

## References

- [1] Kilgarriff, A., Senseval: An exercise in evaluating word sense disambiguation programs. In Proceedings of the First International Conference on Language Resources and Evaluation (LREC 1998), 581-588, Granada, (1998).
- [2] Kilgarriff, A. and Rosenzweig, J., Framework and results for English Senseval. *Computers and the Humanities*, 34(1-2):15-48, (2000).
- [3] Kilgarriff, A., English lexical sample task description. In Proceedings of Senseval-2, Second International Workshop on Evaluating Word Sense Disambiguation Systems, 17-20, Toulouse, (2001).
- [4] Carpuat, M., and Wu, D., Word Sense Disambiguation vs. Statistical Machine Translation. In Proceedings of the 43rd Annual Meeting of the ACL, 387-394, Ann Arbor, (2005).
- [5] Carpuat, M., Su, W., and Wu, D., Augmenting ensemble classification for word sense disambiguation with a Kernel PCA model. In Proceedings of Senseval-3, Third International Workshop on Evaluating Word Sense Disambiguation Systems, Barcelona, (2004).
- [6] Stevenson, M., Word Sense Disambiguation: The case for Combinations of Knowledge Sources, (2003).
- [7] Duda, R. and Hart, P., *Pattern Classification and Scene Analysis*, John Wiley and Sons, New York, (1973).
- [8] Rivest, Ronald L., Learning Decision Lists, *Machine Learning*, 2(3) (1987), 229-246.
- [9] Quinlan, J.R., Induction of Decision Trees, *Machine Learning*, 1, (1986), 81-106.
- [10] Rumelhart, D.E., Hinton, G.E., and Williams, R.J., Learning Internal Representations by Error Propagation. In *Parallel Distributed Processing*, Vol. 1, (1986), 318-362.
- [11] Mitchell, Tom, *Machine Learning*, McGraw Hill, 1997.
- [12] Pedersen, T., A simple approach to building ensembles of naïve bayesian classifiers for word sense disambiguation. In Proceedings of the First Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2000), 63-69, Seattle, 2000.

- [13] Gale, B., Church, K., and Yarowsky, D., A method for disambiguating word senses in a large corpus. *Computers and the Humanities*, 1992, 26:415-439.
- [14] Gale, B., Church, K., and Yarowsky, One sense per discourse. In *Proceedings of the ARPA Workshop on Speech and Natural Language Processing*, 1992, 233-237.
- [15] Mooney, R., Comparative experiments on disambiguating word senses: An illustration of the role of bias in machine learning. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-96)*, University of Pennsylvania, 1996, 82-91.
- [16] Florian, R., Cucerzan, S., Schafer, C. and Yarowsky, D., Combining classifiers for word sense disambiguation. *Natural Language Engineering, Special Issue on Word Sense Disambiguation Systems*, 2002, 8(4):327-341.
- [17] Yarowsky, D., Hierarchical decision lists for word sense disambiguation. *Computers and the Humanities*, 2000, 34(1-2):179-186.
- [18] Agirre, E., and Martinez, D., Exploring automatic word sense disambiguation with decision lists and the web. In *Proceedings of the Coling 2000 Workshop: "Semantic Annotation and Intelligent Annotation"*, Centre Universitaire, Luxembourg, 2000.
- [19] Towell, G. and, Voorhees, E., Disambiguating Highly Ambiguous Text, *Computational Linguistics*, 24(1), 1998, 125-145.
- [20] Black, E., An Experiment in Computational Discrimination of English Word Senses, *IBM Journal of Research and Development*, 1988, 32:185-194.
- [21] Pedersen, T., Evaluating the effectiveness of ensembles of decision trees in disambiguating Senseval lexical samples. In *Proceedings of the ACL 2002 Workshop on Word Sense Disambiguation: Recent Successes and Future Directions*, Philadelphia, 2002.
- [22] Ng, H. T. and Lee, H. B., Integrating multiple knowledge sources to disambiguate word sense: An exemplar-based approach. In *34th Annual Meeting of the Association for Computational Linguistics (ACL 1996)*, 40-47, Santa Cruz, 1996.
- [23] Brown, Peter F., Stephen, D.P., Vincent, J.D.P., and Robert, L.M., Word sense disambiguation using statistical methods. In *Proceedings of the 29th Annual Meeting*, Berkeley, 1991a, 264-270.

