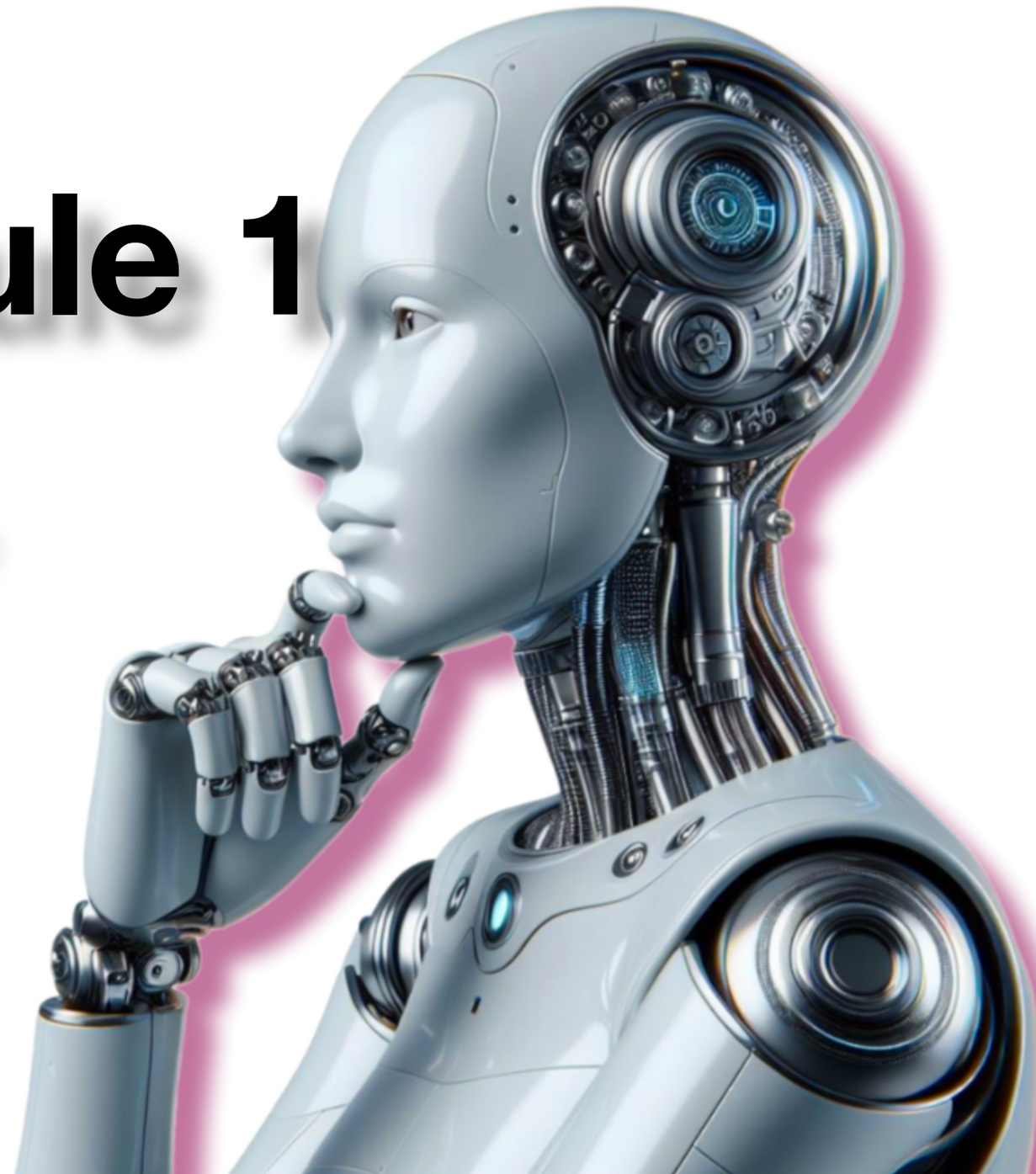
Jeffery. Dirden
ITAI-2372
Professor Bhakthan

Module 1



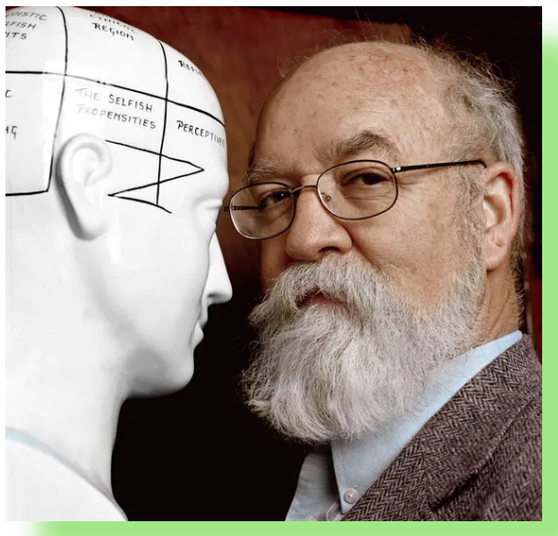
Do machines have free will?

