The Long Road to Agentic AI

Lee Wee Sun
Professor, NUS School of Computing
Deputy Director, NUS Artificial Intelligence Institute (NAII)

NUS National University of Singapore | Artificial Intelligence Institute

Vision: To Be a Global Node of AI Thought Leadership

Established: 25 March 2024

A University-level Platform Institute for AI research

- Conduct fundamental and applied research in AI
- Bring together faculty and domain experts involved in AI research
- Engage with industry and university partners for research collaborations
- Provide hands-on AI research and AI entrepreneurship opportunities for the NUS student community

NUS National University of Singapore | Artificial Intelligence Institute

Vision To be a Global Node of AI Thought Leadership

Mission To establish a novel AI ecosystem that focuses on research excellence that empowers individuals, communities, and businesses, and fosters entrepreneurship, for the public good in Singapore and in the World

Established on 25 March 2024, the NUS Artificial Intelligence Institute (NAII) is a **platform institute** that brings together AI researchers across the whole of NUS. It encompasses **basic and applied research** in AI, along with **research into the societal implications** of AI. The foundational research, combined with deep domain expertise, aims to harness AI for the **public good**.

Foundational AI: Overseen by leading experts from School of Computing, this thrust covers basic research in foundations of AI including theory and systems as well as building statistical models, foundational models, inferencing models, and generative models.

- AI Hardware/Software Systems
- AI Theory
- Reasoning AI
- Resource-efficient AI
- Responsible and Safe AI

Foundational AI Research drives Applications

Applications drive Foundational AI Research

AI+ X: Overseen by leading domain experts from across NUS and AI experts from School of Computing, these domains reflect key strengths of NUS and societal concerns where AI can have a significant impact.

- Education, Humanities, and Social Sciences
- Finance and Economy
- Healthcare and Human Potential
- Logistics, Manufacturing, and Robotics
- Science
- Urban Sustainability

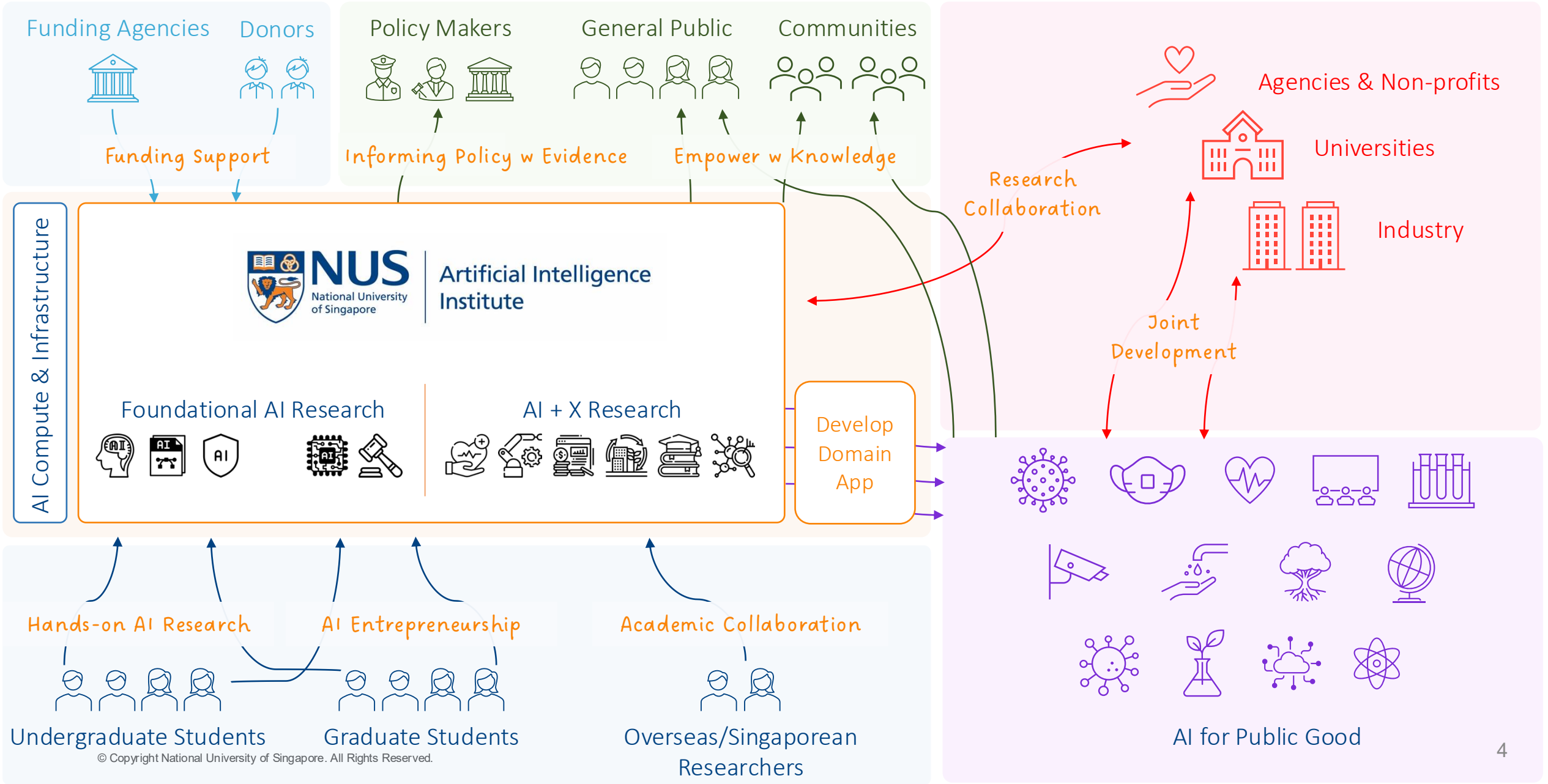
AI Governance and Policy

AI Governance and Policy: Overseen by a prominent Law faculty, this thrust draws on diverse bodies of scholarship to examine the actual and potential impact (positive and negative) of AI.

URL for full member list Over 270 faculty members are actively working in diverse and high-impact areas in AI research across NUS

Vibrant Intellectual Environment

To empower individuals, communities, and businesses



The Long Road to Agentic AI

VOL. LIX. No. 236.] [October, 1950

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

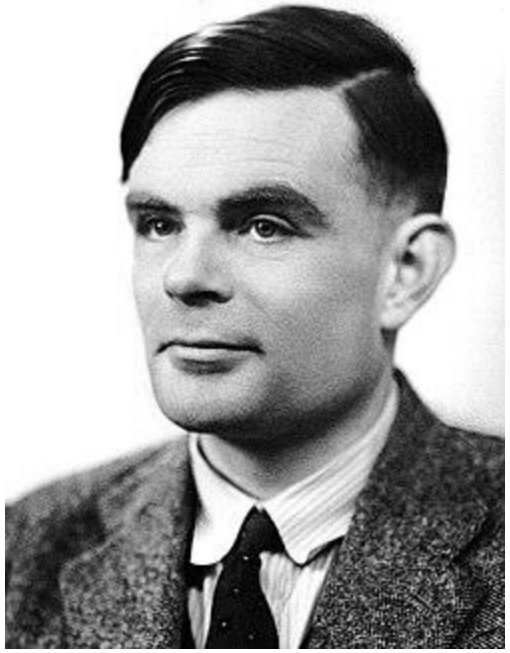
BY A. M. TURING

1. *The Imitation Game.*

I PROPOSE to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning

NUS National University of Singapore

Can Machines Think?

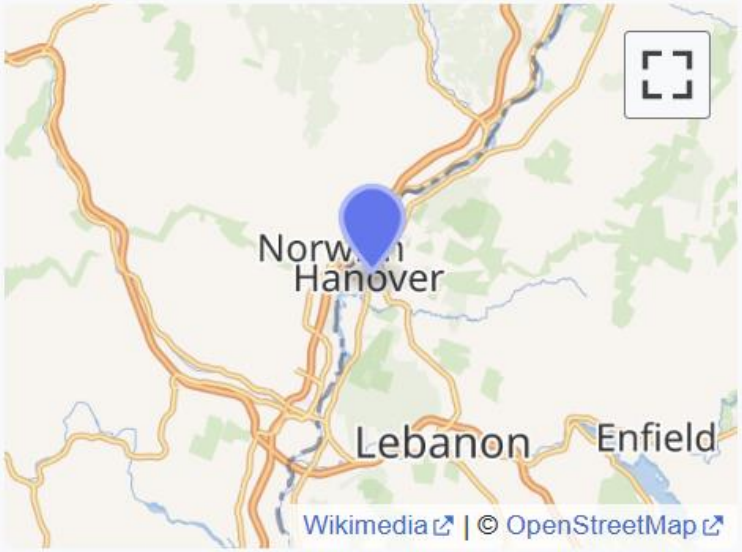


Answer: Check whether machines perform as well as human in conversations

Useful whenever we want to ask whether machines are good enough, e.g. to deploy as agents

Artificial Intelligence

Dartmouth Summer Research Project on Artificial Intelligence



Date	1956
Duration	Eight weeks
Venue	Dartmouth College, Hanover, New Hampshire
Organised by	John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon
Participants	John McCarthy, Marvin Minsky, Nathaniel Rochester, Claude Shannon, and others

We propose that a 2-month, 10-man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.

Image from Wikipedia

AI Winters

From Wikipedia:

There were two major "winters" approximately

1974–1980 and 1987–2000,

and several smaller episodes, including the following:

1974–1980

1987–2000

Image generated by ChatGPT

1966:

failure of machine translation

1969:

criticism of perceptrons (early, single-layer artificial neural networks)

1971–75:

DARPA's frustration with the Speech Understanding Research program at Carnegie Mellon University

1973:

large decrease in AI research in the United Kingdom in response to the Lighthill report

1973–74:

DARPA's cutbacks to academic AI research in general

1987:

collapse of the LISP machine market

1988:

cancellation of new spending on AI by the Strategic Computing Initiative

1990s:

many expert systems were abandoned

1990s:

end of the Fifth Generation computer project's original goals

Deep Blue versus Garry Kasparov

- The ability to play chess was initially considered as the hallmark of intelligence.
- Deep Blue defeated Kasparov in 1997 … large amounts of **computation** is enough for this

Deep Blue vs. Kasparov

Deep Blue
IBM chess computer

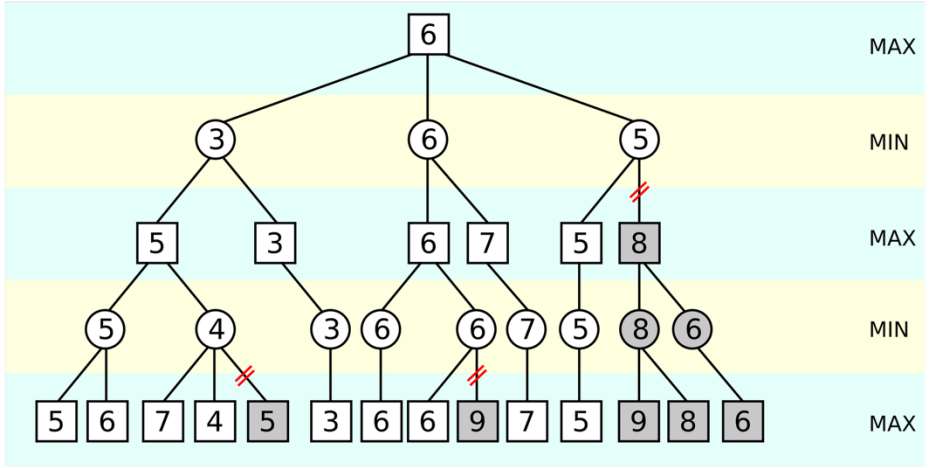
Garry Kasparov
World Chess Champion

First match

- February 10–17, 1996: held in Philadelphia, Pennsylvania
- Result: **Kasparov**–Deep Blue (4–2)
- Record set: First computer program to defeat a world champion in a *classical game* under tournament regulations

Second match (rematch)

- May 3–11, 1997: held in New York City, New York
- Result: **Deep Blue**–Kasparov (3½–2½)
- Record set: First computer program to defeat a world champion in a *match* under tournament regulations



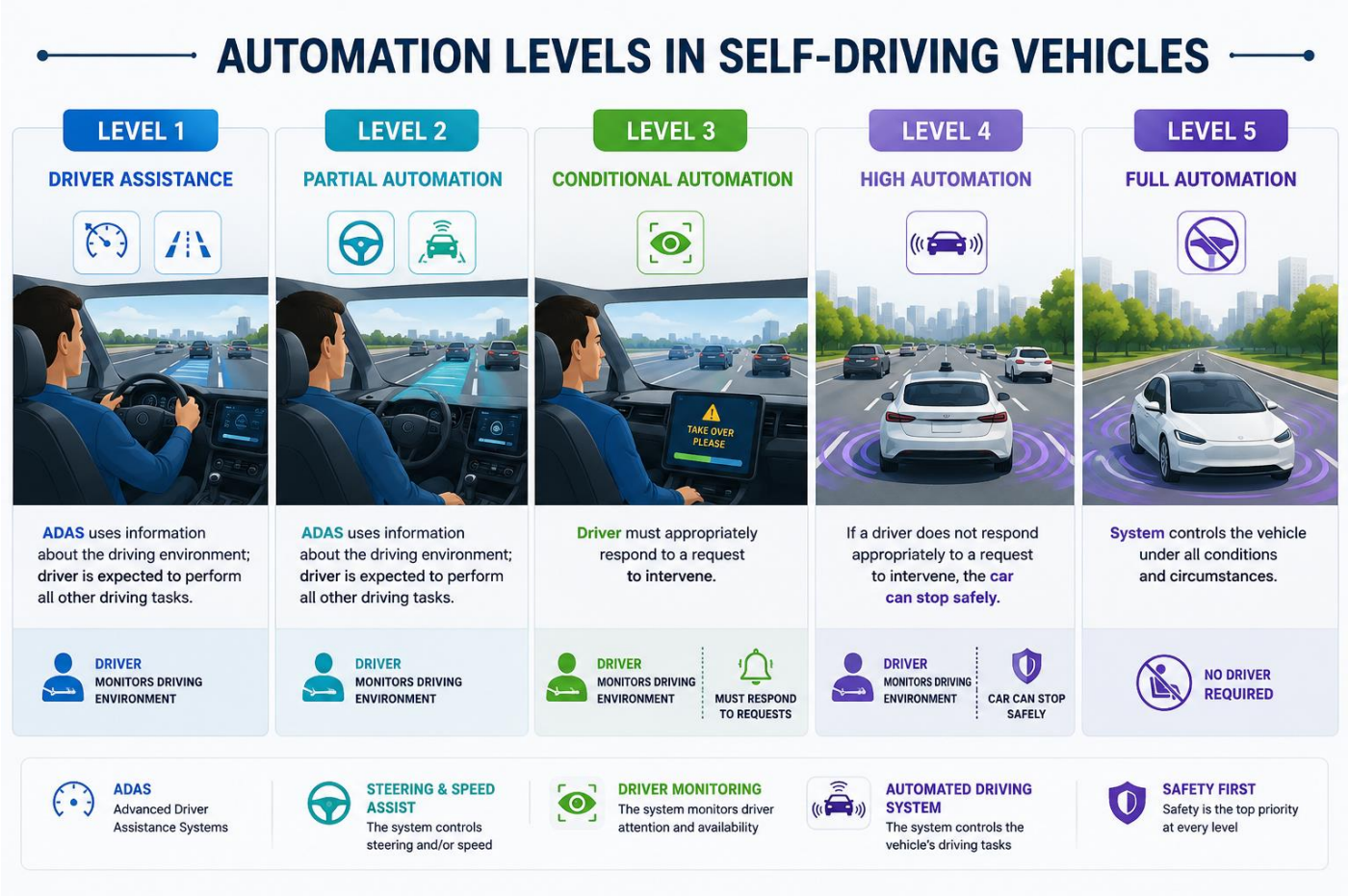
Images from Wikipedia

DARPA Grand Challenge

- In 2005, five vehicles successfully drove 212 km across the desert in the DARPA Grand Challenge
  - Race won by Stanley from Stanford (shown on left)

Image from Wikipedia

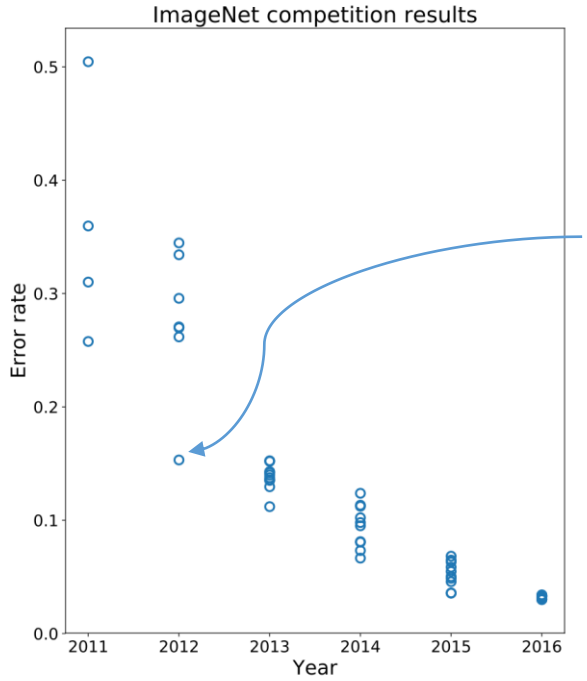
- After **about 20 years**, we are starting to see self-driving taxis in our cities (around Level 4)
  - Substantial amount of mapping and perception data collected
  - Careful engineering



Description from Wikipedia, image by ChatGPT

Neural Networks Makes a Comeback …

- By 2010s, enough **data and compute** became available for neural networks to outperform other methods on many tasks
- Rebranded as deep learning and actually useful to go **deep**
  - Sequential in addition to parallel computation



AlexNet outperformed other methods in the 2012 ImageNet competition by more than 10%

Image from Wikipedia

- In 2017, the transformer architecture was introduced and is currently the dominant deep learning architecture

Attention Is All You Need

Ashish Vaswani* Google Brain avaswani@google.com
Noam Shazeer* Google Brain noam@google.com
Niki Parmar* Google Research nikip@google.com
Jakob Uszkoreit* Google Research usz@google.com
Llion Jones* Google Research llion@google.com
Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu
Łukasz Kaiser* Google Brain lukaszkaiser@google.c
Illia Polosukhin* ‡ illia.polosukhin@gmail.com

Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.8 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.

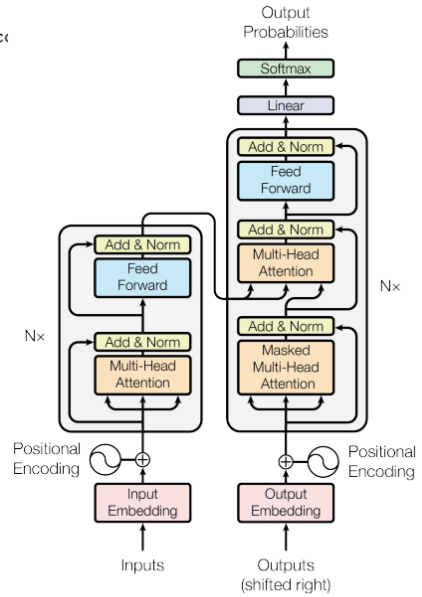


Figure 1: The Transformer - model architecture.

AlphaGo versus Lee Sedol

versus

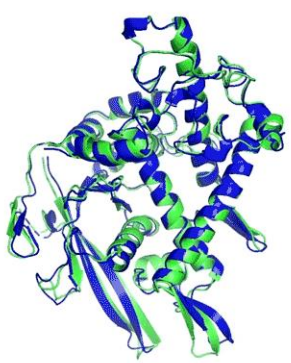
Go is a much harder game and AlphaGo, which combines **deep learning** with **Monte Carlo tree search**, finally defeated the top players around 2016, notably Lee Sedol

Large amounts of data from self-play

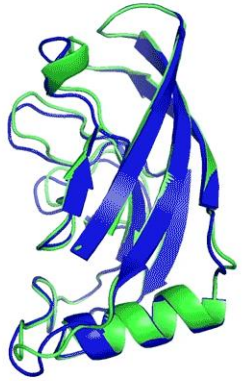
Images from Wikipedia

AlphaFold

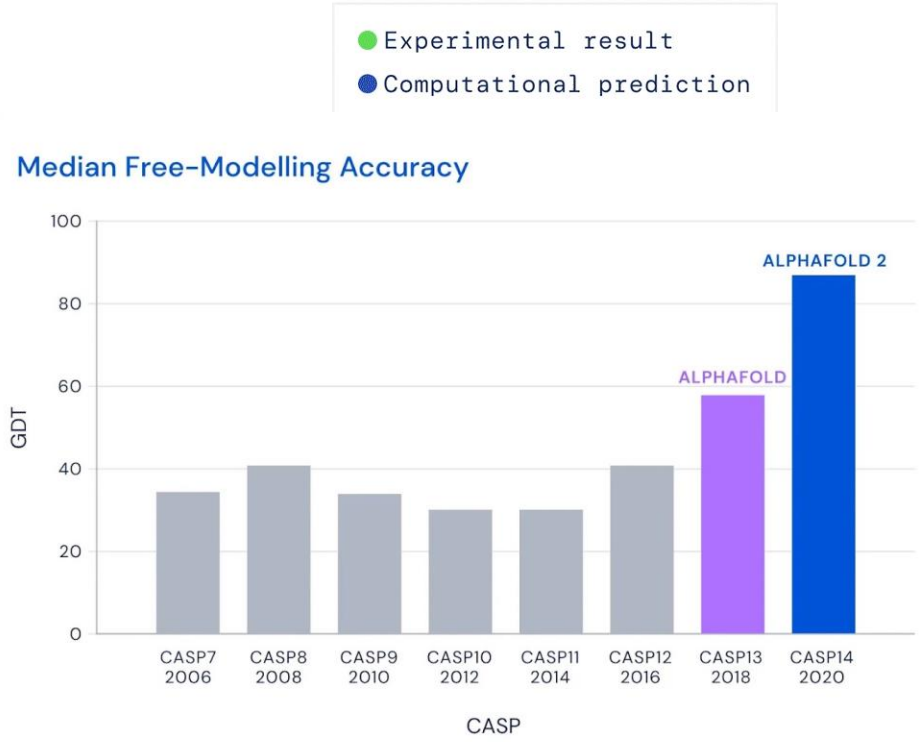
- Fifty-year biology **grand challenge**: Determine the structure of protein from its amino acid sequence
- AlphaFold obtained breakthrough performances at Critical Assessment of protein Structure Prediction (CASP)
  - Trained on ~170,000 protein structures
  - Attention-based neural network
- Led to Nobel Prize for Demis Hassabis and John Jumper



T1037 / 6vr4
90.7 GDT
(RNA polymerase domain)



T1049 / 6y4f
93.3 GDT
(adhesin tip)



Images from https://deepmind.google/blog/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology/

Foundation Models

Traditional AI designed to work in restricted domains

Go World

Foundation models trained on whole WWW

Open World

Foundation Models Unlock Capabilities to Handle Language Ambiguity through Common Sense Reasoning

- **Ambiguity** was a central problem in **natural language** processing
- Commonsense reasoning refers to human-like ability to make reasonable assumptions in ordinary situations typically encountered by humans.

