Language Technologies Institute

Carnegie Mellon University

Multimodal Machine Learning

Lecture 1.1: Introduction

Louis-Philippe Morency

** Co-lecturer: Paul Liang. Original course co-developed with Tadas Baltrusaitis. Spring 2021 and 2022 editions taught by Yonatan Bisk*

Your teaching team This Semester (11-777, Fall 2022)

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

Paul Liang
pliang@cs.cmu.edu
Co-lecturer

Catherine (Yun) Cheng
yuncheng@andrew.cmu.edu
TA

Karthik Ganesan
karthikg@andrew.cmu.edu
TA

Gabriel Moreira
gmoreira@andrew.cmu.edu
TA

Alex Wilf
awilf@andrew.cmu.edu
TA

Yinghuan Zhang
yinghuan@andrew.cmu.edu
TA

Lecture Objectives

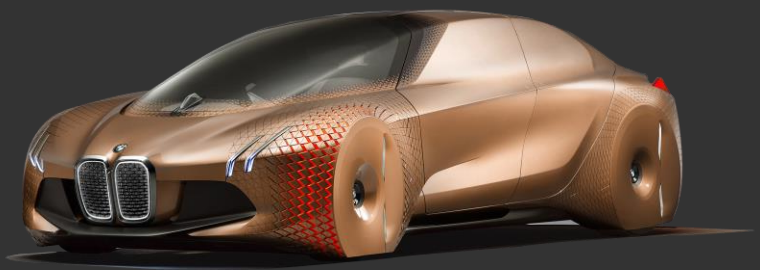
- What is Multimodal?
  - Research-oriented definition
  - Dimensions of modality heterogeneity
  - Modality connections and interactions
- Core technical and conceptual challenges
  - Representation, alignment, reasoning, generation, transference and quantification
- Course schedule

What is Multimodal?

Multimodal AI Technologies

Robots

Personal Vehicles



Ubiquitous

Mobile

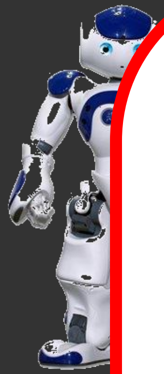
Wearable

Online

YouTube

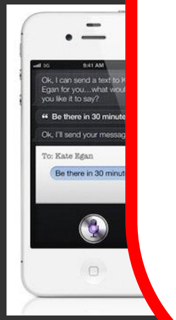
Multimodal AI Technologies

Robots

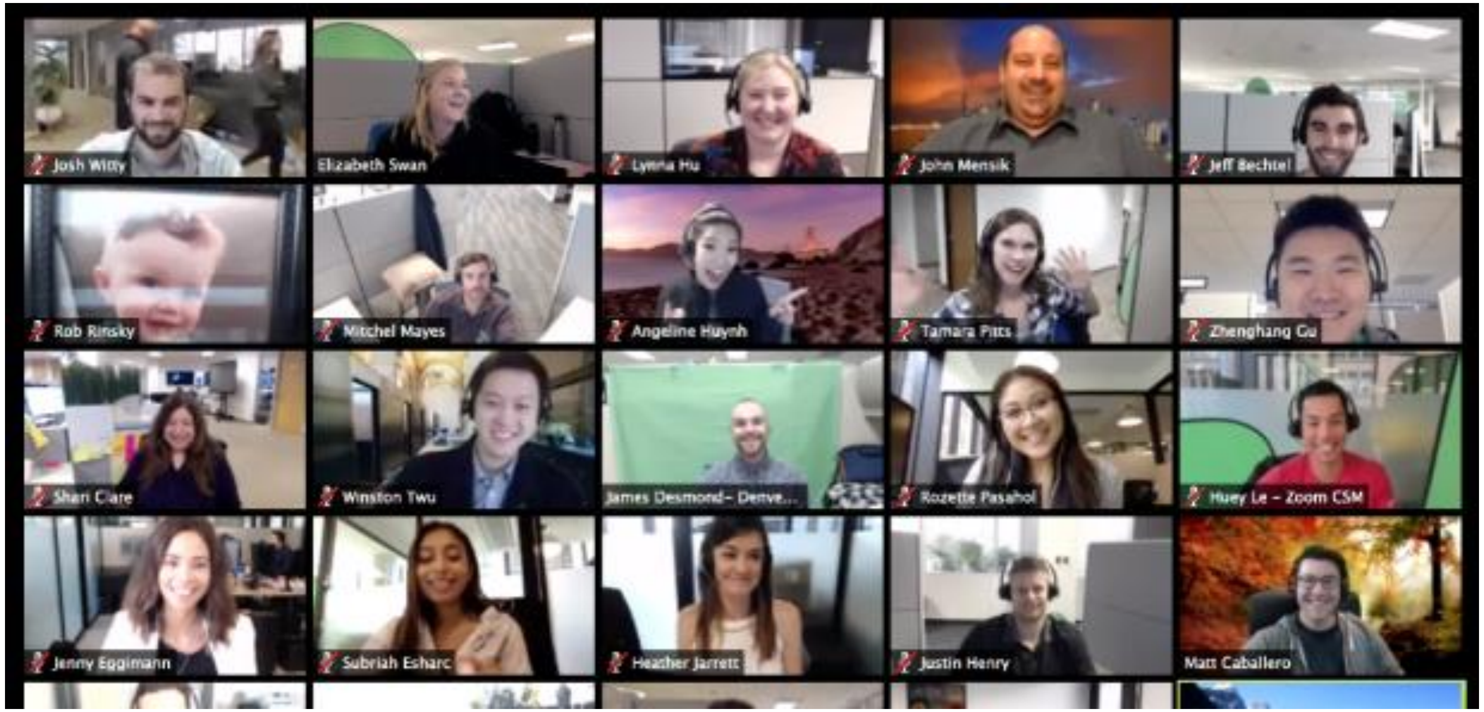


Personal Vehicles

Ubiquitous

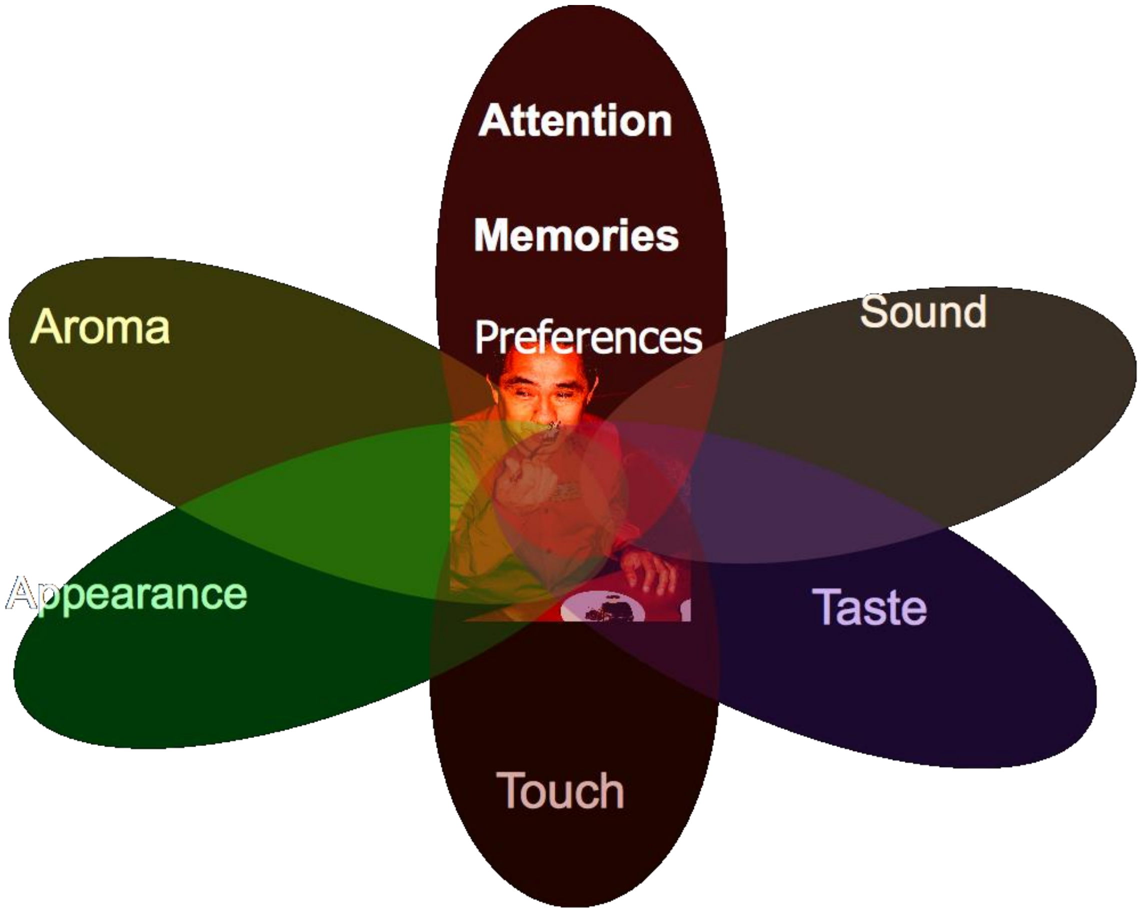


Video Conferencing



Wearable

What is Multimodal?



Sensory Modalities

Multimodal Behaviors and Signals

Language
- **Lexicon**
  - Words
- **Syntax**
  - Part-of-speech
  - Dependencies
- **Pragmatics**
  - Discourse acts

Acoustic
- **Prosody**
  - Intonation
  - Voice quality
- **Vocal expressions**
  - Laughter, moans

Visual
- **Gestures**
  - Head gestures
  - Eye gestures
  - Arm gestures
- **Body language**
  - Body posture
  - Proxemics
- **Eye contact**
  - Head gaze
  - Eye gaze
- **Facial expressions**
  - FACS action units
  - Smile, frowning

Touch
- **Haptics**
- **Motion**

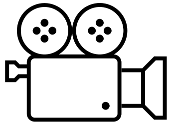
Physiological
- **Skin conductance**
- **Electrocardiogram**

Mobile
- **GPS location**
- **Accelerometer**
- **Light sensors**

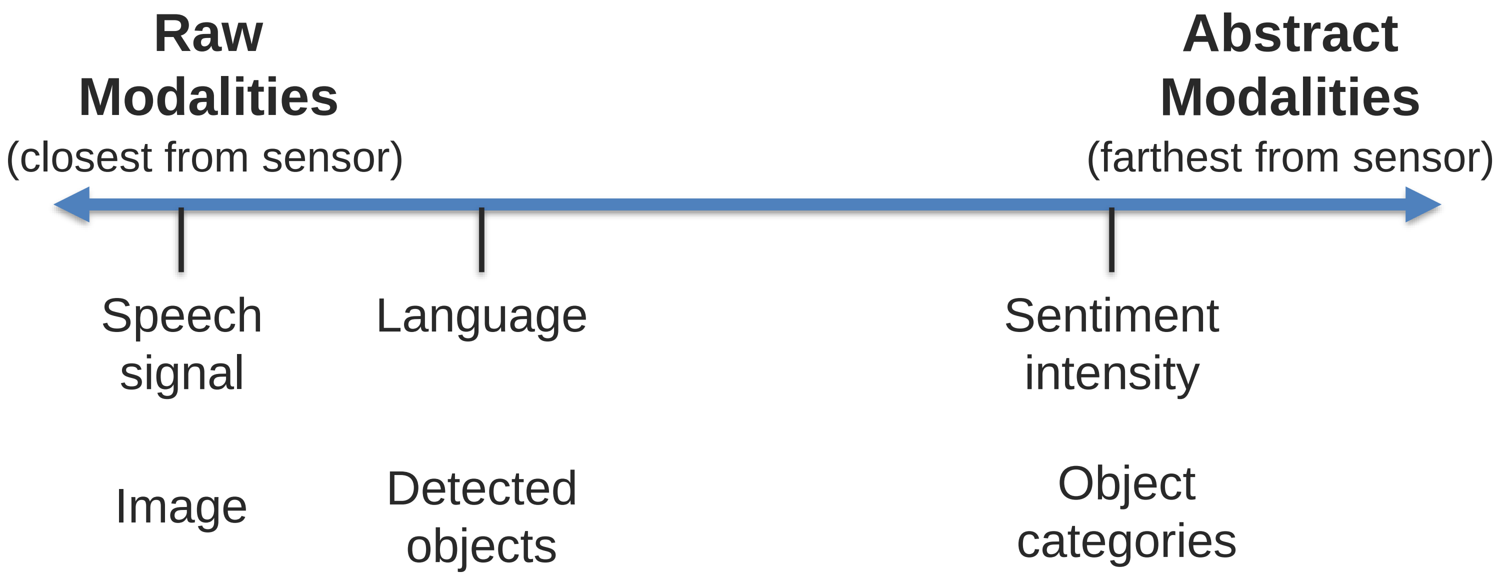
What is a Modality?

