Dan Roth

Research Contributions

Dan Roth is a Founder Professor of Engineering and David F. Linowes Fellow in the Computer Science Department at the University of Illinois at Urbana-Champaign. He is also a Professor at the Beckman Institute and holds faculty positions at the Statistics, Linguistics and ECE Departments and at the graduate School of Library and Information Science.

Roth is a Fellow of the AAAS, ACM, AAAI, and ACL, for his contributions to the foundations of machine learning and inference and for developing learning centered solutions for natural language processing (NLP) problems.

Prof. Roth has made major conceptual and theoretical advances in multiple areas in artificial intelligence (AI), from knowledge representation and reasoning, to probabilistic reasoning, to machine learning and natural language processing. His work has changed how computer scientists develop algorithms and programs for natural language understanding, and how they think about computational modeling of learning and reasoning. Roth's pioneering contributions are uniquely broad and include a theory integrating Learning and Reasoning, formally showing the benefit in jointly studying these important phenomena; developing a highly influential constrained optimization framework that augments the learning of statistical models with declarative constraints; pioneering theoretical and system work establishing the ubiquity of linear classifiers in natural language processing; developing a uniquely broad research program in Natural Language Processing that gave rise to state of the art systems to a large number of NLP tasks, with significant research and commercial impact; the first provably correct lifted probabilistic inference algorithm, and foundational contributions to understanding the complexity of probabilistic reasoning, and more.

Roth's Research is in the area of AI and intelligent access to unstructured data, where he has made seminal contributions spanning multiple aspects—from theoretical questions in Learning and Reasoning, to new models and large scale system development in Natural Language Processing (NLP)—and has contributed significantly to the wide use of machine learning techniques in research and industry. Roth has published over 300 articles in top international journals and highly selective conferences and a book on Textual Entailment. Some of his key research contributions are summarized below.

Learning and reasoning have been long recognized as fundamental phenomena of intelligence. Roth has made seminal contributions to the development of a unifying computational theory for the two phenomena. His theoretical work in the late 1990s on the Learning to Reason [11] framework, laid the foundations for integrating learning into large scale intelligent decision systems by exhibiting the advantage of jointly studying learning and reasoning. These ideas are now mainstream, and the joint study of learning and inference has had a major impact in the context of NLP.

Roth has continued to have significant influence in this area, studying Learning and Inference over **structured output**, viewing inference as a constrained optimization problem. He has contributed to developing theoretical understanding for when to jointly learn and when to decouple learning from joint inference, and the inference framework he has developed has been used broadly in NLP. In particular, a lot of the work done in **global inference in NLP** follows his **Integer Linear Programming (ILP) in NLP** [19**]** framework, incorporating declarative knowledge along with statistical models for making global decisions; along with his Constraints Driven Learning approach [3, 6], these frameworks have been extremely influential and have been used in dozens of NLP papers, including multiple best papers awards in top conferences. Roth's work on inference in NLP has facilitated work on multiple NLP tasks that require global inference including work on Semantic Role Labeling, Dependency Parsing, Co-reference resolution, Events and relations, Sentiment Analysis, Wikification and Entity Linking, Textual Entailment, temporal and quantitative reasoning, and more.

Machine Learning in NLP: Roth was a pioneer of the use of advanced machine learning methods in NLP, with multiple influential contributions. His theoretical work has contributed to developing better understanding of the relations between probabilistic models of classification and discriminative models and, in particular, explained the success of generative models in this area via the learning theoretical notion of Structural Risk

Minimization [16]. His work has established the ubiquity of linear classifiers [17] and has shown that popular models, including naïve Bayes and Hidden Markov Models, are linear models and can be studied and be understood this way. These understandings, now mainstream, have had a significant impact on NLP research and, in particular, provided insight into extensions of these methods to structured learning, now commonly used in NLP. Along with studying theories and algorithms, Roth's group has developed a number of mature tools that have been downloaded and used by thousands of researchers and in industry.

Natural Languages Understanding and Information Extraction Tools: Beyond theoretical paradigms, **p**rogress in NLP relies on the development of computational approaches to multiple fundamental natural language processing, information extraction and semantic processing problems. Roth's group has led the way in developing state-of-the-art solutions to a range semantic processing problems such as Semantic Role Labeling (SRL), Co-reference Resolution, Name Entity Recognition (NER), Entity Linking and Wikification, and many others. These tools have had hundreds of thousands of downloads, and have been used broadly by the research community and commercially. Roth's demos (http://cogcomp.cs.illinois.edu/page/demos/) are also frequently used, including in NLP and Computational Linguistics classes**.**

Structured Learning with Indirect Supervision: over the last few years it became clear that machine learning problems in natural language processing and information extraction require the ability to learn and support inference with respect to *structures*, where the task is to assign values to multiple interdependent variables. Expecting standard supervised machine learning methods to scale up to these task is unrealistic. Roth's group has developed **Indirect Supervision** methods for the structured domain that were highly influential. From the **Constrained Driven Learning** (CoDL) protocol [3, 6] that has led to work by other groups on Posterior Regularization, to work on Structured Output Learning with Indirect Supervision [4] and with latent representation [5]. One of the recent innovations in this direction has stated with his work on **Response Driven Learning** [8] that has influenced all work done today in training semantic parsers.

