Language Technologies Institute

Carnegie Mellon University

Advanced Topics in Multimodal Machine Learning (11-877)

Lecture 1: Introduction

Louis-Philippe Morency, Amir Zadeh, Paul Liang

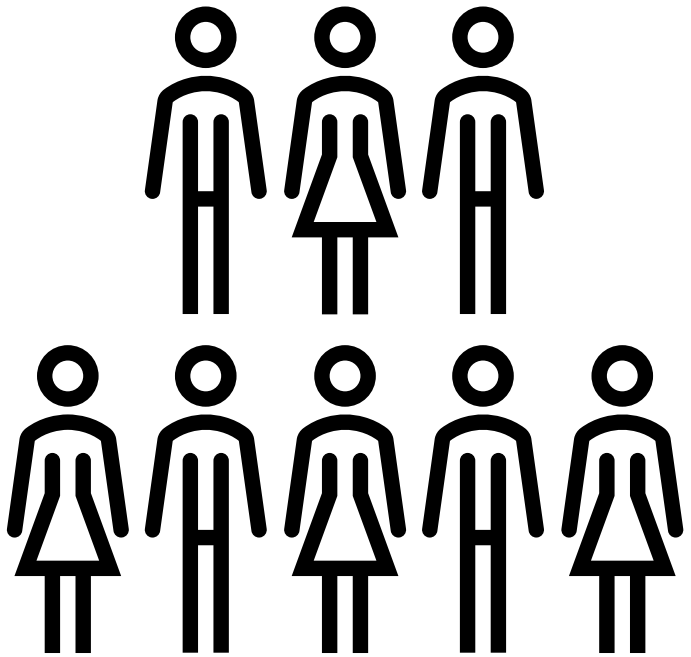
Your Instructors This Semester (11-877)

Louis-Philippe Morency
morency@cs.cmu.edu
Course instructor

Amir Zadeh
abagherz@cs.cmu.edu
Course instructor

Paul Liang
pliang@cs.cmu.edu
Course instructor

Time for Introductions!



Your name, department and programs

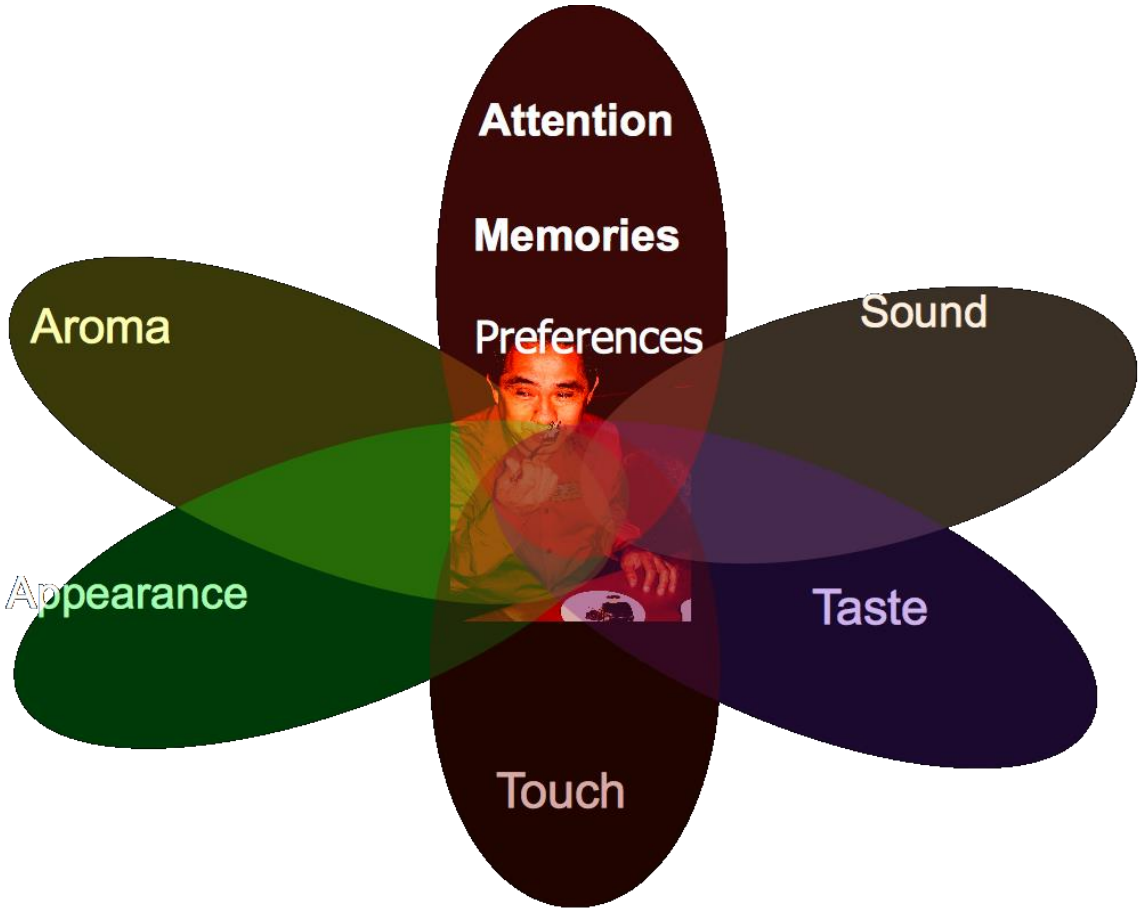
Your favorite modality(ies)!