Modality

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



from a sensor



Examples:

What is Multimodal?

A dictionary definition…

Multimodal: with multiple modalities

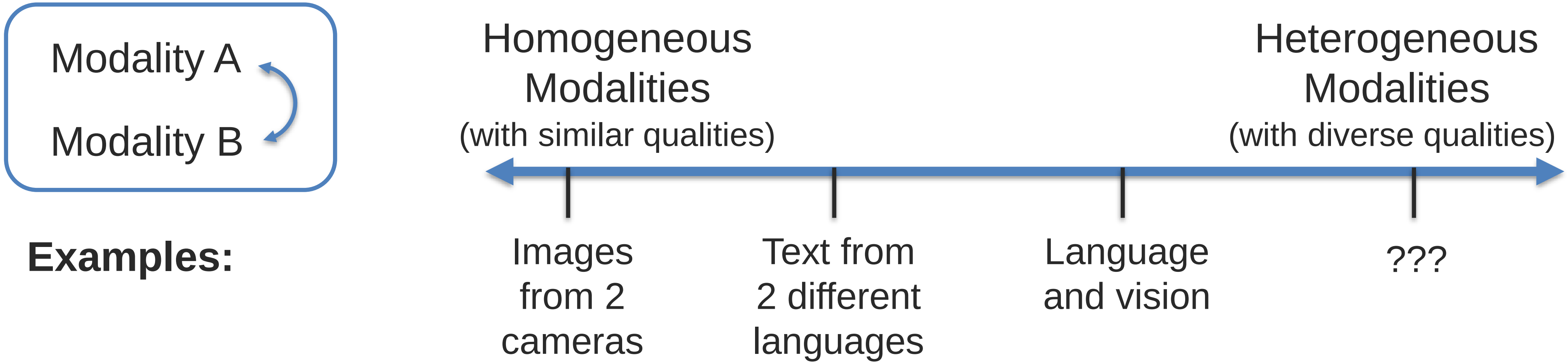
A research-oriented definition…

***Multimodal* is the science of**

heterogeneous and interconnected data

Heterogeneous Modalities

Information present in different modalities will often show diverse qualities, structures and representations.



Abstract modalities are more likely to be homogeneous

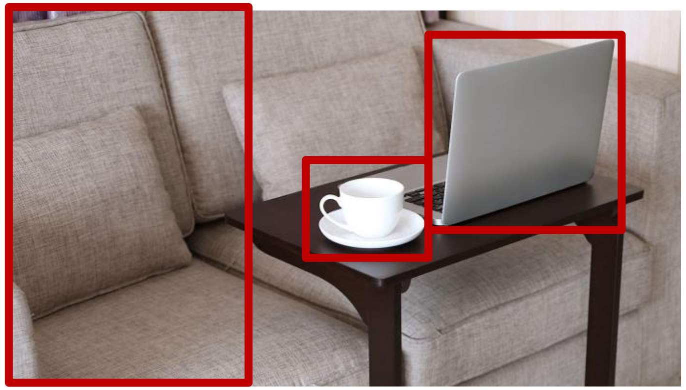
Dimensions of Heterogeneity

Information present in different modalities will often show diverse qualities, structures, and representations.

A teacup on the right of a laptop in a clean room.

Dimensions of Heterogeneity

Information present in different modalities will often show diverse qualities, structures, and representations.



A ***teacup*** *on the* ***right*** *of a* ***laptop*** *in a* ***clean room****.*

1 **Element representations:** discrete, continuous, granularity

{  ,  ,  }

{teacup, right, laptop, clean, room}

Dimensions of Heterogeneity

Information present in different modalities will often show diverse qualities, structures, and representations.

A teacup on the right of a laptop in a clean room.

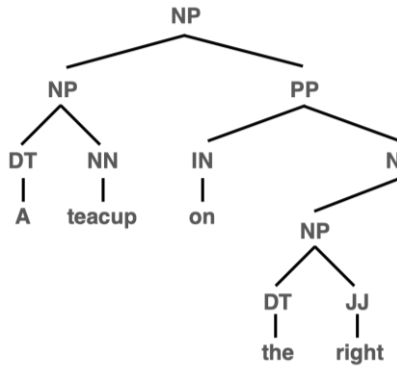
2 **Element distributions:** density, frequency

objects per image

words per minute

Dimensions of Heterogeneity

Information present in different modalities will often show diverse qualities, structures, and representations.



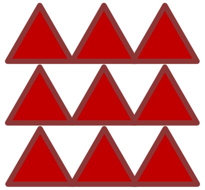
…

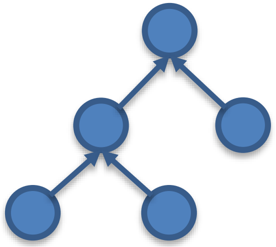
Latent
(implicit)

A teacup on the right …

Explicit
(observable)

3 **Structure:** temporal, spatial, hierarchical, latent, explicit





Dimensions of Heterogeneity

Information present in different modalities will often show diverse qualities, structures, and representations.

A teacup on the right of a laptop in a clean room.

4 **Information:** abstraction, entropy

Dimensions of Heterogeneity

Information present in different modalities will often show diverse qualities, structures, and representations.

A teacup on the right of a laptop in a clean room.

5 **Noise:** uncertainty, signal-to-noise ratio, missing data

teacup → teacip

right → rihjt

Dimensions of Heterogeneity

Information present in different modalities will often show diverse qualities, structures, and representations.

A teacup on the right of a laptop in a clean room.

6 **Relevance:** task relevance, context dependence

→ recreational
→ living room
→ right-handed

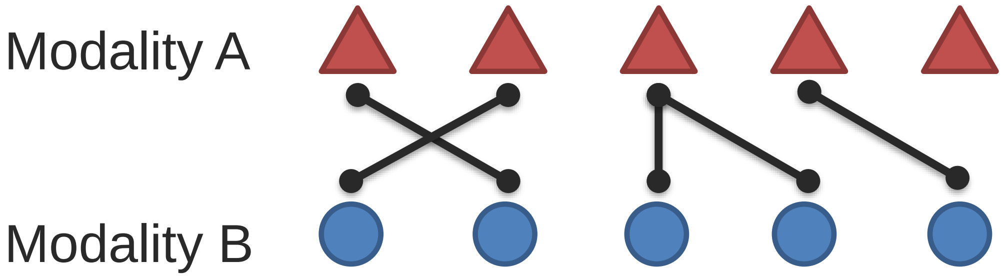
A teacup on the right of a laptop in a clean room.

→ workspace
→ study room

Interconnected Modalities

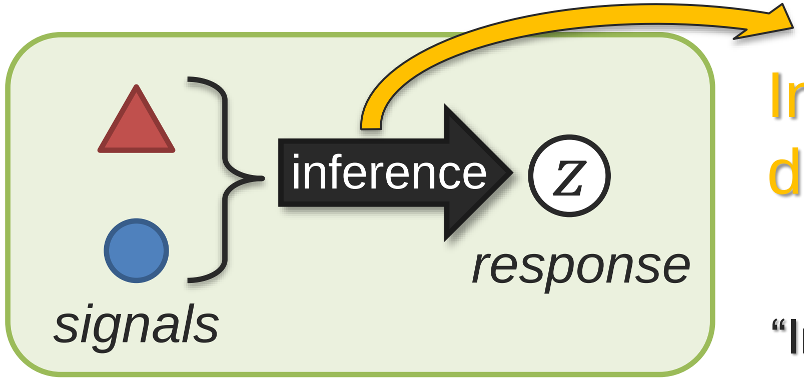
① Modality connections

Modalities are often related and share commonality



② Modality interactions

Modality elements often interact during inference



Interactions happen during <u>inference!</u>

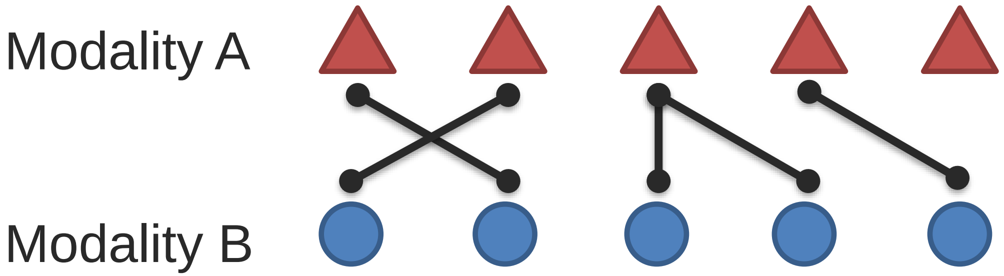
"Inference" examples:

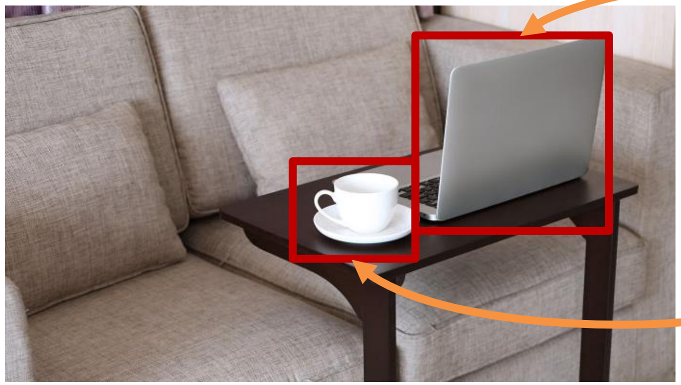
- Behavior perception
- Recognition task
- Modality translation

