



Language  
Technologies  
Institute

Carnegie  
Mellon  
University

# Multimodal Affective Computing

## Lecture 1.1: Introduction

Louis-Philippe Morency  
Jeffrey Girard

# Your Instructors This Semester (11-776)

---



**Louis-Philippe Morency**

[morency@cs.cmu.edu](mailto:morency@cs.cmu.edu)

Office: GHC-5411



**Jeffrey Girard**

[jmgirard@cmu.edu](mailto:jmgirard@cmu.edu)

Office: GHC- 5707



# Your TAs This Semester (11-776)

---



**Alexandria Vail**  
[avail@cs.cmu.edu](mailto:avail@cs.cmu.edu)



**Keith Maki**  
[kmaki@andrew.cmu.edu](mailto:kmaki@andrew.cmu.edu)



# Lecture Objectives

---

- Introductions
- Human multimodal communication
  - Behaviors, multimodal and interpersonal
- Multimodal Affective Computing
  - A historical view
  - Psychological constructs
- Course syllabus and project assignments
  - Grades and course structure
  - Course project



# Human Multimodal Communication



# Human Multimodal Communication : An Early Process

---



[Complements of Jacqueline Nadel]

# Human Multimodal Communication

---

## Multimodal

- Audio
- Visual
- Verbal



- “[...] music online. Guy Kewney is the editor of the technology website News Wireless. Hello, good morning to you.”
- “Good morning.”
- “Were you surprised by this verdict today?”
- “I’m very surprised to see this verdict to come on me. Because I was not expecting that. When I came they told me something else and I’m coming. And they told me something else. Big surprise any way.”
- “A big surprise...”
- “Exactly.”
- “Yeah yeah. With regard to the cost that is involved. Do you think more people will be downloading online?”
- “Actually if you can go everywhere, you gonna see people downloading through the internet and the websites. [...]”

# Human Multimodal Communication

---

## Multimodal

- Audio
- Visual
- Verbal





# Human Multimodal Communication

## Multimodal

- Audio
- Visual
- Verbal



### Verbal

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

### Vocal

- **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

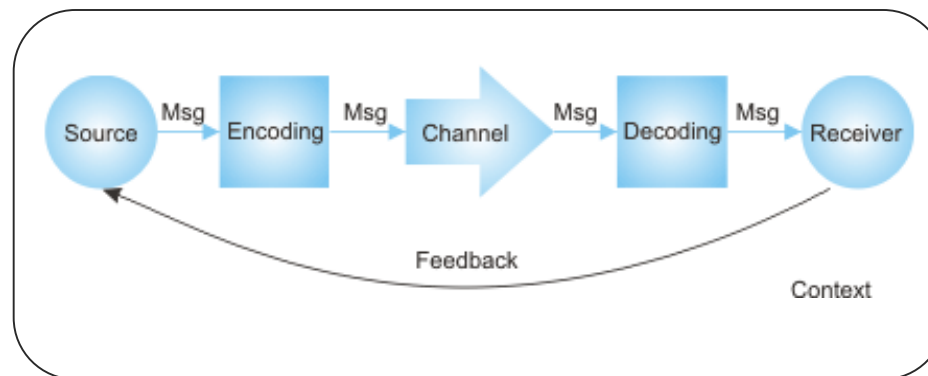
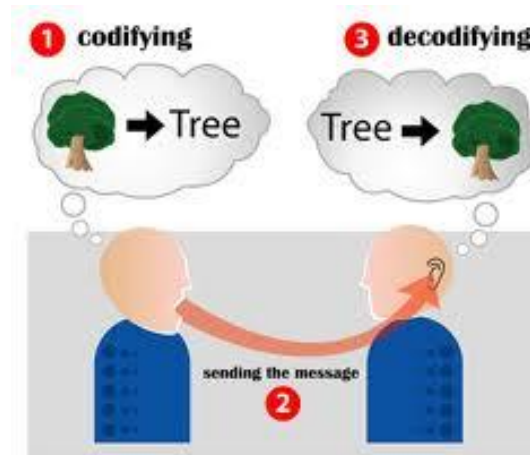


# Interpersonal (Dyadic) Communication

---

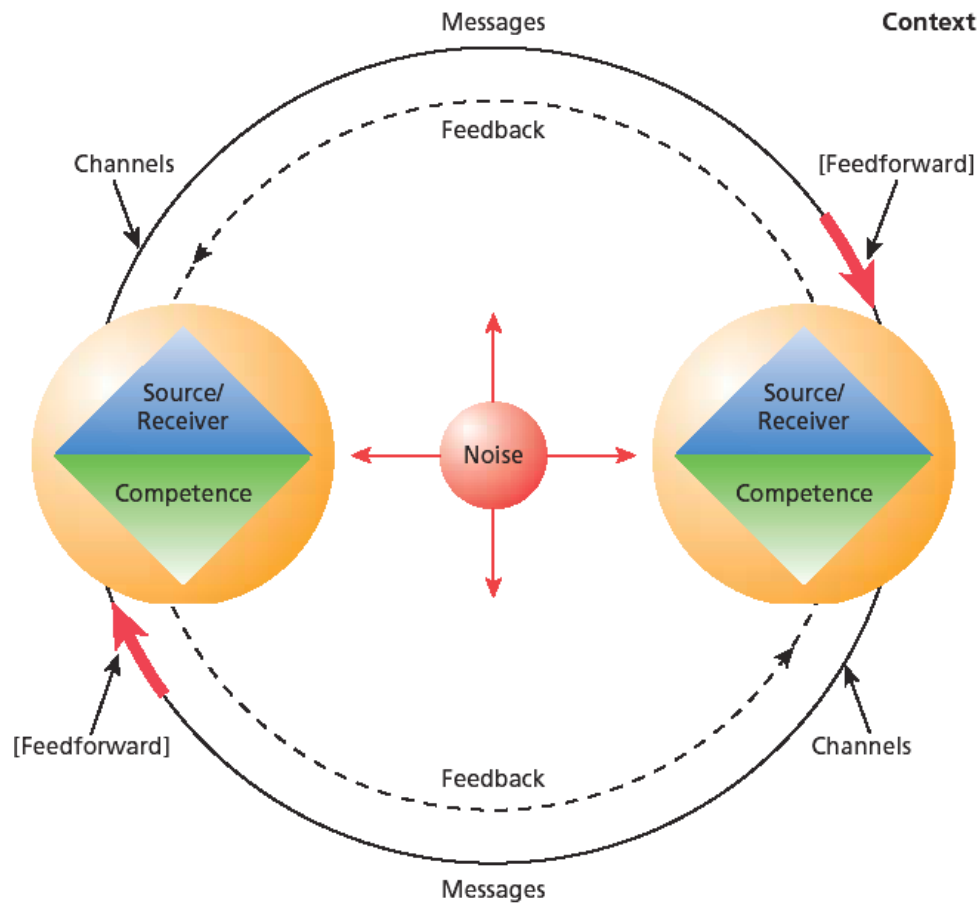


# Communication Process: Encoder-decoder



# Elements of Interpersonal Communication

---



1. Source-Receiver

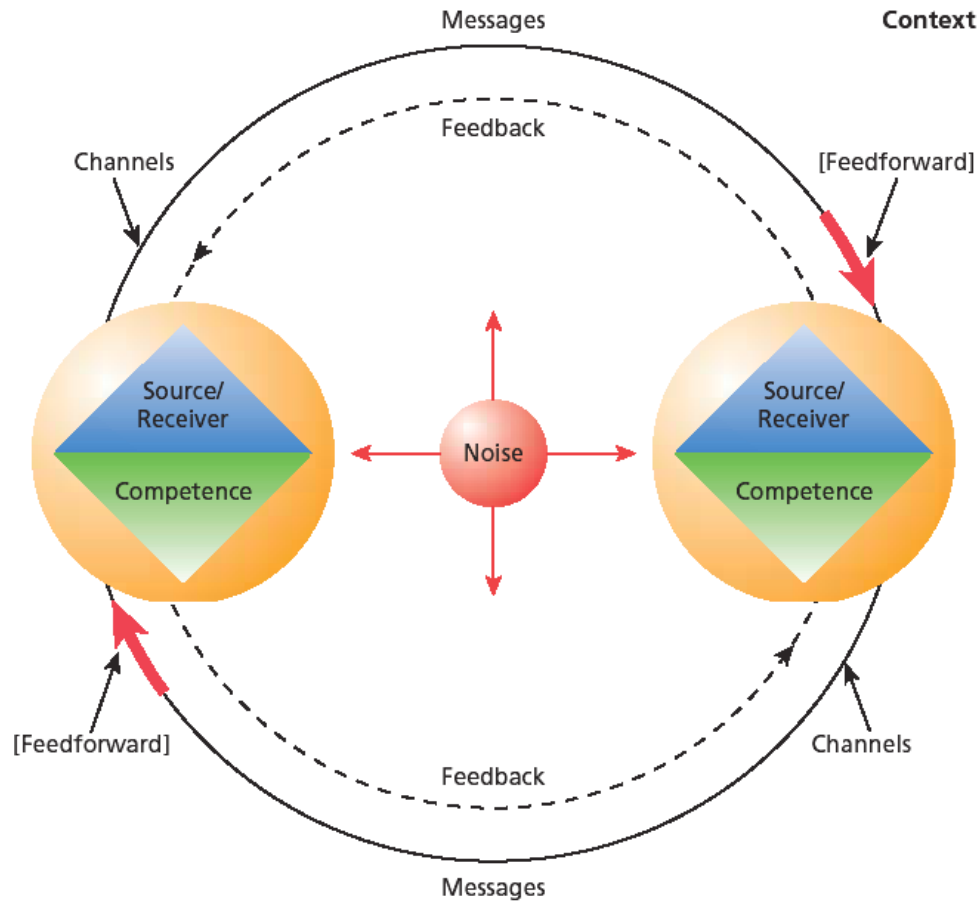
2. Channels

3. Messages

4. Feedback

Messages

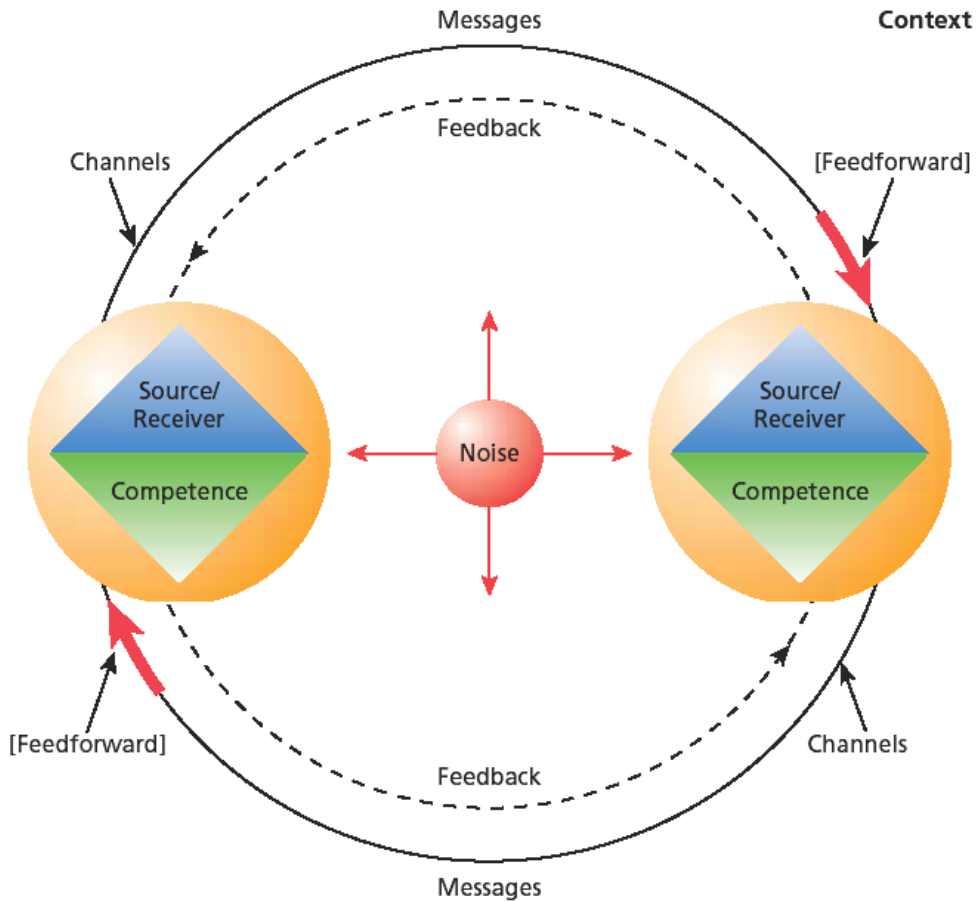
# Elements of Interpersonal Communication



## 5. Types of Noise

- Physical
- Physiological
- Psychological
- Semantic

# Elements of Interpersonal Communication



## 6. Context

- Physical dimension
- Temporal dimension
- Social-psychological dimension
- Cultural context

## 7. Competence

# Diversity in Dyadic Interactions

---



# Multimodal Affective Computing: A Historical View





# Language and Gestures

---



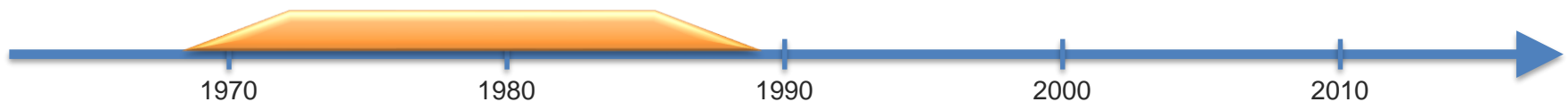
**David McNeill**

University of Chicago

Center for Gesture and Speech Research

*“For McNeill, gestures are in effect the speaker’s thought in action, and integral components of speech, not merely accompaniments or additions.”*

