

Lihong Li

Google Inc.
747 Sixth Street South
Kirkland, WA, USA 98033

lihong@google.com
lihongli.cs@gmail.com
<https://lihongli.github.io>

RESEARCH INTERESTS

My core research interest is in **machine learning for interactive systems that maximizes a utility function by taking actions**, which is in contrast to *prediction*-oriented machine learning like supervised learning. My areas of focus are **reinforcement learning** and **multi-armed bandits**, although I am also interested in related areas such as large-scale online learning with big data, active learning, and planning. In the past, I have applied my work to recommendation, Web search, advertising, conversation systems, and spam detection.

EDUCATION

| | | |
|-------------------|--------|---|
| 01/2005 – 05/2009 | Ph.D. | Computer Science, Rutgers University, USA |
| 09/2002 – 07/2004 | M.Sc. | Computing Science, University of Alberta, Canada |
| 09/1998 – 07/2002 | B.Eng. | Computer Science and Technology, Tsinghua University, China |

RESEARCH & INDUSTRY EXPERIENCE

| | |
|-------------------|---|
| 10/2017 - present | Research Scientist at Google Inc. |
| 03/2017 – 10/2017 | Principal Researcher at Microsoft Research |
| 02/2016 – 02/2017 | Senior Researcher at Microsoft Research |
| 06/2012 – 02/2016 | Researcher at Microsoft Research |
| 09/2010 – 06/2012 | Research Scientist at Yahoo! Research |
| 06/2009 – 08/2010 | Postdoctoral Scientist at Yahoo! Research |
| 06/2008 – 08/2008 | Research Intern at AT&T Shannon Labs |
| 05/2007 – 08/2007 | Research Intern at Yahoo! Research NYC |
| 05/2006 – 08/2006 | Engineering Intern at Google NYC |
| 01/2005 – 05/2009 | Graduate Research Assistant at the Rutgers University |
| 09/2002 – 07/2004 | Research Assistant at the University of Alberta |

SELECTED AWARDS

| | | |
|------|--------|--|
| 2011 | USA | Yahoo! Super Star Team Award (highest team achievement award in the company) |
| 2011 | USA | Notable Paper Award, AISTATS |
| 2011 | USA | Best Paper Award, WSDM |
| 2008 | USA | Best Student Paper Award, ICML |
| 2004 | Canada | Teaching Assistant Award, University of Alberta |

TEACHING/ADVISING EXPERIENCE

- | | |
|--------------------|---|
| Summers since 2013 | Supervised student interns at Microsoft Research and at Google Projects on reinforcement learning, multi-armed bandits, imitation learning and Web search |
| Summers 2010/2011 | Supervised student interns at Yahoo! Labs Projects on anomaly detection in distributed file systems, large-scale prediction models in advertising, and news ranking |
| Spring 2009 | Guest lecturer for a graduate-level course at the Rutgers University Taught the least-squares policy iteration (LSPI) algorithm in the course “Learning and Sequential Decision Making”. |
| 09/2007 – 12/2007 | Co-organizer for a graduate seminar at the Rutgers University Compiled reading materials, arranged weekly meetings, and presented papers for “Planning in Learned Environments” (w/ Michael Littman). |
| 05/2005 – 08/2005 | Organizer for a graduate seminar at the Rutgers University Compiled reading materials, arranged weekly meetings, presented papers, and invited an external speaker for “Abstractions and Hierarchies for Learning and Planning”. |
| 09/2002 – 07/2004 | Teaching Assistant at the University of Alberta Taught seminar sessions and graded homework for the undergraduate course on discrete mathematics: “Formal Systems and Logic in Computing Science”. |

