

Liang Huang, Asst. Prof., Oregon State U.; Principal Scientist, Baidu Research USA

CONTACT INFORMATION	School of EECS, Oregon State University 1148 Kelley Engineering Center (KEC) Corvallis, OR 97331, USA	liang.huang.sh@gmail.com http://eecs.oregonstate.edu/~huanlian 541-737-4694 (o) 541-737-1300 (f)
	Baidu Research USA (2018–2019) 1195 Bordeaux Drive, Sunnyvale, CA	267-738-5325 (c) http://research.baidu.com/People
EDUCATION	University of Pennsylvania Ph.D., Computer and Information Science Advisors: Aravind Joshi and Kevin Knight (external, USC). Committee: Mitch Marcus , Fernando Pereira , Ben Taskar (chair), and Mark Johnson (external, Brown) Thesis: <i>Forest-based Algorithms in Natural Language Processing.</i> (Nominated for the ACM Doctoral Dissertation Award.) Shanghai Jiao Tong University B.S., Computer Science (summa cum laude) Thesis: <i>Probabilistic k-best Earley Parsing for Classical and Modern Chinese.</i>	<i>Sep. 2003 – Dec. 2008</i>
PROFESSIONAL EXPERIENCE	Baidu Research USA <i>Principal Scientist</i> Oregon State University <i>Assistant Professor</i> , Computer Science, School of EECS <i>On Leave as Principal Scientist at Baidu Research Silicon Valley</i> IBM T. J. Watson Research Center <i>Part-Time Research Scientist</i> (manager: Bowen Zhou) City University of New York (CUNY) <i>Assistant Professor</i> , Computer Science, Queens College <i>Doctoral Faculty</i> , Computer Science, The Graduate Center University of Southern California (USC) <i>Research Assistant Professor</i> , Computer Science <i>Computer Scientist</i> , Information Sciences Institute (ISI) Google Research (Mountain View): <i>Research Scientist</i> (manager: Fernando Pereira)	<i>2018–2019</i> <i>2015/9–present</i> <i>2018–2019</i> <i>2014/6–2017/1</i> <i>2012/8–2015/8</i> <i>2012/8–2015/8</i> <i>2010/7–2012/8</i> <i>2009/7–2012/8</i> <i>2009/1–8</i>
RESEARCH INTERESTS	Natural Language Processing: linear-time algorithms for NLP, parsing, semantic parsing, machine translation, neural models for text modeling, parsing, and translation Machine Learning: scalable structured prediction (with inexact inference), deep learning, learning latent structures from unstructured data, online learning and parallelization Computational Biology: linear-time algorithms for RNA secondary structure prediction and protein folding, deep learning for genome sequences <i>Programming Languages:</i> compilers & type theory for NLP; NLP techniques for PL	

Algorithms/Theory: generic dynamic programming, k -best problems, formal language and automata theory, applications of computational geometry in NLP

Psycholinguistics: computational models of human sentence processing

See also: [Google Scholar Profile](#).

VISITING EXPERIENCE Hong Kong University of Science and Technology (HKUST) 2008/10–12
Chinese Academy of Sciences, Institute of Computing Technology (ICT) 2007/10–2008/1
USC Information Sciences Institute (ISI) 2005/5–9, 2005/12–2006/2, and 2006/5–10

HONORS AND AWARDS

Research/Paper Awards

- Best Paper Finalist, SIGMOD 2018
- [Best Paper Honorable Mention, EMNLP 2016](#) (Cross and Huang, 2016b)
- Yahoo! Faculty Research and Engagement Award, 2015
- Google Faculty Research Award, 2013
- Google Faculty Research Award, 2010
- Best Paper Award Finalist: ACL 2010
- Best Paper Award Finalist: EMNLP 2008
- [Best Paper Award, ACL 2008](#) (Huang, 2008)
- Best Paper Award Finalist: ACL 2007

Teaching Awards

- [University Prize for Excellence in Teaching by Graduate Students](#), University of Pennsylvania, 2005
- Nominated for the USC Viterbi School of Engineering Teaching Award, 2012

Programming Contest Awards (ACM/ICPC and OI)

- Champions (as faculty advisor), [ACM/ICPC Southern California Regionals](#) (advanced to the World Finals for the first time in USC history), 2011–2012.
- 4th Place, ACM International Collegiate Programming Contest, Shanghai site, 2000
- Bronze Medal (16th nationwide), National Olympiads in Informatics, China, 1996.

