AA 598B Special Topics

Decision-Making & Control for Safe Interactive Autonomy

Instructor: Prof. Karen Leung

Autumn 2024

https://faculty.washington.edu/kymleung/aa598/

Welcome!

Instructor: Karen Leung

Office hours: Mondays, 11:20AM – 12:20PM, GUG 311B

Email: aa598-safe-autonomy-staff@uw.edu

Note: There are no graders or TAs for this course.

Note: First time offering this course.

Today's plan

- Introduction
  - Motivation and contextualization
- Course logistics
- How to ~~read~~ skim a research paper
- (Fundamentals)

Round of introductions

- Name
- Department
- Year (UG, MS, PhD)
- Research interests
- Where you're from
- Fun fact

So much buzz…  AI vs ML vs DL vs LLM

(5 mins discuss with neighbor)

- **AI**: Ability to reason, learn, and/or predict, mirroring the types of cognitive functions typically associated with human intelligence.
- **ML**: The process of learning a model from data
- **DL**: A subset of ML that uses deep neural networks
- **LLM**: A particular type of DL method designed for tasks involving human language, based on large-scale data.

So much buzz…

AI vs ML vs DL vs LLM

AI

Search algorithms
Probabilistic graphical models
Logic
Control theory
State machine
…

ML

Linear classifiers/regression
K-nearest neighbor
Logistic regression
Decision trees
Support vector machine
RL
…

DL

MLP
RNN
CNN
GNN
Transformers
…

LLM

Gemini
GPT
VILA
Bard
Llama
…

Robotics vs Controls vs Autonomy 🤖💪🦿

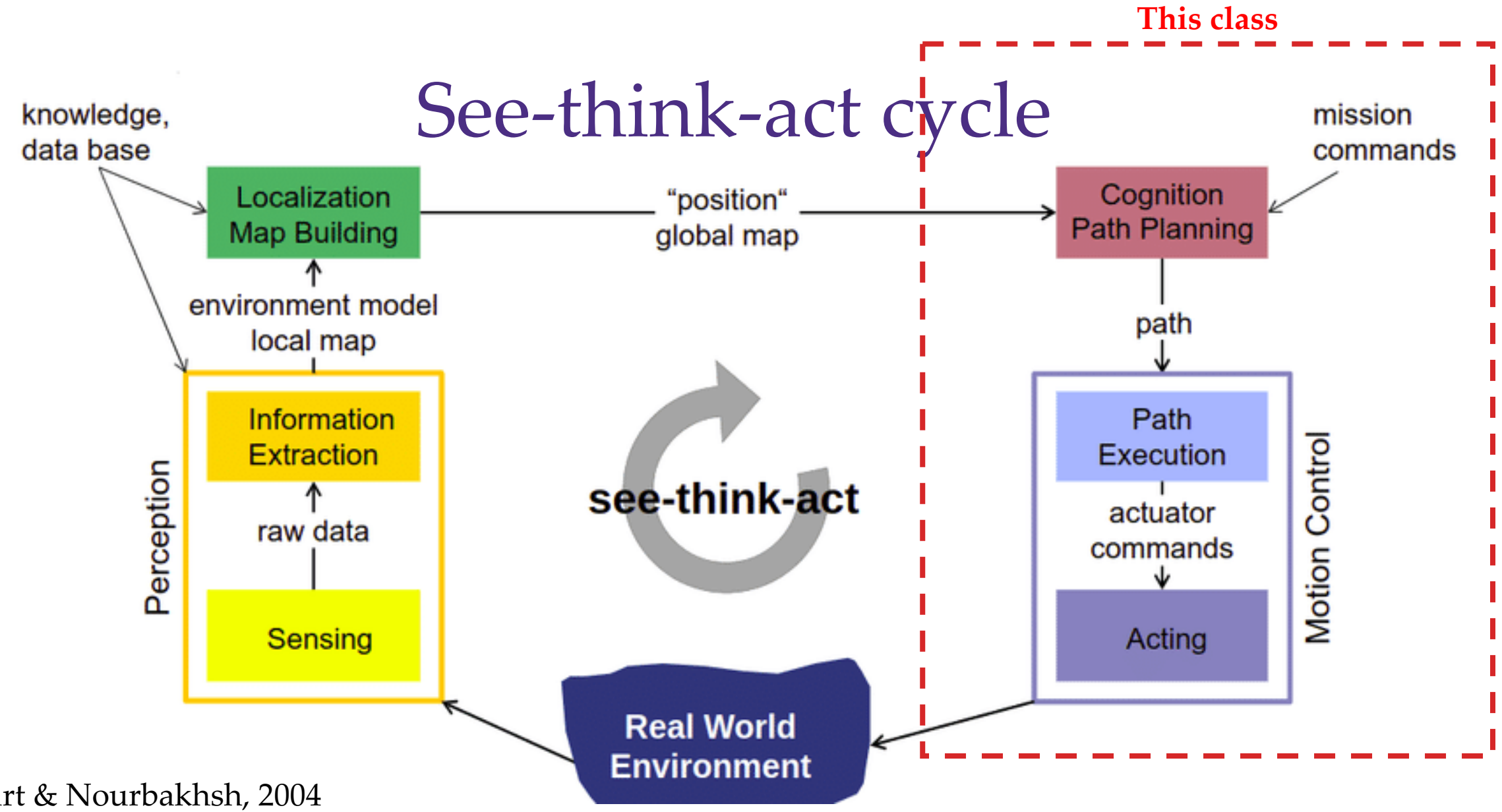
(5 mins discuss with neighbor)

