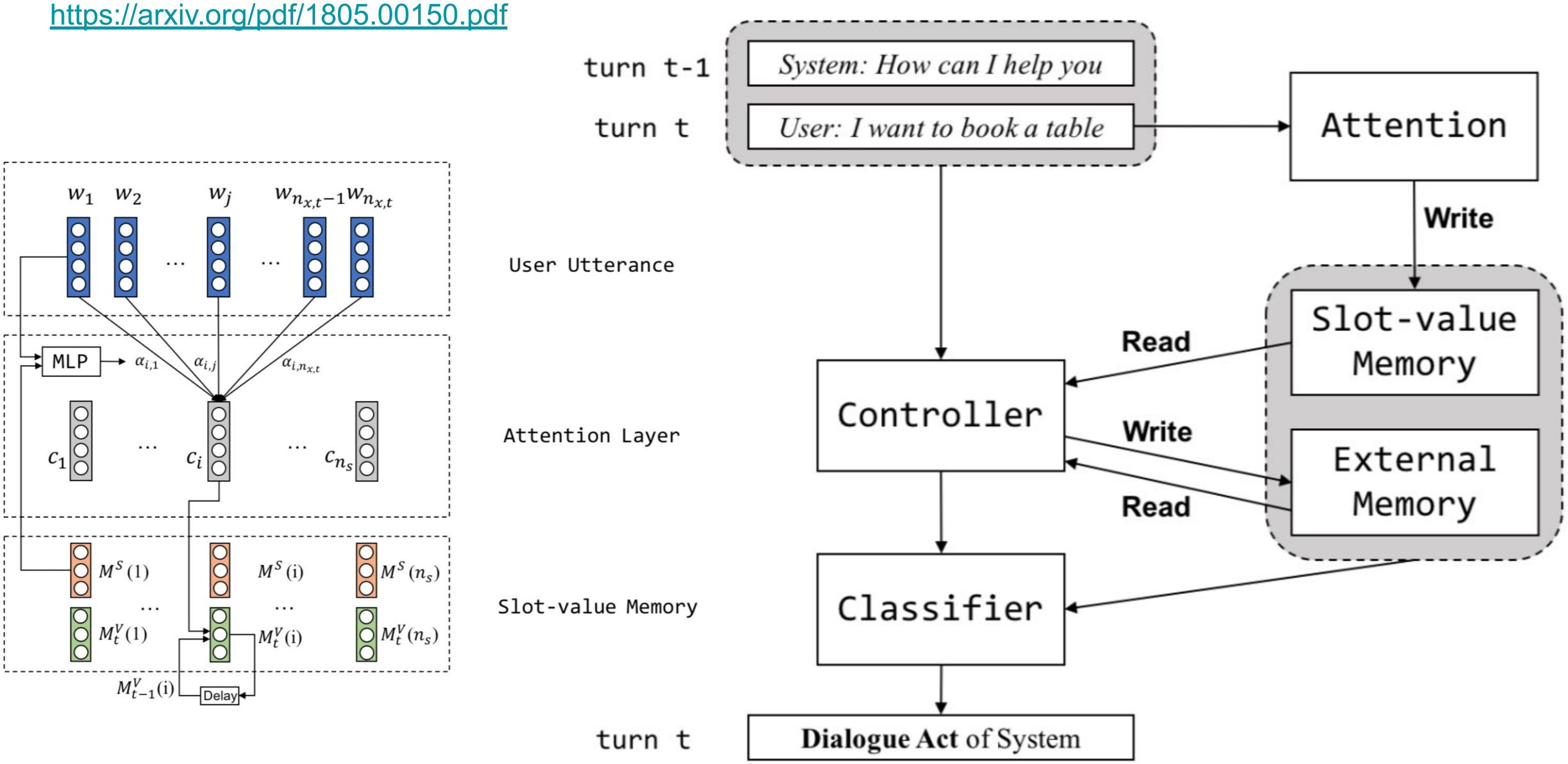


Fig. 3. Memory-augmented Dialogue Management (MAD): At each dialogue turn t, the model takes as input the current user utterance and the previous system response, and predicts the next dialogue act. The slot-value memory is updated with an attentive read of the user utterance by a slot-level attention mechanism while the external memory is read and updated by the controller. The memory controller along with the two memory modules will predict the next dialogue act of the system by a classifier.