- [24] Landes, S., Leacock, C., and Tengi, R., Building a semantic concordance of English, In C. Fellbaum, ed., WordNet: An electronic lexical database and some applications. Cambridge, MA.: MIT Press, 1998.
- [25] Hearst, M., Noun homograph disambiguation using local context in large text corpora. In Proceedings of the 7th Annual Conference of the UW Centre for the New OED and Text Research: Using Corpora, Oxford, 1991.
- [26] Schütze, H., Dimensions of Meaning. In Proceedings of Supercomputing, 1992, 787-796.
- [27] Schütze, H., Word Space. Advances in Neural Information Processing Systems 5, 1993, 895-902.
- [28] Yarowsky, D., Unsupervised word sense disambiguation rivaling supervised methods. In 33th Annual Meeting of the Association for Computational Linguistics (ACL 1995), Cambridge, 1995, 189-196.
- [29] Mihalcea, R. and Moldovan, D., A highly accurate bootstrapping algorithm for word sense disambiguation. International Journal on Artificial Intelligence Tools, 2001a, 10(1-2):5-21.
- [30] Mihalcea, R., Bootstrapping Large Sense Tagged Corpora, in Proceedings of the 3<sup>rd</sup> International Conference on Languages Resources and Evaluations (LREC 2002), Las Palmas, Spain, May 2002.
- [31] Brown, P., Lai, J. C., and Mercer, R., Aligning Sentences in Parallel Corpora. In *Proceedings of ACL-91*, Berkeley CA, 1991b.
- [32] Dagan, I. and Itai, A., Word sense disambiguation using a second language monolingual corpus. Computational Linguistics, 1994, 20(4):563-596.
- [33] Ide, N., Erjavec, T., and Tufis, D., Sense discrimination with parallel corpora. In Proceedings of the ACL 2002 Workshop on Word Sense Disambiguation: Recent Successes and Future Directions, 2002, 56-60.
- [34] Ng, H. T., Wang, B., and Chan, Y. S., Exploiting parallel texts for word sense disambiguation. In 41th Annual Meeting of the Association for Computational Linguistics (ACL 2003), 2003, 455-462.

- [35] Witten, I.H., and Bell, T.C., The zero-frequency problem: Estimating the probabilities of novel events in adaptive text compression. *IEEE Transactions on Information Theory*, 1991, 37(4), 1085-1094.
- [36] Good, I., The populations of frequencies of species and the estimation of population parameters, *Biometrika*, 1953, 40:237-264.
- [37] Gale, W., Good Turing smoothing without tears. Technical report, AT&T Bell Laboratories, Murray Hill, NJ, 1994.
- [38] Yarowsky, D., Word-sense disambiguation using statistical models of Roget's categories trained on large corpora. *Proceedings of 14th International Conference on Computational Linguistics, COLING-92.Nantes*, 454-460, 1992.
- [39] Chapman, R., *Roget's International Thesaurus (Fourth Edition)*. Harper and Row, New York, 1977.
- [40] McCarthy, D., Koeling, R., Weeds, J. and Carroll, J., Finding predominant senses in untagged text. In *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics, Barcelona, Spain*, 2004.
- [41] Schütze, H. and Pedersen, J., Information retrieval based on word senses. In *Fourth Annual Symposium on Document Analysis and Information Retrieval*, 1995, 161-175.
- [42] Pedersen, T. and Bruce, R., Distinguishing word senses in untagged text. In *Proceedings of the Second Conference on Empirical Methods in Natural Language Processing*, 1997, 197–207.
- [43] Schütze, H., Automatic word sense discrimination. *Computational Linguistics*, 1998, 24(1):97–123.
- [44] Purandare, A. and Pedersen, T., Word Sense Discrimination by Clustering Contexts in Vector and Similarity Spaces. In *the Proceedings of the Conference on Computational Natural Language Learning (CoNLL)*, 2004.
- [45] Pedersen, T., and Bruce, R., Knowledge lean word sense disambiguation. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence*, 1998, 800–805.
- [46] Lesk, M., Automatic sense disambiguation using machine readable dictionaries: how to tell a pine cone from an ice cream cone. In *Proceedings of ACM SIGDOC Conference*, 1986, 24-26.