MEDIA COVERAGE

[Simultaneous Translation](#)

covered by numerous news reports, e.g., [IEEE Spectrum](#), [MIT Technology Review](#), [FORTUNE](#), [CNBC](#), [The Register](#), etc.

FUNDING

Total Amount of Funding: \$12.4M; my share: **\$2.02M**.

Extramural:

- Baidu Research Gift, sole PI, \$60k.
- NIH R01, co-I (PI: Robert Tanguay); 2019–2023. \$2.5M, my share: \$190k.
- NSF Small, IIS-1817231, sole PI, 2018–2021, \$400,000.
- Intel. PI. 2018–2019. \$100,000.
- DARPA XAI Project. co-PI. PI: Alan Fern. \$6.5M total (\$400k my share), 2017–2021.

- HP+OSU Seed Fund, “Activity Recognition from Sensors with Minimal Supervision”. PI. \$20,000. 2016-2017.
- Yahoo! Faculty Research and Engagement Award, “Incremental Semantic Parsing with Applications in Question Answering” (sole PI), unrestricted gift, \$25,000, 2015–2016.
- NSF, “EAGER: Collaborative Research: Scaling Up Discriminative Learning for Natural Language Understanding and Translation” (sole PI), \$135,372 for one year, 2014–2015.
- Google Faculty Research Award, “Towards Scalable Discriminative Training for Machine Translation: Fast Decoding, Parallelized Learning with Inexact Search, and Beam Ranking” (sole PI), unrestricted gift, \$87,947 for one year, 2013–2014.
- DARPA Deep Exploration and Filtering (DEFT) Program, “SPARKLER - Scalable Prosodic, Anomaly and Relational Knowledge exploration of Language with Enhanced Robustness” (co-PI), total amount of \$2,000,000 for 4.5 years (Liang Huang budget: \$504,024), 2012–2017.
- Google Faculty Research Award, “Linear-time Dynamic Programming for Parsing and Machine Translation” (PI), unrestricted gift, \$75,000 for one year, 2010–2011.

Intramural:

- OSU EECS Collaboration Initiative, PI: Huang, co-PI: D. Hendrix, one 12-month GRA (salary and intuition, ~\$47,000), 2016.
- CUNY QC Research Enhancement Award, “Type-Driven Incremental Semantic Parsing” (sole PI), \$8,250 for one year, 2015.
- PSC-CUNY Enhanced Research Award, “Scalable Machine Learning for Big-Data Natural Language Processing” (sole PI), \$12,000 for one year, 2013–2014.
- USC Viterbi School of Engineering Research Innovation Fund Award, “Search Error Robust Learning: Theory and Algorithms of Structured Prediction with Inexact Inference” (sole PI), \$6,000 for one year, 2012.

TEACHING
EXPERIENCE

Oregon State University

CS 519, Applied Machine Learning (e-campus)	<i>Spring 2018</i>
CS 519, Graduate Algorithms (MS/MEng-level)	<i>Winter 2018</i>
CS 534, Machine Learning	<i>Fall 2017</i>
CS 519, Natural Language Processing	<i>Spring 2017</i>
CS 519, Graduate Algorithms (MS/MEng-level)	<i>Fall 2016</i>
CS 519, Scientific Writing and Presentation	<i>Spring 2016</i>
CS 480, Translators (Compilers)	<i>Winter 2016</i>
CS 321, Theory of Computation	<i>Fall 2015</i>

City University of New York (CUNY)

Graduate Center: CS 71010, Programming Language Theory	<i>Fall 2014, Fall 2013</i>
Graduate Center: LING 83600/CS 84010, Language Technology	<i>Fall 2014, Spring 2013</i>
Queens College: CS 3813/780: Advanced Programming	<i>Fall 2014, Spring 2014</i>
Graduate Center: CS 87100, Scientific Writing and Presentation	<i>Fall 2013</i>
Graduate Center: CS 84010, Machine Learning	<i>Spring 2013</i>
Queens College: CS 3813/780: Python Programming and Text Processing	<i>Fall 2012</i>

University of Southern California (USC) – all graduate courses

- CS 570, Analysis of Algorithms *Spring 2012*
CS 561, Artificial Intelligence (in Prolog, with K. Sagae) *Spring 2012*
CS 562, Stat. Natural Language Proc. (with Chiang & Knight) *Fall '11, Fall '10, Fall '09*
CS 599, Machine Translation (*new course*, with Chiang and Knight) *Spring 2011*
CS 544, Natural Language Processing (with Hovy et al.) *Spring 2010*
Nominated for the USC Engineering School Teaching Award, 2012.