Interconnected Modalities

① **Modality connections**

Modalities are often related and share commonality



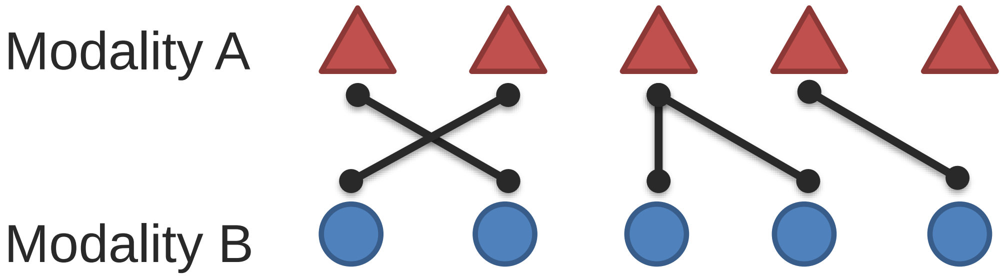


A ***teacup*** *on the right of a* ***laptop*** *in a clean room.*

Interconnected Modalities

1 Modality connections

Modalities are often related and share commonality



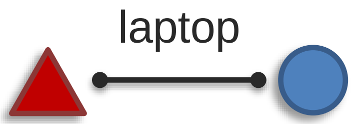
Statistical

Association

e.g., correlation, co-occurrence

Semantic

Correspondence

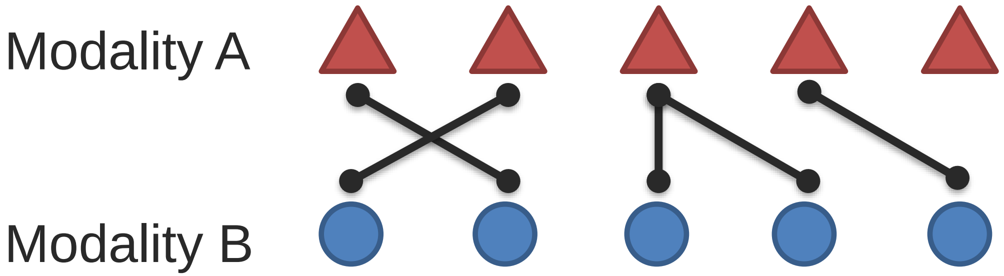


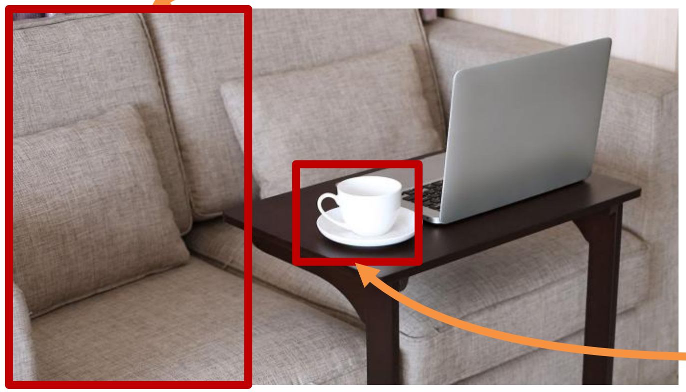
e.g., grounding

Interconnected Modalities

① **Modality connections**

Modalities are often related and share commonality



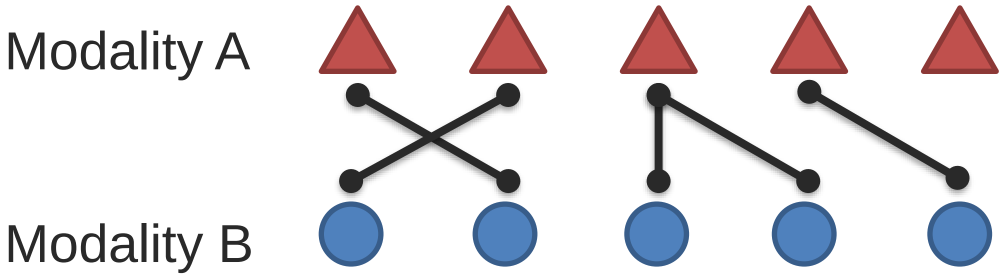


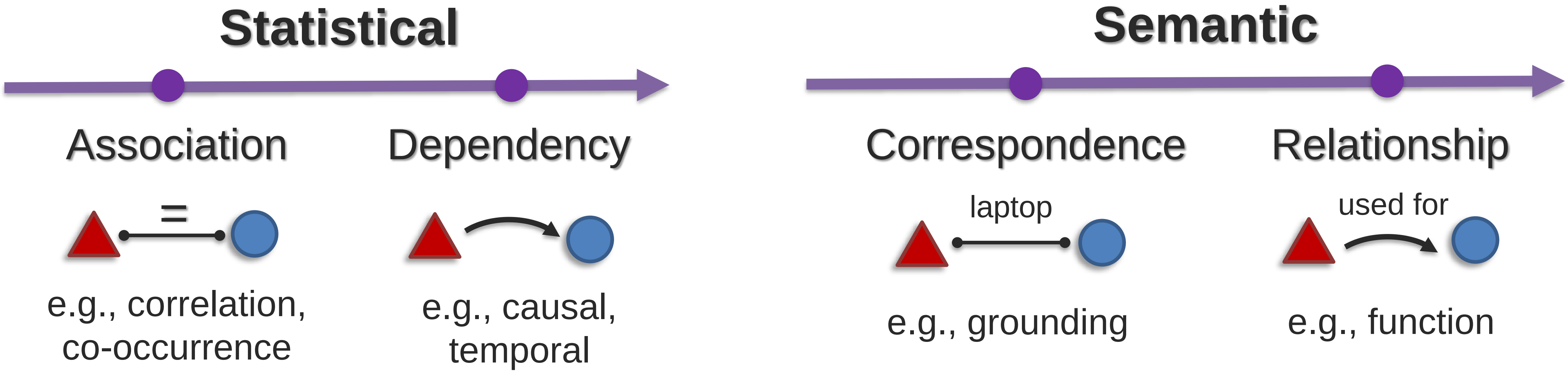
A teacup on the right of a laptop in a ***clean room****.*

Interconnected Modalities

① **Modality connections**

Modalities are often related and share commonality

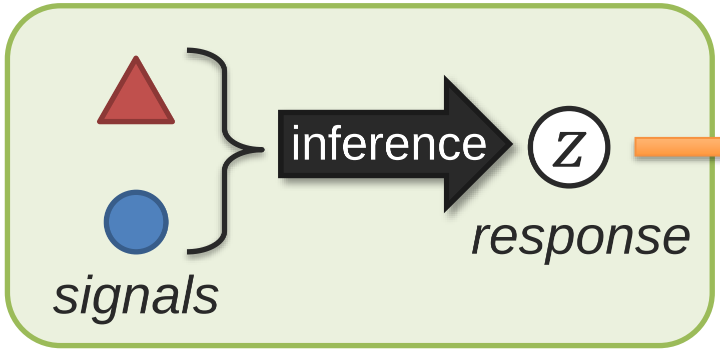




Interconnected Modalities

② Modality interactions

Modality elements often interact during inference



Types of interaction responses?
(a taxonomy)

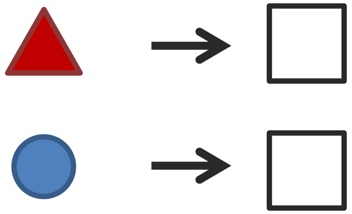
Is this indoors?

inference — *Yes!*

A teacup on the right of a laptop in a clean room.

inference — *Yes!*

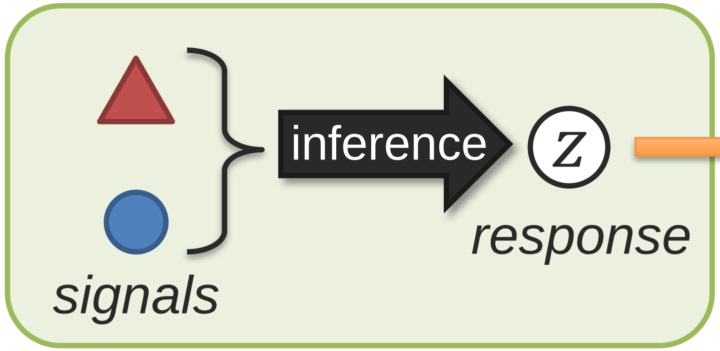
Unimodal redundancy



Interconnected Modalities

② Modality interactions

Modality elements often interact during inference



Types of interaction responses?

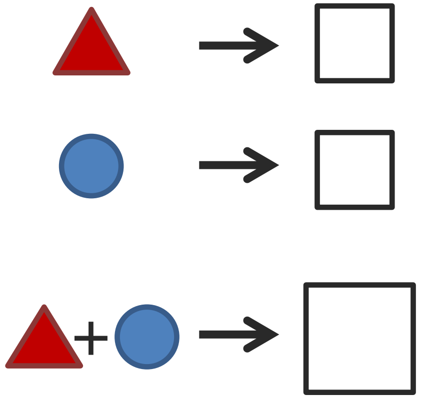
Is this indoors?

A teacup on the right of a laptop in a clean room.

inference

Yes!

Unimodal redundancy

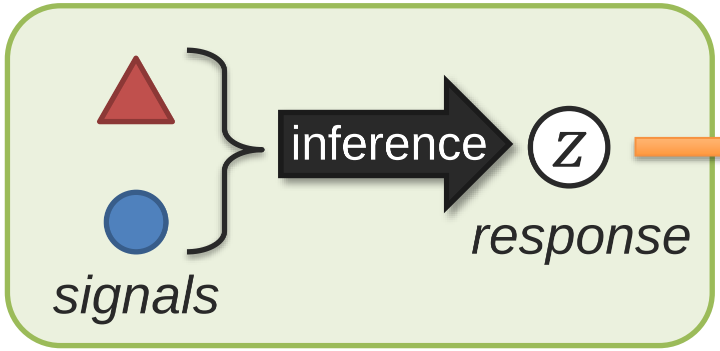


Multimodal enhancement

Interconnected Modalities

② Modality interactions

Modality elements often interact during inference



Types of interaction responses?

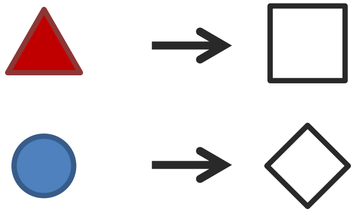
Is this a living room?

inference → *Yes!*

A teacup on the right of a laptop in a clean room.

inference → *No, probably study room.*

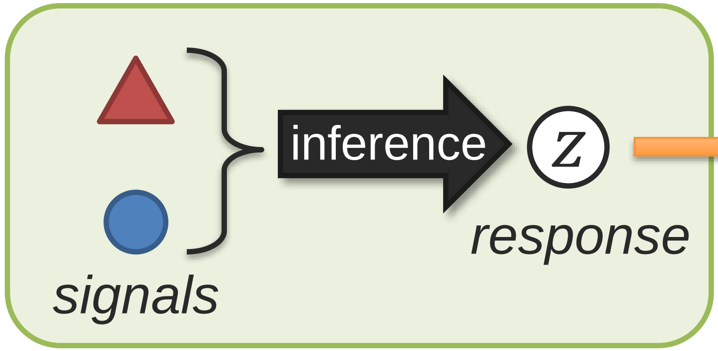
Unimodal Non-redundancy



Interconnected Modalities

② Modality interactions

Modality elements often interact during inference

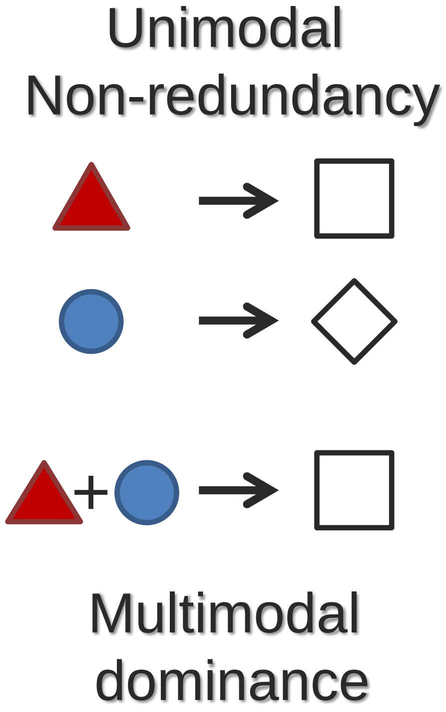


Types of interaction responses?