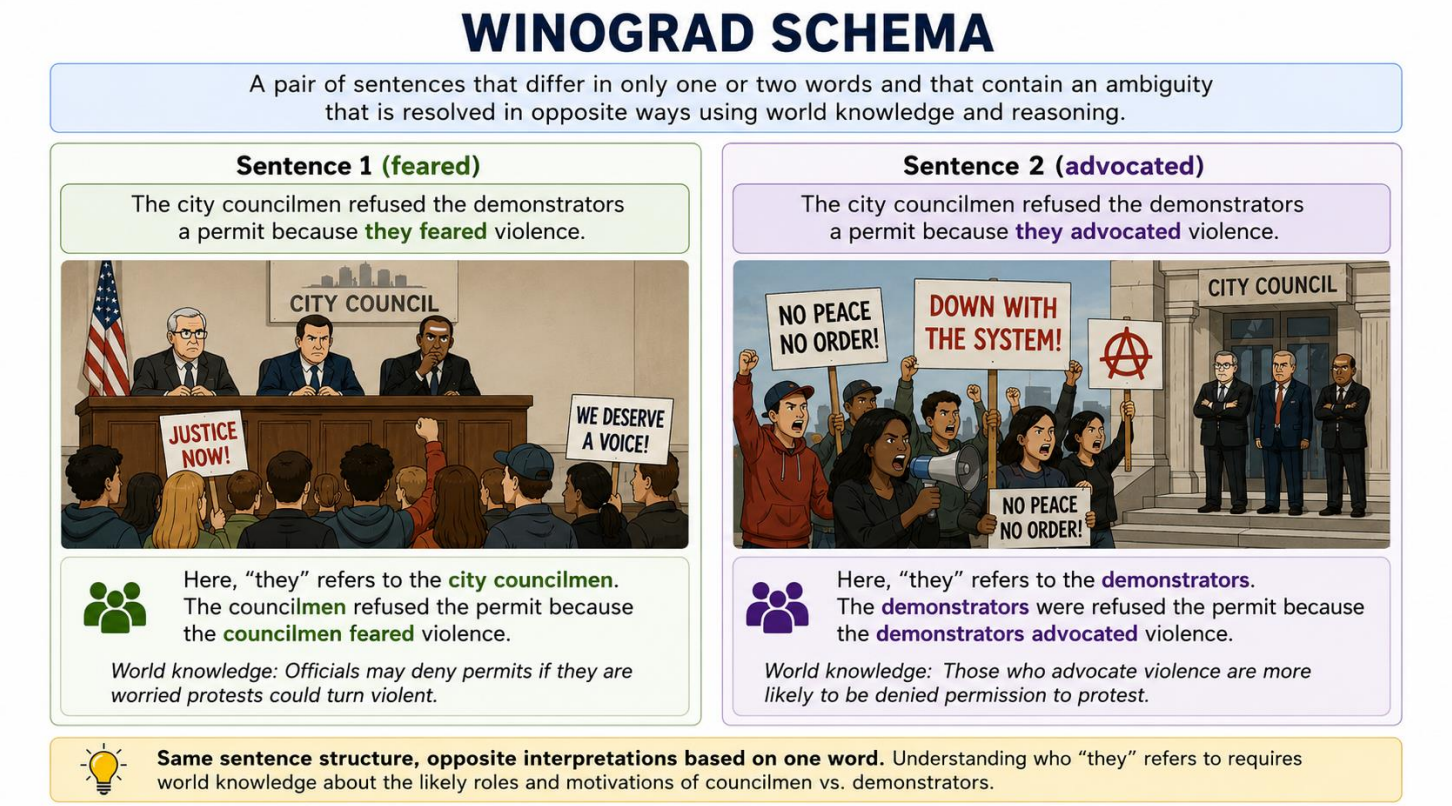


Image generated by ChatGPT

- By 2019, LLMs achieve higher than 90% accuracy on Winograd Schema benchmarks

ChatGPT

Killer app: Chatbot that is able to

- Answer questions
- Compose emails
- Make travel plans
- Summarize text
- Chatbot
- Write code
- Prove theorems
- …

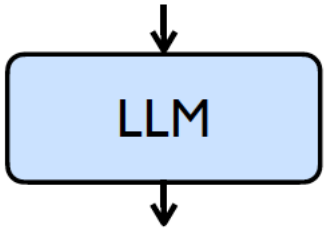
ChatGPT

Developer	OpenAI
Release	November 30, 2022 (3 years ago)[1]
Stable release	June 18, 2026 (34 days ago)[2]
Engine	GPT-5.6
Platform	Cloud computing platforms
Available in	59 languages[3]
Type	Chatbot Large language model Generative pre-trained transformer
License	Proprietary service
Website	chatgpt.com

Exploiting Common Sense World Knowledge of Foundation Models

Large Language Models as Commonsense Knowledge for Large-Scale Task Planning, Zirui Zhao, Wee Sun Lee and David Hsu, NeurIPS 2023

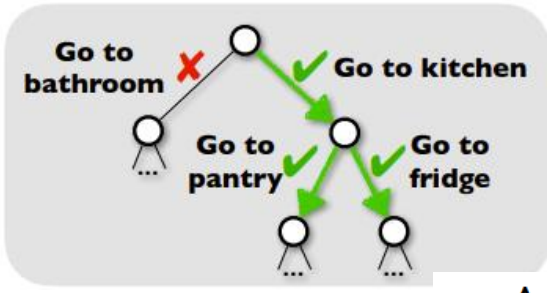
I want to have some fruit please



1. Go to kitchen; 2. Open fridge; 3 …

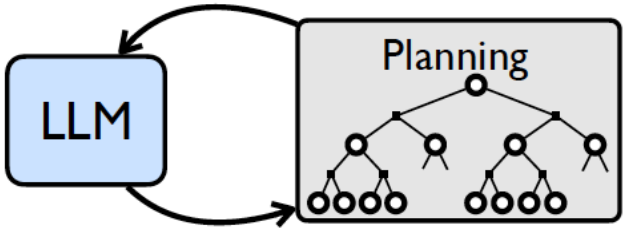
LLM as a policy

E.g., SayCan; Inner Monologue; Voyager; …



Which method should we use?

Action: Go to kitchen

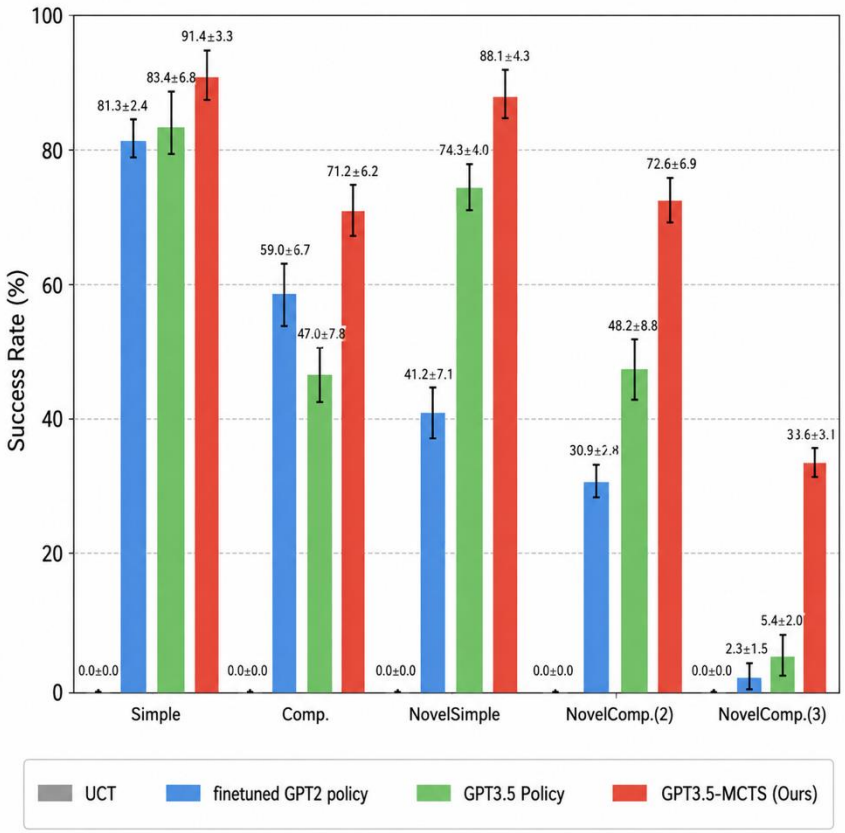


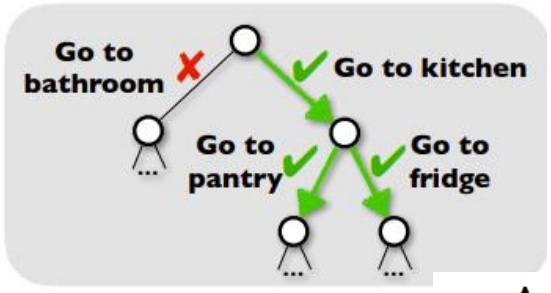
Observation: fridge in the kitchen;
Next state: you at kitchen, …;
Reward: …

