

# Bibliography

- [1] R. K. Mohanty et al., Synset Based Multilingual Dictionary: Insights, Applications and Challenges. GWC 2008: 4th Global WordNet conference, Szeged, Hungary, January 2008
- [2] C. Fellbaum, WordNet: An Electronic Lexical Database(ed.). The MIT Press, 1998,
- [3] S. Jha, D. Narayan, P. Pande, P. Bhattacharyya, A WordNet for Hindi. Workshop on Lexical Resources in Natural Language Processing, Hyderabad, India, January, 2001.
- [4] Vossen, P., P. Diez-Orzas, W. Peters. The Multilingual Design of EuroWordNet. Published in: P. Vossen, N. Calzolari, G. Adriaens, A. Sanfilippo, Y. Wilks (eds.) Proceedings of the ACL/EACL-97 workshop Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications, Madrid, July 12th, 1997. 1-8.
- [5] A. Chakraborty, B. S. Purkayastha, A. Roy, Experiences in building the Nepali Wordnet. Proceedings of the 5th Global WordNet Conference, Mumbai, Narosa Publishing House, India, Mumbai 2010.
- [6] <http://tdil-dc.in>.
- [7] J. Saraswati, R. Shukla, S. Pathade, T. Solanki and P. Bhattacharyya, Challenges in Multilingual Domain-Specific Sense-Marking, Proceedings of the 5th Global WordNet Conference, Mumbai, Narosa Publishing House, India, Mumbai, Jan, 2010.
- [8] [http://www.cfilt.iitb.ac.in/wordnet/webhwm/other/hwm\\_docs\\_2.pdf](http://www.cfilt.iitb.ac.in/wordnet/webhwm/other/hwm_docs_2.pdf), Semantic and lexical relations in Hindi WordNet
- [9] G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. Miller, Introduction to WordNet: An On-line Lexical Database, Princeton University, Cognitive Science laboratory, Technical report, 1993.
- [10] P. Bhattacharyya, IndoWordnet. Proceedings of Lexical Resources Engineering Conference (LREC 2010), Malta, May, 2010
- [11] P. Vossen (ed.). EuroWordNet: A Multilingual Database with Lexical Semantic Networks. Kluwer Academic Publishers, Dordrecht. 1998

- [12] D. Narayan, D. Chakrabarty, P. Pande and P. Bhattacharyya, An Experience in Building the Indo WordNet- a WordNet for Hindi. 1st International Conference on Global WordNet (GWC 02), Mysore, India, 2002.
- [13] M. