References

[Ait02] J.S. Aitken. Learning Information Extraction Rules: An Inductive Logic Programming approach. In *Proceedings of the 15th European Conference on Artificial Intelligence*, pp. 355–359. France, 2002. URL `citeseer.ist.psu.edu/aitken02learning.html`. Http://citeseer.ist.psu.edu/586553.html.

[Alp04] Ethem Alpaydin. *Introduction to Machine Learning*. The MIT Press, Cambridge, Massachusetts, 1st edn., 2004.

[App93] Douglas Appelt, J. Hobbs, J. Bear, D. J. Israel, and M. Tyson. FASTUS: A finite-state processor for information extraction from real-world text. In *Proceedings of International Joint Conference on Atrificial Intelligence (IJCAI-1993)*. Washington DC, 1993.

[Ara03] Arvind Arasu and Hector Garcia-Molina. Extracting structured data from Web pages. In *SIGMOD '03*, pp. 337–348. 2003. URL `http://dx.doi.org/10.1145/872757.872799`.

[Bag97] Amit Bagga and Joyce Yue Chai. A trainable message understanding system. In *CoNLL*, pp. 1–8. 1997. URL `http://citeseer.ist.psu.edu/amit97trainable.html`.

[Ben03] V. Benzaken, G. Castagna, and A. Frisch. CDuce: an XML-centric general-purpose language. In *In ICFP '03, 8th ACM International Conference on Functional Programming*, pp. 51–63. 2003.

[Ben05] V. Benzaken, G. Castagna, and C. Miachon. A full pattern-based paradigm for XML query processing. In *Proceedings of the 7th Int. Symposium on Practical Aspects of Decl. Languages*, 3350. 2005.

[Ber05] A. Berglund, S. Boag, D. Chamberlin, M. Fernández, M. Kay, J. Robie, and J. Siméon. XML Path Language (XPath) 2.0, September 2005. `http://www.w3.org/TR/2005/WD-xpath20-20050915/`.

[Bil05] Philip Bille. A survey on tree edit distance and related problems. *Theor. Comput. Sci.*, 337(1-3):217–239, 2005. ISSN 0304-3975.

[Bir05] S. Bird, Y. Chen, S. B. Davidson, H. Lee, and Y. Zheng. Extending xpath to support linguistic queries. In *Proceedings of Workshop on Programming Languages for XML PLAN-X 2005*. 2005.

[Blo04] Vincent D. Blondel, Anahí Gajardo, Maureen Heymans, Pierre Senellart, and Paul Van Dooren. A measure of similarity between graph vertices. With applications to synonym extraction and web searching. In *SIAM Review*, pp. 647–666. 2004.

[Boa03] S. Boag, D. Chamberlin, M. Fernandez, D. Florescu, J. Robie, J. SimÈon, and M. Stefanescu. XQuery 1.0: An XML Query Language, 2003. `http://www.w3c.org/TR/xquery`.

[BY99] Ricardo A. Baeza-Yates and Berthier A. Ribeiro-Neto. *Modern Information Retrieval*. ACM Press / Addison-Wesley, New York, 1999. ISBN 0-201-39829-X. URL `citeseer.ist.psu.edu/baeza-yates99modern.html`.

[Cal98] Mary E. Califf and Raymond J. Mooney. Relational learning of pattern-match rules for information extraction. In *Working Notes of AAAI Spring Symposium on Applying Machine Learning to Discourse Processing*, pp. 6–11. Menlo Park, CA, 1998. URL `http://citeseer.ist.psu.edu/39151.html`.

[Cal03] Mary E. Califf and Raymond J. Mooney. Bottom-up relational learning of pattern matching rules for information extraction. *Journal of Machine Learning Research*, 4:177–210, 2003. URL `http://www.jmlr.org/papers/volume4/califf03a/califf03a.pdf`.

[Car93] Claire Cardie. A case-based approach to knowledge acquisition for domain-specific sentence analysis. In *Proceedings of the Eleventh National Conference on Artificial Intelligence*, pp. 798–803. AAAI Press, 1993. URL `http://citeseer.ist.psu.edu/cardie93casebased.html`.

[Car03] David Carmel, Yoelle S. Maarek, Matan Mandelbrod, Yosi Mass, and Aya Soffer. Searching XML documents via XML fragments. In *SIGIR '03: Proceedings of the 26th annual int. ACM SIGIR conference*, pp. 151–158. ACM Press, New York,USA, 2003. ISBN 1-58113-646-3.