PROFESSIONAL ACTIVITIES

- Conference Organization
 - Area Chair and/or Senior Program Committee Member
 - * AAAI Conference on Artificial Intelligence (AAAI): 2017, 2018
 - * International Conference on Artificial Intelligence and Statistics (AISTATS): 2017
 - * International Conferences on Machine Learning (ICML): 2012–2017
 - * International Joint Conferences on Artificial Intelligence (IJCAI): 2011, 2016, 2017
 - * Annual Conferences on Neural Information Processing Systems (NIPS): 2014, 2017
 - Workshop Co-chairs/Organizers
 - * Reinforcement Learning Competition (ICML/UAI/COLT’09 Workshop)
 - * PASCAL2 Exploration & Exploitation Challenge (ICML’12 Workshop)
 - * Large-Scale Online Learning and Decision-Making Workshop (Cumberland Lodge, 2012)
 - * IEEE BigData Workshop (DC, USA, 2014)
 - * WWW Workshop on Offline and Online Evaluation of Web-based Services (Florence, Italy, 2015)
 - * SIAM Conference on Optimization — Algorithms for Reinforcement Learning Minisymposium (Vancouver, Canada, 2017)
 - * AI Frontiers (San Jose, CA, USA, November 2017)
 - * From “What If” to “What Next” (NIPS’17 Workshop)
 - Workshop program committee member
 - * Planning and Learning in A Priori Unknown or Dynamic Domains, IJCAI 2005
 - * Abstraction in Reinforcement Learning, ICML/UAI/COLT 2009
 - * Bayesian Optimization, Experimental Design and Bandits, NIPS, 2011
 - * AdML: Online Advertising Workshop, ICML 2012
 - * Bayesian Optimization & Decision Making, NIPS 2012
- Tutorials
 - “Offline Evaluation and Optimization for Interactive Systems: A Practical Guide”, at the *Eighth International Conference on Web Search and Data Mining (WSDM)*, Shanghai, China, February, 2015.
 - “Neural Approaches to Conversational AI”, with Jianfeng Gao and Michel Galley, at the *Fifty-Sixth Annual Meeting of the Association for Computational Linguistics (ACL)*, Melbourne, Australia, July, 2018.
- Referee for funding agencies

- Natural Sciences and Engineering Research Council of Canada (NSERC)
- United States-Israel Binational Science Foundation (BFS)
- Referee for journals
 - ACM Transactions on Intelligent Systems and Technology
 - ACM Transactions on Knowledge Discovery from Data
 - Advances in Complex Systems
 - Artificial Intelligence
 - Artificial Intelligence Communications
 - Computer Speech and Language
 - Data Mining and Knowledge Discovery
 - IEEE Journal of Selected Topics in Signal Processing
 - IEEE Transactions on Automatic Control
 - IEEE Transactions on Knowledge and Data Engineering
 - IEEE Transactions on Neural Networks
 - IEEE Transactions on Wireless Communications
 - Journal of Artificial Intelligence Research
 - Journal of Computer Science and Technology
 - Journal of Machine Learning Research
 - Journal of Selected Topics in Signal Processing
 - Machine Learning
 - Mathematics of Operations Research
 - Neural Computation
 - Neurocomputing
- Referee for conferences (including services as area chair and senior program committee member):
 - AAAI (AAAI Conferences on Artificial Intelligence): 2006, 2008, 2010, 2016 (Demo), 2017 (SPC)
 - AISTATS (International Conferences on Artificial Intelligence and Statistics): 2011, 2017 (SPC)
 - ALT (International Conferences on Algorithmic Learning Theory): 2015
 - COLT (Annual Conferences on Learning Theory): 2010, 2011, 2012, 2015
 - ECML (European Conferences on Machine Learning): 2009
 - KDD (ACM SIGKDD Conferences on Knowledge Discovery and Data Mining): 2012
 - ICML (International Conferences on Machine Learning): 2009–2011, 2012–2017 (AC)
 - IJCAI (International Joint Conferences on Artificial Intelligence): 2007, 2011 (SPC), 2015, 2016 (SPC)
 - NAACL-HLT (Conference of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies): 2018.
 - NIPS (Annual Meetings on Neural Information Processing Systems): 2008–2013, 2014 (AC)
 - STOC (ACM Symposium on Theory of Computing): 2014
 - UAI (Annual Conferences on Uncertainty in Artificial Intelligence): 2010, 2012, 2016
 - UbiComp (International Conferences on Ubiquitous Computing): 2011
 - WSDM (ACM International Conferences on Web Search and Data Mining): 2012, 2013
 - WWW (International Conferences on World Wide Web): 2012
- Open source and dataset contributions
 - Vowpal Wabbit: an open source project started with John Langford and Alexander L. Strehl for fast online learning in large-scale prediction problems. URL: <http://www.hunch.net/~vw>
 - Yahoo! Front Page Today Module User Click Log Dataset: the first large-scale real-life dataset that supports unbiased evaluation of multi-armed bandit algorithms (with help from Wei Chu). URL: <http://webscope.sandbox.yahoo.com/catalog.php?datatype=r>
 - Deep reinforcement learning package in Microsoft CNTK (with Yi Mao et al.)