University of Pennsylvania – all undergraduate courses

- [CSE 399-004 Python Programming](#) (*new course*) *Spring 2006*
Teaching Assistant, CSE 320 Analysis of Algorithms *Spring 2005*
Teaching Assistant, CSE 262 Automata, Complexity & Computability *Fall 2004*
Received [The University Graduate Teaching Award \(2005\)](#).
Recommended by the Department to be the first graduate student to teach a course.

Conference Tutorials: [Scalable Large-Margin Structured Learning: Theory & Algorithms](#)

- ACL 2014, Baltimore, MD (attendance: ~100) *June 2014*
ACL 2015, Beijing, China (attendance: ~100) *July 2015*

Conference Tutorials: [Tree-based and Forest-based Translation](#) (with Yang Liu)

- ACL 2010, Uppsala, Sweden (attendance: ~65) *July 2010*

Conference Tutorials: [Advanced Dynamic Programming for Computational Linguists](#)

- COLING 2008, Manchester, UK (attendance: ~70) *Aug 2008*
NAACL 2009, Boulder, CO (attendance: ~65) *May 2009*

PUBLICATIONS
[\[Google Scholar\]](#)

THESES/DISSERTATIONS

1. Liang Huang (2008). [Forest-Based Algorithms for Natural Language Processing](#). *Ph.D. Dissertation, University of Pennsylvania.*
Nominated for the ACM Doctoral Dissertation Award.
2. Liang Huang (2003). Probabilistic k -best Earley Parsing for Classical and Modern Chinese. *Bachelor's Thesis (with highest distinction), Shanghai Jiao Tong University.*

BOOKS

3. Rujia Liu and Liang Huang (2003). [The Art of Algorithms and Programming Contests](#) (*in Chinese*). Tsinghua University Press. *National best seller in computer science.*

JOURNAL ARTICLES

4. P. Danaee, M. Rouches, M. Wiley, D. Deng, L. Huang, and D. Hendrix (2018). bpRNA: large-scale automated annotation and analysis of RNA secondary structure. *Nucleic Acids Research*, 46 (11), 5381-5394.
5. Wenbin Jiang, Yajuan Lü, Liang Huang and Qun Liu (2015). Automatic Adaptation of Annotations. *Computational Linguistics*, 41 (1), pp. 119-147.
6. Liang Huang, Hao Zhang, Daniel Gildea, and Kevin Knight (2009). [Binarization of Synchronous Context-Free Grammars](#). *Computational Linguistics*, 35 (4). December 2009. Conference version appeared at HLT-NAACL 2006 (see below).
7. Adam Lucas, Liang Huang, Aravind Joshi, and Ken Dill (2007). [Statistical Mechanics of Helix Bundles using a Dynamic Programming Approach](#). *J. Am. Chem. Soc. (JACS)*, 129 (14), pp. 4272-4281.

8. Ken Dill, Adam Lucas, Julia Hockenmaier, Liang Huang, David Chiang, and Aravind Joshi (2007). [Computational Linguistics: a new tool for exploring biopolymer structures and statistical mechanics](#). *Polymer*, 48 (15), pp. 4289-4300.

PREPRINTS ON ARXIV OR BIORXIV

9. Mingbo Ma, Liang Huang, Hao Xiong, Kaibo Liu, Chuanqiang Zhang, Zhongjun He, Hairong Liu, Xing Li, and Haifeng Wang (2018). [STACL: Simultaneous Translation with Integrated Anticipation and Controllable Latency](#). On ArXiv Oct 2018. In submission to TACL.
10. Hairong Liu, Mingbo Ma, Liang Huang, Hao Xiong, Zhongjun He (2018). [Robust Neural Machine Translation with Joint Textual and Phonetic Embedding](#). On ArXiv Oct 2018. In submission to NAACL.
11. Dezhong Deng, Kai Zhao, David Hendrix, David Mathews, and Liang Huang (2018). [LinearFold: Linear-Time Prediction of RNA Secondary Structures](#). On BioRxiv Feb 2018. In submission to Nature Methods.