- **Robotics**: Multidisciplinary field that involves the design, construction, software-hardware integration, and operation of robots.

- **Controls**: Area focused on regulating the behavior of dynamic systems by adjusting inputs based on feedback from sensors.

- **Autonomy**: Ability to perform tasks and make decisions independently without human intervention. Autonomous systems can perceive their environment, reason about the information they gather, make decisions, and act on those decisions.
  - Can use multiple AI algorithms

See-think-act cycle

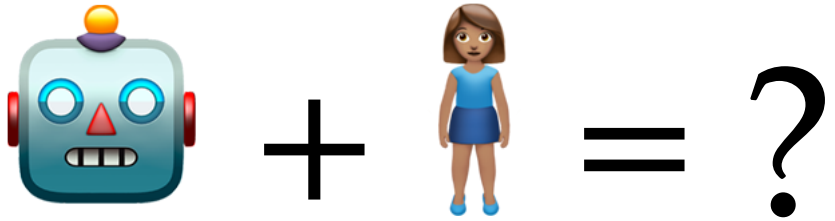


Siegwart & Nourbakhsh, 2004

Core challenges in robotics

- Dealing with uncertainty
  - Modeling and quantifying
  - Need to be robust against it
- Out of distribution scenarios
  - Identifying them
  - Figuring out what to do in those scenarios
- Generalization
  - Transfer skills
  - Adapt to new settings

What happens when you add 🧍‍♀️ into the picture? What new problems do we have?



- Safety! Balancing safety with performance
- Accounting for rare events, or edge cases
- Alignment with human values
- Evaluating performance
- Regulations

This course will cover…

Logistics

Course format

https://faculty.washington.edu/kymleung/aa598/

- Mixture of lectures, paper discussions, guest lectures, student presentations
- Heavy emphasis on **paper reading**
- Submit *short* **paper summaries** and lead one *long* **paper discussion**
- **Small homework** to solidify fundamentals and *empower you* for your course project
- **Course project** accumulating into a presentation (week 10) & report (finals week)

Module structure
Lecture
Collaborative in-class discussion
Long paper discussion
Guest speaker

Grading breakdown

- 15% homework
- 15% *Long* paper discussion
- 30% *Short* paper/talk reviews
- 35% Course project (proposal, presentation, report, peer review)
- 5% Participation

Homework (15%)

- 3 x (small) homework sets
  - 5% each
  - Coding focused
  - Designed to solidify your fundamentals
  - "Building blocks" to empower your course project
  - Roughly corresponds to each module