INVITED TALKS

- Primal-dual Approaches to Reinforcement Learning

- ISMP, Bordeaux, France. July, 2018. (scheduled)
- INFORMS, Taipei, Taiwan. June, 2018. (scheduled)
- CISS, Princeton, NJ, USA. March, 2018. (scheduled)
- Stanford University, Palo Alto, CA, USA. February, 2018. (scheduled)
- Google Brain, Montreal, QC, Canada. September, 2017.
- New York University, New York, NY, USA. May, 2017.
- Simons Institute, Berkeley, CA, USA. February, 2017.
- Reinforcement Learning for Conversational Systems
 - Google Brain, Montreal, QC, Canada. September, 2017.
 - ICML Workshop on Interactive Machine Learning and Semantic Information Retrieval, Sydney, AU. August, 2017.
 - Multidisciplinary Conference on Reinforcement Learning and Decision Making (RLDM), Ann Arbor, MI, USA. June, 2017.
 - Korea Advanced Institute of Science and Technology, Korea. June 2017.
 - Sungkyunkwan University, Suwon, Korea. June 2017.
 - ACML Workshop on Reinforcement Learning, Hamilton, NZ. November, 2016.
 - Global AI Conference, Shanghai, China. November, 2016.
- Off-policy Learning and Off-line Evaluation
 - Graduate School of Business, Stanford University, CA, USA. May, 2017.
 - Oxford University, Oxford, UK. November, 2015.
 - Google DeepMind, London, UK. November, 2015.
 - AdTech LA Meetup, Santa Monica, CA, USA. October, 2015.
 - UW CSE MSR Summer Institute, Union, WA, USA. August, 2015.
 - INRIA SequeL, Lille, France. December, 2014.
 - Criteo, Paris, France. December, 2014.
 - Department of Computing Science, University of Alberta, Edmonton, AB, Canada. November, 2014.
 - KDD Workshop on User Engagement Optimization, New York, NY, USA. August, 2014.
 - AAAI Workshop on Sequential Decision-Making with Big Data, Québec City, QC, Canada. July, 2014.
 - Microsoft Research Latin American Faculty Summit, Viña del Mar, Chile. May, 2014.
 - IEEE Information Theory and Application (ITA) Workshop, San Diego, CA, USA. February, 2014.
 - Distinguished Faculty and Graduate Student Seminar, Department of Statistics, University of Michigan, Ann Arbor, MI, USA. February, 2014.
- Machine Learning in the Bandit Setting: Algorithms, Evaluation, and Case Studies
 - Department of Computer Science, University of South California, Los Angeles, CA, USA. October, 2015.
 - Department of Computer Science, Purdue University, West Lafayette, IN, USA. April, 2014.
 - Joint Statistical Meetings (Statistics in Marketing Track), Montreal, QC, Canada. August, 2013.
 - Tenth National Symposium of Search Engine and Web Mining, Beijing, China. May 2012.
 - Microsoft Research Asia, Beijing, China. May 2012.
 - Department of Machine Intelligence, Peking University, Beijing, China. May 2012.
 - Department of Computer Science and Technology, Tsinghua University, Beijing, China. May 2012.
 - Department of Computer Science and Engineering, University of California, Los Angeles, CA, USA. May 2012.
 - Department of Computer Science and Engineering, University of California, San Diego, CA, USA. May 2012.
 - Department of Computer Science, University of California, Irvine, CA, USA. May 2012.
 - Google Research, Mountain View, CA, USA. April 2012.
 - Microsoft Research, Redmond, WA, USA. April 2012.
 - Adobe Advanced Technology Labs, San Jose, CA, USA. April 2012.
 - Microsoft Research, Mountain View, CA, USA. April 2012.
 - Department of Computer Science, Virginia Tech, Blacksburg, VA, USA. February 2012.
 - Department of Computer Science, Johns Hopkins University, MD, USA. February 2012.
 - Technicolor Research Center, Palo Alto, CA, USA. February 2012.

