

# Integer Linear Programming formulations in Natural Language Processing: References

Dan Roth and Vivek Srikumar

This document is a growing bibliography of papers that use ideas behind constrained conditional models and integer linear programming inference.

- [1] E. Althaus, N. Karamanis, and A. Koller. Computing locally coherent discourses. In *ACL*, pages 399–406, Barcelona, Spain, July 2004.
- [2] R. Barzilay and M. Lapata. Aggregation via set partitioning for natural language generation. In *Proceedings of the Human Language Technology Conference of the NAACL, Main Conference*, pages 359–366, New York City, USA, June 2006. Association for Computational Linguistics.
- [3] J. Berant, I. Dagan, and J. Goldberger. Global learning of focused entailment graphs. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 1220–1229, Uppsala, Sweden, July 2010. Association for Computational Linguistics.
- [4] J. Berant, I. Dagan, and J. Goldberger. Global learning of typed entailment rules. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 610–619, Portland, Oregon, USA, June 2011. Association for Computational Linguistics.
- [5] J. Berant, V. Srikumar, P.-C. Chen, A. V. Linden, B. Harding, B. H. P. Clark, and C. D. Manning. Modeling biological processes for reading comprehension. In *Proceedings of EMNLP*, 2014.
- [6] T. Berg-Kirkpatrick, D. Gillick, and D. Klein. Jointly learning to extract and compress. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 481–490, Portland, Oregon, USA, June 2011. Association for Computational Linguistics.
- [7] P. Bramsen, P. Deshpande, Y. K. Lee, and R. Barzilay. Inducing temporal graphs. In *EMNLP*, pages 189–198, Sydney, Australia, July 2006. Association for Computational Linguistics.
- [8] N. Chambers and D. Jurafsky. Jointly combining implicit constraints improves temporal ordering. In *Proc. of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2008.
- [9] K.-W. Chang, R. Samdani, A. Rozovskaya, N. Rizzolo, M. Sammons, and D. Roth. Inference protocols for coreference resolution. In *CoNLL Shared Task*, pages 40–44, Portland, Oregon, USA, 2011. Association for Computational Linguistics.
- [10] K.-W. Chang, S. Upadhyay, M.-W. Chang, V. Srikumar, and D. Roth. IllinoisSL: A Java library for structured prediction. *arXiv preprint arXiv:1509.07179*, 2015.
- [11] M. Chang, D. Goldwasser, D. Roth, and V. Srikumar. Discriminative learning over constrained latent representations. In *NAACL*, 6 2010.
- [12] M. Chang, L. Ratnikov, N. Rizzolo, and D. Roth. Learning and inference with constraints. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, July 2008.