LLM as a world model

Exploiting Common Sense World Knowledge of Foundation Models

Large Language Models as Commonsense Knowledge for Large-Scale Task Planning, Zirui Zhao, Wee Sun Lee and David Hsu, NeurIPS 2023

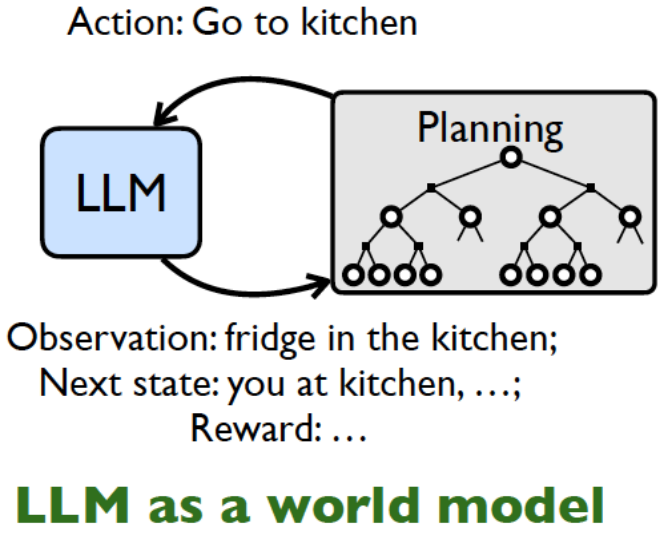




Which method should we use?

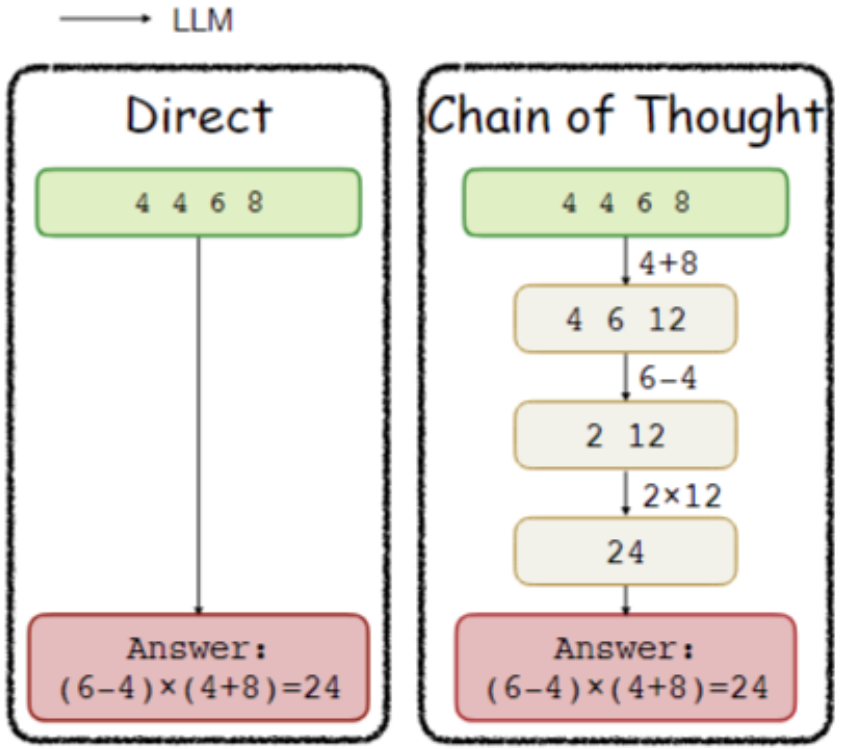
Use both!

- **Build a tree search harness:**
  - **Commonsense world model for observations: Likely to see fridge, pantry, etc. in the kitchen**
  - **Commonsense policy to limit actions and reduce branching: Go to kitchen but not bathroom for fruit**



Reasoning Models

- Progress appears to slow in 2024
- Then OpenAI released reasoning models that spend time thinking by generating a Chain of Thought (CoT) before answering
- For some problems, **reasoning before answering** substantially improves performance
  - Reasoning before taking action usually done in agentic AI systems
- Can be substantially more expensive computationally at inference time



Recipe for Reasoning Models

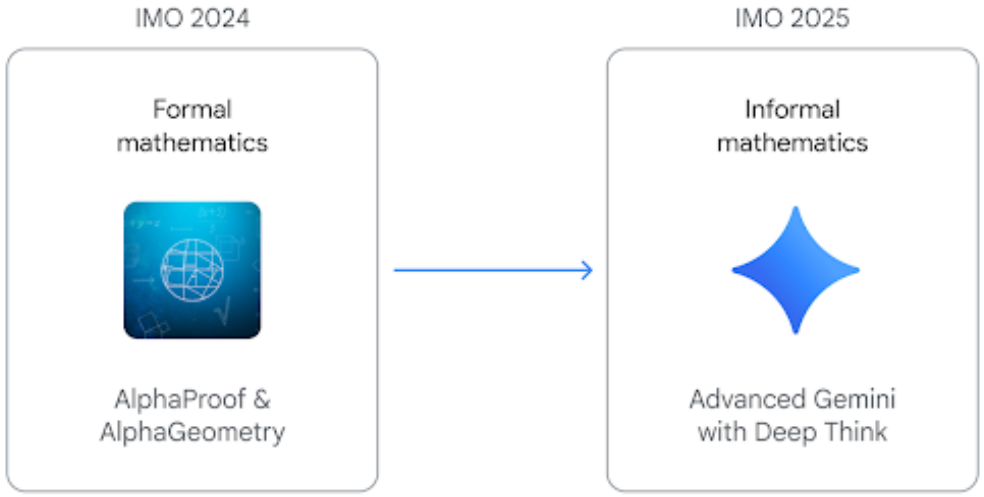


Figure from https://deepmind.google/discover/blog/advanced-version-of-gemini-with-deep-think-officially-achieves-gold-medal-standard-at-the-international-mathematical-olympiad/

- AlphaGeometry and AlphaProof translate IMO problems into formal problems then use search methods for Silver medal in 2024.

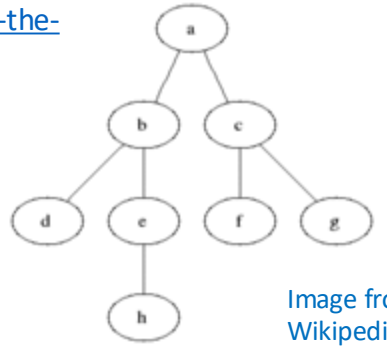


Image from Wikipedia

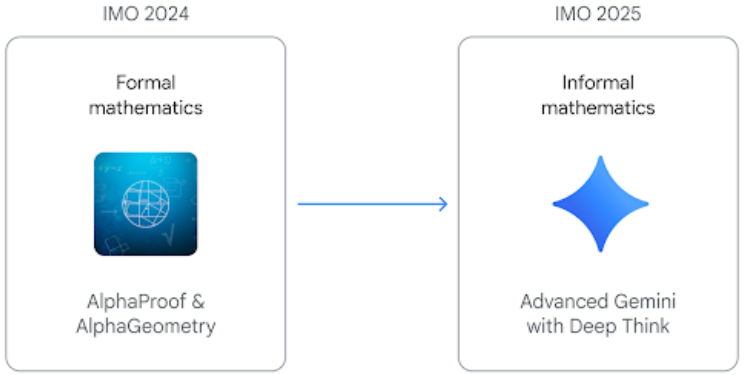
Reinforcement learning fine-tuning strategy for LLM reasoning