- Department of Computing Science, University of Alberta, Edmonton, AB, Canada. June 2011.
- Industrial Affiliates Annual Conference, Department of Statistics, Stanford University, USA. May 2011. With Deepak Agarwal and Bee-Chung Chen.
- Microsoft Silicon Valley Center, Mountain View, CA, USA. March 2011.
- Artificial Intelligence Center, SRI International, Menlo Park, CA, USA. April 2010.
- “Vowpal Wabbit for Extremely Fast Machine Learning”
 - GraphLab Workshop on Big Learning, San Francisco, CA, USA. July, 2012.
 - First data mining meetup on large-scale ML algorithms, San Francisco, CA, USA. August 2011.
- “A Unifying Framework for Computational Reinforcement Learning Theory”
 - ICML Workshop on Planning and Acting with Uncertain Models, Bellevue, WA, USA. June 2011.
 - Department of Computing Science, University of Alberta, Edmonton, AB, Canada. June 2011.
 - Yahoo! Research, Sunnyvale, CA, USA. April 2009.
 - Google Research, New York, NY, USA. April 2009.
 - Yahoo! Research, New York, NY, USA. January 2009.
 - Reasoning and Learning Laboratory, McGill University, McGill, QC, Canada. May 2008.
 - AT&T Shannon Labs, Florham Park, NJ, USA. January 2008.
- “Sparse Online Learning via Truncated Gradient”
 - Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA. November 2009.
 - eBay Research Labs, San Jose, CA, USA. April 2009.
 - Department of Information Analysis & Management, NEC Laboratories America, Cupertino, CA, USA. April 2009.
 - Text Analysis and Machine Learning Group, University of Ottawa, Ottawa, ON, Canada. May 2008.
- “Go as a Testbed for Advancing Reinforcement Learning Research”
 - DARPA Information Processing Technology meeting, Arlington, VA, USA. February 2008.

PUBLICATIONS

Journal Papers

- (J1) M. Dudík, D. Erhan, J. Langford, and L. Li: Doubly robust policy evaluation and optimization. In *Statistical Science*, 29(4):485–511, 2014.
- (J2) J. Bian, B. Long, L. Li, T. Moon, A. Dong, and Y. Chang: Exploiting user preference for online learning in Web content optimization systems. In *ACM Transactions on Intelligent Systems and Technology*, 5(2), 2014.
- (J3) T. Moon, W. Chu, L. Li, Z. Zheng, and Y. Chang: Refining recency search results with user click feedback. In *ACM Transactions on Information Systems*, 30(4), 2012.
- (J4) J. Langford, L. Li, P. McAfee, and K. Papineni: Cloud control: Voluntary admission control for Intranet traffic management. In *Information Systems and e-Business Management*, 10(3):295–308, 2012.
- (J5) L. Li, M.L. Littman, T.J. Walsh, and A.L. Strehl: Knows what it knows: A framework for self-aware learning. In *Machine Learning*, 82(3):399–443, 2011.
- (J6) L. Li and M.L. Littman: Reducing reinforcement learning to KWIK online regression. In the *Annals of Mathematics and Artificial Intelligence*, 58(3–4):217–237, 2010.
- (J7) J. Langford, L. Li, J. Wortman, and Y. Vorobeychik: Maintaining equilibria during exploration in sponsored search auctions. In *Algorithmica*, 58(4):990–1021, 2010.
- (J8) A.L. Strehl, L. Li, and M.L. Littman: Reinforcement learning in finite MDPs: PAC analysis. In the *Journal of Machine Learning Research*, 10:2413–2444, 2009.
- (J9) E. Brunskill, B.R. Leffler, L. Li, M.L. Littman, and N. Roy: Provably efficient learning with typed parametric models. In the *Journal of Machine Learning Research*, 10:1955–1988, 2009.
- (J10) J. Langford, L. Li, and T. Zhang: Sparse online learning via truncated gradient. In the *Journal of Machine Learning Research*, 10:777–801, 2009.
- (J11) T.J. Walsh, A. Nouri, L. Li, and M.L. Littman: Planning and learning in environments with delayed feedback. In the *Journal of Autonomous Agents and Multi-Agent Systems*, 18(1):83–105, 2009.