- [13] M. Chang, L. Ratinov, and D. Roth. Guiding semi-supervision with constraint-driven learning. In *Proc. of the Annual Meeting of the ACL*, pages 280–287, Prague, Czech Republic, Jun 2007. Association for Computational Linguistics.
- [14] M. Chang, L. Ratinov, and D. Roth. Constraints as prior knowledge. In *ICML Workshop on Prior Knowledge for Text and Language Processing*, pages 32–39, July 2008.
- [15] M. Chang, V. Srikumar, D. Goldwasser, and D. Roth. Structured output learning with indirect supervision. In *ICML*, 2010.
- [16] M.-W. Chang, D. Goldwasser, D. Roth, and Y. Tu. Unsupervised constraint driven learning for transliteration discovery. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 299–307, Boulder, Colorado, June 2009. Association for Computational Linguistics.
- [17] M.-W. Chang, L. Ratinov, and D. Roth. Structured learning with constrained conditional models. *Machine learning*, 88(3):399–431, 2012.
- [18] Y. Chang and M. Collins. Exact decoding of phrase-based translation models through lagrangian relaxation. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 26–37, Edinburgh, Scotland, UK., July 2011. Association for Computational Linguistics.
- [19] W. Che, Z. Li, Y. Hu, Y. Li, B. Qin, T. Liu, and S. Li. A cascaded syntactic and semantic dependency parsing system. In *CoNLL*, pages 238–242, Manchester, England, August 2008. Coling 2008 Organizing Committee.
- [20] Y. Choi, E. Breck, and C. Cardie. Joint extraction of entities and relations for opinion recognition. In *Proc. of the 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2006.
- [21] J. Clarke, D. Goldwasser, M. Chang, and D. Roth. Driving semantic parsing from the world’s response. In *CoNLL*, 7 2010.
- [22] J. Clarke and M. Lapata. Constraint-based sentence compression: An integer programming approach. In *Proc. of the COLING/ACL 2006 Main Conference Poster Sessions (ACL)*, 2006.
- [23] J. Clarke and M. Lapata. Modelling compression with discourse constraints. In *Proc. of the Conference on Empirical Methods in Natural Language Processing and on Computational Natural Language Learning (EMNLP-CoNLL)*, 2007.
- [24] J. Clarke and M. Lapata. Global inference for sentence compression: An integer linear programming approach. *Journal of Artificial Intelligence Research (JAIR)*, 31:399–429, 2008.
- [25] H. Daumé III. Cross-task knowledge-constrained self training. In *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*, pages 680–688, Honolulu, Hawaii, October 2008. Association for Computational Linguistics.
- [26] G. de Melo and G. Weikum. Untangling the cross-lingual link structure of wikipedia. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 844–853, Uppsala, Sweden, July 2010. Association for Computational Linguistics.
- [27] J. DeNero and D. Klein. The complexity of phrase alignment problems. In *Proceedings of ACL-08: HLT, Short Papers*, pages 25–28, Columbus, Ohio, June 2008. Association for Computational Linguistics.
- [28] P. Denis and J. Baldridge. Joint determination of anaphoricity and coreference resolution using integer programming. In *Proc. of the Annual Meeting of the North American Chapter of the Association for Computational Linguistics - Human Language Technology Conference (NAACL-HLT)*, 2007.
- [29] P. Denis and P. Muller. Predicting globally-coherent temporal structures from texts via endpoint inference and graph decomposition. In *Proc. of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2011.

- [30] P. Deshpande, R. Barzilay, and D. Karger. Randomized decoding for selection-and-ordering problems. In *Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Proceedings of the Main Conference*, pages 444–451, Rochester, New York, April 2007. Association for Computational Linguistics.
- [31] Q. Do, Y. Chan, and D. Roth. Minimally supervised event causality identification. In *EMNLP*, Edinburgh, Scotland, 7 2011.
- [32] Q. Do, W. Lu, and D. Roth. Joint inference for event timeline construction. In *EMNLP*, 2012.
- [33] K. Filippova and M. Strube. Dependency tree based sentence compression. In *INLG*, 2008.
- [34] K. Filippova and M. Strube. Sentence fusion via dependency graph compression. In *EMNLP*, pages 177–185, Honolulu, Hawaii, October 2008. Association for Computational Linguistics.
- [35] J. R. Finkel and C. D. Manning. The importance of syntactic parsing and inference in semantic role-labeling. In *Proc. of the Annual Meeting of the Association for Computational Linguistics - Human Language Technology Conference, Short Papers (ACL-HLT)*, 2008.
- [36] K. Ganchev, J. Grača, J. Gillenwater, and B. Taskar. Posterior Regularization for Structured Latent Variable Models. *Journal of Machine Learning Research*, 2010.
- [37] U. Germann, M. Jahr, K. Knight, D. Marcu, and K. Yamada. Fast decoding and optimal decoding for machine translation. In *Proceedings of 39th Annual Meeting of the Association for Computational Linguistics*, pages 228–235, Toulouse, France, July 2001. Association for Computational Linguistics.
- [38] D. Goldwasser, R. Reichart, J. Clarke, and D. Roth. Confidence driven unsupervised semantic parsing. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 1486–1495, Portland, Oregon, USA, June 2011. Association for Computational Linguistics.
- [39] J. V. Graca, K. Ganchev, and B. Taskar. Expectation maximization and posterior constraints. In *NIPS*, volume 20, 2007.
- [40] J. Hoffart, M. A. Yosef, I. Bordino, H. Fürstenau, M. Pinkal, M. Spaniol, B. Taneva, S. Thater, and G. Weikum. Robust disambiguation of named entities in text. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 782–792, Edinburgh, Scotland, UK., July 2011. Association for Computational Linguistics.
- [41] R. Iida and M. Poesio. A cross-lingual ilp solution to zero anaphora resolution. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 804–813, Portland, Oregon, USA, June 2011. Association for Computational Linguistics.
- [42] M. Klenner. Grammatical role labeling with integer linear programming. In *EACL*, 2006.
- [43] M. Klenner. Enforcing consistency on coreference sets. In *RANLP*, 2007.
- [44] M. Klenner. Shallow dependency labeling. In *ACL*, pages 201–204, Prague, Czech Republic, June 2007. Association for Computational Linguistics.
- [45] P. Koomen, V. Punyakanok, D. Roth, and W. Yih. Generalized inference with multiple semantic role labeling systems (shared task paper). In I. Dagan and D. Gildea, editors, *Proc. of the Annual Conference on Computational Natural Language Learning (CoNLL)*, pages 181–184, 2005.
- [46] P. Li, Y. Wang, W. Gao, and J. Jiang. Generating aspect-oriented multi-document summarization with event-aspect model. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 1137–1146, Edinburgh, Scotland, UK., July 2011. Association for Computational Linguistics.