❑ TRIVIA: Justine Cassell was a student of David McNeill

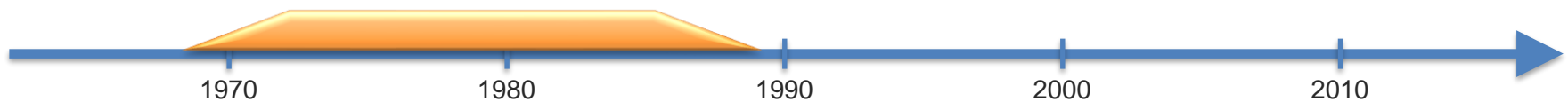


# The McGurk Effect (1976)

---



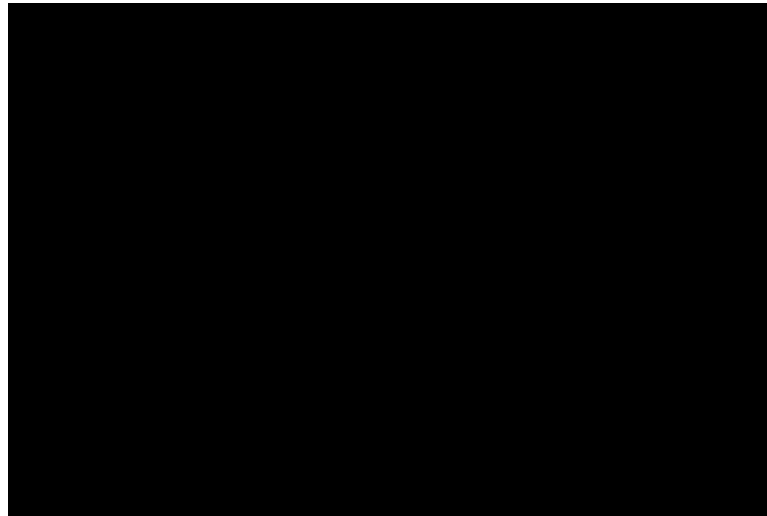
Hearing lips and seeing voices – Nature



## ➤ The “Computational” Era (Late 1980s until 2000)

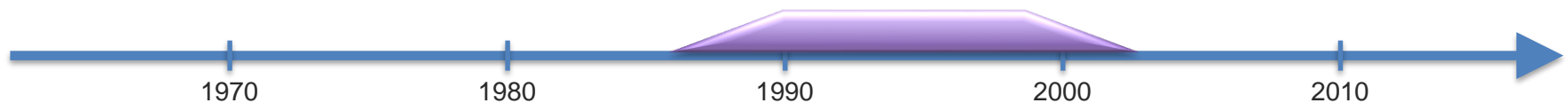
---

### Multimodal/multisensory interfaces (HCI)



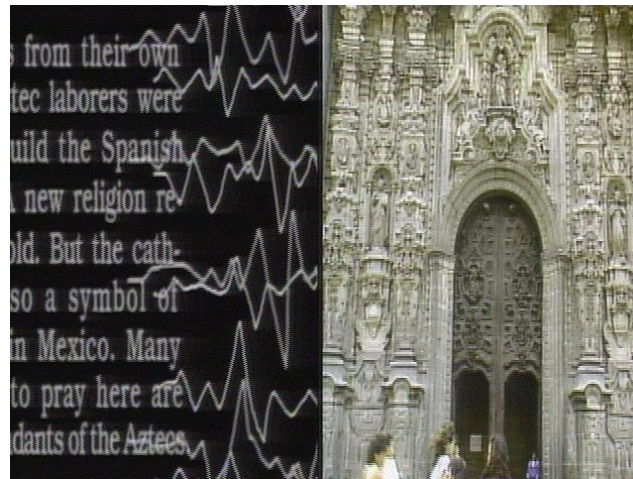
#### pFinder: Real-time Tracking of human body

by C. Wren, A. Azarbayejani, T. Darrell and A. Pentland [1995]



## ➤ The “Computational” Era (Late 1980s until 2000)

### Multimedia Computing

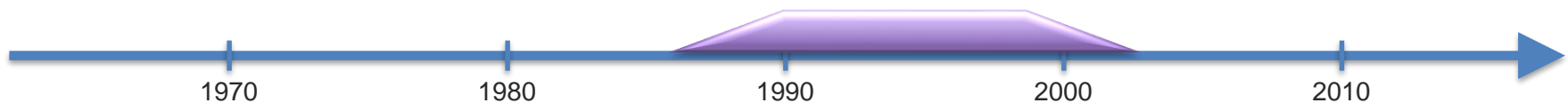


**Carnegie  
Mellon  
University**



[1994-2010]

*“The Informedia Digital Video Library Project automatically combines speech, image and natural language understanding to create a full-content searchable digital video library.”*



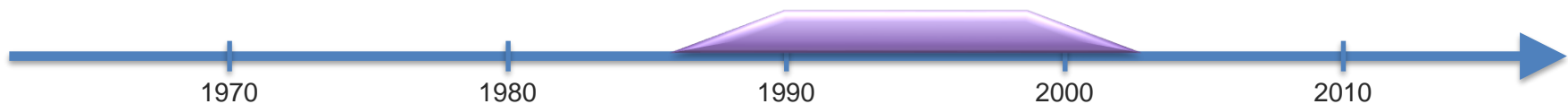
## ➤ The “Computational” Era (Late 1980s until 2000)

---

### Multimedia Computing

#### Multimedia content analysis

- **Shot-boundary detection (1991 - )**
  - Parsing a video into continuous camera shots
- **Still and dynamic video abstracts (1992 - )**
  - Making video browsable via representative frames (keyframes)
  - Generating short clips carrying the essence of the video content
- **High-level parsing (1997 - )**
  - Parsing a video into semantically meaningful segments
- **Automatic annotation (indexing) (1999 - )**
  - Detecting prespecified events/scenes/objects in video



## ➤ The “Computational” Era (Late 1980s until 2000)

---

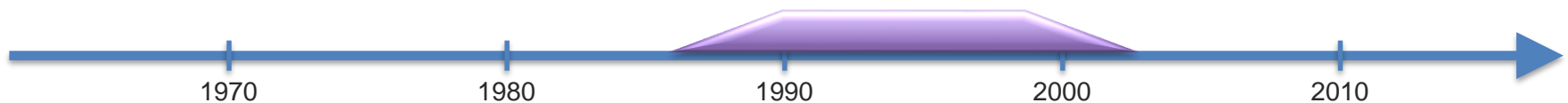
### Affective Computing



Rosalind Picard

***Affective Computing*** is computing that relates to, arises from, or deliberately influences emotion or other affective phenomena.