- (J12) L. Li, V. Bulitko, and R. Greiner: Focus of attention in reinforcement learning. In the *Journal of Universal Computer Science*, 13(9):1246–1269, 2007.
- (J13) L. Li, M. Shao, Z. Zheng, C. He, and Z.-H. Du: Typical XML document transformation methods and an application system (in Chinese). *Computer Science*, 30(2):40–44, February, 2003.

Refereed Conference Papers

- (C1) B. Dai, A. Shaw, N. He, L. Li, and Le Song: Boosting the actor with dual critic. In the *6th International Conference on Learning Representations (ICLR)*, 2018.
- (C2) Z. Lipton, X. Li, J. Gao, L. Li, F. Ahmed, and L. Deng: Efficient dialogue policy learning with BBQ-networks. In the *32nd AAAI Conference on Artificial Intelligence (AAAI)*, 2018.
- (C3) J. Chen, C. Wang, L. Xiao, J. He, L. Li, and L. Deng: Q-LDA: Uncovering latent patterns in text-based sequential decision processes. In *Advances in Neural Information Processing Systems 30 (NIPS)*, 2017.
- (C4) B. Peng, X. Li, L. Li, J. Gao, A. Celikyilmaz, S. Lee, and K.-F. Wong: Composite task-completion dialogue system via hierarchical deep reinforcement learning. In the *2017 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2017.
- (C5) L. Li, Y. Lu, and D. Zhou: Provably optimal algorithms for generalized linear contextual bandits. In the *34th International Conference on Machine Learning (ICML)*, 2017.
- (C6) S. Du, J. Chen, L. Li, L. Xiao, and D. Zhou: Stochastic variance reduction methods for policy evaluation. In the *34th International Conference on Machine Learning (ICML)*, 2017.
- (C7) B. Dhingra, L. Li, X. Li, J. Gao, Y.-N. Chen, F. Ahmed, and L. Deng: Towards end-to-end reinforcement learning of dialogue agents for information access. In the *55th Annual Meeting of the Association for Computational Linguistics (ACL)*, 2017.
- (C8) E. Parisotto, A. Mohamed, R. Singh, L. Li, D. Zhou, and P. Kohli: Neuro-symbolic program synthesis. In the *5th International Conference on Learning Representations (ICLR)*, 2017.
- (C9) T.K. Huang, L. Li, A. Vartanian, S. Amershi, and J. Zhu: Active learning with oracle epiphany. In *Advances in Neural Information Processing Systems 29 (NIPS)*, 2016.
- (C10) J. He, M. Ostendorf, X. He, J. Chen, J. Gao, L. Li, and L. Deng: Deep reinforcement learning with a combinatorial action space for predicting and tracking popular discussion threads. In the *2016 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2016.
- (C11) C.-Y. Liu and L. Li: On the Prior Sensitivity of Thompson Sampling. In the *27th International Conference on Algorithmic Learning Theory (ALT)*, 2016.
- (C12) J. He, J. Chen, X. He, J. Gao, L. Li, L. Deng, and M. Ostendorf: Deep reinforcement learning with a natural language action space. In the *54th Annual Meeting of the Association for Computational Linguistics (ACL)*, 2016.
- (C13) N. Jiang and L. Li: Doubly robust off-policy value evaluation for reinforcement learning. In the *33rd International Conference on Machine Learning (ICML)*, 2016.
- (C14) S. Agrawal, N. R. Devanur, and L. Li: An efficient algorithm for contextual bandits with knapsacks, and an extension to concave objectives. In the *29th Annual Conference on Learning Theory (COLT)*, 2016.
- (C15) M. Zoghi, T. Tunys, L. Li, D. Jose, J. Chen, C.-M. Chin, and M. de Rijke: Click-based hot fixes for underperforming torso queries. In the *39th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, 2016.
- (C16) J. He, J. Chen, X. He, J. Gao, L. Li, L. Deng, and M. Ostendorf: Deep reinforcement learning with an unbounded action space. In the *International Conference on Learning Representations (ICLR), Workshop Track*, 2016.
- (C17) L. Li, R. Munos, and Cs. Szepesvári: Toward minimax off-policy value estimation. In the *18th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2015.
- (C18) L. Li, S. Chen, J. Kleban, and A. Gupta: Counterfactual estimation and optimization of click metrics in search engines: A case study. In the *24th International Conference on World Wide Web (WWW), Companion*, 2015.
- (C19) L. Li, J. Kim, and I. Zitouni: Toward predicting the outcome of an A/B experiment for search relevance. In the *8th International Conference on Web Search and Data Mining (WSDM)*, 2015.
- (C20) L. Li, H. He, and J.D. Williams: Temporal supervised learning for inferring a dialog policy from example conversations. In the *IEEE Spoken Language Technology Workshop (SLT)*, 2014.