Structured learning is difficult partly since it involves an inference step that accounts for interaction between the set of variables one wants to assign values to. Roth's group has shown that this inference can always be formulated as an Integer Linear Program, and has recently developed exciting results providing an **amortized analysis of inference** that is based on new **sensitivity theorems** for Integer Linear Programs [12]. This new line of work is bound to have significant impact since experiments show that it can be used to cut up to 85% of the inference cost in realistic structured prediction problems in natural language processing.

Probabilistic Inference: Roth has made several seminal contributions in Probabilistic Inference. (1) He has proved a now classical result showing that exact inference in Bayesian network is **#P-complete** [15], along with results on the complexity of approximate inference. (2) Roth was the first to develop a general-purpose and exact ***Lifted***, first-order probabilistic inference algorithm [2]. This algorithm takes first-order logic description of a Markov network, and, without propositionalzing (that is, independently of the vocabulary size), it performs exact inference. This algorithm has already had significant impact on the probabilistic inference community, making Lifted Inference a very active research area. This is bound to change the way probabilistic inference is done. (3) Given the intractability of probabilistic inference with standard representations of probability distributions, Roth was among the first to study a new, Multi-Linear Representation of discrete distributions, and the first to propose a learning algorithm for these representations [18]; this has also become a very active research area in the last few years.

Computer Vision: Roth has done early work on Machine Learning in Computer Vision and was among the pioneers of the Part-based (constellation) method in object recognition [1].

Roth continues to make innovative contributions in broad AI areas. His new initiative on **Trustworthiness of Information** stands out [14]. While much work in NLP has focused on determining what a document means, Roth has started to study computational methods and inference techniques to also determine whether we can believe it. This is vital in the era of information overload and rapid publishing and has the potential for significant societal impact. He presented a tutorial on Trustworthiness of Information in AAAI'13.

Text Correction and English as a Second Language (ESL): Roth has pioneered the use of machine learning methods in text correction [9] and, more recently, in supporting second language learning; he has developed some of the best tools around for grammatical correction of text written by ESL writers [20]. Roth's group won

all software competitions in this area in the last few years, including the last two CoNLL shared tasks.

Big Data & Data Science: Roth's work in Machine Learning and Natural Language Processing led to a broad range of interdisciplinary collaborations within and outside the University of Illinois. Key in this direction is his work in the Health Domain. Roth has embarked on a new challenge of developing Natural Language Processing tools for the **Medical Domain.** It has become clear that the ability to "understand" electronic health records and the biomedical literature is at the heart of facilitating intelligent access to medical information, and of supporting better care and medical research. Dealing with medical information has its own challenges, though, and making progress requires a focused effort. Roth's program has already developed some of the best NLP tools available today for the medical domain [10] and he is collaborating with several medical programs to pursue advances in this area. Most of this work was done as part of the SHARPS project http://sharps.org/ and more recently, as part of the BD2K Center of Excellence. In addition, current efforts include collaborations that focus on natural language processing with Harvard Medical School and a local hospital in Urbana, and machine learning & data science collaboration with Mayo Clinic and UIC.

Learning based Programming is a new programming paradigm for specifying computations that require machine learning and probabilistic tools as a way to interact and reason about interactions with messy, naturally occurring, data that is highly variable and ambiguous. Over the last few years Roth has developed two new (probabilistic) programming languages [12] that aim at simplifying the development of software applications utilizing machine learning technology as a way to deal with such data. This is a crucially important direction given that most machine learning applications will be developed eventually by application programmers, but conventional (and even probabilistic) programming languages do not support the natural work flow of developing software applications that interact with naturally occurring data.

Roth has been actively developing other Data Science collaborations, including an active research program with **Psycholinguistics** [7], now in its second period of NIH funding, focusing on language acquisition, and a long term collaboration with Quantitative **Political Scientists** developing Data Analytics methods for the study of societal stability (for which he was twice awarded a Linowes Fellow) and a long term and very successful collaboration with **Education** researchers in the area of English as a Second Language.

Education and Service: Roth is a leader in the Big Data community at the University of Illinois and has led several campus level committees and initiatives in this area. He founded the **Data Science Summer Institute** (DSSI, http://mias.illinois.edu/DSSI2012), probably the first program of its kind in the nation, and directed it for six years, bringing to campus dozens of students—many from minority serving institutions—and top researchers. The institute has been supported by DHS, Yahoo! and local IT companies. DSSI was successful in providing opportunities to strong students from minority institutions to get into the top schools educational pipeline – several of our graduates got into top PhD programs and, as of this academic year, two are faculty in CS and Information Systems departments.