- [47] Cowie, J., Guthrie, J., and Guthrie, L., Lexical disambiguation using simulated annealing. In Proceedings of the 15th [sic] International Conference on Computational Linguistics (Coling 1992), 1992, 359-365.
- [48] Sampson, G., A stochastic approach to parsing. In Proceedings of the 11th International Conference on Computational Linguistics (COLING-86), 1986, 151-155.
- [49] Stevenson, M. and Wilks, Y., The interaction of knowledge sources in word sense disambiguation. *Computational Linguistics*, 2001, 27(3):321-349.
- [50] Pedersen, T. and Banerjee, S., An adapted lesk algorithm for word sense disambiguation using WordNet. In Gelbukh, A., editor, Proceedings of the Third International Conference on Intelligent Text Processing and Computational Linguistics (CICLing-02), 2002.
- [51] Miller, G., Leacock, C., Teng, R., Bunker, R., and Miller, K., Five papers on WordNet. Special Issue of *International Journal of Lexicography*, 1990, 3(4).
- [52] Véronis, J., and Ide, N., An assessment of information automatically extracted from machine readable dictionaries. In Proceedings of the fifth conference of the European chapter of the Association for Computational Linguistics, 1991, 227-232.
- [53] Ide, N., and Véronis, J., Refining taxonomies extracted from machine readable dictionaries. In Hockey, Susan and Ide, Nancy (Eds.) *Research in Humanities Computing II*, Oxford University Press, 1993a, 145-59.
- [54] Ide, N., and Véronis, J., Knowledge extraction from machine readable dictionaries: an evaluation. Third International EAMT Workshop "Machine Translation and the Lexicon", Heidelberg (Germany), 1993b.
- [55] Ide, N. and Véronis, J., Introduction to the special issue on word sense disambiguation: The state of the art. *Computational Linguistics*, 1998, 24(1):1-40.
- [56] Walker, D. E., Knowledge resource tools for accessing large text files. In Sergei Nirenburg (ed.), *Machine Translation: Theoretical and methodological issues*, 254. Cambridge University Press, 1987.
- [57] Lenat, Douglas B. and Guha, Ramanathan V., *Building large knowledge-based systems*. Addison-Wesley, Reading, Massachusetts, 1990.
- [58] Briscoe, Edward J., Lexical issues in natural language processing. In Klein, Ewan H and Veltman, Frank (Eds.). *Natural Language and Speech*. [Proceedings of the

Symposium on Natural Language and Speech, 26-27 November 1991, Brussels, Belgium.] Springer- Verlag, Berlin, 39-68, 1991.

[59] Grishman, R., MacLeod, C., and Meyers, A., COMLEX syntax: Building a computational lexicon. Proceedings of the 15th International Conference on Computational Linguistics, COLING'94, 5-9 August 1994, Kyoto, Japan, 68-272, 1994.

[60] Voorhes, Ellen M., Using WordNet to disambiguate word senses for text retrieval. Proceedings of the 16th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 27 June-1 July 1993, Pittsburgh, Pennsylvania, 171- 180, 1993.

[61] Richardson, R., and Smeaton, Alan F., Automatic word sense disambiguation in a KBIR application. Working paper CA-0595, School of Computer Applications, Dublin City University, Dublin, Ireland, 1994.

[62] Li, X., Szpakowics, S., and Matwin, S., A WordNet-based algorithm for word sense disambiguation. In Proceedings of the 14th International Joint conference on Artificial Intelligence, 1368-1374, Montreal, 1995.

[63] Leacock, C., Chodorow, M., and Miller, G., Using corpus statistics and WordNet relations for sense identification. Computational Linguistics, 24(1):147-165, 1998.