Previous research experience in multimodal

Why are you interested in this course?

What is Multimodal?

What is Multimodal?

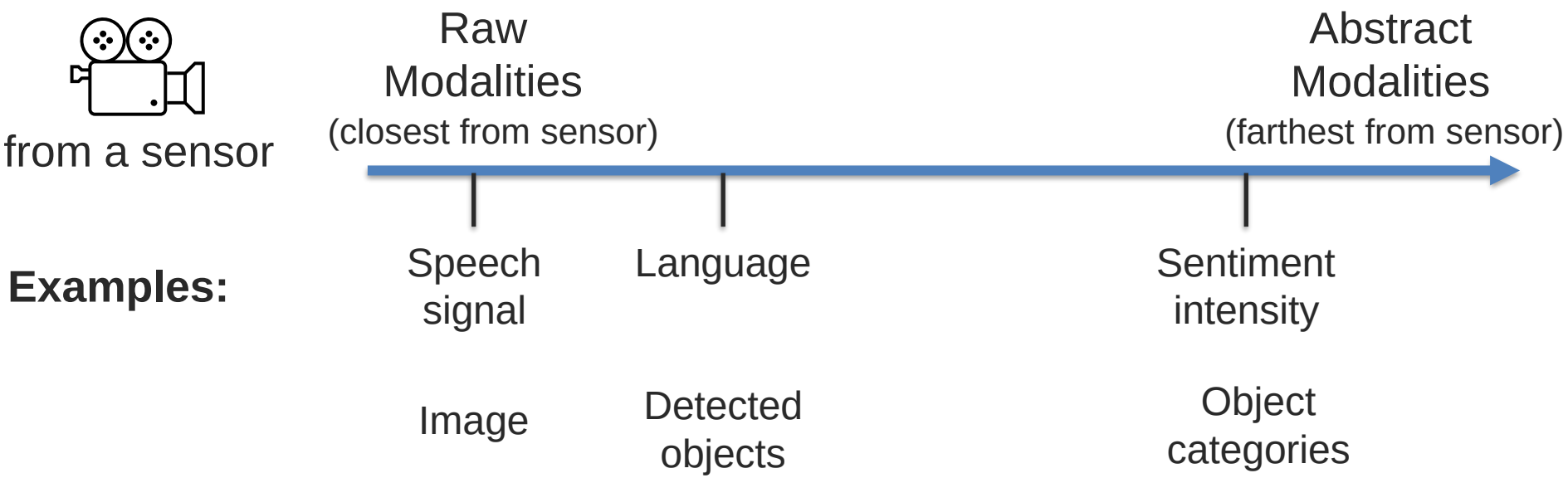


Sensory Modalities

What is Multimodal?

Modality

Modality refers to the way in which something expressed or perceived.