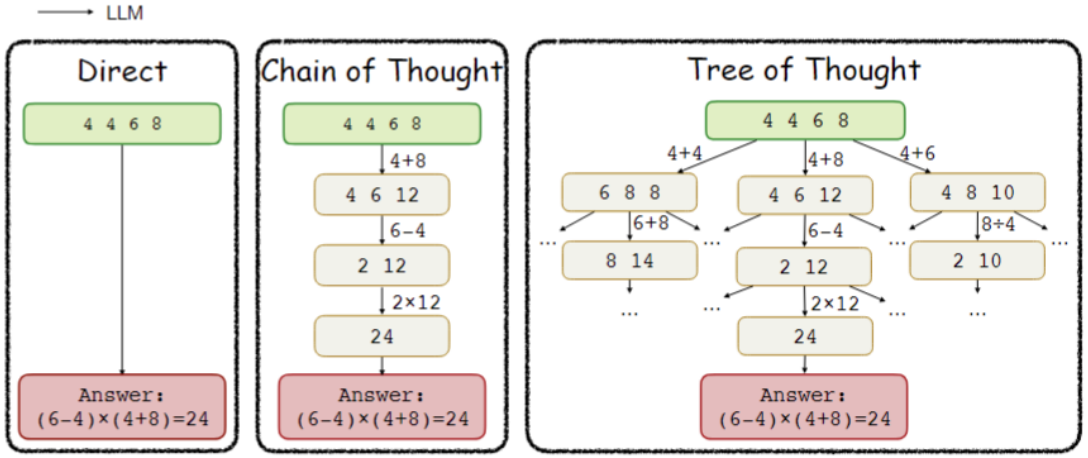
1. Collect dataset of reasoning problems.
2. Annotate chain of thought (CoT) for dataset.
   - Verbalize out execution of the reasoning method step-by-step in CoT.
3. Do supervised fine-tuning (SFT).
4. Refine with reinforcement learning (RL).

- Google's Gemini Deep Think achieved gold medal at IMO 2025 using reasoning model in natural language.

Improving Search with Reasoning Models

LinTree: Improving LLM Reasoning with Explicitly Structured Search Histories, Liwei Kang, Yee Whye Teh, Wee Sun Lee, arXiv:2605.31492.



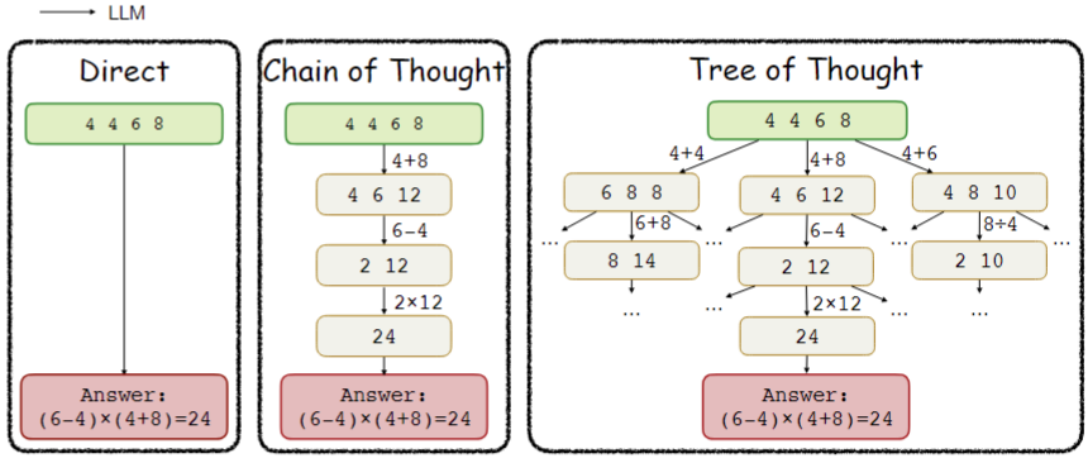


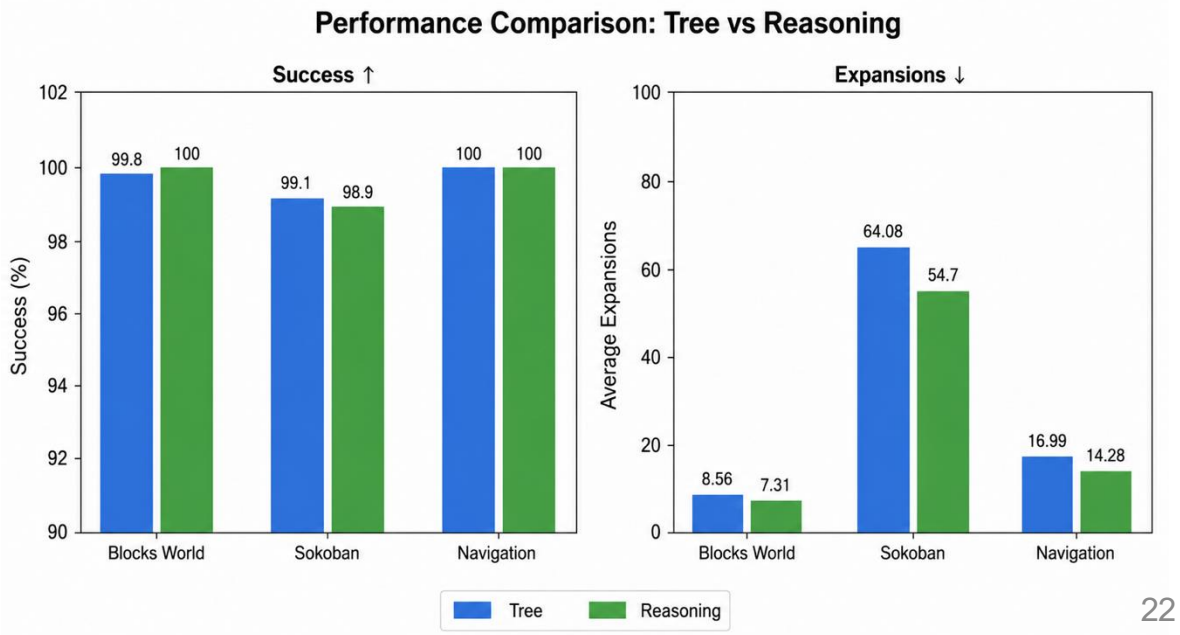
- AI performance in International Mathematical Olympiad
  - Silver in 2024 by translating natural language problem to formal mathematics and solving with tool (likely by doing search)
  - Gold in 2025 using reasoning model in natural language.
- Can we use reasoning model techniques to improve problem solving **within a tool**?
  - E.g. for tree search, can **linearize the tree into a chain** and use reasoning model to generate the chain

Improving Search with Reasoning Models

LinTree: Improving LLM Reasoning with Explicitly Structured Search Histories, Liwei Kang, Yee Whye Teh, Wee Sun Lee, arXiv:2605.31492.

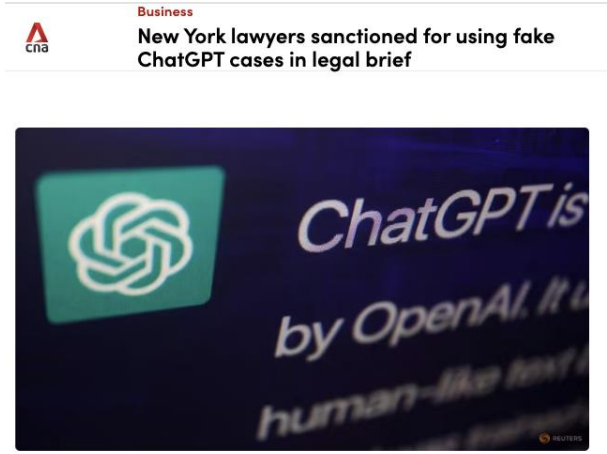
- Reasoning models generates the chain of thought instead of using an external scaffolding to do tree search
  - But can always **linearize the tree into a chain** and use reasoning model
- Is there advantage in using chain instead of tree for more formal search problems? Yes!
  - In tree search, next action conditioned only on current node
  - In CoT, next action conditioned on everything previously generated – can avoid neighbourhoods of states previously generated and deemed not promising
- Reasoning model: around same success as tree search, at around 85% of effort
  - Caveat: need to encode the tree structure in CoT





Hallucination and Non-Hackable Confidence Reward

On the effectiveness of reward functions in reinforcement learning for confidence calibration of large language models, Chee Heng Tan, Zhuoyi Lin, Mehul Motani, Wee Sun Lee, arXiv:2607.04332.