- [47] T. Li and V. Srikumar. Exploiting sentence similarities for better alignments. In *Proceedings of EMNLP*, 2016.
- [48] H. Lin and J. Bilmes. Multi-document summarization via budgeted maximization of submodular functions. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 912–920, Los Angeles, California, June 2010. Association for Computational Linguistics.
- [49] A. Martins, N. Smith, M. Figueiredo, and P. Aguiar. Dual decomposition with many overlapping components. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 238–249, Edinburgh, Scotland, UK., July 2011. Association for Computational Linguistics.
- [50] A. F. T. Martins. *The geometry of constrained structured prediction: applications to inference and learning of natural language syntax*. PhD thesis, Carnegie Mellon University and Instituto Superior Técnico, 2012.
- [51] A. F. Martins, M. A. Figueiredo, P. M. Aguiar, N. A. Smith, and E. P. Xing. Ad3: alternating directions dual decomposition for map inference in graphical models. *Journal of Machine Learning Research*, 16:495–545, 2015.
- [52] A. F. Martins, N. A. Smith, and E. P. Xing. Concise integer linear programming formulations for dependency parsing. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1-Volume 1*, pages 342–350. Association for Computational Linguistics, 2009.
- [53] E. Mayfield and C. Penstein Rosé. Recognizing authority in dialogue with an integer linear programming constrained model. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 1018–1026, Portland, Oregon, USA, June 2011. Association for Computational Linguistics.
- [54] R. McDonald. A study of global inference algorithms in multi-document summarization. In *ECIR*, 2007.
- [55] I. Meza-Ruiz and S. Riedel. Jointly identifying predicates, arguments and senses using markov logic. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 155–163, Boulder, Colorado, June 2009. Association for Computational Linguistics.
- [56] G. Murray, G. Carenini, and R. Ng. Interpretation and transformation for abstracting conversations. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 894–902, Los Angeles, California, June 2010. Association for Computational Linguistics.
- [57] V. Punyakanok, D. Roth, W. Yih, and D. Zimak. Semantic role labeling via integer linear programming inference. In *Proc. the International Conference on Computational Linguistics (COLING)*, pages 1346–1352, Geneva, Switzerland, August 2004.
- [58] V. Punyakanok, D. Roth, W. Yih, and D. Zimak. Learning and inference over constrained output. In *Proc. of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2005.
- [59] V. Punyakanok, D. Roth, W. Yih, D. Zimak, and Y. Tu. Semantic role labeling via generalized inference over classifiers (shared task paper). In H. T. Ng and E. Riloff, editors, *Proc. of the Annual Conference on Computational Natural Language Learning (CoNLL)*, pages 130–133, 2004.
- [60] S. Riedel and J. Clarke. Incremental integer linear programming for non-projective dependency parsing. In *EMNLP*, pages 129–137, Sydney, Australia, July 2006. Association for Computational Linguistics.
- [61] S. Riedel and A. McCallum. Fast and robust joint models for biomedical event extraction. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 1–12, Edinburgh, Scotland, UK., July 2011. Association for Computational Linguistics.