[Cas05] G. Castagna. Patterns and types for querying XML. In *DBPL - XSYM 2005 joint keynote talk*. 2005.

[Cha01] Don Chamberlin, Jonathan Robie, and Daniela Florescu. Quilt: An XML query language for heterogeneous data sources. *LNCS*, 1997:1–11, 2001. URL `citeseer.ist.psu.edu/chamberlin00quilt.html`.

[Che92] Hsinchun Chen and Kevin J. Lynch. Automatic construction of networks of concepts characterizing document databases. In *IEEE Transactions on Systems, Man and Cybernetics*, vol. 22(5), pp. 885–902. 1992.

[Che01] Weimin Chen. New algorithm for ordered tree-to-tree correction problem. *J. Algorithms*, 40(2):135–158, 2001. ISSN 0196-6774.

[Chi02] Hai Leong Chieu and Hwee Tou Ng. A maximum entropy approach to information extraction from semi-structured and free text. In *AAAI 2002*. 2002. URL `http://citeseer.ist.psu.edu/chieu02maximum.html`.

[Cir01a] Fabio Ciravegna. Adaptive information extraction from text by rule induction and generalisation. In *Proceedings of the 17th International Joint Conference on Artificial Intelligence (IJCAI)*. Seattle, USA, 2001.

[Cir01b] Fabio Ciravegna. (LP)^2, an adaptive algorithm for information extraction from Web-related texts. In *IJCAI-2001 Workshop on Adaptive Text Extraction and Mining*. Seattle, USA, 2001. URL `http://www.smi.ucd.ie/ATEM2001/proceedings/ciravegna-atem2001.pdf`.

[Cir02] Fabio Ciravegna and Alberto Lavelli. LearningPinocchio: Adaptive information extraction for real world applications. In *Proceedings of the 2nd Workshop on Robust Methods in Analysis of Natural Language Data (ROMAND 2002)*. Frascati, Italy, 2002. URL `http://www.dcs.shef.ac.uk/~fabio/paperi/romand2002.zip`.

[Cla99] J. Clark and S. DeRose. XML Path Language (XPath), 1999. `http://www.w3.org/TR/Xpath`.

[Coh95] William W. Cohen. Fast effective rule induction. In Armand Prieditis and Stuart Russell, eds., *Proc. of the 12th International Conference on Machine Learning*, pp. 115–123. Morgan Kaufmann, Tahoe City, CA, Jul. 9–12, 1995. ISBN 1-55860-377-8. URL `citeseer.ist.psu.edu/cohen95fast.html`.

[Coh96] D. A. Cohn, Z. Ghahramani, and M. I. Jordan. Active learning with statistical models. *JOURNAL OF ARTIFICIAL INTELLIGENCE RESEARCH*, 4:129, 1996. URL `http://www.citebase.org/abstract?id=oai:arXiv.org:cs/9603104`.

[Col96] R. Collier. *Automatic template creation for information extraction, an overview*. Tech. rep., University of Sheffield, 1996. URL `http://citeseer.ist.psu.edu/collier96automatic.html`.

[Cor90] Thomas H. Cormen, Charles E. Leiserson, and Ronald L. Rivest. *Introduction to algorithms*. MIT Press and McGraw-Hill, 1990.

[Cre01] Valter Crescenzi, Giansalvatore Mecca, and Paolo Merialdo. RoadRunner: Towards automatic data extraction from large Web sites. In *Proceedings of 27th International Conference on Very Large Data Bases*, pp. 109–118. 2001. URL `http://citeseer.ist.psu.edu/crescenzi01roadrunner.html`.

[Cre03] Valter Crescenzi and Giansalvatore Mecca. *On Automatic Information Extraction from Large Web Sites*. Tech. Rep. DIA-76-2003, University "Roma Tre", 2003. URL `http://www.dia.uniroma3.it/db/roadRunner/publications/dia-76-2003.ps.gz`.

[Cun01] H. Cunningham, D. Maynard, V. Tablan, C. Ursu, and K. Bontcheva. Developing language processing components with gate, 2001. URL `citeseer.ist.psu.edu/cunningham01developing.html`.

[Cur02] James R. Curran and Marc Moens. Improvements in automatic thesaurus extraction. In *Proceedings of the Workshop of the ACL Special Interest Group on the Lexicon (SIGLEX)*, pp. 59–66. Association for Computational Linguistics, 2002.