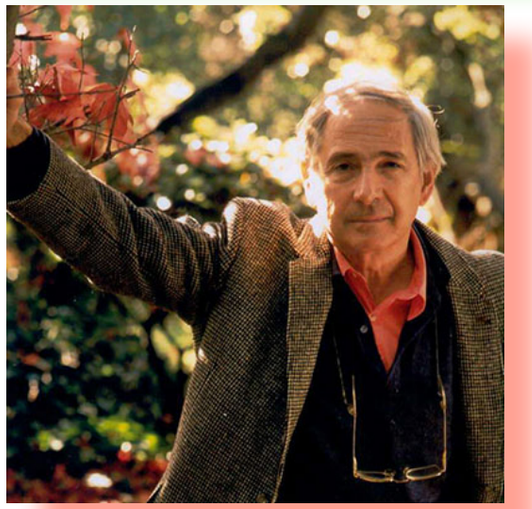
Introduction

- Free will has been seen as a human quality associated with autonomy and making choices.
- This question in AI calls into question our concept of predictable algorithms and autonomous action.
- This talk investigates whether AI can have free will by examining arguments from Daniel Dennett and John Searle.
- While some say that AI can make decisions that approach free will, others believe it lacks true independence.
- Understanding this argument is critical to AI ethics, legal accountability, and system design.

Philosophical Perspectives from Daniel Dennett & John Searle

- Dennett believes that free will is not absolute, but rather compatible with inevitability.
- AI, like humans, can work under limits while still making advanced decisions.
- AI systems that can analyze information, consider effects, and alter their behavior could have a functioning kind of free will.





- Searle's "Chinese Room" argument indicates that AI may comprehend symbols without real understanding.
- Free will needs awareness and intention, which AI lacks.
- Even the most advanced AI still follows preset instructions rather than making decisions on its own.

Philosopher	Perspective on AI & Free will
Daniel Dennet	AI can demonstrate an actual version of free will through complicated decision-making.
John Searle	AI lacks actual free will since it cannot comprehend its own actions.

Analysis and Implications

- Dennett's viewpoint: AI's ability to make decisions implies a limited type of free will.

- Strength: Complements real-world AI applications where decision-making is important.

- Weakness: disregards the lack of self-awareness in AI.

- Searle's viewpoint: AI lacks knowledge and independence, which means it lacks true free will.

- Strength: distinguishes human thought processes from computer processing.

- Weakness: Underestimates AI's ability to imitate complicated decision-making.

- Legal and ethical responsibilities: If AI lacks free will, should its designers be held responsible for its actions?

- AI in Business and Society: Decision-making is critical in autonomous systems such as banking, healthcare, and self-driving cars. Do we consider their choices as their own or planned?

- Future of AI Design: Should we strive to create AI with free will-like characteristics, or is this ethically dangerous?

- In conclusion, I believe that AI decision-making can resemble free will but lacks actual autonomy and intention. For the record, from what I understand, AI does not have emotions or feel anything in its current state. The debate is critical to AI ethics, accountability, and governance. The future of AI may hinge on whether we value practical decision-making above complete autonomy in system design.

Cited Sources

- Hksspr. (2024b, June 4). *Interview with Daniel Dennett: AI, Free Will, and the Future of Humankind - HKS Student Policy Review*. HKS Student Policy Review. https://studentreview.hks.harvard.edu/interview-with-daniel-dennett-ai-free-will-and-the-future-of-humankind/

- The Editors of Encyclopaedia Britannica. (2024, December 30). Chinese room argument | Definition, Machine Intelligence, John Searle, Turing Test, Objections, & Facts. Encyclopedia Britannica. https://www.britannica.com/topic/Chinese-room-argument

- HKS Student Policy Review. (2024, April 20). Interview with Daniel Dennett: AI, Free Will, and the Future of Humankind (Part 1) [Video]. YouTube. https://www.youtube.com/watch?v=OED3VKPZOsU

- Searle, J. (1980). The Chinese room. In *Minds, Brains, and Programs*. https://rintintin.colorado.edu/~vancecd/phil201/Searle.pdf

-