- (C21) A. Agarwal, D. Hsu, S. Kale, J. Langford, *L. Li*, and R.E. Schapire: Taming the monster: A fast and simple algorithm for contextual bandits. In *the 31st International Conference on Machine Learning (ICML)*, 2014.
- (C22) E. Brunskill and *L. Li*: PAC-inspired option discovery in lifelong reinforcement learning. In *the 31st International Conference on Machine Learning (ICML)*, 2014.
- (C23) E. Brunskill and *L. Li*: Sample complexity of multi-task reinforcement learning. In *the 29th Conference on Uncertainty in Artificial Intelligence (UAI)*, 2013.
- (C24) M. Dudík, D. Erhan, J. Langford, and *L. Li*: Sample-efficient nonstationary-policy evaluation for contextual bandits. In *the 28th Conference on Uncertainty in Artificial Intelligence (UAI)*, 2012.
- (C25) *L. Li*, W. Chu, J. Langford, T. Moon, and X. Wang: An unbiased offline evaluation of contextual bandit algorithms with generalized linear models. In *Journal of Machine Learning Research - Workshop and Conference Proceedings 26: On-line Trading of Exploration and Exploitation 2*, 2012.
- (C26) V. Navalpakkam, R. Kumar, *L. Li*, and D. Sivakumar: Attention and selection in online choice tasks. In *the 20th International Conference on User Modeling, Adaptation and Personalization (UMAP)*, 2012.
- (C27) H. Wang, A. Dong, *L. Li*, Y. Chang, and E. Gabrilovich: Joint relevance and freshness learning From click-throughs for news search. In *the 21st International Conference on World Wide Web (WWW)*, 2012.
- (C28) O. Chapelle and *L. Li*: An empirical evaluation of Thompson sampling. In *Advances in Neural Information Processing Systems 24 (NIPS)*, 2012.
- (C29) M. Dudík, J. Langford, and *L. Li*: Doubly robust policy evaluation and learning. In *the 28th International Conference on Machine Learning (ICML)*, 2011.
- (C30) W. Chu, M. Zinkevich, *L. Li*, A. Thomas, and B. Tseng: Unbiased online active learning in data streams. In *the 17th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, 2011.
- (C31) D. Agarwal, *L. Li*, and A.J. Smola: Linear-time algorithms for propensity scores. In *the 14th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2011.
- (C32) A. Beygelzimer, J. Langford, *L. Li*, L. Reyzin, and R.E. Schapire: Contextual bandit algorithms with supervised learning guarantees. In *the 14th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2011. **Co-winner of the Notable Paper Award.**
- (C33) W. Chu, *L. Li*, L. Reyzin, and R. Schapire: Linear contextual bandit problems. In *the 14th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2011.
- (C34) *L. Li*, Wei Chu, John Langford, and Xuanhui Wang: Unbiased offline evaluation of contextual-bandit-based news article recommendation algorithms. In *the 4th ACM International Conference on Web Search and Data Mining (WSDM)*, 2011. **Winner of the Best Paper Award.**
- (C35) A.L. Strehl, J. Langford, *L. Li*, and S. Kakade: Learning from logged implicit exploration data. In *Advances in Neural Information Processing Systems 23 (NIPS)*, 2011.
- (C36) M. Zinkevich, M. Weimer, A.J. Smola, and *L. Li*: Convergence rates of parallel online learning via stochastic gradient descent. In *Advances in Neural Information Processing Systems 23 (NIPS)*, 2011.
- (C37) T. Moon, *L. Li*, W. Chu, C. Liao, Z. Zheng, and Y. Chang: Online learning for recency search ranking using real-time user feedback (short paper). In *the 19th ACM Conference on Information and Knowledge Management (CIKM)*, 2010.
- (C38) *L. Li*, W. Chu, J. Langford, and R.E. Schapire: A contextual-bandit approach to personalized news article recommendation. In *the 19th International Conference on World Wide Web (WWW)*, 2010.
- (C39) Y. Xie, Y. Zhang, and *L. Li*: Neuro-fuzzy reinforcement learning for adaptive intersection traffic signal control. In *the Annual Meeting of Transportation Research Board (TRB)*, 2010.
- (C40) *L. Li*, J.D. Williams, and S. Balakrishnan: Reinforcement learning for spoken dialog management using least-squares policy iteration and fast feature selection. In *the 10th Annual Conference of the International Speech Communication Association (INTERSPEECH)*, 2009.
- (C41) C. Diuk, *L. Li*, and B.R. Leffler: The adaptive k -meteorologists problem and its application to structure learning and feature selection in reinforcement learning. In *the 26th International Conference on Machine Learning (ICML)*, 2009.
- (C42) J. Asmuth, *L. Li*, M.L. Littman, A. Nouri, and D. Wingate: A Bayesian sampling approach to exploration in reinforcement learning. In *the 25th International Conference on Uncertainty in Artificial Intelligence (UAI)*, 2009.