- [62] N. Rizzolo and D. Roth. Modeling Discriminative Global Inference. In *Proc. of the First International Conference on Semantic Computing (ICSC)*, pages 597–604, Irvine, California, September 2007. IEEE.
- [63] N. Rizzolo and D. Roth. Learning based java for rapid development of nlp systems. In *LREC*, Valletta, Malta, 5 2010.
- [64] D. Roth. Learning based programming. 2005.
- [65] D. Roth and W. Yih. Integer linear programming inference for conditional random fields. In *Proc. of the International Conference on Machine Learning (ICML)*, pages 737–744, 2005.
- [66] D. Roth and W. Yih. Global inference for entity and relation identification via a linear programming formulation. In L. Getoor and B. Taskar, editors, *Introduction to Statistical Relational Learning*. MIT Press, 2007.
- [67] A. Rush, D. Sontag, M. Collins, and T. Jaakkola. On dual decomposition and linear programming relaxations for natural language processing. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 1–11. Association for Computational Linguistics, 2010.
- [68] K. Sagae, Y. Miyao, and J. Tsujii. Hpsg parsing with shallow dependency constraints. In *ACL*, pages 624–631, Prague, Czech Republic, June 2007. Association for Computational Linguistics.
- [69] R. Samdhani, M. Chang, and D. Roth. Unified expectation maximization. In *NAACL*, 6 2012.
- [70] W. Seeker, I. Rehbein, J. Kuhn, and J. Van Genabith. Hard constraints for grammatical function labelling. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 1087–1097, Uppsala, Sweden, July 2010. Association for Computational Linguistics.
- [71] V. Srikumar, G. Kundu, and D. Roth. On amortizing inference cost for structured prediction. In *EMNLP*, 2012.
- [72] V. Srikumar and D. Roth. A joint model for extended semantic role labeling. In *EMNLP*, Edinburgh, Scotland, 2011.
- [73] M. Tomasoni and M. Huang. Metadata-aware measures for answer summarization in community question answering. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 760–769, Uppsala, Sweden, July 2010. Association for Computational Linguistics.
- [74] T. Tsai, C. Wu, Y. Lin, and W. Hsu. Exploiting full parsing information to label semantic roles using an ensemble of ME and SVM via integer linear programming. In *CoNLL*, pages 233–236, Ann Arbor, Michigan, June 2005. Association for Computational Linguistics.
- [75] G. Wisniewski, A. Allauzen, and F. Yvon. Assessing phrase-based translation models with oracle decoding. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 933–943, Cambridge, MA, October 2010. Association for Computational Linguistics.
- [76] K. Woodsend, Y. Feng, and M. Lapata. Title generation with quasi-synchronous grammar. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 513–523, Cambridge, MA, October 2010. Association for Computational Linguistics.
- [77] K. Woodsend and M. Lapata. Learning to simplify sentences with quasi-synchronous grammar and integer programming. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 409–420, Edinburgh, Scotland, UK., July 2011. Association for Computational Linguistics.
- [78] Y. Wu, Q. Zhang, X. Huang, and L. Wu. Structural opinion mining for graph-based sentiment representation. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 1332–1341, Edinburgh, Scotland, UK., July 2011. Association for Computational Linguistics.