Beyond DSSI and outside Illinois, Roth's service contributions and educational initiatives have significantly enriched the AI, Machine Learning, and Natural Language Processing communities. Roth was among the founders of the Association for Computational Linguistics' (ACL's) **Special Interest Group on Natural Language Learning**, serving as its president and secretary, as program chair of their CoNLL conference and, currently, as a board member. In this role, he helped start and run the **CoNLL** shared task, an annual software competition that has had an enormous impact by developing methodologies, data, and benchmarks for a large number of NLP tasks, setting the standards in that area. He also chaired **ACL**, the main Computational Linguistics and NLP conference, and **AAAI**, the main AI conference, and has organized several symposia on Machine Learning and Natural Language processing, contributing to better interaction and flow of ideas between these communities.

Roth has graduated 30 Ph.D. students, 22 MS students, and over 30 undergraduate students; many of the undergrads have moved to top graduate programs, two of them were nationally recognized with a honorable mention and one as a finalist in the Computing Research Association's Outstanding Undergraduate Award.

Roth is currently the Editor-in-Chief of the Journal of Artificial Intelligence Research (JAIR), one of the top Journals in AI, after serving as an Associate Editor-in-Chief 2013-2014. He was the program chair of AAAI'11, CoNLL'02 and of ACL'03, has served as an area chair and senior program committee member for all major conferences in his research areas, and has been on the editorial board of several journals in his research areas.

Prof. Roth has given keynote talks in major conferences, including AAAI, The Conference of the American Association Artificial Intelligence; EMNLP, The Conference on Empirical Methods in Natural Language Processing, ECML & PKDD, the European Conference on Machine Learning and the Principles and Practice of Knowledge Discovery in Databases and EACL, the European Conference of the Association of Computational Linguistics (EACL) and the Chinese Conference on NLP. He has also presented tutorials in universities and conferences including at AAAI, ACL and the European ACL and has won several teaching and paper awards.

Selected References:

1. S. Agarwal, A. Awan and D. Roth, "Learning to Detect Objects in Images via a Sparse, Part-Based Representation", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20 (11), 2004.
2. R. Braz, E. Amir and D. Roth, "Lifted First-Order Probabilistic Inference", IJCAI'05, 2005.
3. M. Chang and L. Ratinov and D. Roth, "Guiding Semi-Supervision with Constraint-Driven Learning", ACL 2007.
4. M. Chang and V. Srikumar and D. Goldwasser and D. Roth, "Structured Output Learning with Indirect Supervision", ICML 2010
5. M. Chang and D. Goldwasser and D. Roth and V. Srikumar, "Discriminative Learning over Constrained Latent Representations", NAACL 2010.
6. M. Chang, L. Ratinov and D. Roth, "Structured Learning with Constrained Conditional Models", Machine Learning Journal, vol. 88 (3), 2012.
7. M. Connor, C. Fisher and D. Roth "Starting from Scratch in Semantic Role Labeling: Early Indirect Supervision", A Chapter invited to "Cognitive Aspects of Computational Language Acquisition", Afra Alishahi, Thierry Poibeau, Anna Korhonen, Editors. Springer. 2012.
8. J. Clarke, D. Goldwasser, M. Chang and D. Roth, "Driving Semantic Parsing from the World's Response", CoNLL'10, The Annual Conference on Computational Natural Language Learning, 2010.
9. A. R. Golding and D. Roth, "A Winnow-Based Approach to Spelling Correction", Machine Learning, Special issue on Machine Learning and Natural Language Processing, Vol. 34 (1/3), 1999.
10. P. Jindal and D. Roth, "Using Domain Knowledge and Domain-Inspired Discourse Model for Coreference Resolution for Clinical Narratives", JAMIA, J. of American Medical Informatics Association, Vol. 20 (2), 2013.
11. R. Khardon and D. Roth, "Learning to Reason", Journal of the Association for Computing Machinery, Vol. 44 (5), 1997.
12. P. Kordjamshidi and D. Roth and H. Wu, "Saul: Towards Declarative Learning Based Programming" IJCAI'15, 2015.
13. G. Kundu and V. Srikumar and D. Roth, "Margin-based Decomposed Amortized Inference", ACL 2013
14. J. Pasternack and D. Roth, "Latent Credibility Analysis", WWW'13, 2013.
15. D. Roth, "On the Hardness of Approximate Reasoning", Artificial Intelligence, Vol. 82, 1996.
16. D. Roth, "Learning in Natural Language", IJCAI'99, 1999.
17. D. Roth, "Learning to Resolve Natural Language Ambiguities: A Unified Approach" AAAI'98, 1998,
18. D. Roth and R. Samdani, "Learning Multi-Linear Representations", Machine Learning, Volume 76 (2), 2009.
19. D. Roth and W. Yih, "A Linear Programming Formulation for Global Inference in Natural Language Tasks", CoNLL'04: The 8th Conference on Natural Language Learning, 2004.
20. A. Rozovskaya ad D. Roth, "Building a State-of-the-Art Grammatical Error Correction System", Transactions of the Association for Computational Linguistics, Vol 2, 2014.