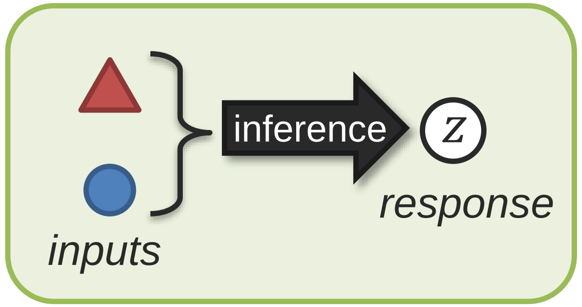
Is this a living room?

A teacup on the right of a laptop in a clean room.

inference → ***Yes!***



Taxonomy of Interaction Responses – A Behavioral Science View



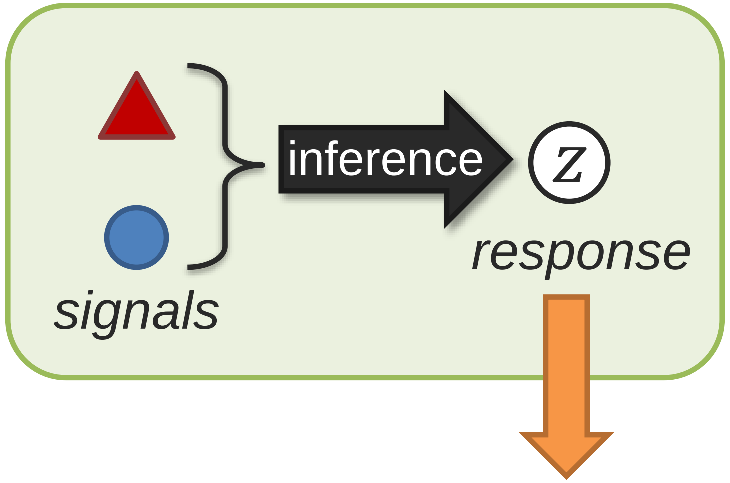
Multimodal Communication

	signal	response	signal	response	
Redundancy	a	→ □	a+b	→ □	Equivalence
	b	→ □	a+b	→ □ (larger)	Enhancement
Nonredundancy	a	→ □	a+b	→ □ and ○	Independence
	b	→ ○	a+b	→ □	Dominance
			a+b	→ □ (or □)	Modulation
			a+b	→ △	Emergence

Partan and Marler (2005). *Issues in the classification of multimodal communication signals*. American Naturalist, 166(2)

Dimensions of Modality Interactions

What are the dimensions for **digitally-represented** modalities?

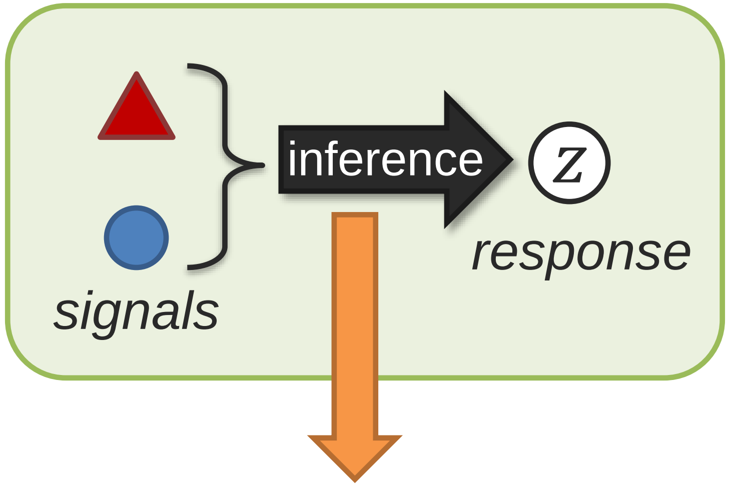


1 Interaction Responses:

- Redundancy
- Non-redundancy
- Dominance
- Emergence…

Dimensions of Modality Interactions

What are the dimensions for **digitally-represented** modalities?

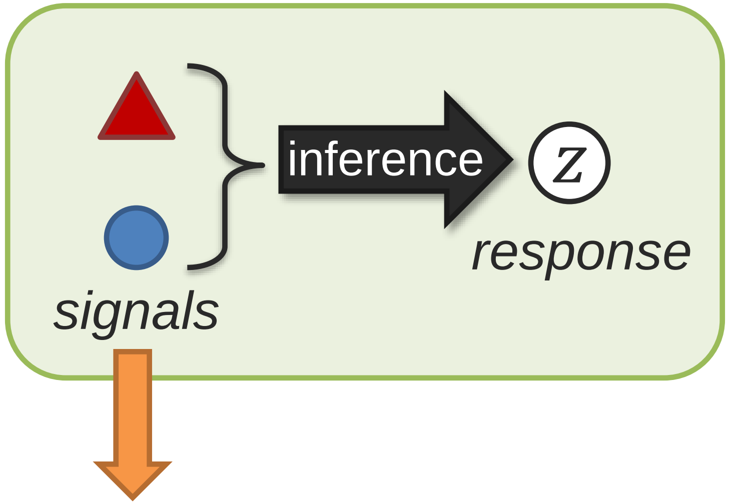


2 Interaction Mechanics:

- Additive
- multiplicative
- Nonlinear
- Causal,
- Logical, …

Dimensions of Modality Interactions

What are the dimensions for **digitally-represented** modalities?

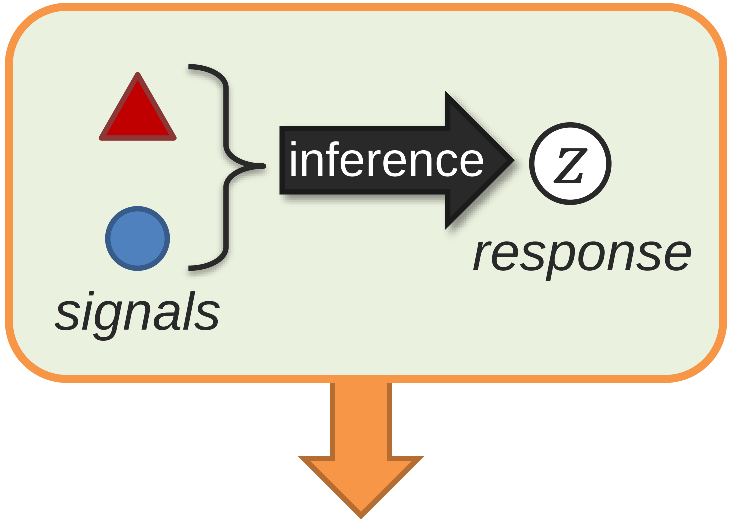


3 Input modalities:

- Unimodal
- Bimodal
- Trimodal
- High-modal, ...

Dimensions of Modality Interactions

What are the dimensions for **digitally-represented** modalities?



4 Context:

- Structure context
- Task relevance
- Context dependence
- High-modal, ...

What is Multimodal?

Multimodal **is the science of**

heterogeneous and interconnected data ☺

Multimodal Machine Learning

What is Multimodal Machine Learning?

Multimodal Machine Learning (ML) is the study of computer algorithms that learn and improve through the use and experience of data from multiple modalities

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

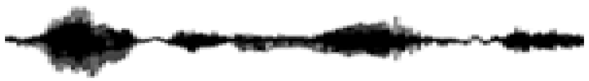
Multimodal AI is a superset of **Multimodal ML**

Multimodal Machine Learning

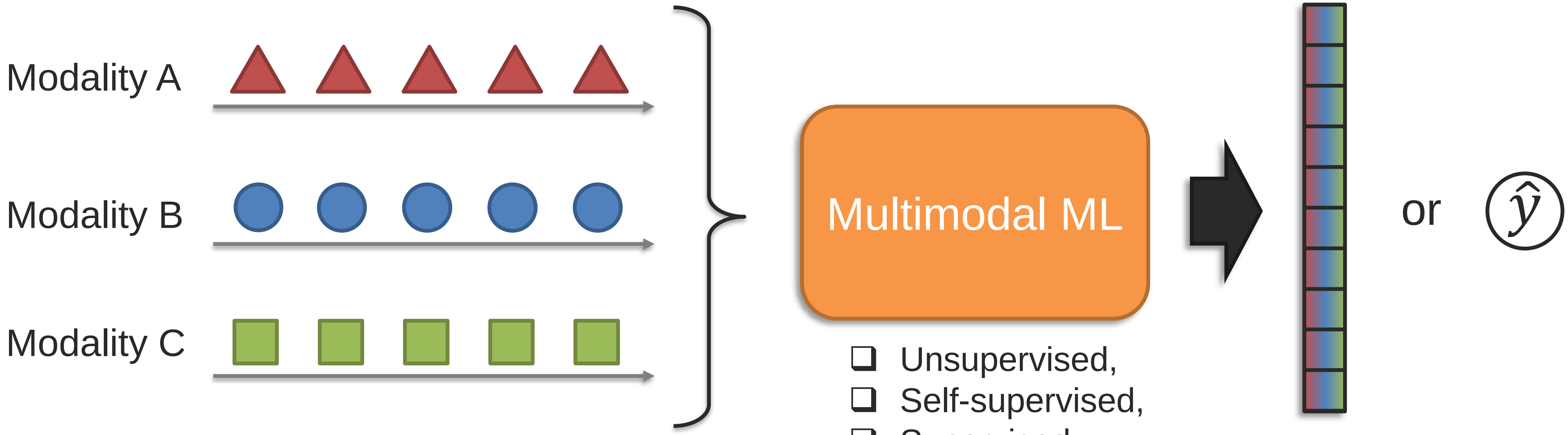
Language I really like this tutorial

Vision

Acoustic



Multimodal Machine Learning



- ❑ Unsupervised,
- ❑ Self-supervised,
- ❑ Supervised,
- ❑ Reinforcement,

…

Multimodal Machine Learning

*What are the **core multimodal technical challenges,** understudied in conventional machine learning?*

Multimodal Technical Challenges – Surveys, Tutorials and Courses

2016

Multimodal Machine Learning: A Survey and Taxonomy

Tadas Baltrusaitis, Chaitanya Ahuja and Louis-Philippe Morency

(Arxiv 2017, IEEE TPAMI journal, February 2019)

https://arxiv.org/abs/1705.09406

Tutorials: CVPR 2016, ACL 2016, ICMI 2016, …

Graduate-level courses:

Multimodal Machine learning (11th edition)

https://cmu-multicomp-lab.github.io/mmml-course/fall2020/

Advanced Topics in Multimodal Machine learning

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

2022

Fundamentals of Multimodal ML: A Taxonomy & Open Challenges

Paul Liang, Amir Zadeh and Louis-Philippe Morency

- ☑ 6 core challenges
- ☑ 50+ taxonomic classes
- ☑ 600+ referenced papers

Tutorials: CVPR 2022, NAACL 2022, …

Updated graduate-level course:

Multimodal Machine learning (12th edition)

Fall 2022 semester

Challenge 1: Representation

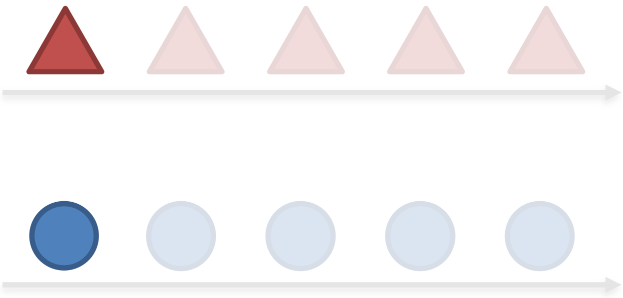
Definition: Learning representations that reflect cross-modal interactions between individual elements, across different modalities

➡ This is a core building block for most multimodal modeling problems!