REFEREED CONFERENCE PAPERS (37 in top-conferences on [csrankings.org](#))

12. Renjie Zheng, Mingbo Ma and Liang Huang (2018). Multi-Reference Training with Pseudo-Reference Generation for Text Generation. In *Proceedings of EMNLP 2018*.
13. Yilin Yang, Liang Huang and Mingbo Ma (2018). Break the Beam Search Curse: A Study of Rescoring Methods and Stopping Criteria for Neural Machine Translation. In *Proceedings of EMNLP 2018*.
14. Jiaji Huang, Yi Li, Wei Ping and Liang Huang (2018). Large Margin Neural Language Model. In *Proceedings of EMNLP 2018*.
15. Wen Zhang, Lei Shen, Yang Feng, Liang Huang and Qun Liu (2018). Cube Pruning for Neural Machine Translation. In *Proceedings of EMNLP 2018*.
16. Juneki Hong and Liang Huang (2018). Linear-Time Constituency Parsing with RNNs and Dynamic Programming. In *Proceedings of ACL 2018*.
17. B. McCamish, V. Ghadakchi, A Termehchy, B. Touri, and L. Huang (2018). The Data Interaction Game. In *Proceedings of SIGMOD 2018*.

Best Paper Finalist.

18. Tianze Shi, Liang Huang, and Lillian Lee (2017). Fast(er) Exact Decoding and Global Training for Transition-Based Dependency Parsing via A Minimal Feature Set. In *Proceedings of EMNLP*.
19. Kai Zhao and Liang Huang (2017). Joint Syntactic and Discourse Parsing with Recurrent Neural Models. In *Proceedings of EMNLP*.
20. Liang Huang, Kai Zhao, and Mingbo Ma (2017). When to Finish? Optimal Beam Search for Neural Text Generation (modulo beam size). In *Proceedings of EMNLP*.
21. Mingbo Ma, Liang Huang, Bing Xiang, and Bowen Zhou (2017). Group Sparse CNNs for Question Classification with Answer Sets. In *Proceedings of ACL*.
22. Mingbo Ma, Kai Zhao, Liang Huang, Bing Xiang and Bowen Zhou, Jointly Trained Sequential Labeling and Classification by Sparse Attention Neural Networks. In *Proceedings of Interspeech*.
23. Kai Zhao, Liang Huang, and Mingbo Ma (2016). Textual Entailment with Structured Attentions and Composition. In *Proceedings of COLING*.
24. James Cross and Liang Huang (2016b). [Span-Based Constituency Parsing with a Structure-Label System and Provably Optimal Dynamic Oracles](#). In *Proc. of EMNLP*.

Best Paper Honorable Mention.

25. James Cross and Liang Huang (2016a). Incremental Parsing with Minimal Features Using Bi-Directional LSTM. In *Proceedings of ACL*.

26. Reza Ghaeini, Xiaoli Fern, Liang Huang and Prasad Tadepalli (2016). Event Nugget Detection with Bidirectional Recurrent Neural Networks. In *Proceedings of ACL*.
27. Feifei Zhai and Liang Huang (2015). A Pilot Study Towards End-to-End MT Training. In *Proceedings of MT Summit XV*.
28. Feifei Zhai, Liang Huang and Kai Zhao (2015). Search-Aware Tuning for Hierarchical Phrase-based Decoding. In *Proceedings of EMNLP 2015*.
29. Mingbo Ma, Liang Huang, Bing Xiang, and Bowen Zhou (2015). Dependency-based Convolutional Neural Networks for Sentence Embedding. In *Proceedings of ACL 2015*.
30. Kai Zhao and Liang Huang (2015). Type-Driven Incremental Semantic Parsing with Polymorphism. In *Proceedings of NAACL 2015*.
31. Haitao Mi and Liang Huang (2015). Shift-Reduce Constituency Parsing with Dynamic Programming and POS Tag Lattice. In *Proceedings of NAACL 2015*.
32. I. Naim, Y. Song, Q. Liu, L. Huang, H. Kautz, J. Luo, and D. Gildea (2015). [Discriminative Unsupervised Alignment of Natural Language Instructions with Corresponding Video Segments](#). In *Proceedings of NAACL 2015*.
33. Lemao Liu and Liang Huang (2014). Search-Aware Tuning for Machine Translation. In *Proceedings of EMNLP 2014*.
34. Heng Yu, Haitao Mi, Liang Huang and Qun Liu (2014). A Structured Language Model for Incremental Tree-to-String Translation In *Proceedings of COLING 2014*.
35. Kai Zhao, Liang Huang, Haitao Mi, and Abe Ittycheriah (2014). Hierarchical MT Training using Max-Violation Perceptron. In *Proceedings of ACL 2014*.
36. Heng Yu, Liang Huang, Haitao Mi and Kai Zhao (2013). [Max-Violation Perceptron and Forced Decoding for Scalable MT Training](#). In *Proceedings of EMNLP 2013*.
37. Kai Zhao, James Cross, and Liang Huang (2013). [Dynamic Programming for Optimal Best-First Shift-Reduce Parsing](#). In *Proceedings of EMNLP 2013*.
38. Hao Zhang, Kai Zhao, Liang Huang, and Ryan McDonald (2013). [Online Learning for Inexact Hypergraph Search](#). In *Proceedings of EMNLP 2013*.
39. Yoav Goldberg, Kai Zhao, and Liang Huang (2013). [Efficient Implementation of Beam-Search Incremental Parsers](#). In *Proceedings of ACL 2013*.
40. Qi Li, Heng Ji, and Liang Huang (2013). [Joint Event Extraction via Structured Prediction with Global Features](#). In *Proceedings of ACL 2013*.
41. Kai Zhao and Liang Huang (2013). [Minibatch and Parallelization for Large-Margin Structured Learning](#). In *Proceedings of NAACL 2013*.
42. Liang Huang, Suphan Fayong, and Yang Guo (2012). [Structured Perceptron with Inexact Search](#). In *Proceedings of NAACL 2012*.
43. Ashish Vaswani, Liang Huang and David Chiang (2012). [Smaller Alignment Models for Better Translations: Unsupervised Word Alignment with the \$\ell_0\$ -norm](#). In *Proceedings of ACL 2012*.
44. Ashish Vaswani, Haitao Mi, Liang Huang and David Chiang (2011). [Rule Markov Models for Fast Tree-to-String Translation](#). In *Proceedings of ACL 2011*.
45. Liang Huang and Haitao Mi (2010). [Efficient Incremental Decoding for Tree-to-String Translation](#). In *Proceedings of EMNLP 2010*.
46. Liang Huang and Kenji Sagae (2010). [Dynamic Programming for Linear-time Incremental Parsing](#). In *Proceedings of ACL 2010*.
Nominated for the Best Paper Award.
47. Haitao Mi, Liang Huang, and Qun Liu (2010). Machine Translation with Lattices and Forests. In *Proceedings of COLING 2010*.
48. Liang Huang, Wenbin Jiang, and Qun Liu (2009). [Bilingually-Constrained \(Monolingual\) Shift-Reduce Parsing](#). In *Proceedings of EMNLP 2009*.