[Dam64] Fred J. Damerau. A technique for computer detection and correction of spelling errors. *Commun. ACM*, 7(3):171–176, 1964. ISSN 0001-0782.

[Del94] Sylvain Delisle, Ken Barker, Jean-Francois Delannoy, Stan Matwin, and Stan Szpakowicz. From text to Horn clauses: Combining linguistic analysis and machine learning. In *10th Canadian AI Conf.* 1994. URL `http://citeseer.ist.psu.edu/delisle94from.html`.

[Deu99] Alin Deutsch, Mary F. Fernandez, D. Florescu, A.Y. Levy, and D. Suciu. A query language for XML. *Computer Networks*, 31(11-16):1155–1169, 1999.

[Emb98] David W. Embley, Douglas M. Campbell, Randy D. Smith, and Stephen W. Liddl. Ontology-based extraction and structuring of information from data-rich unstructured documents. In *Conference on Information and Knowledge Management (CIKM)*, pp. 52–59. 1998. URL `http://citeseer.ist.psu.edu/embley98ontologybased.html`.

[Fer04] David Ferrucci and Adam Lally. Uima: an architectural approach to unstructured information processing in the corporate research environment. *Nat. Lang. Eng.*, 10(3-4):327–348, 2004. ISSN 1351-3249.

[Fin04a] Aidan Finn and Nicholas Kushmerick. Information extraction by convergent boundary classification. In *AAAI-2004 Workshop on Adaptive Text Extraction and Mining*. San Jose, USA, 2004. URL `http://www.cs.ucd.ie/staff/nick/home/research/download/finn-aaai04-atem.pdf`.

[Fin04b] Aidan Finn and Nicholas Kushmerick. Multi-level boundary classification for information extraction. In *ECML 2004*, pp. 111–122. 2004.

[Fre98] Dayne Freitag. Toward general-purpose learning for information extraction. In Christian Boitet and Pete Whitelock, eds., *Proc. 36th Annual Meeting of the Association for Computational Linguistics*, pp. 404–408. San Francisco, CA, 1998. URL `http://citeseer.ist.psu.edu/freitag98toward.html`.

[Fre99] Dayne Freitag and Andrew K. McCallum. Information extraction with HMMs and shrinkage. In *Proceedings of the AAAI-99 Workshop on Machine Learning for Information Extraction*. 1999. URL `http://citeseer.ist.psu.edu/freitag99information.html`.

[Fre00a] Dayne Freitag and Nicholas Kushmerick. Boosted wrapper induction. In *AAAI/IAAI*, pp. 577–583. 2000. URL `http://citeseer.ist.psu.edu/freitag00boosted.html`.

[Fre00b] Dayne Freitag and Andrew K. McCallum. Information extraction with HMM structures learned by stochastic optimization. In *AAAI/IAAI*, pp. 584–589. 2000. URL `http://citeseer.ist.psu.edu/freitag00information.html`.

[Für99] Johannes Fürnkranz. Separate-and-conquer rule learning. *Artificial Intelligence Review*, 13(1):3–54, 1999. URL `http://citeseer.ist.psu.edu/26490.html`.

[Gus97]   Dan Gusfield. *Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology.* Cambridge University Press, January 1997. ISBN 0521585198.

[Han02]   Siegfried Handschuh, Steffen Staab, and Fabio Ciravegna. S-CREAM—semi-automatic creation of metadata. In Asuncion Gomez-Perez and V. Richard Benjamins, eds., *Proc. 13th International Conference on Knowledge Engineering and Management.* 2002.

[Hea98]   M. A. Hearst. Automated discovery of Wordnet relations. In C. Fellbaum, ed., *Wordnet An Electronic Lexical Database*, pp. 131–151. MIT Press, Cambridge, MA, 1998.

[Hos01]   H. Hosoya and P.C. Pierce. Regular expression patern matching for XML. In *In POPL '01, 25th Symposium on Principles of Prog. Languages.* 2001.

[Hos03]   Haruo Hosoya and Benjamin C. Pierce. XDuce: A statically typed XML processing language. *ACM Trans. Inter. Tech.*, 3(2):117–148, 2003. ISSN 1533-5399.

[Joa98]   T. Joachims. Making large-scale support vector machine learning practical. In A. Smola B. Schölkopf, C. Burges, ed., *Advances in Kernel Methods: Support Vector Machines.* MIT Press, Cambridge, MA, 1998. URL `citeseer.ist.psu.edu/joachims98making.html`.