[64] Hawkins, P., DURHAM: A Word Sense Disambiguation System. PhD thesis, Laboratory for Natural Language Engineering, Department of Computer Science, University of Durham, Durham, 1999.

[64] Fellbaum, C., Palmer, M., Trang Dang, H., Delfs, L., Wolff, S., Manual and Automatic Semantic Annotation with WordNet. In "Proceedings of the NAACL Workshop on WordNet and Other Lexical Resources: Applications, Customizations". Carnegie Mellon University, Pittsburg, PA, 2001.

[65] Resnik, P., Disambiguating Noun Groupings with Respect to WordNet Senses. Proceedings of the Third Workshop on Very Large Corpora, Cambridge, Massachusetts, 54-68, 1995.

[66] Agirre, E. and Rigau, G., Word sense disambiguation using conceptual density. In Proceedings of the 16th International Conference on Computational Linguistics (Coling 1996), 16-22, 1996.

[67] Mihalcea, R. and Moldovan, D., A method for word sense disambiguation of

unrestricted text. In 37th Annual Meeting of the Association for Computational Linguistics (ACL 1999), 152-158, Maryland, 1999.

[68] Lin, D., Using syntactic dependency as local context to resolve word sense ambiguity. In 35th Annual Meeting of the Association for Computational Linguistics and 8<sup>th</sup> Conference of the European Chapter of the Association for Computational Linguistics (ACL/EACL 1997), pages 64-71, Madrid, 1997.

[69] Lin, D., Automatic retrieval and clustering of similar words. In COLINGACL'98, Montreal, 1998.

[70] Lin, D., Word sense disambiguation with a similarity-smoothed case library. *Computers and the Humanities*, 34(1-2):147-152, 2000.

[71] Jiang, Jay J., and Conrath, David W., Semantic similarity based on corpus statistics and lexical taxonomy. In: *Proceedings of ROCLING X*, Taiwan, 1997.

[72] Agirre, E. and Martinez, D. Knowledge sources for word sense disambiguation. In Matousek, V., Mautner, P., Moucek, R., and Tauser, K., editors, *Proceedings of the Fourth International Conference TSD 2001*, Plzen, *Notes in Computer Science*, 1-10, Berlin, 2001.

[73] Haynes, S., Semantic tagging using WordNet examples. In *Proceedings of Senseval-2, Second International Workshop on Evaluating Word Sense Disambiguation Systems*, 79- 82, Toulouse, 2001.

[74] Banerjee, Satanjeev, and Pedersen, Extended Gloss Overlaps as a Measure of Semantic Relatedness. In: *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence IJCAI-2003*, Acapulco, Mexico, 2003.

[75] Litkowski, K., Senseval: The CL research experience. *Computers and the Humanities*, 34(1-2):153-158, 2000

[76] Litkowski, K., Use of machine readable dictionaries for word-sense disambiguation in Senseval-2. In *Proceedings of Senseval-2, Second International Workshop on Evaluating Word Sense Disambiguation Systems*, 107-110, Toulouse, 2001.

[77] Mihalcea, R. and Moldovan, D., Word sense disambiguation based on semantic density. In *Proceedings of the Coling-ACL'98 Workshop "Usage of WordNet in Natural Language Processing Systems"*, 16-22, Montreal, 1998.