49. Wenbin Jiang, Liang Huang, and Qun Liu (2009). Automatic Adaptation of Annotation Standards: Chinese Word Segmentation and POS Tagging: A Case Study. In *Proceedings of ACL-IJCNLP 2009*.
50. Liang Huang (2008). [Forest Reranking: Discriminative Parsing with Non-Local Features](#). In *Proceedings of ACL 2008*.
Received the Best Paper Award.
 Recognized in ACL 2012 as the **most-cited paper** published in 2008 within the [ACL Anthology](#) (most NLP conferences and journals).
51. Liang Huang (2008). [Advanced Dynamic Programming in Semiring and Hypergraph Frameworks](#). In *Proceedings of COLING 2008*. Survey paper to accompany the conference tutorial.
52. Haitao Mi and Liang Huang (2008). [Forest-based Translation Rule Extraction](#). In *Proceedings of EMNLP 2008*.
Nominated for the Best Paper Award.
53. Haitao Mi, Liang Huang and Qun Liu (2008). [Forest-based Translation](#). In *Proceedings of ACL 2008*.
54. Wenbin Jiang, Liang Huang, Qun Liu, and Yajuan Lü (2008). A Cascaded Linear Model for Joint Chinese Word Segmentation and Part-of-Speech Tagging. In *Proceedings of ACL 2008*.
55. Liang Huang and David Chiang (2007). [Forest Rescoring: Faster Decoding with Integrated Language Models](#). In *Proceedings of ACL 2007*.
Nominated for the Best Paper Award.
56. Liang Huang, Kevin Knight, and Aravind Joshi (2006). [Statistical Syntax-Directed Translation with Extended Domain of Locality](#). In *Proceedings of AMTA 2006*.
57. Hao Zhang, Liang Huang, Dan Gildea and Kevin Knight (2006). [Synchronous Binarization for Machine Translation](#). In *Proc. of NAACL 2006*.
 Journal version appeared in *Computational Linguistics* (2009). (see above)
58. Stephanie Weirich and Liang Huang (2005). [A Design for Type-Directed Programming in Java](#). In *Electronic Notes in Theoretical Computer Science*, **138** (2), 117-136.
59. L. Huang, Y. Peng, Z. Wu, Z. Yuan, H. Wang and H. Liu (2003). Pseudo Context-Sensitive Models for Parsing Isolating Languages: Classical Chinese – A Case Study. In *Proceedings of the International Conference on Intelligent Text Processing and Computational Linguistics (CICLING)*. (poster)

