CHIA-YING (JACKIE) LEE

One Front Street, Suite 1850, San Francisco CA 94111
http://people.csail.mit.edu/chiaying/

+1-617-960-7405   **chiaying@csail.mit.edu**

EDUCATION

MASSACHUSETTS INSTITUTE of TECHNOLOGY, Cambridge, MA
Ph.D. in Computer Science, Sep 2014

- Thesis topic: Discovering Linguistic Structures in Speech: Models and Applications.
- Advisor: Dr. James Glass.
- Thesis Committee: Prof. Regina Barzilay and Prof. Victor Zue
- Research interests: machine learning, pattern recognition, unsupervised statistical inference, natural language processing, speech recognition, and human-computer interaction.

MASSACHUSETTS INSTITUTE of TECHNOLOGY, Cambridge, MA
Master of Science in Computer Science, June 2010

- Thesis topic: Closed-loop Auditory-based Representation for Robust Speech Recognition.
- Advisors: Dr. James Glass and Prof. Oded Ghitza.

NATIONAL TAIWAN UNIVERSITY, Taipei, Taiwan
Bachelor of Science in Electrical Engineering, June 2008

- Class rank 3rd out of 232.
- GPA~ 91.88/100.
- Dean's List, 2006-2008 (since inception), and five-time inclusion on the Honor Roll.
- Recipient, Presidential Award, reserved for top 5% of students in each department, 2004-2008.

WORK EXPERIENCE

- *Machine Learning Engineer, Manhattan Engineering Incorporated (a.k.a. Kite)*
  - *Feb 2015 – present:* At Kite, we are inventing the new way of how people will build software in the future. Our goal is to improve programmers' productivity and software quality. As a machine learning engineer, the projects that I've carried out at Kite are described as follows. Developed statistical models to find common usage patterns in computer programs. Utilized natural language processing techniques to automatically detect code quality. Designed and implemented a recommendation system for suggesting possible ways of fixing program errors. Built a ranking system for code example suggestion. Developed an algorithmic framework to predict solutions to errors in computer programs. I was the eighth employee of the company.
- *Research Assistant, Spoken Language Systems Group, Computer Science Artificial Intelligence Lab (CSAIL), MIT*
  - *July 2011 – Sep 2014:* Developed unsupervised models for learning linguistic structures including the phonetic, syllabic, and the lexical inventory of a language directly from speech data.
  - *Nov. 2009 – June 2011*: Leveraged human resources available on crowdsourcing platforms to transcribe large speech corpus with low costs. Designed efficient mechanisms to enable fast transcription convergence and applied machine learning techniques to automate quality control.
  - *Jan. 2009 – May 2010:* Integrated an efferent-inspired auditory model into speech feature generation processes to addressed the problem of unseen noise for robust speech recognition.
- *Visiting Research Assistant, Natural Language Processing Group, Information Science Institute, University of Southern California, June 2013 - August 2013:* Worked on improving translation quality of Swahili to English by exploiting the rich morphological property embedded in the Swahili language. Designed and implemented new inference algorithms to train IBM Model 1 and the HMM alignment model with lattices as inputs, which allowed learning alignments between Swahili morpheme lattices and English sentences.

- *Research Intern, Microsoft Research, Redmond, May 2012 - August 2012*: Developed a joint model of a Recurrent Neural Network (RNN) and a Deep Belief Network (DBN) for training speech recognition systems without using any phone-level information. Designed and implemented the corresponding training algorithm.
- *Research Intern, Google, New York, June 2010 - August 2010*: Developed a methodology to enable HMM-based synthesizers trained with languages with rich resources to synthesize speech in languages with sparse resources.