❑ TRIVIA: Rosalind Picard came from the same group as pFinder (MIT, Sandy Pentland)



## ➤ The “Interaction” Era (2000s)

---

### Modeling Human Multimodal Interaction



#### **AMI Project** [2001-2006, IDIAP]

- 100+ hours of meeting recordings
- Fully synchronized audio-video
- Transcribed and annotated



#### **CHIL Project** [Alex Waibel]

- Computers in the Human Interaction Loop
- Multi-sensor multimodal processing
- Face-to-face interactions

❑ TRIVIA: Samy Bengio started at IDIAP working on AMI project



## ➤ The “Interaction” Era (2000s)

---

### Modeling Human Multimodal Interaction



#### **CALO Project** [2003-2008, SRI]

- Cognitive Assistant that Learns and Organizes
- Personalized Assistant that Learns (PAL)
- Siri was a spinoff from this project

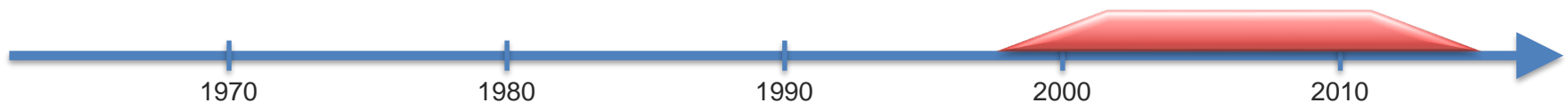


Social Signal Processing Network

#### **SSP Project** [2008-2011, IDIAP]

- Social Signal Processing
- First coined by Sandy Pentland in 2007
- Great dataset repository: <http://sspnet.eu/>

❑ TRIVIA: LP's PhD research was partially funded by CALO ☺





## ➤ The “Interaction” Era (2000s)

---

Many new challenges and multimodal corpora !!

### Audio-Visual Emotion Challenge (AVEC, 2011- )



- Emotional dimension estimation
- Standardized training and test sets
- Based on the SEMAINE dataset

### Emotion Recognition in the Wild Challenge (EmotiW 2013- )



- Emotional dimension estimation
- Standardized training and test sets
- Based on the SEMAINE dataset



## ➤ The “deep learning” era (2010s until ...)

---

### Representation learning (a.k.a. deep learning)

- Multimodal deep learning [ICML 2011]
- Multimodal Learning with Deep Boltzmann Machines [NIPS 2012]
- Visual attention: Show, Attend and Tell: Neural Image Caption Generation with Visual Attention [ICML 2015]

### Key enablers for multimodal deep learning research:

- New large-scale multimodal datasets
- Faster computer and GPUS
- High-level visual features
- “Dimensional” linguistic features



# Multimodal Affective Computing

**Robots**



**Virtual Humans**



**Ubiquitous**



**Mobile**



**Online**



**Technologies able to analyze, recognize and predict human subtle communicative behaviors in social context.**



# Wide Applicability

Medical  
Education  
Online



Psychological signals



Suicide prevention



Autistic children



Group learning analytics



Virtual Learning Peer



Public speaking training



Opinion mining



Social influence



Negotiation outcomes

## Phenomena

### Pathology

- Distress
- Autism

### Social

- Empathy
- Dominance

### Emotion

- Sentiment
- Frustration

### Cognitive

- Attention
- Curiosity

### Personality

- Assertive
- Trusting



# Multimodal Affective Computing

## Behaviors

### Verbal

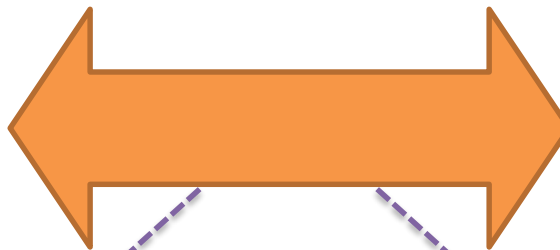
- **Lexicon**
  - Spoken words
- **Pragmatics**
  - Discourse acts

### Vocal

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

### Visual

- **Body language**
  - Head gestures
- **Facial expressions**
  - Smile, frowning



## Phenomena

### Pathology

- Distress
- Autism

### Social

- Empathy
- Dominance

### Emotion

- Sentiment
- Frustration

### Cognitive

- Attention
- Curiosity

### Personality

- Assertive
- Trusting

### Statistical analysis

- Variance analysis
- Reliability tests

### Prediction models

- Bayesian networks
- Markov fields

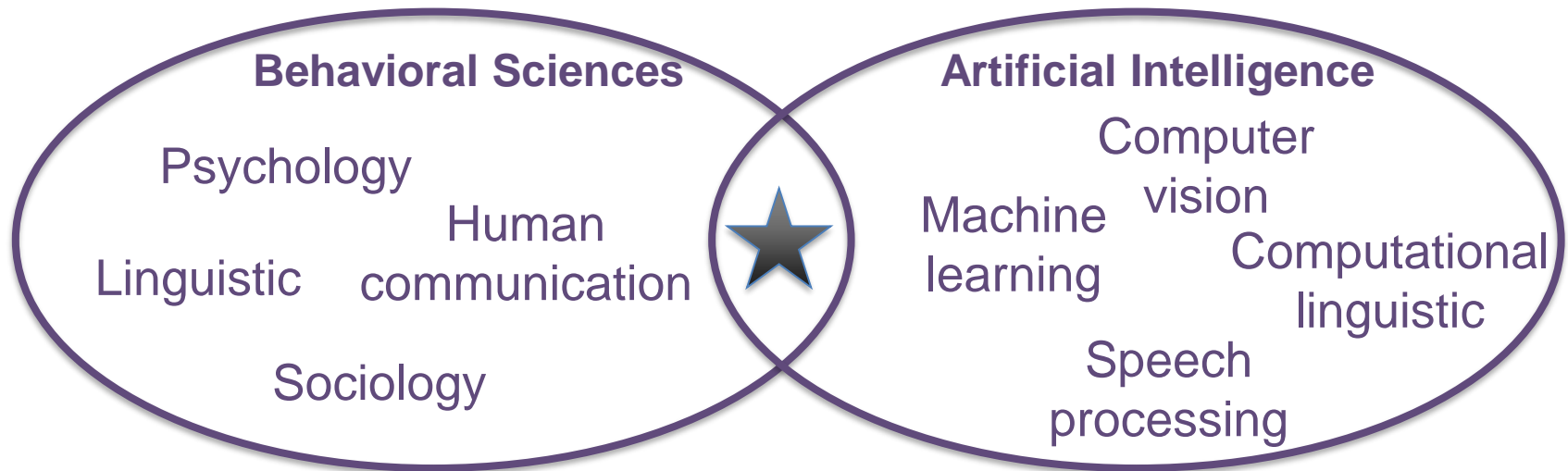
### Deep learning

- Bayesian networks
- Markov fields

## Computation

# Multimodal Affective Computing: Behavioral Sciences and Artificial Intelligence

---



**New tools to study  
human factors with  
technologies**

**Brings new or  
understudied  
learning challenges**



# Multimodal Conferences





# Multimodal Conferences – ICMI 2019

---

[Home](#)[Important Dates](#)[People](#)[Author Instructions](#)[Call for Papers](#)[Call for Workshops](#)[Call for Grand Challenges](#)[Conference Venue](#)[About Suzhou](#)[ICMI Archive](#)

## 21th ACM International Conference on Multimodal Interaction



**Suzhou, Jiangsu, China. October 14-18, 2019**

The 21st ACM International Conference on Multimodal Interaction (ICMI 2019) will be held in Suzhou, Jiangsu, China. ICMI is the premier international forum for multidisciplinary research on multimodal human-human and human-computer interaction, interfaces, and system development. The conference focuses on theoretical and empirical foundations, component technologies, and combined multimodal processing techniques that define the field of multimodal interaction analysis, interface design, and system development. ICMI 2019 will feature a single-track main conference which includes: keynote speakers, technical full and short papers (including oral and poster presentations), special sessions, demonstrations, exhibits and doctoral spotlight papers.