REFEREED WORKSHOP PAPERS

60. Liang Huang (2007). Binarization, Synchronous Binarization, and Target-side Binarization. In *Proceedings of NAACL Workshop on Syntax and Structure in Statistical Translation (SSST)*.
61. Liang Huang and David Chiang (2005). [Better k-best Parsing](#). In *Proceedings of the 9th International Workshop on Parsing Technologies (IWPT)*.
62. Liang Huang, Hao Zhang and Daniel Gildea (2005). [Machine Translation as Lexicalized Parsing with Hooks](#). In *Proceedings of the 9th International Workshop on Parsing Technologies (IWPT)*.
63. Stephanie Weirich and Liang Huang (2004). A Design for Type-Directed Programming in Java. In *Proceedings of the Workshop on Object-Oriented Developments (WOOD)*. Extended version as Penn/CIS Technical Report MS-CIS-04-11.
64. L. Huang, Y. Peng, H. Wang, and Z. Wu (2002). PCFG Parsing for Restricted Classical Chinese Texts. In *Proceedings of the COLING Workshop on Chinese Processing (SIGHAN)*, Taipei.

- POSTDOCS
- [Lemao Liu](#) (2013/11–2014/7): machine translation, structured prediction. Now Research Scientist at NICT, Japan.
 - [Feifei Zhai](#) (2014/9–2015/8): machine translation, unsupervised learning. Now Research Staff Member at IBM Watson.
- PHD THESES SUPERVISED (4)
- [Ashish Vaswani](#) (USC, co-advised by D. Chiang). Defended June 2014. Unsupervised learning, small models, deep learning. **Now Sr. Research Scientist at Google Brain.**
 - [James Cross](#) (CUNY→Oregon State). Defended December 2016. Parsing with Recurrent Neural Networks. **Now Research Scientist at Facebook.**
 - [Kai Zhao](#) (CUNY→Oregon State). Defended June 2017. structured prediction, parsing, machine translation. **Now Research Scientist at Google NYC.**
 - [Mingbo Ma](#) (CUNY→Oregon State). Defended September 2018. **Now Research Scientist at Baidu Research USA**
- PHD STUDENTS (ACTIVE)
- [Yilin Yang](#) (Oregon State). Since Fall 2017.
 - [Renjie Zheng](#) (Oregon State). Since Winter 2018.
- M.S. THESES SUPERVISED (3)
- [Luyao Zhang](#) (2015–2016). Now CS instructor at Oregon State.
 - He Zhang (2017–2018). Now Research Engineer at Baidu Research USA.
 - Kaibo Liu (2017–2018). Now Research Engineer at Baidu Research USA.
- M.S. THESIS STUDENTS (ACTIVE)
- Matthew Meyn
 - Liang Zhang
- VISITING STUDENTS & SUMMER INTERNS
- [Renjie Zheng](#) (2017/10-12): converted to PhD student at OSU.
 - Heng Yu (Ph.D. student from CAS/ICT, 2013/7–11): machine translation, latent variable structured prediction. Now Research Scientist at Alibaba.
 - Zhuoran Yu (M.S. student from NYU-Poly, 2013/6–2013/9): parsing algorithms. Now Software Engineer at Google Research, NYC.
 - Licheng Fang (PhD, Rochester), Summer 2011. Now Software Engineer at Google.
 - [Alexander Rush](#) (PhD, MIT), Summer 2010. Now Asst. Professor at Harvard.
 - [Yoav Goldberg](#) (PhD, Ben Gurion), Summer 2010 (with Knight & Chiang). Now Senior Lecturer at Bar Ilan University.
- VISITING SCHOLARS & PROFESSORS
- Dr. Haitao Mi, CAS/ICT, China, 2010–2011. Now Research Scientist at IBM.
- PHD COMMITTEE SERVICE
- External Committee Member/Examiner:
- Francesco Sartorio, Università di Padova (advisor: Giorgio Satta), expected Feb. 2015
 - Licheng Fang, University of Rochester (advisor: Dan Gildea), defended Aug. 2013
- Internal:
- Liping Liu, OSU (advisor: Tom Diettrich)
 - Reza Ghaeini, OSU (advisor: Xiaoli Fern)
 - Hamed Shahbazi, OSU (advisor: Xiaoli Fern)

- Rasha Obeidat, OSU (advisor: Xiaoli Fern)
- Zahra Iman, OSU (advisor: Scott Sanner)
- Xiannian Fan, CUNY (advisor: Changhe Yuan), defended Apr. 2016
- Jie Chu, CUNY (advisor: Ping Ji), expected 2015
- Pengfei Lu, CUNY (advisor: Matt Huenerfauth), defended Oct. 2013
- Zheng Chen, CUNY (advisor: Heng Ji), defended Sep. 2012.