Individual elements:

Modality A

Modality B



It can be seen as a "local" representation

or

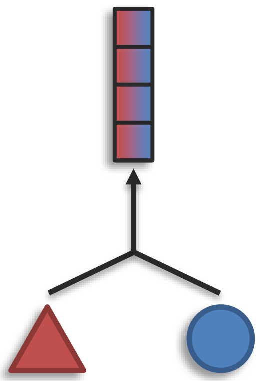
representation using holistic features

Challenge 1: Representation

Definition: Learning representations that reflect cross-modal interactions between individual elements, across different modalities

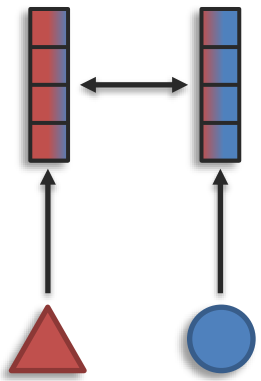
Sub-challenges:

Fusion



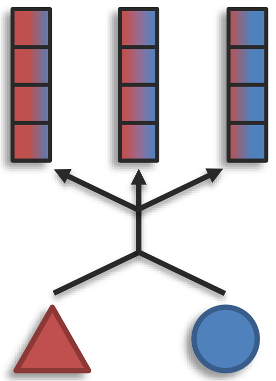
modalities > # representations

Coordination



modalities = # representations

Fission



modalities < # representations

Challenge 2: Alignment

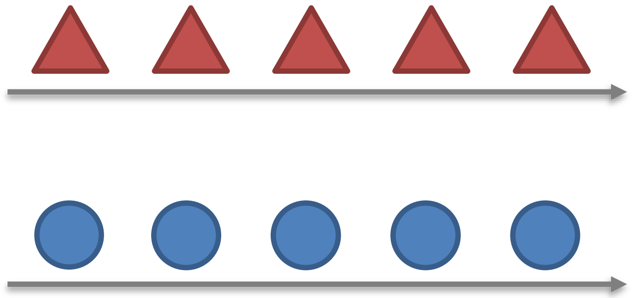
Definition: Identifying and modeling cross-modal connections between all elements of multiple modalities, building from the data structure

➡ Most modalities have internal structure with multiple elements

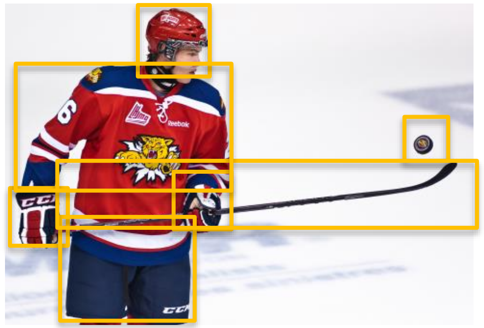
Elements with temporal structure:

Modality A

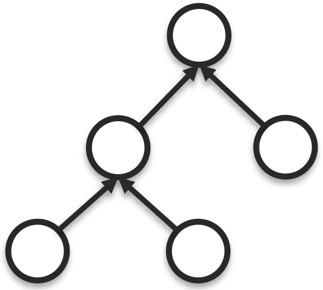
Modality B



Other structured examples:



Spatial



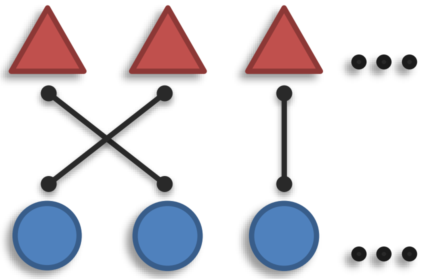
Hierarchical

Challenge 2: Alignment

Definition: Identifying and modeling cross-modal connections between all elements of multiple modalities, building from the data structure

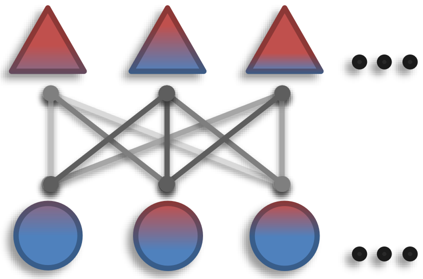
Sub-challenges:

Connections



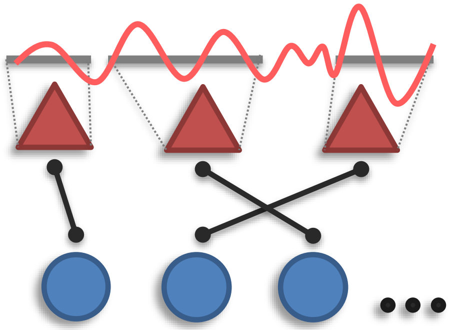
Explicit alignment
(e.g., grounding)

Aligned Representation



Alignment + representation
(aka, contextualized representation)

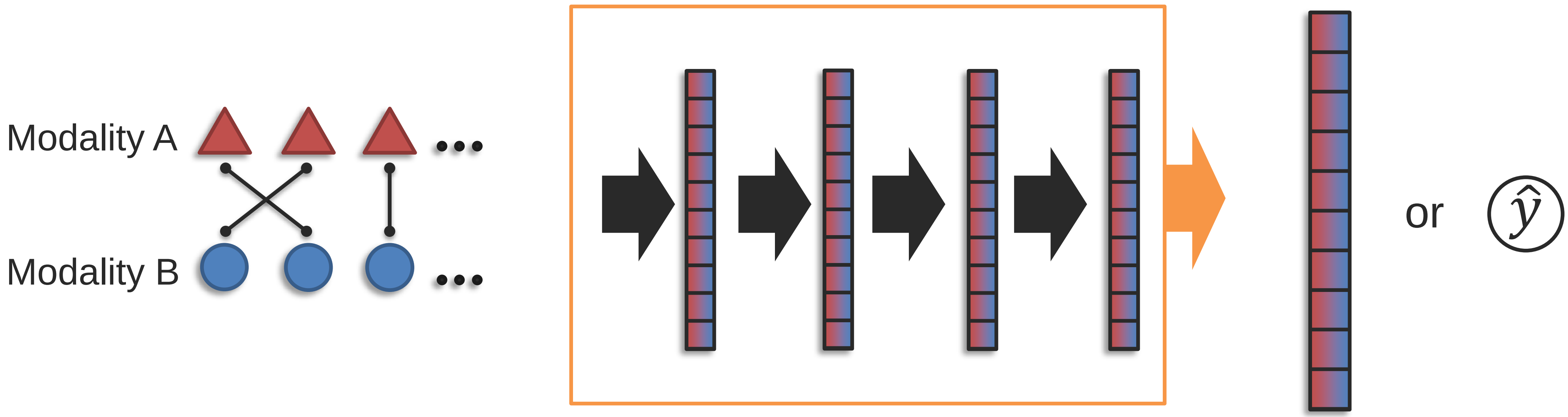
Elements



Segmentation of individual elements

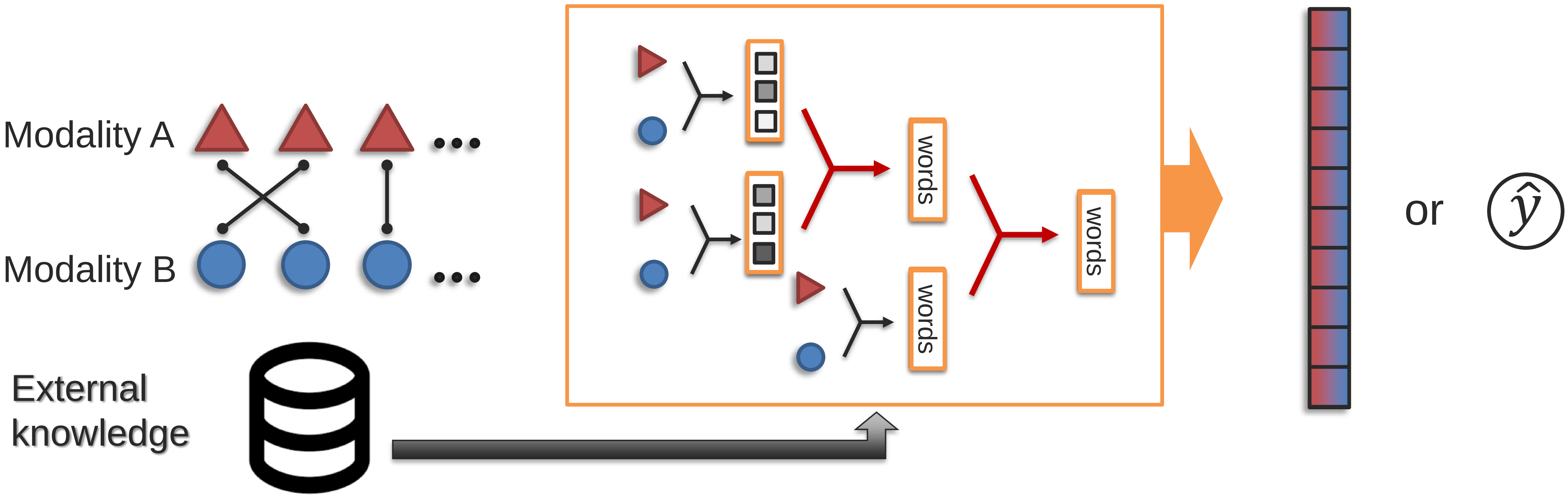
Challenge 3: Reasoning

Definition: Combining knowledge, usually through multiple inferential steps, exploiting multimodal alignment and problem structure



Challenge 3: Reasoning

Definition: Combining knowledge, usually through multiple inferential steps, exploiting multimodal alignment and problem structure

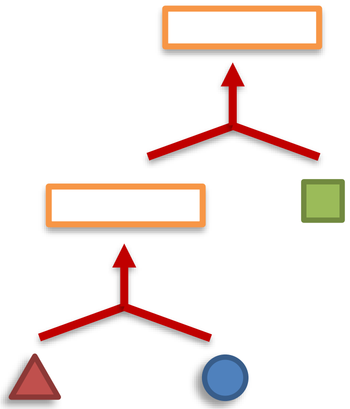


Challenge 3: Reasoning

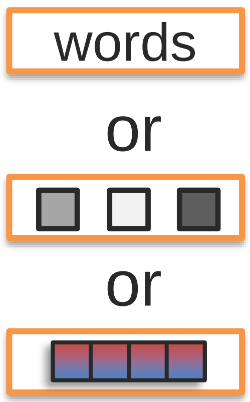
Definition: Combining knowledge, usually through multiple inferential steps, exploiting multimodal alignment and problem structure

Sub-challenges:

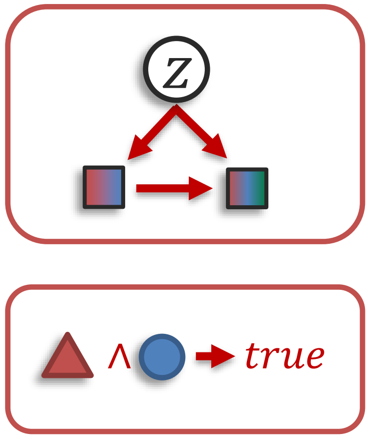
Structure Modeling



Intermediate concepts



Inference Paradigm



External Knowledge

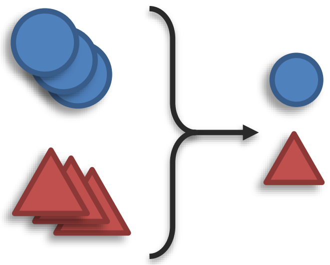
Challenge 4: Generation

Definition: Learning a generative process to produce raw modalities that reflects cross-modal interactions, structure and coherence

