

AGI

PRESENTED BY

EMIL SHAHBAZOV

CEPHER ONANO

What is AGI?

- ▶ Imagine a robot that can learn math, cook, and play games — like a smart human friend.
- ▶ AGI means a machine that can learn anything, just like a person.
- ▶ - Unlike 'regular AI', AGI not only has the ability to make decisions on its own, but also learn on its own. AGI is supposed to have the depth and versatility of human cognition.
- ▶ Different from regular AI which only knows how to do one task well.
- ▶ Can transfer its abilities and knowledge across domains.

How is AGI different from narrow AI?

- ▶ Narrow AI: Siri, chess bots, facial recognition — each good at one thing.
- ▶ AGI: Can learn across domains, reason abstractly, and adapt to new tasks

Key Components of AGI

- ▶ Perception and understanding – core human cognition,
- ▶ Learning and adaptation mechanisms – reinforcement l., unsupervised and transfer learning
- ▶ Memory and knowledge representation
- ▶ Reasoning, planning, decision making

History of AGI –

- ▶ 1956: Dartmouth Conference proposes thinking machines.
- ▶ AI winters: Times when progress slowed due to hype vs reality.
- ▶ Now: Renewed interest with brain-inspired models and large language models.

Roadmap to AGI –

- ▶ Needs powerful hardware, rich and continuous data, and flexible algorithms.
- ▶ Four main AGI architectures: Symbolic, Emergentist, Hybrid, Universalist.
- ▶ Incremental vs revolutionary approach
- ▶ Building blocks: perception, understanding, learning, reasoning and decision making.
- ▶ Integrate disparate AI technologies into AGI synthesis
- ▶ Collaborative efforts and open research.

Examples of promising AI

- ▶ CogNet and OpenCog: Brain-like cognition projects.
- ▶ DeepMind's Adaptive Agents: Learn in 3D environments.
- ▶ PaLM-E: Combines vision, language, and robotics.
- ▶ MindMatrix – neuroscience + CS
- ▶ Cognitive Nexus – rigid rule based system, not scalable
- ▶ CogGenius – genetic algorithms (natural selection)

Consciousness & AGI

- ▶ Could AGI feel something? Could it be aware?
- ▶ Indicator Properties ;Recurrent Processing, Global Workspace Theory, Attention Schema, Predictive Processing theories, High Order Thought, Agency & Embodiment.
- ▶ Indicator properties help assess possible AI consciousness.

Limits of AGI

- ▶ Organisms recognize 'affordances' — opportunities to act.
- ▶ AI struggles with meaning-making and goal-setting from within.
- ▶ Some argue AGI might never fully match natural intelligence.

Ethics & Control

- ▶ How do we ensure AGI is safe?
- ▶ What if it feels pain? Who controls it?
- ▶ Ethics in AGI = technical & moral challenge.
- ▶ Transparency/accountability – make sure AI systems are transparent in their reasoning and thinking, and not hiding anything.
- ▶ AGI consciousness vs Human consciousness – Philosophical considerations of what it means to be sentient, and truly 'intelligent'

Where Are We Headed?

- ▶ AGI is still in development.
- ▶ It needs both strong tech and ethical guidance.
- ▶ Let's build minds wisely.

References

- ▶ 1. Goertzel, B. (2014). Artificial General Intelligence: Concept, State of the Art, and Future Prospects. Journal of Artificial General Intelligence.
- ▶ 2. Roli, A., Jaeger, J., & Kauffman, S. (2022). How Organisms Come to Know the World: Fundamental Limits on AGI. Frontiers in Ecology and Evolution. <https://doi.org/10.3389/fevo.2021.806283>
- ▶ 3. Butlin, P., Long, R., Bengio, Y., et al. (2023). Consciousness in Artificial Intelligence: Insights from the Science of Consciousness. arXiv:2308.08708 [cs.AI]

AGI Component Summary Goertzel (Overview)

AGI Components Across Mainstream And Experimental Systems

	AI System	Perception & Understanding	Learning & Adaptation	Memory & Knowledge	Reasoning & Planning	Decision Making
1	ChatGPT (GPT-4)	✓	✓	●	●	●
2	DeepMind's Adaptive Agent	✓	✓	●	●	✓
3	PaLM-E (Google)	✓	✓	●	●	●
4	Tesla FSD (Full Self Driving)	✓	✓	●	●	✓
5	Anthropic Claude	✓	✓	●	●	●
6	OpenAI Codex	✗	✓	●	●	●
7	Meta CICERO (diplomacy agent)	✓	✓	●	✓	✓
8	CogNet / OpenCog	✓	✓	✓	✓	✓
9	MindMatrix	✓	●	●	●	●
10	Cognitive Nexus	●	✗	●	●	✗
11	CogGenius	●	✓	●	●	●

Roli et al. Evaluation Summary – Limits to AGI

Roli Et Al. AGI Philosophy-Based Evaluation					
	AI System	Self-Defined Goals (Agency)	Meaning-Making (Affordances)	Beyond Algorithmic Limits	Biological Self-Organization (Closure)
1	ChatGPT (GPT-4)	✗	✗	✗	✗
2	DeepMind's Adaptive Agent	●	✗	✗	✗
3	PaLM-E (Google)	●	✗	✗	✗
4	Tesla FSD (Full Self Driving)	●	✗	✗	✗
5	Anthropic Claude	✗	✗	✗	✗
6	OpenAI Codex	✗	✗	✗	✗
7	Meta CICERO	●	✗	✗	✗
8	CogNet / OpenCog	●	●	✗	✗
9	MindMatrix	●	✗	●	✗
10	Cognitive Nexus	✗	✗	✗	✗
11	CogGenius	●	●	✗	✗

Scientific Theories of Consciousness Summary

	AI System	Recurrent Processing	Global Workspace	Higher-Order Thought (HOT)	Attention Schema (AST)	Predictive Processing	Agency & Embodiment
1	ChatGPT (GPT-4)	●	●	✗	✗	✓	✗
2	DeepMind's Adaptive Agent	✓	●	✗	✗	✓	●
3	PaLM-E (Google)	●	●	✗	✗	✓	●
4	Tesla FSD	●	✗	✗	✗	✓	●
5	Anthropic Claude	●	●	✗	✗	✓	✗
6	OpenAI Codex	✗	✗	✗	✗	✓	✗
7	Meta CICERO	●	✓	✗	✗	✓	●
8	CogNet / OpenCog	✓	✓	✗	✗	✓	●
9	MindMatrix	●	●	✗	✗	●	●
10	Cognitive Nexus	✗	✗	✗	✗	●	✗
11	CogGenius	●	●	✗	✗	✓	●