Submission: Canvas

- Recommended due dates – *strongly encouraged*
- Prioritize OH activity on current module
- Must submit everything by midnight Dec 13th (Friday finals week) **NO EXCEPTIONS**

Long paper discussion (15%)

- In groups, you lead a paper discussion session on an assigned paper
  - **~35 mins** long
  - **Summarize** the paper using the PACES method
    - Problem, Approach, Claim, Evaluation, Substantiation
  - **Archaeologist**: What is the positioning/context of the paper
  - **Researcher**: What new problems or methods does this paper open?
  - **Practitioner**: What is the practical impact of the paper?
  - **Open discussion**
- 1 long paper discussion day per module
  - 2 paper discussions per session

Submission: In-person (no late submissions)

Short paper/talk summaries (30%)

- Write paper summaries for each long discussion paper
  - In preparation for long paper discussions
  - Excludes the module you are leading a long paper discussion on
- Write a talk summary + reflection for each guest speaker
  - Encourage reflection on on-going research problems
- 7 opportunities (4 paper, 3 talks)
  - Submit 6 only (5% each)

Submission: Canvas (no late submissions)

Course project (35%)

- Project proposal (5%)
- In-class presentation in week 10 (10%) + peer review (5%)
- Report due in finals week (15%)
  - 1 week after presentation
  - *Research project*—A research project that is connected to topics covered in this course.
  - *Literature survey*—A deep dive into several papers on a chosen topic area, including your inclusion criteria, motivating questions, and insights.
- Report to be written in conference-paper format

Submission: In-person & Canvas

(no late submissions)

Due dates	Deliverables
Week 10 (Mon)	Group A presentations
Week 10 (Wed)	Group B presentations
Finals week (Mon)	Group A reports
Finals week (Wed)	Group B reports

Participation (5%)

- Active participation (attendance and asking questions) is key to your and your peer's success and learning experience.
  - In-class paper discussions
  - Asking questions during presentations and talks
- Expected to attend class *in-person* during scheduled time
- If you have challenges regularly attending class, please reach out (aa598-safe-autonomy-staff@uw.edu)
- Deliberate in maintaining small class size to encourage more meaningful interactions between peers

How is this course different from other courses?

- **Core courses**
  - Often required for a major and cover well-established fundamental topics
  - Provide the fundamentals necessary to explore more advanced topics
  - Not necessarily representative of instructor's research or state-of-the-art
  - Homework/textbook driven
- **Technical electives**
  - Continuation on core courses and dives deeper into specific topics
  - Middle ground between core and special topics
  - Perhaps a special topics in a previous lifetime
- **Special topics *(this course)***
  - Reflective of instructor's research interests
  - Focus on state-of-the-art research and techniques
  - Less homework/textbook driven, more paper reading and open-ended discussions
  - Assumes familiarity in foundational topics

What you will gain from taking this course*

- Develop research skills!
  - Technical *(related to safe interactive autonomy)*
    - Formalize problems related to human-robot interaction
    - Develop human behavior prediction models
    - Formulate a robot planning problem that accounts for the presence of human agents
    - Devise control techniques that account for uncertainty in human behavior
    - Propose new approaches to decision-making and control for safe interactive autonomy
  - Practical skills *(translates to other research settings)*
    - Reading and reviewing papers
    - Implementing algorithms
    - Group discussions, collaboration, and evaluation
    - Attending talks
  - Communication *(translates to any technical career path)*
    - Presenting and summarizing complex ideas
    - Technical writing
    - Public speaking

*Assuming you put in the work and actively engage in class discussion

Is this course right for me?

- Assumes *strong* familiarity in
  - Machine learning
  - Deep learning
  - Optimization and optimal control
  - Mathematical abstraction
  - Programming (Python)
- Strong emphasis on *self-learning*
- This course is capacity constrained
  - If you have any concerns about whether you are well-prepared, please reach out. I would be happy to discuss your background with you.
  - The plan is to have the course offered again next year, so if you feel that waiting would better align with your current skill set, you'll have the chance to enroll later.