Sub-challenges:

Information: (content)

Summarization



Reduction

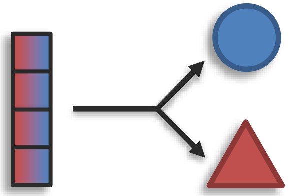
□ > □

Translation

Maintenance

□ = □

Creation

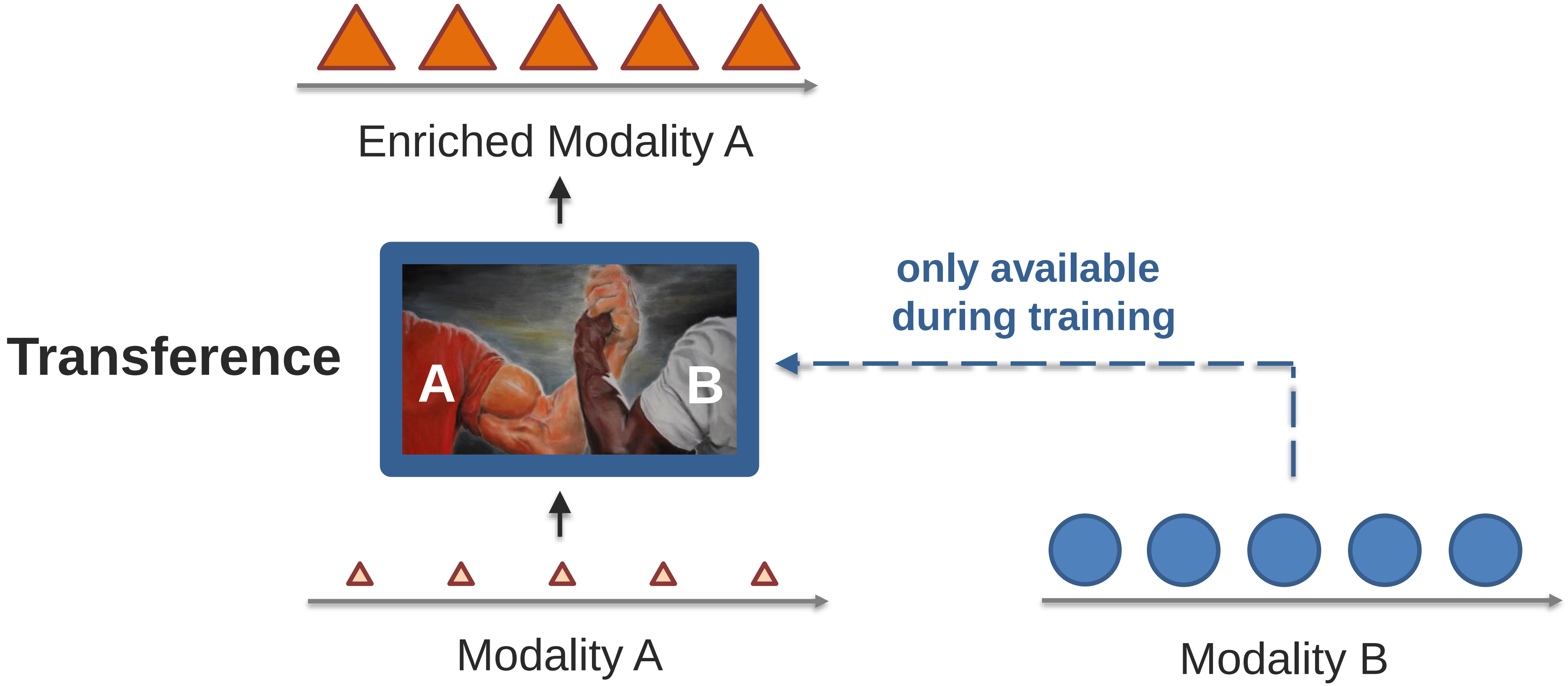


Expansion

□ < □

Challenge 5: Transference

Definition: Transfer knowledge between modalities, usually to help the target modality which may be noisy or with limited resources

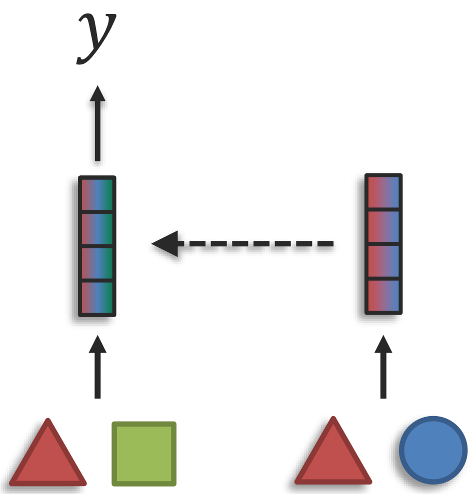


Challenge 5: Transference

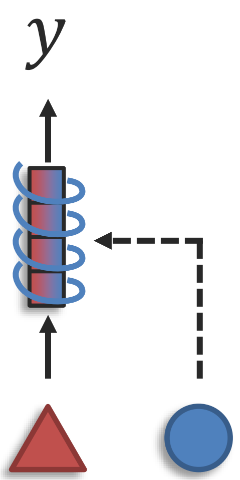
Definition: Transfer knowledge between modalities, usually to help the target modality which may be noisy or with limited resources

Sub-challenges:

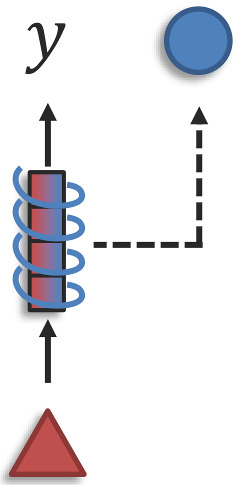
Transfer



Co-learning
via representation



Co-learning
via generation

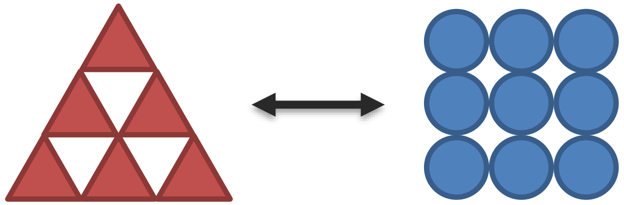


Challenge 6: Quantification

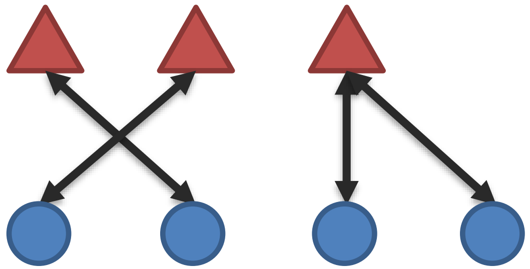
Definition: Empirical and theoretical study to better understand heterogeneity, cross-modal interactions and the multimodal learning process

Sub-challenges:

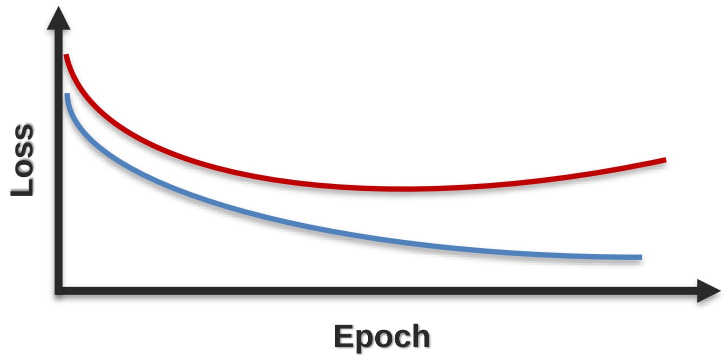
Heterogeneity



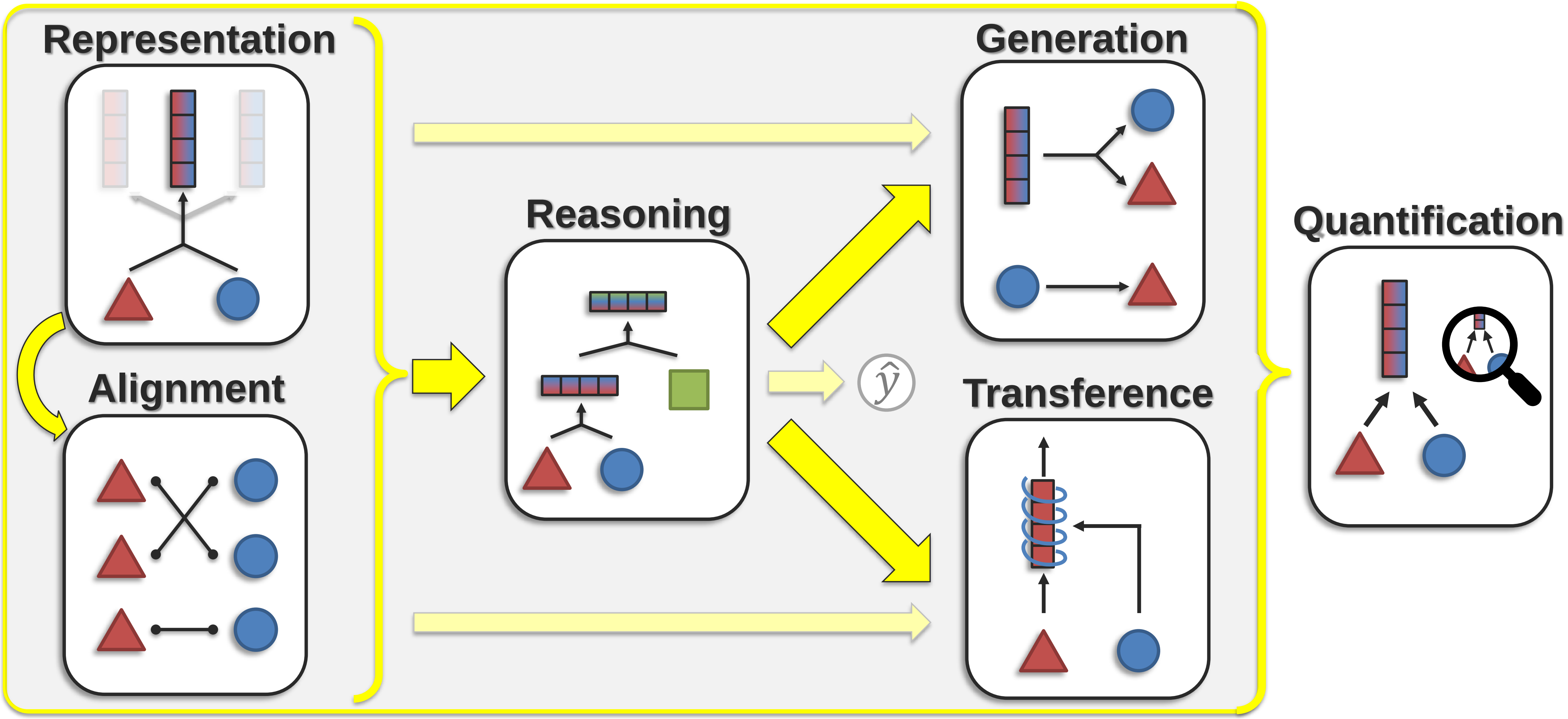
Interactions



Learning



Core Multimodal Challenges



Lecture Schedule

Classes	Tuesday Lectures	Thursday Lectures
Week 1 8/30 & 9/1	**Course introduction** • Multimodal core challenges • Course syllabus	**Multimodal applications and datasets** • Research tasks and datasets • Team projects
Week 2 9/6 & 9/8 Read due: 9/9	**Basic concepts: neural networks** • Loss functions and neural networks • Gradient and optimization	**Unimodal representations** • Dimensions of heterogeneity • Visual representations
Week 3 9/13 & 9/15 Read due: 9/16 Proj. Due: 9/14	**Unimodal representations** • Language representations • Signals, graphs and other modalities	**Multimodal representations** • Cross-modal interactions • Multimodal fusion
Week 4 9/20 & 9/22 *Proj. due: 9/25*	**Multimodal representations** • Coordinated representations • Multimodal fission	**Multimodal alignment** • Explicit alignment • Multimodal grounding
Week 5 9/27 & 9/29 Read due: 9/30	***Project hours (Research ideas)***	**Aligned representations** • Self-attention transformer models • Masking and self-supervised learning
Week 6 10/4 & 10/6 *Proj. due: 10/9*	**Multimodal aligned representations** • Multimodal transformers • Video and graph representations	**Multimodal Reasoning** • Structured and hierarchical models • Memory models