PUBLICATIONS

- *Unsupervised Lexicon Discovery from Acoustic Input.* Chia-ying Lee, Timothy J. O'Donnell, James Glass. Transactions of ACL, 2015.
- *Discovering Linguistic Structures in Speech: Models and Applications.* Chia-ying Lee. *Ph.D. Thesis, MIT, July 2014.*
- *One-shot Learning of Generative Speech Concepts*. Brenden Lake*, Chia-ying Lee*, James Glass, Joshua Tenenbaum. Cognitive Science Society (CogSci), Quebec City, Canada, 2014.
  * The first two authors contributed equally to this work.
- *Joint Learning of Phonetic Units and Word Pronunciations for ASR*. Chia-ying Lee, Yu Zhang, James Glass. EMNLP, Seattle, USA, 2013.
- *A Nonparametric Bayesian Approach to Acoustic Model Discovery*. Chia-ying Lee, James Glass. *ACL, Jeju, Korea, 2012*.
- *A Transcription Task for Crowdsourcing with Automatic Quality Control*. Chia-ying Lee, James Glass. *Interspeech, Florence, Italy, 2011*.
- *An Efferent-inspired Auditory Model Front End for Speech Recognition*. Chia-ying Lee, James Glass, Oded Ghitza. *Interspeech, Florence, Italy, 2011*.
- *Closed-loop Auditory-based Representation for Robust Speech Recognition.* Chia-ying Lee. *Master Thesis, MIT, May 2010.*
- *Collecting Voices from the Cloud*. Ian McGraw, Chia-ying Lee, Lee Hetherington, Stephanie Seneff, James Glass. *The Language Resources and Evaluation Conference (LREC), Malta, Italy, 2010.*
- *Empirical Mode Decomposition Descriptor for Plane Closed Curves*. Soo-Chang Pei, Yu-Zhe Hsiao, Chia-ying Lee. *International Conference on Multimedia and Expo (ICME), Cancun, Mexico, 2009*.

PROFESSIONAL SERVICES

- *Reviewer,* IEEE Transactions on Audio, Speech, and Language Processing
- *Reviewer,* IEEE Signal Processing Letters
- *Reviewer,* Journal of Neural Computation
- R*eviewer,* International Journal of Pattern Recognition and Artificial Intelligence

LEADERSHIP EXPERIENCE

- *Co-president, Graduate Women in Course 6, MIT, Feb 2011 – Feb 2012* ~ Planned and organized events for female graduate students in the department of EECS, MIT with two other co-presidents. Organized approximately one event per month to create opportunities for female graduate students to reach out.
- *President, CSAIL Student Committee, CSAIL, MIT, Aug. 2009 - May 2010* ~ Led the committee to raise funding and organize lab-wide events to encourage graduate students to socialize beyond their own research groups with an aim of making the lab a more friendly environment.
- *Chief Organizer, CSAIL Student Workshop, CSAIL, MIT 2009* ~ Led a team of six people to raise funding, plan and organize in 2009 the CSAIL student workshop (CSW), a lab-wide event attended by more than one hundred lab members, including graduate students, post-docs and professors.

- *Event Manager, MIT Republic of China Student Association, June 2009 - May 2010 ~* Organized social and cultural events for Taiwanese students at MIT with a focus on providing opportunities for the Taiwanese students to connect with each other and build a strong community network.
- *Vice President, Sunny Kite Youth Club, NTU, Aug 2006 – May 2007 ~* Led a group that has designed and delivered educational programs and mentoring to 80 at-risk children.

TEACHING EXPERIENCE

- *Introduction to Artificial Intelligence, Spring 2011, MIT*
- *Mentor of Undergraduate Research Opportunity Project, Summer 2011, CSAIL, MIT*
- *Mentor of Undergraduate Research Opportunity Project, Spring 2010, CSAIL, MIT*

INVITED TALKS

- *A Non-parametric Bayesian Approach to Acoustic Discovery, April 2012, Johns Hopkins University*
- *A Non-parametric Bayesian Approach to Acoustic Discovery, May 2012, MIT*
- *Zero-Resource Speech Pattern and Sub-Word Unit Discovery, October 2012, SANE*
- *Bayesian Approaches to Acoustic Model and Pronunciation Lexicon Discovery, July 2013, ISI, USC*
- *Unsupervised Bayesian Approach to Training Automatic Speech Recognizers, October 2013, MIT*
- *Discovering Linguistic Structures from Speech: Models and Applications, November 2015, Berkeley*
- *Discovering Linguistic Structures from Speech: Models and Applications, January 2016, SRI*

PROGRAMMING LANGUAGES

- C++, Go, C, Python, MATLAB, JavaScript, and Church.