



Course Overview

CS3263

Foundations of Artificial Intelligence

Sem 2, AY2023-24

This Lecture
May Be
Recorded!

Syllabus

Responsible AI

Week	Topics	Week	Topics
1	Overview and Introduction – AI and Society - Toward Responsible AI and Trustworthy Systems	7	Rational Decision Making - Decision Analysis, Value alignment
2	Knowledge Representation - Logic, Production Systems, Semantic Networks.	8	Mid-Term Test [14 March 2024]
3 Q	Problem Solving in Complex Environments - Local Search, Online Search	9	Decision Making over Time - Markov Decision Processes
4	Problem Solving with Limited Resources - Constraint Satisfaction Problems	10 Q	Human Factors - Cognitive Modeling, Human AI collaboration
5	Representing Uncertainty - Bayesian Networks, Conditional Independence	11	Decision Making with Learning - Reinforcement Learning
6 Q	Reasoning under Uncertainty - Probabilistic Inference, Causal Modeling	12	Intro to Generative AI - Generative Models, Responsible AI Decision Making
R		13 Q	Project Presentations

Overview

Q: Quiz week

3

What Courses Can I Take After This?

CS2019S (Intro AI+ML)



CS3263 (FAI)

CS3264 (FML)

CS4244

Knowledge Representation
and Reasoning

CS4246

AI Planning and
Decision Making

CS4248

Natural Language
Processing

CS4243

Computer Vision and
Pattern Recognition

AI Electives

Teaching Team

Name	Role
LEONG Tze Yun	Lecturer
KONG Lingdong	Tutor
MA Haozhe	Tutor
ONG Hanyang	Tutor
WU Jiele	Tutor

Note: Please ask technical and course organization questions through **Discussions on CANVAS!**

Teaching Team – In the New AI Era

- Effective and responsible use of:

- Generative AI Foundational Models
- Search Engines

- Examples:

- LLAMA 2
- ChatGPT
- PaLM 2
- Bing
- ...

Be Careful When you Learn from Them!
Be Aware of the Limitations and Dangers!

...

Note: Please attribute **References** or **Sources** from Generative AI Models or Search Engines!

Course Logistics

- Classes (weekly)

- Thursday 1000 - 1200 SR1 Lecture
- Mondays and Tuesdays TBA Tutorial

- Grading policy

- Homework, participation 20 + 10%
- Quizzes 20%
- Test [1000 to 1200 Thu 14 Mar 2024 *in-person*] 25%
- Project [presentation in week 13; report due Fri 19 Apr 2024] 25%
- No final exam for this module!



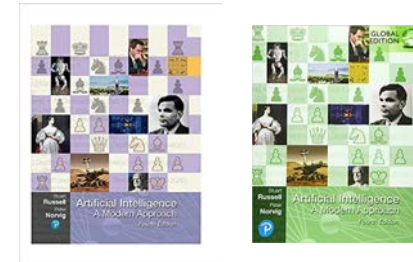
Course Logistics

- Core Contents
 - Course materials and announcements on Canvas
 - Discussions on Canvas.
- Tutorials and Assignments
 - Attempt tutorial questions before class; presentations and discussions in class
 - Team Assignments: Written questions + Programming – Prerequisite: Python
- Project
 - Project Guidelines to be issued in Week 3
 - Thematic or Self-defined topic in teams of 2-3
 - 1 page proposal due after Recess Week

Reference Books

- Main reference book:

- (RN) Russell, S. and P. Norvig, Artificial intelligence: A modern approach. 4th ed. (Global ed.) 2021: Pearson. [Kindle Edition] (Alternate: 3rd ed.)
[Table of contents for 4th ed.: <http://aima.cs.berkeley.edu/contents.html>]



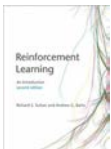
- Reference books:



- (M) Murphy, K. P. (2023). Probabilistic machine learning: Advanced topics, MIT Press.[Book website: <http://probml.github.io/book2>]



- (NJ) Neapolitan, R.E. and X. Jiang, Artificial intelligence : with an introduction to machine learning. Second edition. ed. Chapman & Hall/CRC artificial intelligence and robotics series. 2018, Boca Raton: CRC Press, Taylor & Francis Group. xiii, 466 pages.
[Book website: <https://doi.org/10.1201/b22400>]