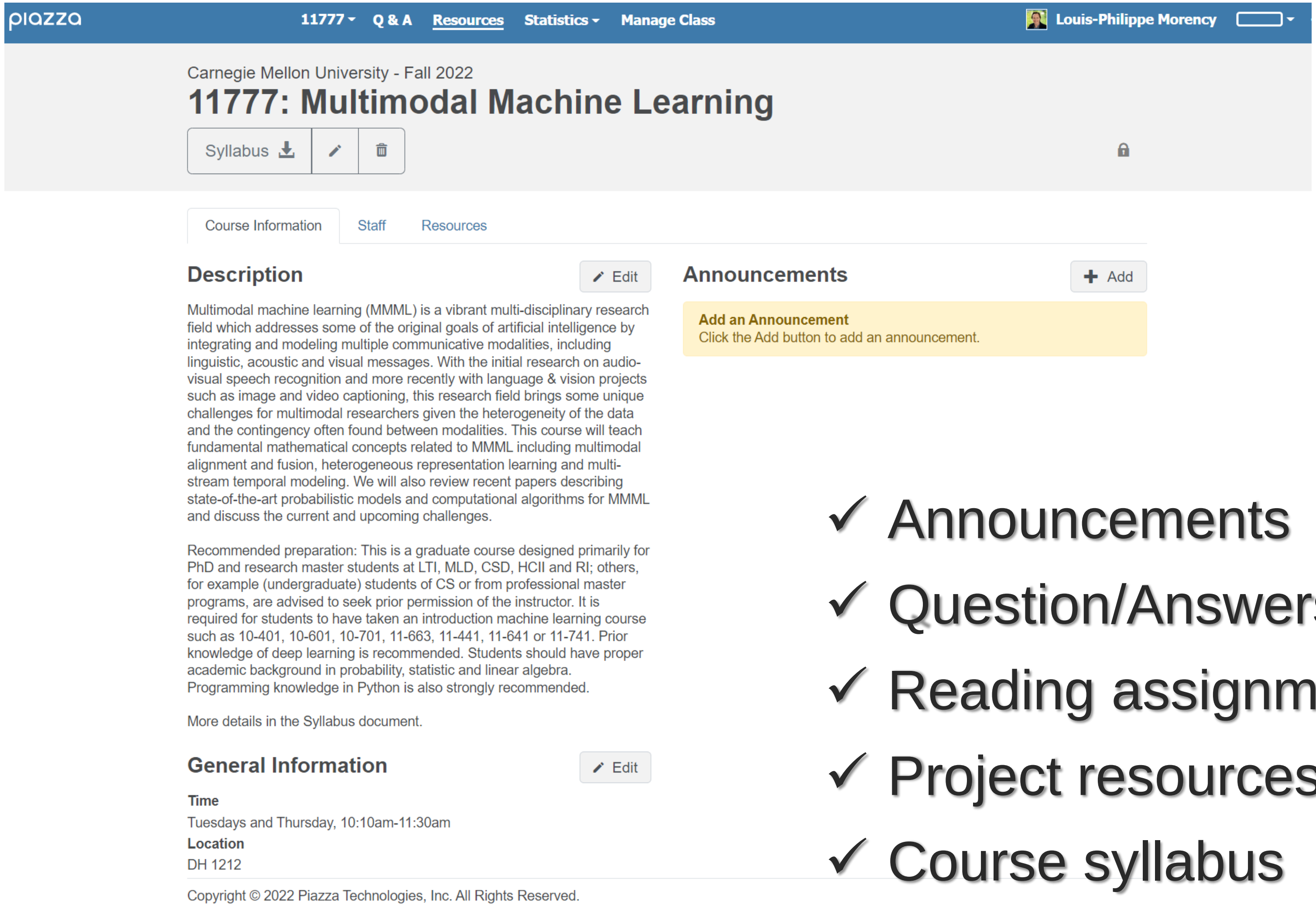
Lecture Schedule

Classes	Tuesday Lectures	Thursday Lectures
Week 7 10/11 & 10/13 Read due: 10/14	**Multimodal Reasoning** • Reinforcement learning • Discrete structure learning	**Multimodal Reasoning** • Logical and causal inference • External knowledge
Week 8 10/18 & 10/20	***Fall Break – No lectures***	
Week 9 10/25 & 10/27 *Proj. due: 10/30*	**Generation** • Translation, summarization, creation • Generative models: VAEs	**Generation** • GANs and diffusion models • Model evaluation and ethics
Week 10 11/1 & 11/3	***Project presentations (midterm)***	***Project presentations (midterm)***
Week 11 11/8 & 11/10 Read due: 11/12	**Transference** • Modality transfer • Multimodal co-learning	**Quantification** • Heterogeneity and interactions • Biases and fairness
Week 12 11/15 & 11/17 Read due: 11/21	***Project hours (Research ideas)***	**New research directions** • Recent approaches in multimodal ML

Lecture Schedule

Classes	Tuesday Lectures	Thursday Lectures
Week 13 11/22 & 11/24	***Thanksgiving Week – No Class –***	
Week 14 11/30 & 12/2	***Language, Vision, and Actions*** • Motion and navigation • Robots and embodied AI	***Multimodal Applications*** • Healthcare and affective computing • Artificial social intelligence
Week 15 12/6 & 12/8 *Proj. due: 12/11*	***Project presentations (final)***	***Project presentations (final)***

Piazza https://piazza.com/cmu/fall2022/11777/info



piazza 11777 Q & A Resources Statistics Manage Class Louis-Philippe Morency

Carnegie Mellon University - Fall 2022

11777: Multimodal Machine Learning

Syllabus

Course Information Staff Resources

Description Edit

Multimodal machine learning (MMML) is a vibrant multi-disciplinary research field which addresses some of the original goals of artificial intelligence by integrating and modeling multiple communicative modalities, including linguistic, acoustic and visual messages. With the initial research on audio-visual speech recognition and more recently with language & vision projects such as image and video captioning, this research field brings some unique challenges for multimodal researchers given the heterogeneity of the data and the contingency often found between modalities. This course will teach fundamental mathematical concepts related to MMML including multimodal alignment and fusion, heterogeneous representation learning and multi-stream temporal modeling. We will also review recent papers describing state-of-the-art probabilistic models and computational algorithms for MMML and discuss the current and upcoming challenges.

Recommended preparation: This is a graduate course designed primarily for PhD and research master students at LTI, MLD, CSD, HCII and RI; others, for example (undergraduate) students of CS or from professional master programs, are advised to seek prior permission of the instructor. It is required for students to have taken an introduction machine learning course such as 10-401, 10-601, 10-701, 11-663, 11-441, 11-641 or 11-741. Prior knowledge of deep learning is recommended. Students should have proper academic background in probability, statistic and linear algebra. Programming knowledge in Python is also strongly recommended.

More details in the Syllabus document.

General Information Edit

Time

Tuesdays and Thursday, 10:10am-11:30am

Location

DH 1212

Announcements Add

Add an Announcement

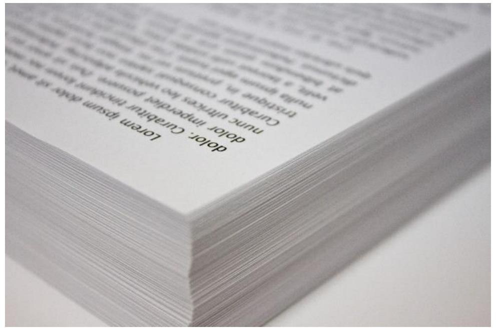
Click the Add button to add an announcement.

- ✓ Announcements
- ✓ Question/Answers
- ✓ Reading assignments
- ✓ Project resources
- ✓ Course syllabus

Course Syllabus

Three Course Learning Paradigms

Course lecture participation
(16% of your grade)



Reading assignments
(12% of your grade)

$$
\begin{aligned}
i_t &= \sigma\left(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i\right)\\
f_t &= \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f\right)\\
c_t &= f_t c_{t-1} + i_t \tanh\left(W_{xc}x_t + W_{hc}h_{t-1} + b_c\right)\\
o_t &= \sigma\left(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o\right)\\
h_t &= o_t \tanh(c_t)
\end{aligned}
$$

Course project assignments
(72% of your grade)

Course Recommendations and Requirements

1. Ready to read about 6 papers this semester !
   - Curated list of research papers for the 6 reading assignments
   - Summarize one paper and contrast it with other papers
2. Already taken a machine learning course
   - Strongly recommended for students to have taken an introduction machine learning course
   - 10-401, 10-601, 10-701, 11-663, 11-441, 11-641 or 11-741
3. Motivated to produce a high-quality course project
   - Projects are designed to enhance state-of-the-art algorithms
   - Four project assignments, to help scaffold the project tasks

Course Project Guidelines

- Dataset should have at least two modalities:
  - Natural language and visual/images
- Teams of 3, 4 or 5 students
- The project should explore algorithmic novelty
- Possible venues for your final report:
  - ACL 2023, IJCAI 2023, ICML 2023, ICMI 2023
- We will discuss on Thursday about project ideas
- GPU resources available:
  - Amazon AWS and Google Cloud Platform

$$
\begin{aligned}
i_t &= \sigma\left(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i\right)\\
f_t &= \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f\right)\\
c_t &= f_t c_{t-1} + i_t \tanh\left(W_{xc}x_t + W_{hc}h_{t-1} + b_c\right)\\
o_t &= \sigma\left(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o\right)\\
h_t &= o_t \tanh(c_t)
\end{aligned}
$$

Course Project Timeline

Pre-proposal *(due Wednesday Sept. 14)*

- Define your dataset, research task and teammates

First project assignment *(due Sunday Sept. 25)*

- Study related work to your selected research topic

Second project assignment *(due Sunday Oct 9)*

- Experiment with unimodal representations

Midterm project assignment (due Sunday Oct 30)

- Implement and evaluate state-of-the-art model(s)

Final project assignment (due Sunday Dec. 11)

- Implement and evaluate new research ideas

# Equal Contribution by All Teammates!

- Each team will be required to create a GitHub repository which will be accessible by TAs
- Each report should include a description of the task from each teammate
- Please let us know soon if you have concerns about the participation levels of your teammates

Process for Selecting your Course Project

- **Thursday 9/1:** Lecture describing available multimodal datasets and research topics
- **Tuesday 9/6:** Let us know your dataset preferences for the course project
- **Thursday 9/8:** During the later part of the lecture, we will have an interactive period to help with team formation. More details to come
- **Wednesday 9/14:** Pre-proposals are due. You should have selected your teammates, dataset and task