[Kle98]   Philip N. Klein. Computing the edit-distance between unrooted ordered trees. In *ESA '98: Proceedings of the 6th Annual European Symposium on Algorithms*, pp. 91–102. Springer-Verlag, London, UK, 1998. ISBN 3-540-64848-8.

[Laf01a] John Lafferty, Andrew K. McCallum, and Fernando Pereira. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *ICML.* 2001. URL `http://citeseer.ist.psu.edu/lafferty01conditional.html`.

[Laf01b] Mathieu Lafourcade. Lexical sorting and lexical transfer by conceptual vectors. In *Proceedings of the First International Workshop on MultiMedia Annotation.* Tokyo, Japan, 2001.

[Laf01c] Mathieu Lafourcade and Violaine Prince. Relative synonymy and conceptual vectors. In *Proceedings of the NLPRS.* Tokyo, Japan, 2001.

[Lap02]   C. Laprun, J. Fiscus, J. Garofolo, and S. Pajot. A practical introduction to AT-LAS. In *Proceedings of the Third International Conference on Language Resources and Evaluation(LREC).* 2002. URL `http://www.nist.gov/speech/atlas/download/lrec2002-atlas.pdf`.

[Lev66]   Vladimir I. Levenshtein. *Binary codes capable of correcting deletions, insertions, and reversals.* Tech. Rep. 8, 1966.

[Lev06]   Roger Levy and Galen Andrew. Tregex and tsurgeon: tools for querying and manipulating tree data structures. In *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006).* 2006.

[Lin98]   Dekang Lin. An information-theoretic definition of similarity. In *Proceedings of the Fifteenth International Conference on Machine Learning*, pp. 296–304. Madison, WI, 1998.

[Lit88]   Nick Littlestone. Learning quickly when irrelevant attributes abound: A new linear-threshold algorithm. *Machine Learning*, 2:285–318, 1988.

[Man99]   Christopher D. Manning and Hinrich Schütze. *Foundations of Statistical Natural Language Processing.* MIT Press, Cambridge, MA 2000, 1999.

[Mar93]   Mitchell P. Marcus, Beatrice Santorini, and Mary Ann Marcinkiewicz. Building a large annotated corpus of english: The penn treebank. *Computational Linguistics*, 19(2):313–330, 1993.

[McC00]   Andrew K. McCallum, Dayne Freitag, and Fernando Pereira. Maximum entropy Markov models for information extraction and segmentation. In *ICML.* 2000. URL `http://www-2.cs.cmu.edu/afs/cs/user/dayne/www/ps/memm.ps`.

[McC03a] Andrew McCallum and Ben Wellner. Object consolidation by graph partitioning with a conditionally-trained distance metric. In *KDD Workshop on Data Cleaning, Record Linkage, and Object Consolidation.* 2003. URL `http://csaa.byu.edu/kdd03-papers/mccallum-wellner.ps`.

[McC03b] Andrew K. McCallum and David Jensen. A note on the unification of information extraction and data mining using conditional-probability, relational models. In *IJCAI'03 Workshop on Learning Statistical Models from Relational Data*. 2003. URL `http://www.cs.umass.edu/~mccallum/papers/iedatamining-ijcaiws03.pdf`.

[Mus01] Ion Muslea, Steven Minton, and Craig A. Knoblock. Hierarchical wrapper induction for semistructured information sources. *Autonomous Agents and Multi-Agent Systems*, 4(1/2):93–114, 2001. URL `http://citeseer.ist.psu.edu/muslea01hierarchical.html`.

[Nah00] Un Yong Nahm and Raymond J. Mooney. Using information extraction to aid the discovery of prediction rules from text. In *Proceedings of the Sixth International Conference on Knowledge Discovery and Data Mining (KDD-2000) Workshop on Text Mining*. Boston, MA, 2000. URL `http://citeseer.ist.psu.edu/nahm00using.html`.

[Neu99] G. Neumann and G. Mazzini. *Domain-adaptive Information Extraction*. Tech. rep., Deutsches Forschungszentrum für Künstliche Intelligenz, 1999.

[Nob99] Chikashi Nobata and Satoshi Sekine. Towards automatic acquisition of patterns for information extraction. In *International Conference of Computer Processing of Oriental Languages*. 1999. URL `http://citeseer.ist.psu.edu/nobata99towards.html`.