## Deadline in May 2019



# Multimodal Conferences – ACII 2019

ACII 2019	Calls	Important Dates	Attend	People
				
<b>8th International Conference on Affective Computing &amp; Intelligent Interaction (ACII 2019)</b> <b>3rd-6th September, 2019</b> <b>Cambridge, United Kingdom</b> <p>The Bi-annual Conference of the Association for the Advancement of Affective Computing (AAAC, formerly known as Humaine) is the premier international forum for research on affective and multimodal human-machine interaction and systems. ACII 2019 will be held in Cambridge, the home of the University of Cambridge, founded in 1209 and one of the top five universities in the world.</p>			<b>231 Days</b> until Sep 3rd, 2019.  <b>Newly added:</b> <b>14/01/2019:</b> <a href="#">Call for Special Session Proposals</a> <b>25/10/2018:</b> <a href="#">Keynote Speaker</a> <b>07/08/2018:</b> <a href="#">Call for Sponsorship</a> <a href="#">Call for Demos</a> <a href="#">Call for Doctoral Consortium Participation</a> <a href="#">Call for Tutorial Proposals</a>	

**Deadline in May 2019**

# Multimodal Conferences – EMNLP 2019

---

EMNLP-IJCNLP 2019

[Program](#) [Organization](#)



## News

---

**December 16, 2018.** The list of accepted [workshops](#) is out.

**November 25, 2018.** The important dates have been announced.

## Welcome!

---

[SIGDAT](#), the Association for Computational Linguistics (ACL) special interest group on linguistic data and corpus-based approaches to NLP, and the [AFNLP](#), the Asian Federation of Natural Language Processing, invite you to participate in the Conference on Empirical Methods in Natural Language Processing (EMNLP) & International Joint Conference on Natural Language Processing (IJCNLP) 2019 in Hong Kong. EMNLP-IJCNLP 2019 will be held at the [Asia World Expo](#) in Hong Kong from November 3rd through November 7th, 2019.

---

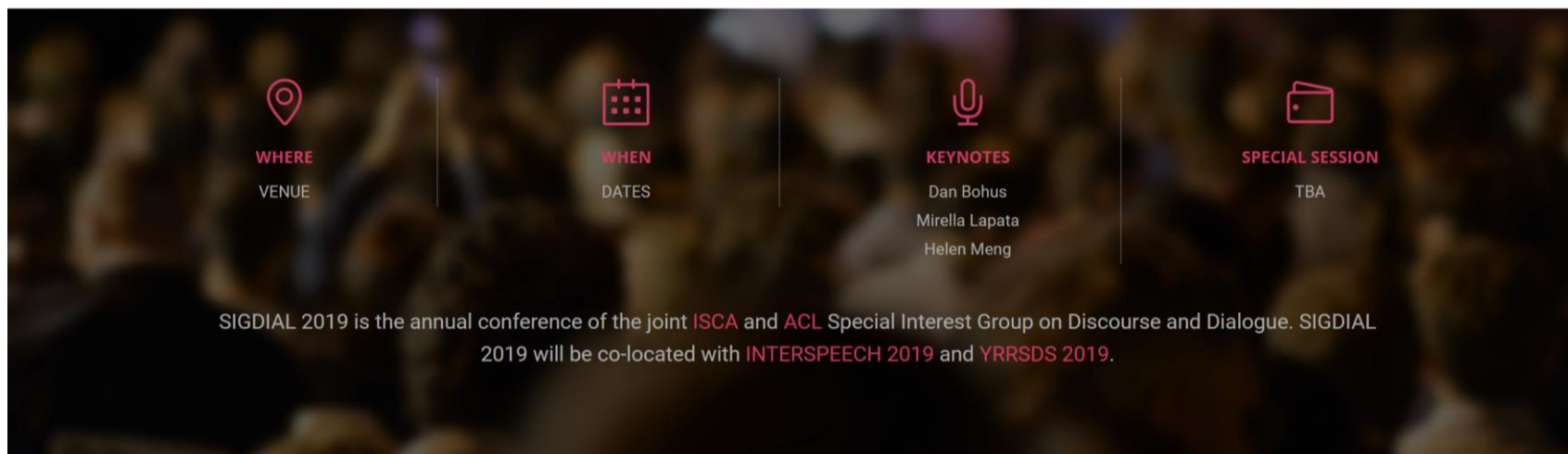
## Deadline in May 2019

# Multimodal Conferences – SigDial 2019

---

## SIGDIAL 2019

Stockholm, Sweden  
September 11-13, 2019



The graphic features a dark background with a blurred image of a crowd. It is divided into four vertical sections by thin white lines. Each section has a red icon at the top, a red title, and white text below. The sections are: 1. WHERE (location pin icon) VENUE; 2. WHEN (calendar icon) DATES; 3. KEYNOTES (microphone icon) listing Dan Bohus, Mirella Lapata, and Helen Meng; 4. SPECIAL SESSION (folder icon) TBA. At the bottom, a paragraph of white text provides context about the conference's affiliation with ISCA and ACL, and its co-location with INTERSPEECH 2019 and YRRSDS 2019.

WHERE	WHEN	KEYNOTES	SPECIAL SESSION
VENUE	DATES	Dan Bohus Mirella Lapata Helen Meng	TBA

SIGDIAL 2019 is the annual conference of the joint **ISCA** and **ACL** Special Interest Group on Discourse and Dialogue. SIGDIAL 2019 will be co-located with **INTER\_SPEECH 2019** and **YRRSDS 2019**.

## Deadline in May 2019

# Multimodal Conferences – FG 2019



The screenshot shows the official website for the 14th IEEE International Conference on Automatic Face and Gesture Recognition (FG2019). The header features the conference logo on the left, which is a stylized face with a crown, and navigation links (HOME, ABOUT, PARTICIPATE, KEYNOTE SPEAKERS) on the right. The main banner displays the dates "May 14-18, 2019" and the location "Lille, France" over a photograph of a historic European cityscape with a prominent clock tower. Below the banner, there are three sections: "ABOUT" with a blue bar, "RECENT POSTS" listing "Demo submission deadline postponed", and a "Tweets" section showing a tweet from @IEEEFG2019 and a retweet by user Ikbel.