- (SB) Sutton, R. S. and A. G. Barto. Reinforcement Learning: An introduction. 2nd ed. MIT Press, 2018, 2020
[Book website: <http://incompleteideas.net/book/the-book.html>]
[e-Book for personal use: <http://incompleteideas.net/book/RLbook2020.pdf>]

Overview



Additional Resources

- You will also find good tutorials, tools, publications at:
 - Conference in Uncertainty in Artificial Intelligence (UAI)
 - <https://www.auai.org>
 - American Association for Artificial Intelligence Conference (AAAI)
 - <https://www.aaai.org>
 - International Joint Conference on Artificial Intelligence (IJCAI)
 - <https://www.ijcai.org>
 - Neural Information Processing Systems Conference (NeurIPS)
 - <https://nips.cc>
 - International Conference on Automated Planning and Scheduling (ICAPS)
 - <https://www.icaps-conference.org>
 - International Conference on Autonomous Agents and Multiagent Systems (AAMAS)
 - <https://www.ifaamas.org>
 - International Conference on Artificial Intelligence and Statistics (AISTATS)
 - <https://aistats.org>
 - ...



Honour Code

- NUS Code of Student Conduct:
 - (A) Academic, Professional and Personal Integrity
 - (B) Respect for People
 - (C) Respect for and Compliance with the Law and with Campus Policies and Regulations
 - (D) Responsibility towards Maintaining the Campus as a Place Conducive for Learning and Living
- This module will teach you how to apply and develop powerful **Responsible AI** technologies for the betterment of humankind
- If you are unable or unwilling to respect and abide by the Honour code, please **DO NOT** take this module!



NUS Course Materials: Ethical Behaviour and Respecting Copyright

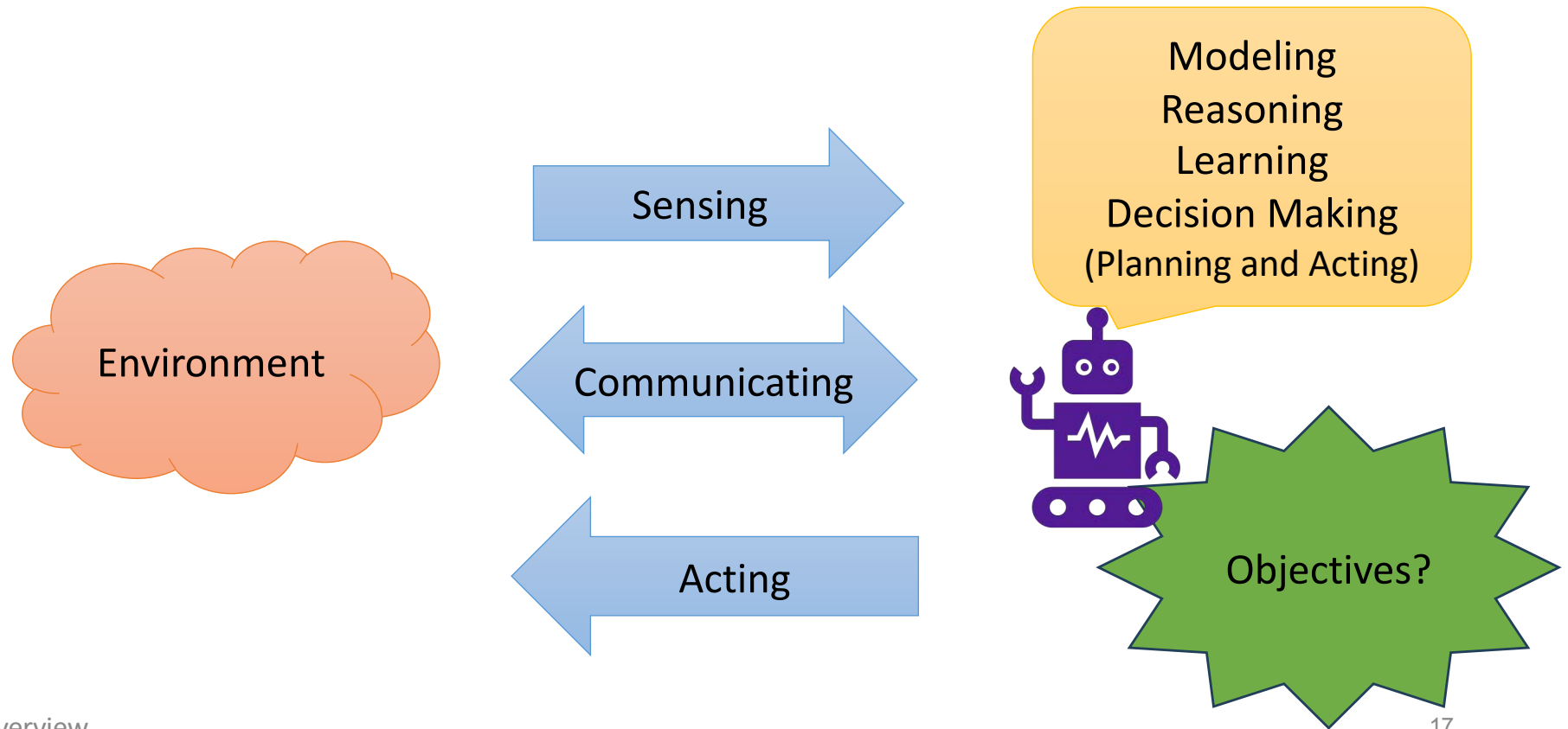
All course participants (including permitted guest students) who have access to the course materials on LumiNUS or any approved platforms by NUS for delivery of NUS modules are not allowed to re-distribute the contents in any forms to third parties without the explicit consent from the module instructors or authorized NUS officials