[Ope05] OpenThesaurus. OpenThesaurus - Deutscher Thesaurus, 2005. `http://www.openthesaurus.de`.

[Pes03] Leonid Peshkin and Avi Pfeffer. Bayesian information extraction network. In *IJCAI*. 2003. URL `http://citeseer.ist.psu.edu/565989.html`.

[Phi02] William Phillips and Ellen Riloff. Exploiting strong syntactic heuristics and co-training to learn semantic lexicons. In *Proceedings of the 2002 Conference on Empirical Methods in NLP*. 2002.

[Qui95] J. Ross Quinlan and R. Mike Cameron-Jones. Induction of logic programs: FOIL and related systems. *New Generation Computing*, 13(3,4):287–312, 1995. URL `http://citeseer.ist.psu.edu/quinlan95induction.html`.

[Req03] XQuery and XPath Full-Text Requirements, 2003. `http://www.w3.org/TR/2003/WD-xquery-full-text-requirements-20030502/`.

[Ril98] Ellen Riloff and Mark Schmelzenbach. An empirical approach to conceptual case frame acquisition. In *Proceedings of the Sixth Workshop on Very Large Corpora*. 1998. URL `http://citeseer.ist.psu.edu/riloff98empirical.html`.

[Rob81] D. F. Robinson and L. R. Foulds. Comparison of phylogenetic trees. *Mathematical Biosciences*, 53(1-2):131–147, February 1981. URL `http://dx.doi.org/10.1016/0025-5564(81)90043-2`.

[Rob98] Jonathan Robie. The design of XQL. 1998. `http://www.ibiblio.org/xql/xql-design.html`.

[Rot01] Dan Roth and Wen-tau Yih. Relational learning via propositional algorithms: An information extraction case study. In *IJCAI*. 2001. URL `http://citeseer.ist.psu.edu/roth01relational.html`.

[Rug97] Gerda Ruge. Automatic detection of thesaurus relations for information retrieval applications. *Foundations of Computer Science: Potential - Theory - Cognition*, LNCS 1337:499–506, 1997.

[Sch01] Tobias Scheffer, Christian Decomain, and Stefan Wrobel. Active hidden Markov models for information extraction. In *Proceedings of the International Symposium on Intelligent Data Analysis*. 2001. URL `http://citeseer.ist.psu.edu/sche01active.html`.

[Sch02] Tobias Scheffer, Stefan Wrobel, Borislav Popov, Damyan Ognianov, Christian Decomain, and Susanne Hoche. Learning hidden Markov models for information extraction actively from partially labeled text. *Künstliche Intelligenz*, (2), 2002. URL `http://kd.cs.uni-magdeburg.de/~scheffer/papers/kisi.ps.gz`.

[Sen03] Pierre P. Senellart and Vincent D. Blondel. Automatic discovery of similar words. In Michael Berry, ed., *Survey of Text Mining. Clustering, classification, and retrieval*, pp. 25–44. Springer Verlag, Berlin, 2003.

[Sie04] C. Siefkes. A shallow algorithm for correcting nesting errors and other well-formedness violations in XML-like input. In *Extreme Markup Languages*. 2004. URL `http://www.siefkes.net/papers/eml/EML2004.pdf`.

[Sie05] Christian Siefkes. Incremental information extraction using tree-based context representations. In *CICLing 2005*, LNCS 3406. Springer, 2005. URL `http://www.siefkes.net/papers/incremental-ie.pdf`.

[Sin05] Peter Siniakov and Christian Siefkes. An overview and classification of adaptive approaches to information extraction. *Journal on Data Semantics*, IV:172–212, 2005. URL `http://siniakov.de.vu/articles/overview-ie.pdf`. LNCS 3730.

[Sin06a] Peter Siniakov. Querying XML documents with multi-dimensional markup. In *Proceedings of the 5th Workshop on NLP and XML (NLPXML-2006)*, pp. 43–50. Trento, Italy, 2006. URL `http://siniakov.de.vu/articles/PU-NLPXML.pdf`.

[Sin06b] Peter Siniakov. Recognition of synonyms by a lexical graph. In *Proceedings of the Workshop on Adaptive Text Extraction and Mining (ATEM 2006)*, pp. 33–40. Trento, Italy, 2006. URL `http://siniakov.de.vu/articles/lex_graph_ATEM.pdf`.