14th IEEE International Conference on Automatic Face and Gesture Recognition  
**FG2019**

HOME ABOUT PARTICIPATE KEYNOTE SPEAKERS

May 14-18, 2019  
Lille, France

ABOUT

RECENT POSTS

- Demo submission deadline postponed

Tweets by @IEEEFG2019

IEEE FG2019 Retweeted

Ikbel

**New deadline in September 2019**



# Multimodal Conferences – ACM Multimedia 2019



## Home

Welcome to the [27th ACM International Conference on Multimedia](#)

ACM Multimedia is the premier international conference in the area of multimedia within the field of computer science. Multimedia research focuses on integration of the multiple perspectives offered by different digital modalities including images, text, video, music, sensor data, spoken audio. Since 1993, ACM Multimedia has been bringing together researchers and practitioners from academia and industry to present innovative research results and discuss recent advancements. A special part of the conference is the interactive art program, which explores the boundaries of computer science and art.

The [ACM Multimedia 2019](#) Conference will be held at the [NICE ACROPOLIS Convention Center](#) in [Nice, France](#) from [21-25 October 2019](#).

Theme: [Engaging users with multimedia](#)

## ACMMM19 on Twitter

### Tweets by @ACMMM19

**ACMMM19**  
@ACMMM19  
ACM MultiMedia website is up and running smoothly again. Upcoming deadlines are workshop proposals 15 Feb [acmmm.org/2019/call-for-...](#) and call for volunteers [acmmm.org/2019/call-for-...](#) #ACMMM19

Jan 11, 2019

**ACMMM19**  
@ACMMM19

## Deadline in April 2019 (or 2020)



# IEEE Transactions on Affective Computing

The screenshot shows the IEEE Xplore Digital Library interface. At the top, there's a navigation bar with 'Browse', 'My Settings', 'Get Help', and 'Subscribe'. Below this is a search bar with a dropdown menu set to 'All' and a search icon. A note in the search bar states: 'Enter keywords or phrases (Note: Searches metadata only by default. A search for 'smart grid' = 'smart AND grid')'. Below the search bar, there's a checkbox for 'Search within Publication' and links for 'Advanced Search' and 'Other Search Options'. The main content area is titled 'IEEE Transactions on Affective Computing' and includes a 'Browse Journals & Magazines > IEEE Transactions on Affective ...' breadcrumb. A navigation bar below the title has tabs for 'Popular', 'Early Access', 'Current Issue', 'Past Issues', 'About Journal', and 'Submit Your Manuscript'. The 'Submit Your Manuscript' tab is highlighted in orange. Below the navigation bar, there's a description of the journal: 'The IEEE Transactions on Affective Computing is a cross-disciplinary and international archive journal aimed at disseminating results of research on the design of systems that can recognize, interpret, and simulate human emotions and related affective phenomena. Aims & Scope >'. To the right of the description are three boxes showing journal metrics: 'Impact Factor' (4.585), 'Eigenfactor' (0.00308), and 'Article Influence Score' (1.272). At the bottom, there are two sections: 'Latest Published Articles' and 'Popular Articles'.

IEEE Xplore®  
Digital Library

> Institutional Sign In

IEEE

Browse ▾ My Settings ▾ Get Help ▾ Subscribe

All ▾ Enter keywords or phrases (Note: Searches metadata only by default. A search for 'smart grid' = 'smart AND grid') 🔍

■ Search within Publication Advanced Search | Other Search Options ▾

Browse Journals & Magazines > IEEE Transactions on Affective ... ?

## IEEE Transactions on Affective Computing

Add Journal To My Alerts

Popular Early Access Current Issue Past Issues About Journal Submit Your Manuscript

The *IEEE Transactions on Affective Computing* is a cross-disciplinary and international archive journal aimed at disseminating results of research on the design of systems that can recognize, interpret, and simulate human emotions and related affective phenomena.  
[Aims & Scope >](#)

4.585  
Impact Factor

0.00308  
Eigenfactor

1.272  
Article Influence Score

📄 Latest Published Articles ★ Popular Articles

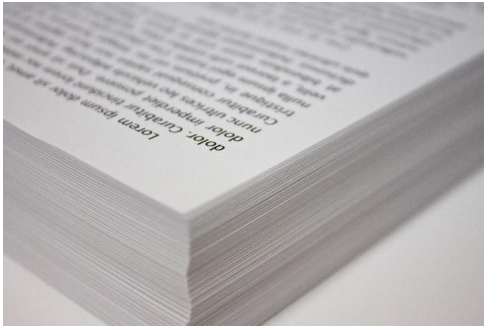
# Course Syllabus

---



# Three Course Learning Paradigms

---



Research paper reading  
and group discussion  
(30% of your grade)

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

Course project assignments  
and homework  
(70% of your grade)



Course lectures  
(including guest lectures)



# Course Structure

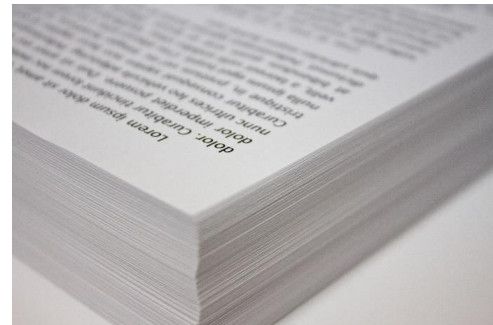
---

## Tuesdays



Course lectures

## Thursdays



Group discussion and  
student presentations



# Course Recommendations and Requirements

---

- 1 Ready to read about 20 papers this semester !
  - Each week, 1-3 papers to complement the lectures
  - Questions will be posted to guide your reading
- 2 Background in probability, statistics and linear algebra
  - Previous experience in machine learning is expected.
  - Programming knowledge in Python, R and/or Matlab
- 3 Motivated to produce a high-quality course project
  - Three course project assignments
  - Designed to answer a new research question



# Course Outline

---

Classes	Tuesday	Thursday
<b>Week 1</b> 1/15 & 1/17	<b>Introduction and communication models</b> <ul style="list-style-type: none"><li>• Human communication dynamics</li><li>• Signals and communicative messages</li><li>• Communication models (Brunswick's model)</li></ul>	<b>Applications and datasets</b> <ul style="list-style-type: none"><li>• Topics in affective computing</li><li>• Affective computing databases</li><li>• Affective computing data sources</li></ul>
<b>Week 2</b> 1/22 & 1/24	<b>Measuring psychological constructs</b> <ul style="list-style-type: none"><li>• Links between constructs and measurement</li><li>• Self-report and observational measurement</li><li>• Trustworthiness and measurement validation</li></ul>	<b>Student-led discussion</b> <ul style="list-style-type: none"><li>• Student 1</li><li>• Student 2</li><li>• Student 3</li></ul>
<b>Week 3</b> 1/29 & 1/31 <i>*preproposal*</i>	<b>Theories behind psychological constructs</b> <ul style="list-style-type: none"><li>• Theories of affect and emotion</li><li>• Theories of personality</li><li>• Theories of psychopathology</li></ul>	<b>Student-led discussion</b> <ul style="list-style-type: none"><li>• Student 1</li><li>• Student 2</li><li>• Student 3</li></ul>
<b>Week 4</b> 2/05 & 2/07	<b>Nonverbal communicative messages</b> <ul style="list-style-type: none"><li>• Prosody and vocal expressions</li><li>• Gesture, gaze, posture and proxemics</li><li>• Acoustic and visual representations</li></ul>	<b>Student-led discussion</b> <ul style="list-style-type: none"><li>• Student 1</li><li>• Student 2</li><li>• Student 3</li></ul>