A response by ChatGPT, an AI chatbot developed by OpenAI, is seen on its website in this illustration picture taken February 9, 2023. REUTERS/Florence Lo/Illustration

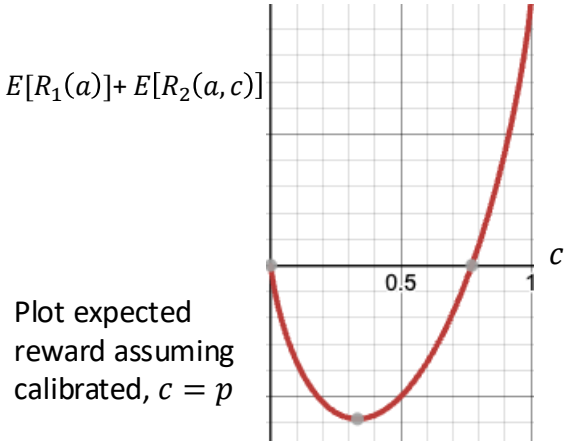
Business

New York lawyers sanctioned for using fake ChatGPT cases in legal brief

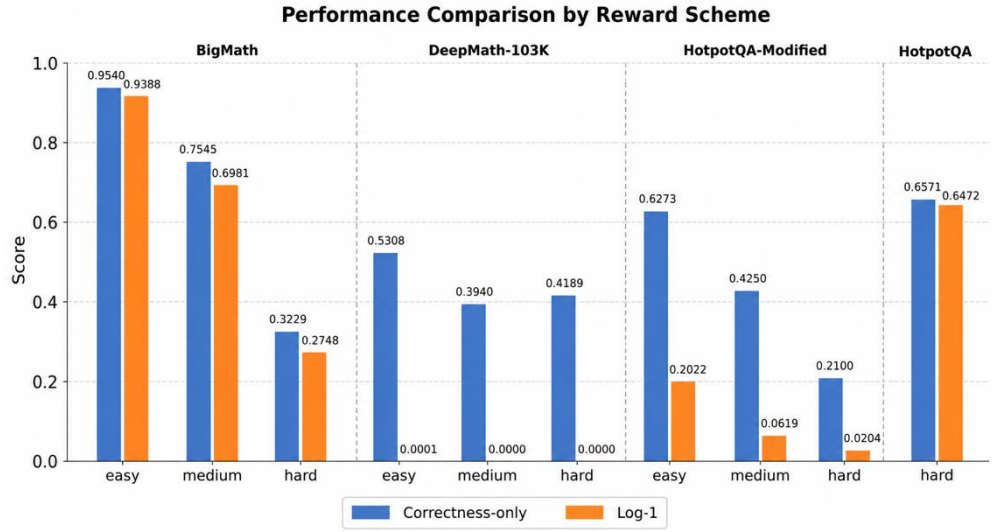
NEW YORK : A U.S. judge on Thursday imposed sanctions on two New York lawyers who submitted a legal brief that included six fictitious case citations generated by an artificial intelligence chatbot, ChatGPT.

U.S. District Judge P. Kevin Castel in Manhattan ordered lawyers Steven Schwartz, Peter LoDuca and their law firm Levidow, Levidow & Oberman to pay a $5,000 fine in total.

- Hallucination remains a problem
- One approach is to ask the model to generate its confidence c alongside the answer a.
  - Goal: Use reinforcement learning. Natural reward components include
    - Indicator function of correctness: $R_1(a) = IsCorrect(a)$ with $E[R_1(a)] = p$
    - Log likelihood, which is known to be **calibrated** when expected value maximized:
      - $R_2(a,c) = isCorrect(a)\log c + \big(1 - isCorrect(a)\big)\log(1-c))$ with
      - $E[R_2(a,c)] = p\log c + (1-p)\log(1-c)$ maximized at $c = p$
  - Natural method is to sum two reward functions: $Log-1(a,c) = R_1(a) + R_2(a,c)$



Plot expected reward assuming calibrated, $c = p$



Hallucination and Non-Hackable Confidence Reward

On the effectiveness of reward functions in reinforcement learning for confidence calibration of large language models, Chee Heng Tan, Zhuoyi Lin, Mehul Motani, Wee Sun Lee, arXiv:2607.04332.

Non-Hackable Confidence Reward Scheme

1. **Interpretability**: $f(c)$ increasing on $(0,1)$ and $g(c)$ decreasing on $(0,1)$
2. **Proper Scoring**: Let

   $R(c,p) = pf(c) + (1-p)g(c).$

   For all $p \in (0,1)$,

   $p = argmax_c R(c,p)$
3. **Best Effort**: Let $R_{max}(c) = R(c,c).$

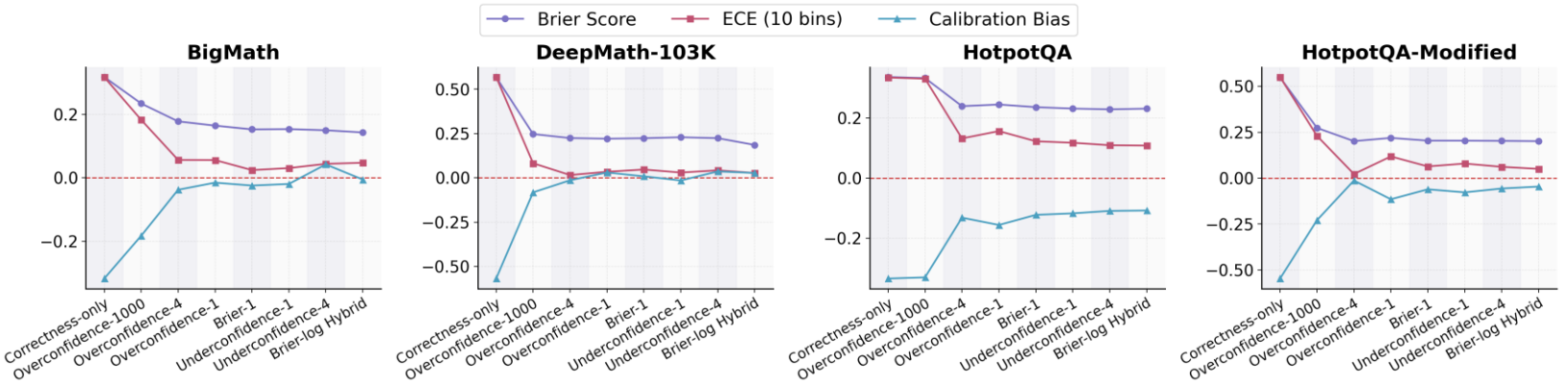
   Then $R_{max}(c)$ is increasing on $(0,1)$