INVITED TALKS

“Better k -best Parsing, Hypergraphs, and Dynamic Programming.”

- USC Information Sciences Institute (NL seminar), Marina del Rey, CA *June 2005*
- Google Research, Mountain View, CA *Oct 2005*
- New York University (NYCNLP series), New York, NY *Nov 2005*
- University of Rochester (big picture series), Rochester, NY *Nov 2005*
- Microsoft Research, Redmond, WA [\[video\]](#) *Dec 2005*
- University of Alberta (AI Seminar), Edmonton, Canada *Oct 2006*
- Universität Potsdam (CL Kolloquium), Potsdam, Germany *June 2007*
- Microsoft Research Asia, Beijing *July 2007*
- Institute of Automation, Chinese Academy of Sciences, Beijing *Nov 2007*

“Fast Decoding with Synchronous Grammars and n -gram Models.”

- Microsoft Research, Redmond [\[video\]](#) *Dec 2006*

“Binarizing Synchronous Grammars for Machine Translation.”

- Institute of Computing Tech., Chinese Academy of Sciences, Beijing *July 2007*
- Hong Kong University of Science and Technology (HKUST) *July 2007*

“Forest-based Algorithms in Natural Language Processing.”

- Google Research, Mountain View, CA [\[video\]](#) *Mar 2008*
- Johns Hopkins University (CLSP Seminar), Baltimore, MD [\[video\]](#) *Apr 2008*
- Carnegie Mellon University (LTI Seminar), Pittsburgh, PA *May 2009*
- Yahoo! Research, Santa Clara, CA *Aug 2009*

“Tree-based and Forest-based Translation.”

- BBN Technologies, Cambridge, MA *Oct 2008*
- Hong Kong University of Science and Technology (HKUST) *Nov 2008*
- The Chinese University of Hong Kong (CUHK) *Nov 2008*
- University of California at Berkeley (host: Dan Klein) *Feb 2009*
- Pomona CS Colloquium + Harvey Mudd (hosts: Bruce and Kauchak) *Oct 2010*

“Linear-time Dynamic Programming for Incremental Parsing.”

- University of California at San Diego (host: Roger Levy) *May 2010*
- Google Research, Mountain View (host: Hiyan Alshawi) *July 2010*
- Johns Hopkins University (CLSP Seminar, host: Fred Jelinek) [\[video\]](#) *Sep 2010*
- MIT CSAIL (host: Michael Collins) *Oct 2010*
- IBM Research, T.J. Watson (hosts: Salim Roukos and Bing Zhao) *May 2011*
- AT&T Labs Research, Florham Park, NJ *Jan 2013*
- Educational Testing Service (ETS), Princeton, NJ *May 2013*

- Facebook, Palo Alto (NLP Faculty Summit) *Oct 2015*

- “Structured Learning with Inexact Inference.”
 - Columbia University (host: Michael Collins) *Apr 2012*
 - University of Massachusetts, Amherst *Mar 2013*
 - Baidu Inc., Beijing *Jan 2014*

- “Search-Aware Tuning for Machine Translation.”
 - Bloomberg Research *Nov 2014*
 - Columbia University *Oct 2014*

- “Large-Scale Discriminative Training for Machine Translation.”
 - USC Information Sciences Institute (USC/ISI) [video] *Sep 2013*
 - Johns Hopkins University *Nov 2013*
 - Microsoft Research Asia, Beijing *Jan 2014*
 - Microsoft Research, Redmond, WA [video] *Feb 2014*

- “Linear-time Language Understanding and Learning.”
 - Carnegie Mellon University (Faculty Candidate Talk) *Feb 2012*
 - TTI Chicago *Mar 2012*
 - University of Rochester (CS Colloquium) *Oct 2012*
 - Stony Brook University (CS Colloquium) *Sep 2013*
 - Tsinghua University *Jan 2014*
 - University of Washington, Seattle, WA (UW-MSR Joint Symposium) *Feb 2014*
 - University of Oregon, Eugene, OR *Feb 2015*

- “Linear-Time Structure Prediction in Language and Biology.”
 - University of Rochester, CS Colloquium *Feb 2017*