- (C43) L. Li, M.L. Littman and C.R. Mansley: Online exploration in least-squares policy iteration. In *the 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2009.
- (C44) L. Langford, L. Li, and T. Zhang: Sparse online learning via truncated gradient. In *Advances in Neural Information Processing Systems 21 (NIPS)*, 2009.
- (C45) L. Li: A worst-case comparison between temporal difference and residual gradient. In *the 25th International Conference on Machine Learning (ICML)*, 2008.
- (C46) L. Li, M.L. Littman, and T.J. Walsh: Knows what it knows: A framework for self-aware learning. In *the 25th International Conference on Machine Learning (ICML)*, 2008. **Co-winner of the Best Student Paper Award. A Google Student Award winner at the New York Academy of Sciences Symposium on Machine Learning, 2008.**
- (C47) R. Parr, L. Li, G. Taylor, C. Painter-Wakefield, and M.L. Littman: An analysis of linear models, linear value function approximation, and feature selection for reinforcement learning. In *the 25th International Conference on Machine Learning (ICML)*, 2008.
- (C48) E. Brunskill, B.R. Leffler, L. Li, M.L. Littman, and N. Roy: CORL: A continuous-state offset-dynamics reinforcement learner. In *the 24th Conference on Uncertainty in Artificial Intelligence (UAI)*, 2008.
- (C49) L. Li and M.L. Littman: Efficient value-function approximation via online linear regression. In *the 10th International Symposium on Artificial Intelligence and Mathematics (AI&Math)*, 2008.
- (C50) J. Wortman, Y. Vorobeychik, L. Li, and J. Langford: Maintaining equilibria during exploration in sponsored search auctions. In *the 3rd International Workshop on Internet and Network Economics (WINE)*, LNCS 4858, 2007.
- (C51) T.J. Walsh, A. Nouri, L. Li, and M.L. Littman: Planning and learning in environments with delayed feedback. In *the 18th European Conference on Machine Learning (ECML)*, LNCS 4701, 2007.
- (C52) R. Parr, C. Painter-Wakefield, L. Li, and M.L. Littman: Analyzing feature generation for value-function approximation. In *the 24th International Conference on Machine Learning (ICML)*, 2007.
- (C53) A.L. Strehl, L. Li, E. Wiewiora, J. Langford, and M.L. Littman: PAC model-free reinforcement learning. In *the 23rd International Conference on Machine Learning (ICML)*, 2006. **Best Student Poster Award winner at the New York Academy of Sciences Symposium on Machine Learning, 2006.**
- (C54) A.L. Strehl, L. Li, and M.L. Littman: Incremental model-based learners with formal learning-time guarantees. In *the 22nd Conference on Uncertainty in Artificial Intelligence (UAI)*, 2006.
- (C55) L. Li, T.J. Walsh, and M.L. Littman: Towards a unified theory of state abstraction for MDPs. In *the 9th International Symposium on Artificial Intelligence and Mathematics (AI&Math)*, 2006.
- (C56) L. Li, M.L. Littman: Lazy approximation for solving continuous finite-horizon MDPs. In *the 20th National Conference on Artificial Intelligence (AAAI)*, 2005.
- (C57) L. Li, V. Bulitko, and R. Greiner: Batch reinforcement learning with state importance (extended abstract). In *the 15th European Conference on Machine Learning (ECML)*, LNCS 3201, 2004.
- (C58) V. Bulitko, L. Li, R. Greiner, and I. Levner: Lookahead pathologies for single agent search (poster paper). In *the 18th International Joint Conference on Artificial Intelligence (IJCAI)*, 2003.
- (C59) I. Levner, V. Bulitko, L. Li, G. Lee, and R. Greiner: Towards automated creation of image interpretation systems. In *the 16th Australian Joint Conference on Artificial Intelligence*, LNCS 2903, 2003.
- (C60) L. Li, V. Bulitko, R. Greiner, and I. Levner: Improving an adaptive image interpretation system by leveraging. In *the 8th Australian and New Zealand Intelligent Information System Conference*, 2003.