Theorem: Non-Hackable Confidence over (a,b) if

1. Exists $h(c)$ such that

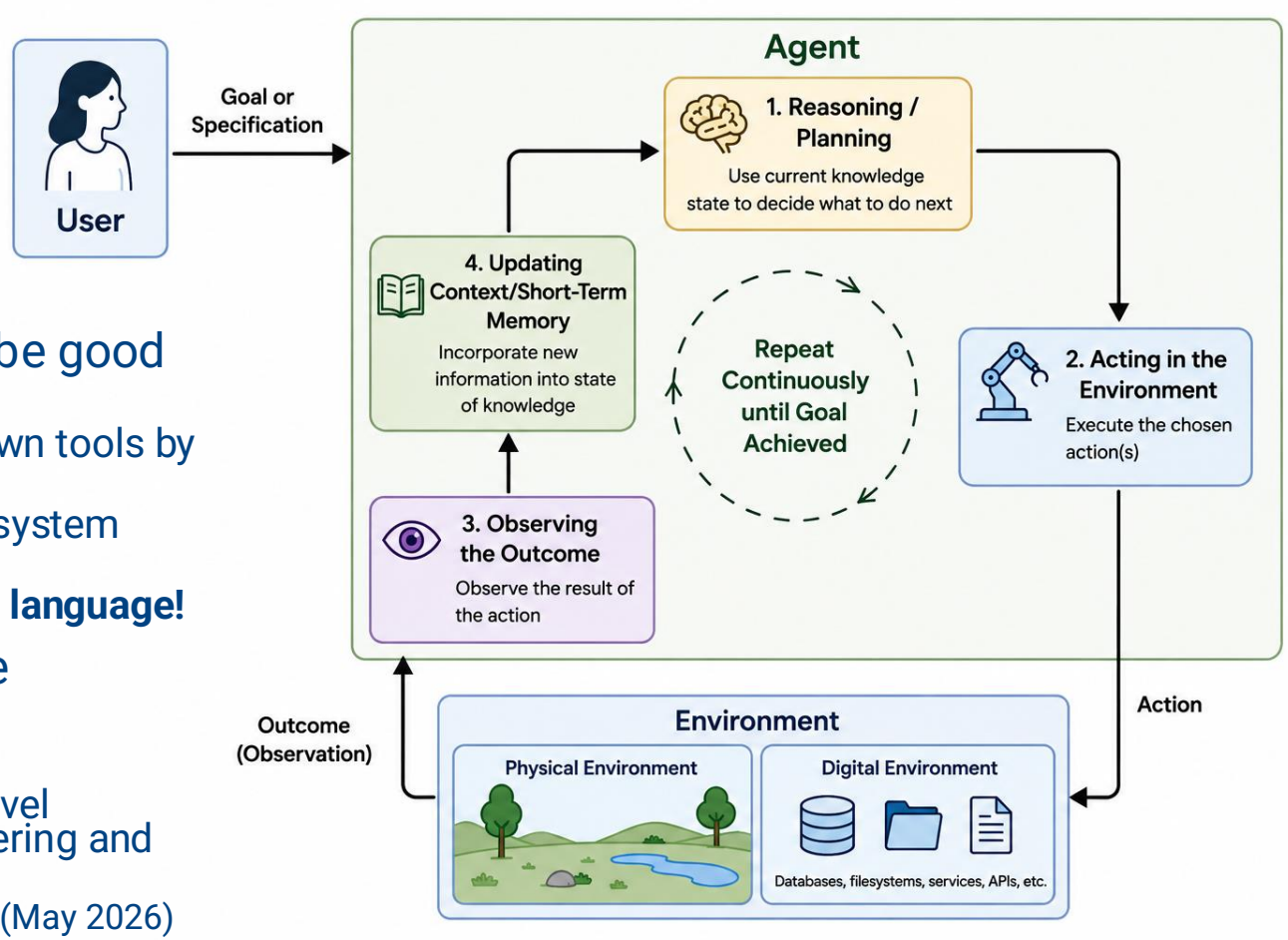
   $h(c) = \frac{f'(c)}{c-1} = \frac{g'(c)}{c}$
2. $h(c) < 0$
3. $f(a^+) \geq g(a^+)$

Can construct many schemes with different trade-offs when resources are limited



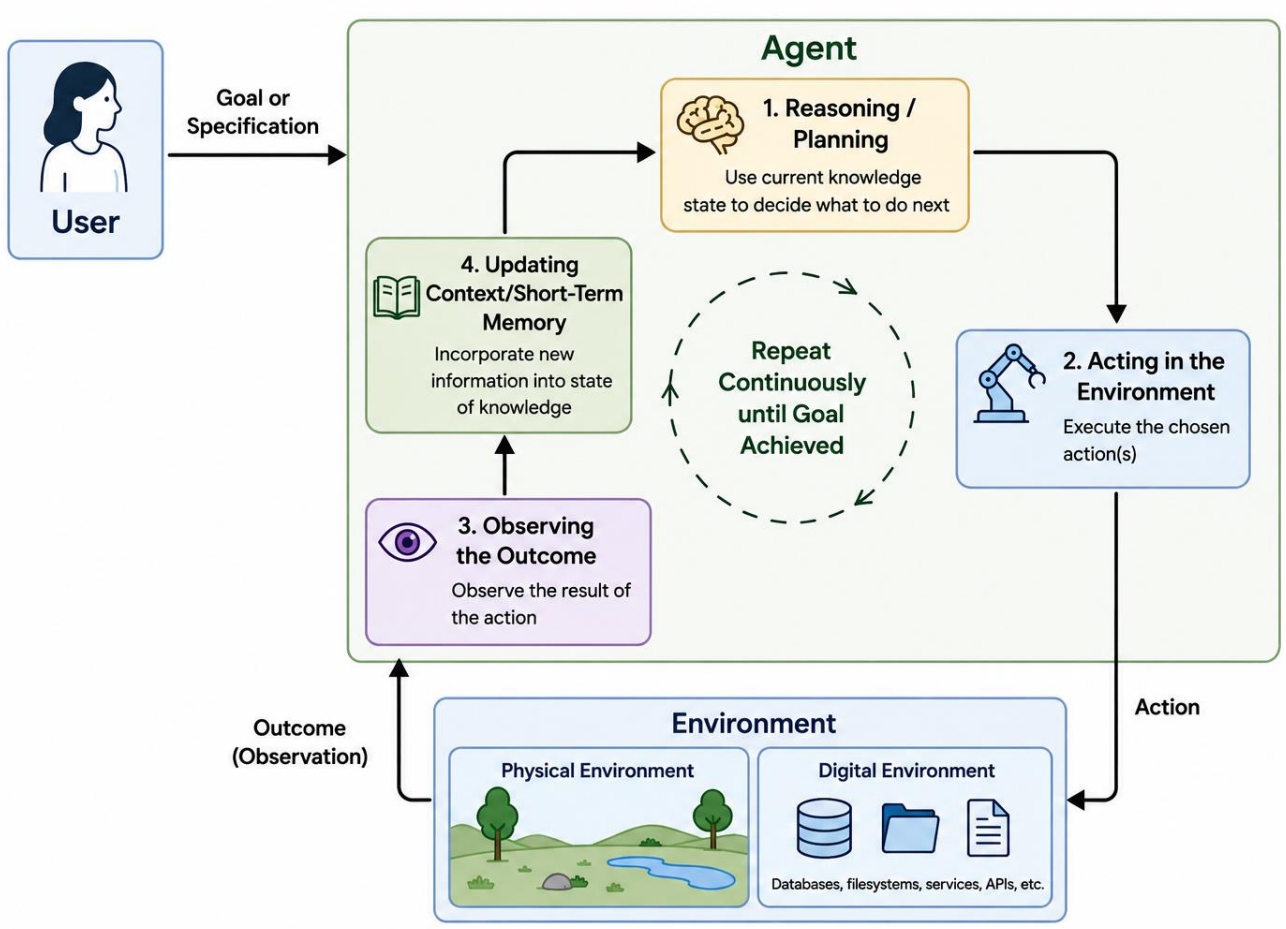
Agentic AI

- Agents act autonomously in the environment to help achieve user's goal
- LLM by itself insufficient
  - Need tools to carry out actions
  - Need to update memory after observation
- But in multiple domains, LLMs now trained to be good enough to
  - Use tools effectively, both known tools and unknown tools by reading the tool descriptions
  - Be effectively steered using general guidelines in system prompts and procedurally through skill prompts
  - Understand tool descriptions, skills, etc. in **natural language!**
- In some digital environments, e.g. for software engineering
  - Harnesses now provide complete tools
  - LLM reasoning is now good enough that human level performance can be achieved with the help of steering and tools from a good harness
    - E.g. > 80% of code in Anthropic authored by Claude (May 2026)



Agentic AI

- Is this a good time to build agents for your area?
  - Going forward, expect rapid progress for environments where
    - Complete tools can be provided
    - Enough data available to train capable AI reasoners, so that steering them through prompts is effective
  - For some environments, e.g. physical environments, not yet clear how to get enough data
    - How to make progress in these domains?
  - Lots of research on reasoning, memory, reliability, efficiency/cost, etc.
    - Expect rapid progress



NUS National University of Singapore

ai.nus.edu.sg

naii@nus.edu.sg

COM3 #B1-08, 11 Research Link, S(119391)

NUS National University of Singapore | Artificial Intelligence Institute