- “Marrying Dynamic Programming and Recurrent Neural Networks.”
 - Keynote at EMNLP 2017 Workshop on Structured Prediction *Sep 2017*
 - Facebook AI Research *Nov 2017*

- “Linear-Time Prediction of RNA Secondary Structures.”
 - Center for Genome Research and Biocomputing, Oregon State University *Feb 2017*
 - University of Rochester Medical School, Bioinformatics Cluster *June 2018*
 - PingWest SYNC 2018 Silicon Valley Innovation Conference *Aug 2018*
 - EterRNA Conference 2018, Stanford University Medical School *Aug 2018*
 - Keynote at First West Coast NLP Summit (@Facebook) *Sep 2018*

- “Breakthrough in Simultaneous Translation.”
 - Google AI *Nov 2018*
 - Stanford NLP Seminar *Jan 2019*

- NSF Panelist, 2014, 2015, 2017 × 2, 2019.
- Grant Proposal Reviewer, Foreign NSFs: Canadian (NSERC), 2017; Hong Kong Research Grants Council, 2017; The Netherlands (NWO), 2017; Israeli (ISF), 2015.

PROFESSIONAL
SERVICE

- Program Co-Chair, International Conference on Parsing Technologies (IWPT 2013)
- Area Chair, ACL 2019 (sentence-level semantics)
- Area Chair, EMNLP 2018 (Syntax and Parsing)
- Senior Area Chair, ACL 2018 (Syntax and Parsing)
- Area Chair, IJCNLP 2017 (Tagging and Parsing)
- Area Chair, EMNLP 2016 (Segmentation, Tagging, and Parsing)
- Area Chair, ACL 2014 (Machine Translation), and ACL 2012 (Syntax and Parsing)
- Senior Program Committee, IJCAI 2016 and IJCAI 2013
- Standing Review Committee, Transactions of the ACL (TACL), 2014–2016
- Textbook Reviewer: *Cambridge University Press*, 2016; *Oxford University Press*, 2010
- Journal Reviewer for
 - *Computational Linguistics*, 2008–present
 - *Transactions of the Association of Computational Linguistics (TACL)*, 2012–
 - *Journal of Artificial Intelligence Research (JAIR)*, 2014, 2016
 - *Journal of Natural Language Engineering*, 2011
 - *Computational Intelligence*, 2011
 - *ACM Transactions on Intelligent Systems and Technology*, 2010–2011
 - *Journal of Computer Science and Technology*, 2010
 - *IEEE Transactions on Audio, Speech and Language Processing*, 2008
 - *ACM Transactions on Asian Language Information Processing*, 2008, 2011
- Conference PC Member for
 - ACL, annually 2006–present (top conference)
 - HLT-NAACL, annually 2007–present (top conference)
 - EMNLP, annually 2007–present (top conference)
 - NIPS 2016, 2017 (top conference)
 - ICML 2016, 2017 (top conference)
 - ICLR 2016, 2017
 - AAAI 2015 (top conference)
 - IWPT 2011; MT Summit 2009; AMTA 2008; CoNLL 2007–2009

INTERNAL SERVICE

- Data Science Hiring Committee, College of Engineering, Oregon State U., 2016–2017.
- CS Undergrad. Curriculum Committee, School of EECS, Oregon State U., 2015–2016.
- Curriculum Committee, CUNY Graduate Center CS Program, 2012–2015.
- Co-Organizer, [CUNY NLP Seminar Series](#), 2012–2015.
- Co-Organizer and Coach, CUNY/QC Programming Contests, 2013–2015.
- PhD Admissions and Fellowship Committee, USC CS Dept, 2010–2012.
- Co-Organizer and Coach, USC Programming Contests (ACM/ICPC), 2009–2012.

Champions, Southern California Regionals, 2011 (first time in USC history)

- Co-Organizer, NACLO: Computational Linguistics Olympiad, USC/ISI site, 2009–2012.

OTHER
INFORMATION

- Citizen of the People’s Republic of China. Permanent Resident of the US.
- Born April 1981, Shanghai, China.
- Human Languages: [Wu](#) (mother tongue), Mandarin (native); English (fluent); French, Spanish and Italian (basic); Japanese (basic conversational).
- Computer Languages: Python, Haskell, OCaml, \LaTeX , Java, C/C++, Pascal, Prolog.
- Hobbies/Interests: Historical Geography, Historical Linguistics, History of Mathematics, Classical Chinese (esp. poetry), Classical Music, Ping-Pong, Badminton, Football (Soccer not “handegg”!), Go, English poetry (esp. Shakespearean sonnet and Robert Frost), etc.

LAST UPDATED December 2018.