Multimodal: from multiple modalities

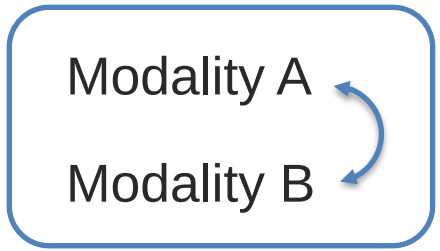
Examples of Modalities

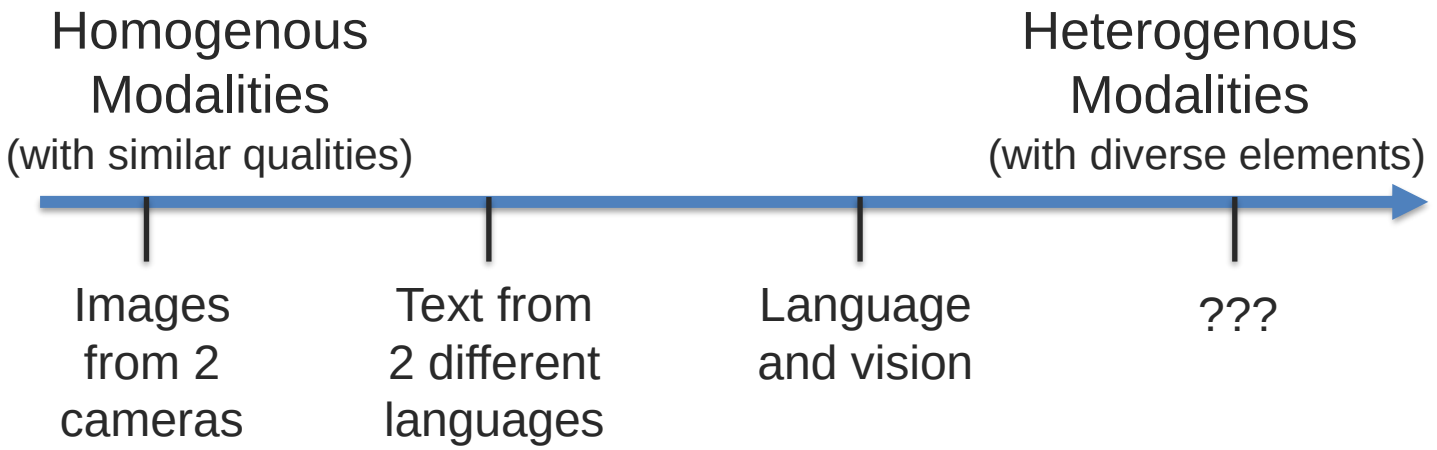
- Natural language (both spoken or written)
- Visual (from images or videos)
- Auditory (including voice, sounds and music)
- Haptics / touch
- Smell, taste and self-motion
- Physiological signals
  - Electrocardiogram (ECG), skin conductance
- Other modalities
  - Infrared images, depth images, fMRI

What is Multimodal?

Heterogeneity

Information present in the different modalities will often show diverse qualities and elements.





Examples:

What is Multimodal?

Multimodal Machine Learning is the study of computer algorithms that learn and improve through the use and experience of multimodal data

Multimodal Artificial Intelligence studies computer agents able to demonstrate intelligence capabilities such as understanding, reasoning and planning, through multimodal experiences, and data

***Multimodal* is the science of heterogenous data ☺**

What are the main Dimension of Multimodal Heterogeneity?

https://miro.com/app/dashboard/

Course Syllabus

Learning Objectives

1. Study recent technical achievements in multimodal research
2. Improve critical and creative thinking skills
3. Understand future research challenges in multimodal
4. Explore new research ideas in multimodal learning

Two Versions: 6-credits and 12-credits

- 6-credit version:
  - Reading assignments
  - Small group discussions
  - Synopsis leads
- 12-credit version
  - Same 6-credit expectations + a high-quality research project:
    - Proposal + literature review
    - Midterm and final reports
    - Weekly updates (during team meetings with instructor)

Course Topics *(subject to change, based on student interests and course discussions)*

Week 1 (1/21): Introductions

Week 2 (1/28): Cross-modal interactions

Week 3 (2/4): Multimodal Co-learning

Week 4 (2/11): Pre-training paradigm

Week 5 (2/18): Multimodal reasoning

Week 6 (2/25): Memory and long-term interactions

Week 7 (3/4): *No classes – Spring break*

Week 8 (3/11): *No classes – Spring break*

Course Topics *(subject to change, based on student interests and course discussions)*

Week 9 (3/18): Brain and multimodal perception

Week 10 (3/25): Beyond language and vision

Week 11 (4/1): Subjectivity and dataset biases

Week 12 (4/8): *No classes – CMU Carnival*

Week 13 (4/15): Fairness and real-world constraints

Week 14 (4/22): Multimodal generalization

Week 15 (4/29): Multimodal with low-resources

What are Your preferences?