# Course Outline

---

Classes	Tuesday	Thursday
<b>Week 5</b> 2/12 & 2/14	<b>Verbal communicative messages</b> <ul style="list-style-type: none"><li>• Speech and dialogue acts</li><li>• Boundaries, fillers and disfluencies</li><li>• Turn-taking and backchannel</li></ul>	<b>Student-led discussion</b> <ul style="list-style-type: none"><li>• Student 1</li><li>• Student 2</li><li>• Student 3</li></ul>
<b>Week 6</b> 2/19 & 2/21 <i>*proposal report*</i>	<b>Proposal presentations</b>	<b>Proposal presentations</b>
<b>Week 7</b> 2/26 & 2/28	<b>Statistical foundations</b> <ul style="list-style-type: none"><li>• Sampling and sampling error</li><li>• Point and interval estimation</li><li>• Statistical hypothesis testing</li></ul>	<b>Student-led discussion</b> <ul style="list-style-type: none"><li>• Student 1</li><li>• Student 2</li><li>• Student 3</li></ul>
<b>Week 8</b> 3/05 & 3/07	<b>Linear statistical modeling</b> <ul style="list-style-type: none"><li>• Understanding the linear model</li><li>• Generalizing the linear model</li><li>• Prediction using GLMNET</li></ul>	<b>Student-led discussion</b> <ul style="list-style-type: none"><li>• Student 1</li><li>• Student 2</li><li>• Student 3</li></ul>

# Course Outline

---

Classes	Tuesday	Thursday
<b>Week 9</b> 3/12 & 3/14	<b>NO CLASS</b>	<b>NO CLASS</b>
<b>Week 10</b> 3/19 & 3/21 <i>*midterm</i> <i>homework*</i>	<b>Probabilistic predictive modeling</b> <ul style="list-style-type: none"><li>• Bayes rule and Naïve Bayes classifier</li><li>• Dynamic Bayesian networks</li><li>• Hidden Markov model</li></ul>	<b>Student-led discussion</b> <ul style="list-style-type: none"><li>• Student 1</li><li>• Student 2</li><li>• Student 3</li></ul>
<b>Week 11</b> 3/26 & 3/28	<b>Discriminative predictive modeling</b> <ul style="list-style-type: none"><li>• Neural networks and SVM classifiers</li><li>• Maximum entropy model</li><li>• Markov random fields</li></ul>	<b>Student-led discussion</b> <ul style="list-style-type: none"><li>• Student 1</li><li>• Student 2</li><li>• Student 3</li></ul>
<b>Week 12</b> 4/02 & 4/04 <i>*midterm report*</i>	<b>Midterm presentations</b>	<b>Midterm presentations</b>



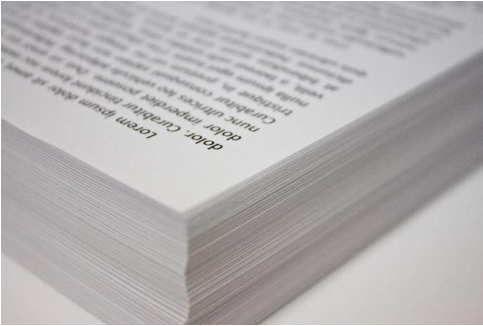
# Course Outline

---

Classes	Tuesday	Thursday
<b>Week 13</b> 4/09 & 4/11	<b>Multimodal deep learning</b> <ul style="list-style-type: none"><li>• Multimodal representations</li><li>• Attention and modality alignment</li><li>• Temporal and multimodal fusion</li></ul>	<b>NO CLASS</b>
<b>Week 14</b> 4/16 & 4/18	<b>Medical applications</b> <ul style="list-style-type: none"><li>• Assessment in the clinical process</li><li>• Biomarkers and behavioral indicators</li><li>• Validation in the medical sciences</li></ul>	<b>Student-led discussion</b> <ul style="list-style-type: none"><li>• Student 1</li><li>• Student 2</li><li>• Student 3</li></ul>
<b>Week 15</b> 4/23 & 4/25	<b>Educational applications</b> <ul style="list-style-type: none"><li>• Multimodal learning analytics</li><li>• Intelligent tutoring systems</li><li>• </li></ul>	<b>Student-led discussion</b> <ul style="list-style-type: none"><li>• Student 1</li><li>• Student 2</li><li>• Student 3</li></ul>
<b>Week 16</b> 4/30 & 5/02 <i>*final report*</i>	<b>Final presentations</b>	<b>Final presentations</b>

# Course Grades

---



- Discussion participation (including leading one session) 15%
- Reading assignments 20%

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

- Proposal presentation and report 15%
- Mid-term presentation, homework and report 25%
- Final project presentation and report 25%



# Reading Assignments and Group Discussions

---

- 1-3 research papers assigned per week
    - Selected to complement the lectures
  - You need to read carefully the assigned papers.
  - Questions will be posted to guide your reading
  - Submit your answers before Wednesday noon
    - Submit via Gradescope (1 submission per student)
    - Each answer should be at least one paragraph long
    - Answers will be shared with discussion leaders
  - Discussion on Thursdays, student-led
- <https://piazza.com/cmu/spring2019/11776/home>



# Student Discussion Leading

---

- Teams of 2 or 3 students will lead each discussion
  - Sign-up sheet available at the end of this lecture
  - Each student will lead once this semester (as part of a team)
- The discussion leaders should bring something new and interactive to the discussion
  - Examples: relevant datasets, simple demos, new figures or cheatsheets, challenging questions, and novel applications.
- The reading assignment responses of all other students will be shared with the discussion leaders
  - Will help to identify discussion topics (e.g., disagreements, confusions, new ideas)
- TAs are there to help you with this process!

## 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

# Course Project

---

Use Affective Computing techniques to answer a new research question!