Project Preferences – Due Tuesday 9/6

- Post your project preferences:
  - List of your ranked preferred projects
    - Use alphanumeric code of each dataset
    - Detailed dataset list in the "Lecture1.2-datasets“ slides
  - Previous unimodal/multimodal experience
  - Available CPU / GPU resources
- For topics or datasets not in the list:
  - Include a description with links (for other students)

Lecture Schedule

Classes	Tuesday Lectures	Thursday Lectures
Week 1 8/30 & 9/1	**Course introduction** • Multimodal core challenges • Course syllabus	**Multimodal applications and datasets** • Research tasks and datasets • Team projects
Week 2 9/6 & 9/8 Read due: 9/9	**Basic concepts: neural networks** • Loss functions and neural networks • Gradient and optimization	**Unimodal representations** • Dimensions of heterogeneity • Visual representations
Week 3 9/13 & 9/15 Read due: 9/16 Proj. Due: 9/14	**Unimodal representations** • Language representations • Signals, graphs and other modalities	**Multimodal representations** • Cross-modal interactions • Multimodal fusion
Week 4 9/20 & 9/22 *Proj. due: 9/25*	**Multimodal representations** • Coordinated representations • Multimodal fission	**Multimodal alignment and groun** • Explicit alignment • Multimodal grounding
Week 5 9/27 & 9/29 Read due: 9/30	***Project hours (Research ideas)***	**Aligned representations** • Self-attention transformer models • Masking and self-supervised learning
Week 6 10/4 & 10/6 *Proj. due: 10/9*	**Multimodal aligned representations** • Multimodal transformers • Video and graph representations	**Multimodal Reasoning** • Structured and hierarchical mo • Memory models

Project preferences due on Tuesday 9/6

Pre-proposals due on Wednesday 9/14

First assignment due on Sunday 9/25

Second assignment due on Sunday 10/9

Lecture Schedule

Classes	Tuesday Lectures	Thursday Lectures
Week 7 10/11 & 10/13 Read due: 10/14	**Multimodal Reasoning** • Reinforcement learning • Discrete structure learning	**Multimodal Reasoning** • Logical and causal inference • External knowledge
Week 8 10/18 & 10/20	***Fall Break – No lectures***	
Week 9 10/25 & 10/27 *Proj. due: 10/30*	**Generation** • Translation, summarization, creation • Generative models: VAEs	**Generation** • GANs and diffusion models • Model evaluation and ethics
Week 10 11/1 & 11/3	***Project presentations (midterm)***	***Project presentations (midterm)***
Week 11 11/8 & 11/10 Read due: 11/12	**Transference** • Modality transfer • Multimodal co-learning	**Quantification** • Heterogeneity and interactions • Biases and fairness
Week 12 11/15 & 11/17 Read due: 11/21	***Project hours (Research ideas)***	**New research directions** • Recent approaches in multimodal ML

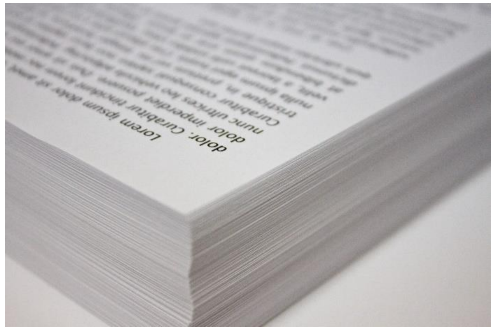
Midterm assignment due on Sunday 10/30

Lecture Schedule

Classes	Tuesday Lectures	Thursday Lectures
Week 13 11/22 & 11/24	***Thanksgiving Week – No Class –***	
Week 14 11/30 & 12/2	***Language, Vision, and Actions*** • Motion and navigation • Robots and embodied AI	***Multimodal Applications*** • Healthcare and affective computing • Artificial social intelligence
Week 15 12/6 & 12/8 *Proj. due: 12/11*	***Project presentations (final)***	***Project presentations (final)***

Final assignment due on Sunday 12/11

Course Grades



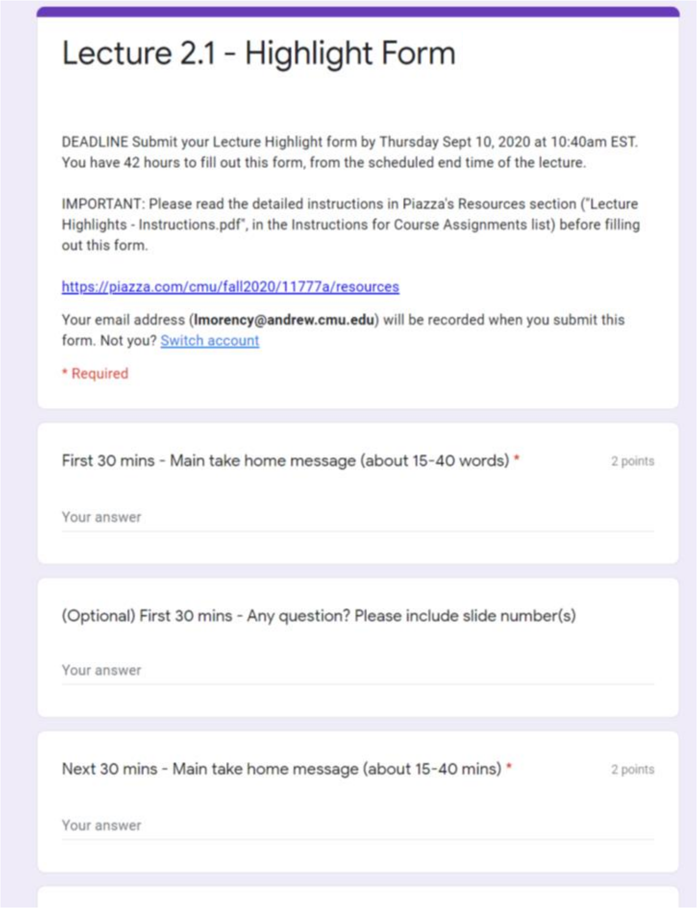
$$i_t = \sigma\left(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i\right)$$
$$f_t = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f\right)$$
$$c_t = f_t c_{t-1} + i_t \tanh\left(W_{xc}x_t + W_{hc}h_{t-1} + b_c\right)$$
$$o_t = \sigma\left(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o\right)$$
$$h_t = o_t \tanh(c_t)$$

- Lecture highlights 16%
- Reading assignments 12%

- Project preferences/pre-proposal 2%
- First project assignment 10%
- Second project assignment 10%
- Mid-term project assignment
  - Report and presentation 20%
- Final project assignment
  - Report and presentation 30%

Lecture Highlight Form

Starting Week 2 !!


Lecture 2.1 - Highlight Form

DEADLINE Submit your Lecture Highlight form by Thursday Sept 10, 2020 at 10:40am EST. You have 42 hours to fill out this form, from the scheduled end time of the lecture.

IMPORTANT: Please read the detailed instructions in Piazza's Resources section ("Lecture Highlights - Instructions.pdf", in the Instructions for Course Assignments list) before filling out this form.

https://piazza.com/cmu/fall2020/11777a/resources

Your email address (**lmorency@andrew.cmu.edu**) will be recorded when you submit this form. Not you? Switch account

* Required

First 30 mins - Main take home message (about 15-40 words) * 2 points

Your answer

(Optional) First 30 mins - Any question? Please include slide number(s)

Your answer

Next 30 mins - Main take home message (about 15-40 mins) * 2 points

Your answer

Similar to note-taking during lectures

- For each course segment (30mins):
  2 sentences describing the main points

Help you summarizing the lecture

- What is the main take-away message from the lecture
  Short paragraph (15-40 words)

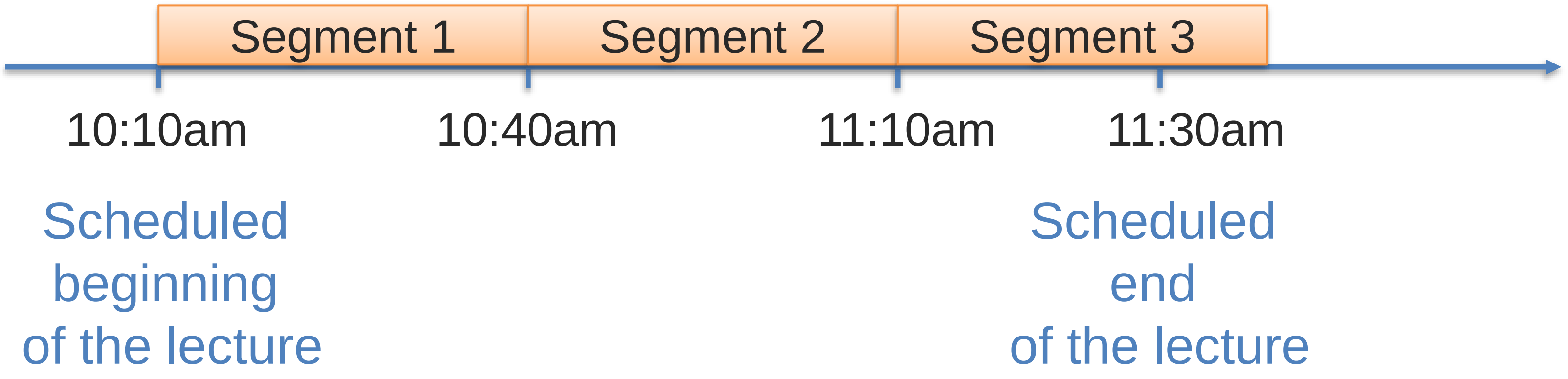
Ask questions about the lecture

- Will be answered either online or at the next lecture

Submitted same day as lecture (before 11:59pm)

- Students are encouraged to attend lectures in person

Lecture Highlight Form - Segments



- Segment 1 starts at 10:10am, even if the lecture starts slightly later.
- Segment 3 ends whenever the lecture ends
- Slides happening around the segment borders (+/- 5min of 10:40am and 11:10am) can be included in either neighboring segment.

First Reading Assignment – Week 2

- Study groups: 9-10 students per group (randomly, in Piazza)
- 3 paper options are available
  - **Each student should pick one paper option!**
    - Google Sheets were created to help balance the papers between group members
  - Then you will create a short summary to help others [1 point]
- Discussions with your study group
  - Read other’s summaries. Ask questions!
  - Write follow-up posts comparing the papers and suggesting ideas [1 point]
    - At least one follow-up post for every paper you did not read

First Reading Assignment – Week 2

Four main steps for the reading assignments

1. **Monday 8pm:** Official start of the assignment
2. **Wednesday 8pm:** Select your paper
3. **Friday 8pm:** Post your summary
4. **Monday 8pm:** Post your follow-up posts

Detailed instructions posted on Piazza

https://piazza.com/cmu/fall2022/11777/resources

Late Submissions and Wildcards

- Each student has **6** late submission wildcards
  - For lecture highlight forms or reading assignments
- Each project team has **2** late submission wildcards
  - For any of the project assignments
- Total number of wildcards: 8 (6 individual and 2 team-level)
- Each wildcard gives 24-hour extension
  - No partial credits for the wildcards
  - Automatically calculated (no need to contact us apriori)

See details about late submission policy in syllabus

https://piazza.com/cmu/fall2021/11777/resources

Spring 2023 Edition of the MMML Course !

Yonatan Bisk
ybisk@cs.cmu.edu

https://yonatanbisk.com/

Daniel Fried
dfried@cs.cmu.edu

https://dpfried.github.io/