https://arxiv.org/pdf/2309.11696.pdf

Memory-Augmented LLM Personalization with Short- and Long-Term Memory Coordination

Kai Zhang[1]**, Fubang Zhao**[2]**, Yangyang Kang**[2]**, Xiaozhong Liu**[1]

[1]Worcester Polytechnic Institute
[2]Damo Academy
kzhang8@wpi.edu, fubang.zfb@alibaba-inc.com, yangyang.kangyy@alibaba-inc.com, xliu14@wpi.edu

Abstract

Large Language Models (LLMs), such as GPT3.5, have exhibited remarkable proficiency in comprehending and generating natural language. However, their unpersonalized generation paradigm may result in suboptimal user-specific outcomes. Typically, users converse differently based on their knowledge and preferences. This necessitates the task of enhancing user-oriented LLM which remains unexplored. While one can fully train an LLM for this objective, the resource consumption is unaffordable. Prior research has explored memory-based methods to store and retrieve knowledge to enhance generation without retraining for new queries. However, we contend that a mere memory module is inadequate to comprehend a user's preference, and fully training an LLM can be excessively costly. In this study, we propose a novel computational bionic memory mechanism, equipped with a parameter-efficient fine-tuning schema, to personalize LLMs. Our extensive experimental re-

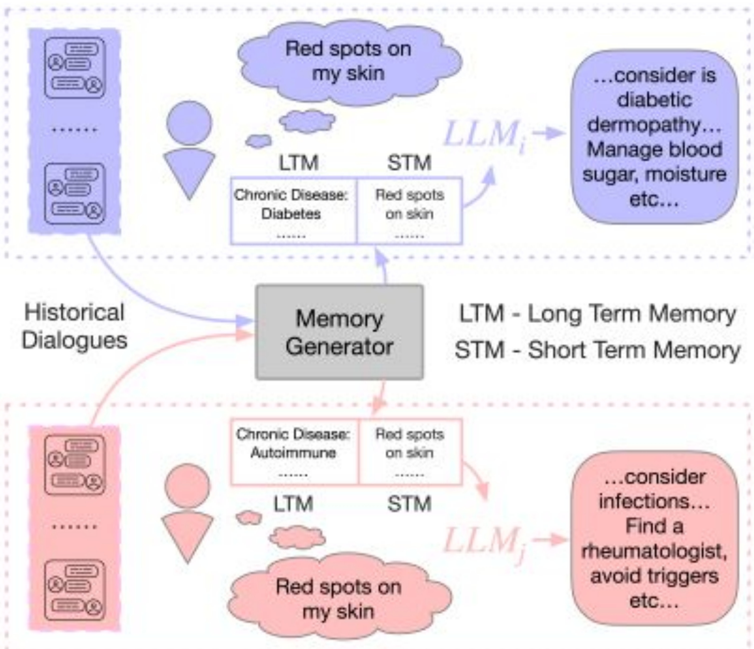


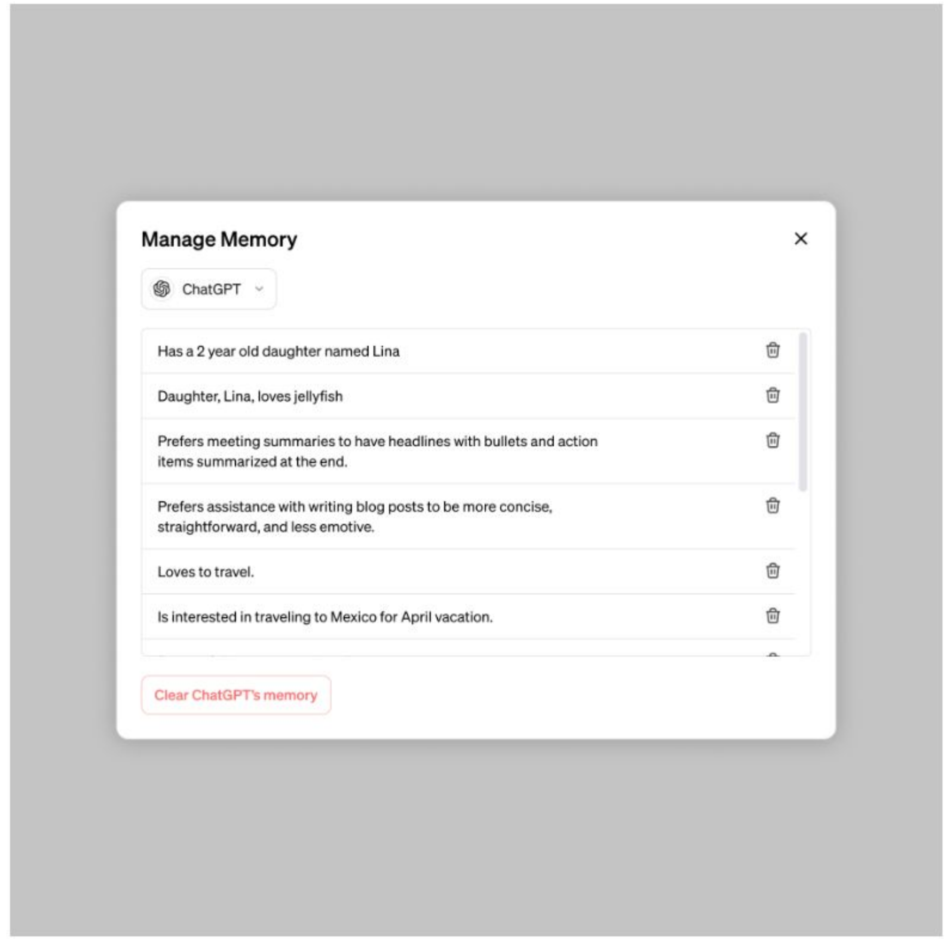
Figure 1: Personalized responses for different users in terms

AI Assistants getting personal

Memory and new controls for ChatGPT

We're testing the ability for ChatGPT to remember things you discuss to make future chats more helpful. You're in control of ChatGPT's memory.

February 13, 2024



https://openai.com/blog/memory-and-new-controls-for-chatgpt

AI4kids (non profit) personalized learning app

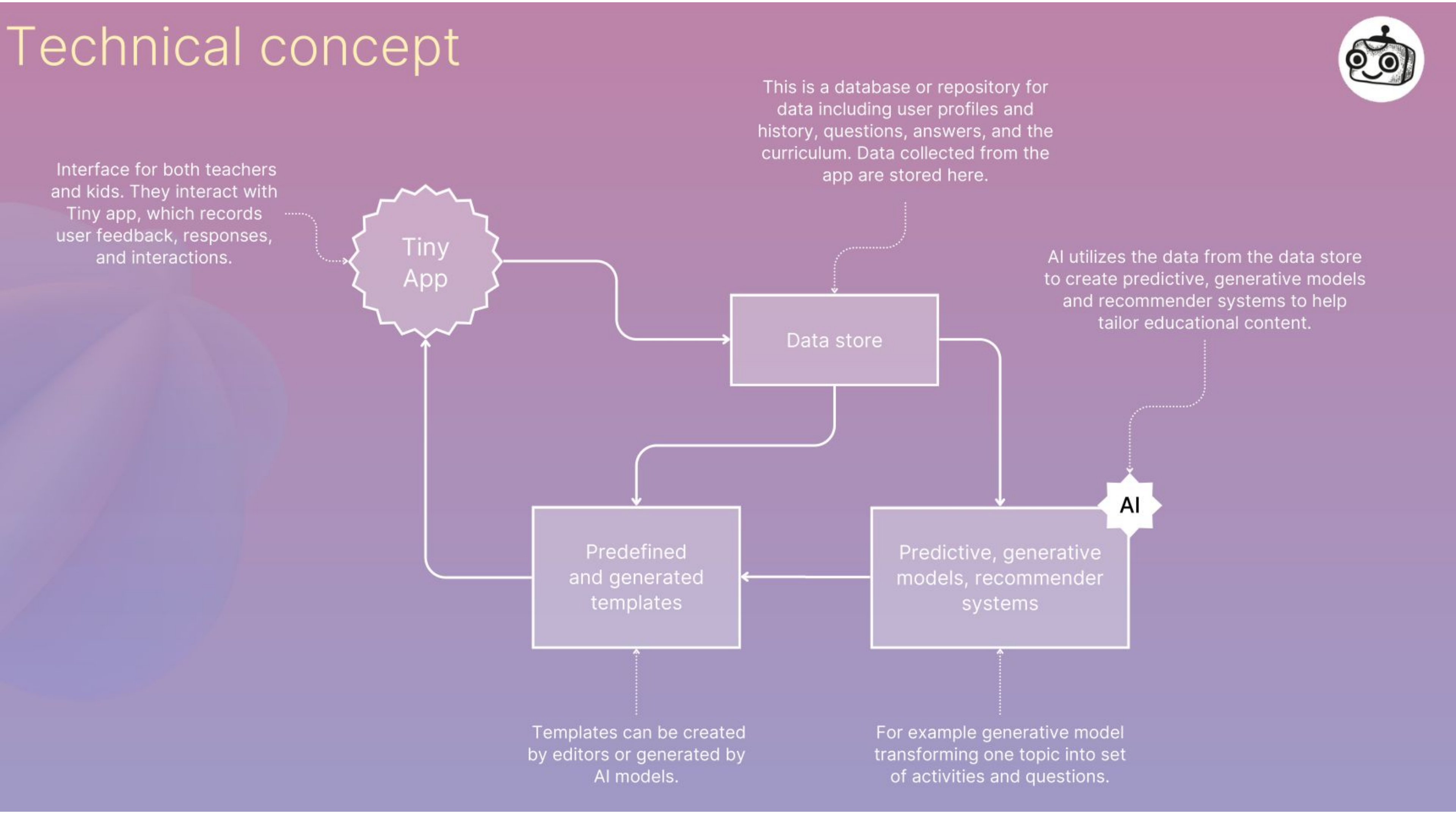
Supported by Google.org Impact Challenge: Tech for Social Good.