- [78] Gaustad, T., Linguistic Knowledge and Word Sense Disambiguation, PhD Thesis, 2004.
- [79] Hirst, G., Semantic Interpretation and the Resolution of Ambiguity. Cambridge University Press, Cambridge, 1987.
- [80] Dahlgren, K., McDowell, J., and Stabler, Edward P., Knowledge representation for commonsense reasoning with text. *Computational Linguistics*, 15(3):149-170, 1989.
- [81] Cotton, S., Edmonds, P., Kilgarriff, A., and Palmer, M., SENSEVAL-2: Second International Workshop on Evaluating Word Sense Disambiguation Systems, Toulouse, 2001.
- [82] Mihalcea, R., Word sense disambiguation with pattern learning and automatic feature selection, *Journal of Natural Language and Engineering (JNLE)*, 8(4):343-358, 2002.
- [83] Stetina, J., Kurohashi, S., and Nagao, M., General word sense disambiguation method based on a full sentential context. In *Proceedings of the Coling-ACL'98 Workshop "Usage of WordNet in Natural Language Processing Systems"*, 1-8, Montreal, 1998.
- [84] Collins, M., Head-Driven Statistical Models for Natural Language Parsing, PhD thesis, Computer and Information Science Department, University of Pennsylvania, Philadelphia, 1999.
- [85] Martinez, D., Agirre, E., and Marquez, L., Syntactic features for high precision word sense disambiguation, In *Proceedings of the 19th International Conference on Computational Linguistics (Coling 2002)*, 626-632, Taipei, 2002.
- [86] Lin, D., Principle-based parsing without overgeneration, In *31th Annual Meeting of the Association for Computational Linguistics (ACL 1993)*, 112-120, Columbus, 1993.
- [87] Dunning, Ted, Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19(1):61-74, 1993.
- [88] Weeds, J., A Comparison of Lin's Similarity Measure and Lee's Alpha-Skew Divergence Measure, 2002.



- [89] Resnik, P., Using information content to evaluate semantic similarity in a taxonomy. In: Proceedings of the 14th International Joint Conference on Artificial Intelligence Montreal, 1995.
- [90] Hirst, G., and St-Onge, D., Lexical chains as representations of context for the detection and correction of malapropisms. In: Fellbaum, Christiane, ed., WordNet: An Electronic Lexical Database, MIT Press, 1998.
- [91] Wu, Z. and Palmer, M., Verb semantics and lexical selection. In Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics, Las Cruces, New Mexico, 1994.
- [92] Patwardhan, S., Banerjee, S., and Pedersen, T., Using measures of semantic relatedness for word sense disambiguation. In Proceedings of the Fourth International Conference on Intelligent Text Processing and Computational Linguistics, Mexico City, 2003.
- [93] Alam, H. et al., Extending A Broad-Coverage Parser for a General NLP Toolkit, 2002. 454-460.
- [94] Yarowsky, D., One sense per collocation. Proceeding of ARPA Human Language Technology Workshop, Princeton, New Jersey, 266-271, 1993.
- [95] Yarowsky, D., Decision lists for lexical ambiguity resolution: application to accent restoration in Spanish and French. Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics, Las Cruces, New Mexico, 88-95, 1994.
- [96] Yarowsky, D., A comparison of corpus-based techniques for restoring accents in Spanish and French text. Proceedings of the 2nd Annual Workshop on Very Large Text Corpora. Las Cruces, 19-32, 1994.
- [99]de Garis, H. (1996). CAM-BRAIN: The Evolutionary Engineering of a Billion Neuron Artificial Brain by 2001 Which Grows/Evolves at Electronic Speeds Inside a Cellular Automata Machine (CAM), In Lecture Notes in Computer Science - Towards Evolvable Hardware, Vol. 1062, pages 76-98. Springer Verlag, pages. 76-98.
- [100]D\_avila, J. (1999) Exploring the Relationship Between Neural Network Topology and Optimal Training Set by Means of Genetic Algorithms. In International Conference on Artificial Neural Networks and Genetic Algorithms, pages 307-311. Springer-Verlag.

- [101]D\_avila, J. (2000) A New Metric for Evaluating Genetic Optimization of Neural Networks. In Proceedings of the First IEEE Symposium on Combinations of Evolutionary Computation and Neural Networks. pages 52-58. IEEE
- [102]Fullmer, B., Miikkulainen, R., (1991). Using marker- based genetic encoding of neural networks to evolve finite-state behavior. In Proceedings of the first European Conference on Artificial Life, pages. 253-262.
- [103]Kitano, H., (1994). Designing Neural Networks using Genetic Algorithm with Graph Generation System. In Complex Systems, 4. pages 461-476.
- [104]Weiss, S., Kulikowski, C. (1991). Computer Systems that Learn. Classification and Prediction Methods from Statistics. In Neural Nets, Machine Learning, and Expert Systems. Morgan Kaufmann.
- [105]Yao, X. (1993). A review of evolutionary artificial neural networks. In International Journal of Intelligent Systems, 4. pages 203-222.