[Sko03] Marios Skounakis, Mark Craven, and Soumya Ray. Hierarchical hidden Markov models for information extraction. In *IJCAI*. 2003. URL `http://www.biostat.wisc.edu/~craven/papers/ijcai03.pdf`.

[Sod95] Stephen Soderland, David Fisher, Jonathan Aseltine, and Wendy Lehnert. CRYSTAL: Inducing a conceptual dictionary. In Chris Mellish, ed., *Proc. 14th International Joint Conference on Artificial Intelligence*, pp. 1314–1319. San Francisco, 1995. URL `http://citeseer.ist.psu.edu/soderland95crystal.html`.

[Sod97] Stephen Soderland. *Learning Text Analysis Rules for Domain-specific Natural Language Processing*. Ph.D. thesis, University of Massachusetts, Amherst, 1997. URL `http://citeseer.ist.psu.edu/279256.html`.

[Sod99] Stephen Soderland. Learning information extraction rules for semi-structured and free text. *Machine Learning*, 34(1–3):233–272, 1999. URL `http://citeseer.ist.psu.edu/soderland99learning.html`.

[Stu95] Stuttgart-Tübingen Tagset (STTS), 1995. URL `http://www.sfs.uni-tuebingen.de/Elwis`.

[Sud01] Kiyoshi Sudo, Satoshi Sekine, and Ralph Grishman. Automatic pattern acquisition for Japanese information extraction. In *HLT2001*. 2001. URL `http://citeseer.ist.psu.edu/sudo01automatic.html`.

[Sut05] Charles Sutton and Andrew McCallum. Composition of conditional random fields for transfer learning. In *Proceedings of HLT/EMNLP 2005*. 2005.

[Tai79] Kuo-Chung Tai. The tree-to-tree correction problem. *J. ACM*, 26(3):422–433, 1979. ISSN 0004-5411.

[Tho99] Cynthia A. Thompson, Mary Elaine Califf, and Raymond J. Mooney. Active learning for natural language parsing and information extraction. In *Proc. 16th International Conf. on Machine Learning*, pp. 406–414. 1999. URL `http://citeseer.ist.psu.edu/thompson99active.html`.

[Tre04] TreeTagger, 2004. `http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/`.

[Tur01] Peter D. Turney. Mining the Web for synonyms: PMI–IR versus LSA on TOEFL. *Lecture Notes in Computer Science*, 2167:491–502, 2001. URL `citeseer.ist.psu.edu/turney01mining.html`.

[Tur06] Jordi Turmo, Alicia Ageno, and Neus Català. Adaptive information extraction. *ACM Comput. Surv.*, 38(2), 2006. ISSN 0360-0300. URL `http://portal.acm.org/citation.cfm?id=1132957`.

[txt]   txt2html. `http://txt2html.sourceforge.net/`.

[Wee04]   Julie Weeds, David Weir, and Diana McCarthy. Characterising measures of lexical distributional similarity, 2004. URL `citeseer.ist.psu.edu/653949.html`.

[Wei05]   Sholom M. Weiss, Nitin Indurkhya, Tong Zhang, and Fred J. Damerau. *Text Mining*. Springer, New York, 2005.

[Wil93]   Craig A. Will. Comparing human and machine performance for natural language information extraction: results for english microelectronics from the muc-5 evaluation. In *MUC5 '93: Proceedings of the 5th conference on Message understanding*, pp. 53–67. Association for Computational Linguistics, Morristown, NJ, USA, 1993. ISBN 1-55860-336-0.

[Wit04]   A. Witt. Multiple hierarchies: new aspects of an old solution. In *Extreme Markup Languages 2004*. 2004. URL `www.mulberrytech.com/Extreme/Proceedings/html/2004/Witt01/EML2004Witt01.html`.

[Wor05]   WordNet, 2005. `http://wordnet.princeton.edu/w3wn.html`.

[Zav03]   Jakub Zavrel and Walter Daelemans. Feature-rich memory-based classification for shallow NLP and information extraction. In Jürgen Franke, Gholamreza Nakhaeizadeh, and Ingrid Renz, eds., *Text Mining, Theoretical Aspects and Applications*, pp. 33–54. Springer Physica, 2003. URL `http://cnts.uia.ac.be/cnts/ps/20040106.3653.zd03.pdf`.

[Zha89]   K. Zhang and D. Shasha. Simple fast algorithms for the editing distance between trees and related problems. *SIAM J. Comput.*, 18(6):1245–1262, 1989. ISSN 0097-5397.