Examples of Disallowed Actions

- No Posting on any websites (except for the materials explicitly allowed by your lecturer in the respective module)
- No selling of material
- No sharing of questions/answers which could lead to cheating/plagiarism

So, What Will I Learn in this Course?

AI: Building A Rational Agent



Example: Smart Assistants



Example: Strategic Decision Making

- Defeating the World Champion in Go
- AlphaGo uses learning and Monte Carlo Tree Search to defeat world champion Lee Sedol
 - https://www.youtube.com/watch?v=8tq1C8spV_g



Example: Good AI Fighting Cybercrime (and Evil AI)



Example: AI in Finance and Business

Example: AI in Health Care

Home and Community

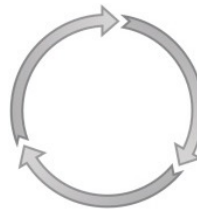
Policy, Regulation
and Governance

Clinic and Hospital

**Global Health,
Population Health,
Public Health**

Workflow and Process Optimization

Prevention and Care Support



**Scientific discovery
and experimentation**

Monitoring and Communication
Overview

Diagnosis and Treatment

Example: AI in Space Exploration

AI for Good!



Foundations of Artificial Intelligence

- Knowledge Representation
- Problem solving in complex environments
- Constraint satisfaction problems
- Reasoning under Uncertainty
- Rational Decision Making
- Human factors
- Decision Making over Time
- Integrating Learning and Decision Making
- Generative AI Models
- And other new trends ...

In CS2109S, you have learned:
Basic and Adversarial Search
Machine learning

AI and Society
Responsible AI

Ask ChatGPT:

Why should I study
Foundations of Artificial Intelligence?

Answer:

- In essence, while large language models and other tools make AI more accessible, they're akin to using a calculator without understanding arithmetic.
- To harness the full potential of AI, drive innovation, and ensure ethical and effective use, a foundational knowledge remains indispensable.

Source: A chat with ChatGPT 4, 12 Aug 2023

(Demystifying) Magic of Artificial Intelligence

- Histories and Trends:
 - Symbolic AI
 - Neural AI
 - Statistical AI
 - Cognitive and Behavioral AI
 - Generative AI
- Toward Responsible AI - Human-aware AI
 - AI working for, working with, and working alongside humans



Beyond Technical Knowledge

- **Domain challenges**
 - Involving deep domain knowledge and operational issues
 - Interacting conditions, processes, and goals
- **User challenges**
 - Different skill levels and preferences
 - Varying usage patterns and cognitive biases
- **Economic challenges**
 - High implementation costs
 - Unclear market viability and scalability
- **System challenges**
 - Uncertain and changing information, processes, environments
 - Evolving IT and communication systems

A Grand Challenge!