- Potential data sources:
  - Existing multimodal datasets
  - Online videos (e.g., YouTube)
  - Your own video dataset [message us soon!]
  - Run a new study [you may need IRB approval]
- Your data should include multimodal and/or social behaviors

# Three main steps of your Course Project

---

- 1 Literature Review and Qualitative Analysis
  - Better understand phenomena and behaviors of interest
  - Understand related theories and previous models
- 2 Statistical Analysis and Hypothesis Testing
  - Evaluate the quality of your annotations and features
  - Develop hypotheses from Step 1 and test them
- 3 Predictive Modeling and Multimodal ML
  - Build computational models to predict the constructs
  - Integrate information from verbal, vocal and visual behaviors



# Course Project - Examples

---

## Modeling Domain Knowledgeability through Non-Verbal Cues – by Abdelwahab Bourai (Fall 2016)



- *Data from University Challenge series from BBC*
- Long paper published at ICMI 2017

# Course Project - Examples

---

## Multimodal Emotion Recognition for Children – by Behnaz Nojavanasghari (Fall 2015)

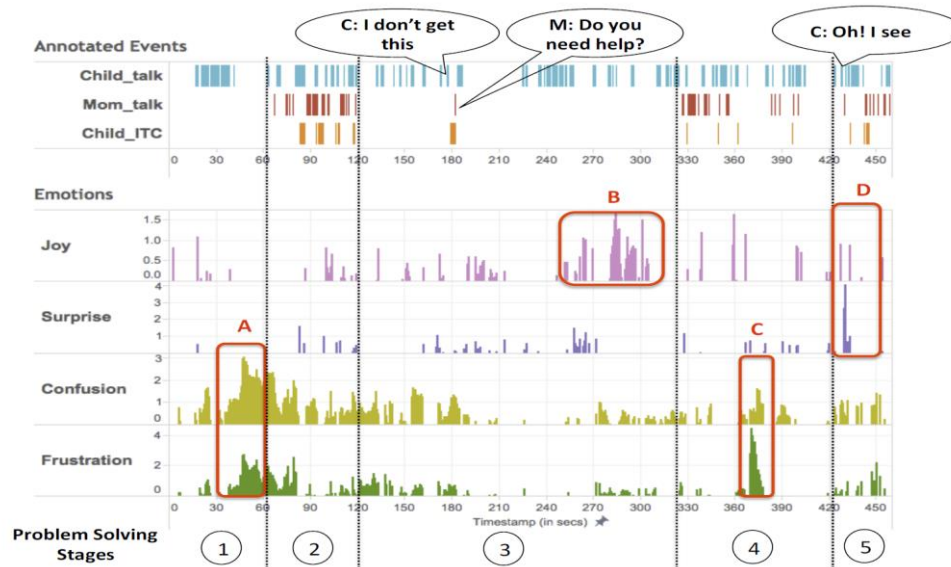


- Data from KidsReact channel on YouTube
- Long paper published at ICMI 2016



# Course Project - Examples

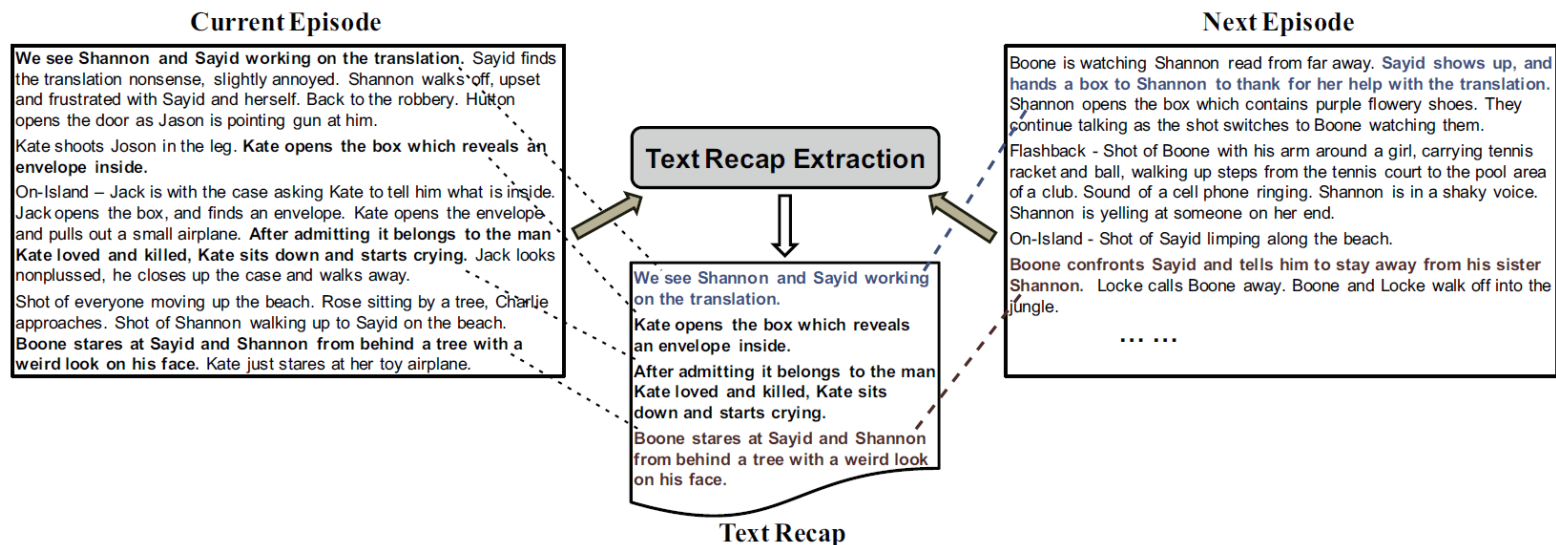
Riding emotional rollercoaster: Intent-to-connect behaviors in young child's math problem solving activities – by Lujie Chen (Fall 2015)



- Data from a new study (with her own son)
- Published at Educational Data Mining conference (2016)

# Course Project - Examples

## Generating Visual Synopsis Using Text-based Plot Summary – by Hongliang Yu and Shikun Zhang (Fall 2015)



- Data from the TV series *Lost*
- Long paper published at EMNLP 2016



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

# Course Project

---

1. Pre-proposal (in 2 weeks)
  - Define your topic, dataset/study and teammates
2. Proposal presentation and report (in 5 weeks)
  - Defining your research questions
  - Literature review and qualitative analysis
3. Midterm homework (after week 8)
  - Statistical analysis on *our* pre-processed dataset
4. Midterm presentation and report (in 11 weeks)
  - Statistical analysis of your constructs
  - Hypothesis testing based on qualitative analysis
5. Final presentation and report (in 15 weeks)
  - Prediction models using results from statistical analysis
  - Multimodal behavior integration and modeling

# Course Project Guidelines

---

- Preferred course project should have:
  - At least 2 modalities (language, visual, acoustic)
  - Social / interpersonal interactions
- Teams of 2, 3 or 4 students
  - No individual projects
- Possible venues for your final report:
  - ACII 2019, SigDial 2019, ICMI 2019
- We will discuss on Thursday about project ideas

# Process for Selecting your Course Project

---

- **Thursday 1/17:** Lecture describing ideas and datasets for course project
- **Tuesday 1/22:** Submit a short paragraph listing your top 3 choices (in Piazza)
- **Thursday 1/24:** After the discussion, we will have some time to facilitate team formation
- **Sunday 2/3:** Pre-proposals are due. You should have selected your teammates, dataset (or study) and research questions