- LLMs as personal tutors?
- LLMs evaluating teaching progress?
- LLMs generating learning tasks for children?
- How to help both kids and teachers?
- How to make AI safe for kids?

First results: https://github.com/slavivo/open_evaluation

Who is interested? Let me know! I am looking for advisory board member with experience in AI safety.

Integrating AI Assistants into Classroom Learning



MVP—basic roadmap 2024

	Research in schools							
	April	May	June	July	August	September	November	December
Student part	**POC prototype (dumb)** • manual preparation of exercise for the first part of curriculum • based on existing tools (questionnaires etc.) • data collection • custom build POC using low code **Knowledge validation** • how to generate questions? AI vs. humans		**Prototype testing** • in schools **MVP design (student's part)** • pre-final stage	**Reflection & self-evaluation** • reflective conversation design • questions to support self-evaluation		**Reflection & self-evaluation part testing**	**Putting all parts together**	**FINAL MVP prototype**
Teacher part				**POC prototype (dumb)** • similar to student part	**MVP design (teacher's part)**	**Prototype testing** • similar to student part		
Research POC	**AI research POC for student part (AI tutor)** • research will continue throughout the duration of the project		**AI research POC for teachers part (AI teaching assistant)** • research will continue throughout the duration of the project					

POC—research proof of concept focusing on AI technologies that enable the MVP to be scalable and applicable in school learning environments; in particular, we need to take the first steps towards an AI tutor for kids and an AI teaching assistant for teachers.
Possible research topics: generating good questions for students given an educational material, task or activity to assess their understanding; evaluating students' answers and extracting their learning progress, giving appropriate feedback, presenting learning progress of individual kids to the teacher, notifying the teacher that some kid needs attention

Other interesting papers

ChatGPT: https://arxiv.org/pdf/2206.02336.pdf
https://arxiv.org/pdf/2206.07699.pdf

https://arxiv.org/pdf/2201.11903.pdf

PALM: https://arxiv.org/pdf/2204.02311.pdf

GATO: https://arxiv.org/pdf/2205.06175.pdf?fs=e&s=cl

Model free RL: https://arxiv.org/pdf/2208.07860.pdf

Diffusion world models: https://arxiv.org/pdf/2402.03570.pdf

Mamba: https://openreview.net/forum?id=AL1fq05o7H

CodeIt: https://arxiv.org/pdf/2402.04858.pdf

Algorithm discovery:
https://www.nature.com/articles/s41586-023-06887-8

https://www.jmlr.org/papers/volume24/21-0449/21-0449.pdf

Thanks for attention

https://open.substack.com/pub/pavelkordik/p/latest-ai-developments-explained

kordik@gmail.com (meta inbox)

kordikp@fit.cvut.cz

pavel.kordik@recombee.com

pavel@experts.ai

pavel@prg.ai

pavel.kordik@aidetem.cz