Book and Book Chapters

- (B1) K. Hofmann, L. Li, and F. Radlinski: Online Evaluation for Information Retrieval. *Foundations and Trends in Information Retrieval*, 10(1):1–107, 2016. ISBN 978-1-68083-163-4.
- (B2) L. Li: Sample complexity bounds of exploration. In Marco Wiering and Martijn van Otterlo, editors, *Reinforcement Learning: State of the Art*, Springer Verlag, 2012.
- (B3) M. Shao, L. Li, Z. Zheng, and C. He: Practical Programming in XML. *Tsinghua University Press*, Beijing, China, December, 2002. ISBN 7-900643-85-0.

Theses

- (T1) *L. Li*: A unifying framework for computational reinforcement learning theory. *Doctoral dissertation*, Department of Computer Science, Rutgers University, New Brunswick, NJ, USA, May, 2009.
- (T2) *L. Li*: Focus of attention in reinforcement learning. *MSc thesis*, Department of Computing Science, University of Alberta, Edmonton, Alberta, Canada, July, 2004.
- (T3) *L. Li*: Design and implementation of an agent communication module based on KQML. *Bachelor degree thesis*, Department of Computer Science and Technology, Tsinghua University, Beijing, China, June, 2002.

Other Papers

- (O1) X. Li, Z.C. Lipton, B. Dhingra, *L. Li*, J. Gao, Y.-N. Chen: A user simulator for task-completion dialogues. MSR technical report, December 2016.
- (O2) E. Brunskill and *L. Li*: The online discovery problem and its application to lifelong reinforcement learning. CoRR abs/1506.03379, June 2015.
- (O3) D. Yankov, P. Berkhin, and *L. Li*: Evaluation of explore-exploit policies in multi-result ranking systems. Microsoft Journal on Applied Research, volume 3, pages 54–60, 2015. Also available as Microsoft Research Technical Report #MSR-TR-2015-34, May 2015.
- (O4) Z. Qin, V. Petricek, N. Karampatziakis, *L. Li*, and J. Langford: Efficient online bootstrapping for large scale learning. *NIPS Workshop on Big Data*, December, 2013. Also available as Microsoft Research Technical Report #MSR-TR-2013-132.
- (O5) *L. Li* and O. Chapelle: Regret bounds for Thompson sampling (Open Problems). In the *Twenty-Fifth Annual Conference on Learning Theory (COLT)*, 2012
- (O6) *L. Li* and M.L. Littman: Prioritized sweeping converges to the optimal value function. Technical report DCS-TR-631, Department of Computer Science, Rutgers University, May 2008.
- (O7) A.L. Strehl, *L. Li*, and M.L. Littman: PAC reinforcement learning bounds for RTDP and Rand-RTDP. *AAAI technical report WS-06-11*, pages 50-56, July 2006.
- (O8) *L. Li* and M.L. Littman: Lazy approximation: A new approach for solving continuous finite-horizon MDPs. Technical report DCS-TR-577, Department of Computer Science, Rutgers University, May 2005.
- (O9) *L. Li*, V. Bulitko, and R. Greiner: Focus of attention in sequential decision making. *AAAI technical report WS-04-08*, pages 43-48, July 2004.