https://miro.com/app/dashboard/

Reading Assignments

- Two main parts:
  - **Assigned reading papers:** Read the assigned papers and summarizing the main take-away points of each paper
    - Optional: if you have clarification questions about the papers
  - **Research question probes:** Reflect on the question probes related to the reading papers and prepare discussion points.
    - Students should also scout for extra papers, blog posts or other resources related to these question probes
- 11 readings assignments, with usually 2 assigned papers

Small Group Discussions

- Two groups of 8-10 students, one instructor per group

- Round-table discussions:
  - Understanding papers: focus on clarifying any questions or misunderstandings related to the two research papers (15-20 mins)

  - Research discussions: Discuss the research question probes. Each student is expected to actively participate in this discussion.

Discussion Synopsis Leads

- 2 leads per session, one for each small group

- The main tasks of the discussion synopsis leads are
  - **Active support:** Leads are expected to read the assigned papers with extra details, to assist other students if clarification questions
  - **Note-taking:** Leads should take detailed notes during discussions. Notes are for internal use, not shared outside the course.
  - **Synopsis:** Both leads will meet to create one coherent synopsis from both discussions. These synopses will be public on the course website.

Grading Scheme for 6-credit Version

- Reading assignment 40%
  - 5 points per assignment
  - Top 8 scores kept for final grade
- Participation and discussions 32%
  - 4 points per discussion session
  - Top 8 scores kept for final grade
- Discussion synopsis leads
  - 14 points for each synopsis (including note-keeping & support)
  - Top 2 scores kept for final grade

Research Course Project (12-credit version)

- ✓ Similar in spirit to a 6-credit independent study project
- ✓ Project teams of 2 or 3 students
- ✓ Final report should be like a research paper
- ✓ Expected to explore new research ideas
- ✓ Regular meetings with instructors

Course Project Timeline

- **Project preferences** (Due Monday 1/24 at 8m ET) – Online form to share your interests about research projects and help with team matching.
- **Pre-proposal** (Due Thursday 2/3 at 8pm ET) – You should have selected your teammates, have ideas about your dataset and task.
- **Proposal and Literature Review** (Due Monday 2/21 at 8pm ET) – Description of your research ideas and review of relevant papers.
- **Midterm report** (Due Monday 3/14 at 8pm ET) – Intermediate report documenting the initial results exploring new research ideas.
- **Final report** (Due Monday 5/2 at 8pm ET) – Final report describing explored research ideas, with experimental results and discussion.

Weekly Project Updates

- Meetings on a weekly basis, 20-30 minutes per meeting
- Online document for updates
  - Either Google Slides (preferred) or Google Docs
  - Same document for the whole semester
  - Weekly updates should be informal, focusing only on the highlights
    - For example, one new slide per week, with 3-4 main items
- Meetings designed to get feedback from all instructors regularly

Grading Scheme for 12-credit Version

- Grading breakdown of the 6-unit version will be scaled to 50%.

- The second 50% comes from the course project:
  - Pre-proposal and project preferences 5%
  - Proposal and literature review 15%
  - Midterm report 20%
  - Final report 30%
  - Weekly updates 30%
    - 3 points per update, top 10 scores kept for final grade

Absences and Late Submissions

- Lectures are not recorded, students expected to attend live
  - If you plan to miss more than one lecture this semester, let us know as soon as possible.
- Reading assignment wildcards (3 per students)
  - 24-hours extension, max 1 per week
- Project assignment wildcards (2 per teams)
  - 24-hours extension, can be used together

Course Websites

- Piazza
  - For course announcements and assignments

  https://piazza.com/cmu/spring2022/11877/info
- CMU Canvas
  - For assignment submissions and grading

  https://canvas.cmu.edu/courses/28476
- Course website
  - A general public version of the course information
    - Discussion synopsis will be posted here

  https://cmu-multicomp-lab.github.io/adv-mmml-course/spring2022/

Assignments for This Coming Week

Week 2 reading assignment (Due Wednesday 1/26 at 11pm ET)

- Detailed instructions will be posted on Piazza

For students taking the 12-credit version:

1. Project preference form (to help with team matching)
   - Google Form will be shared on Piazza
2. Schedule availability for project update meetings
   - When2meet form will be shared on Piazza

Both due this Tuesday 1/25 at 8pm ET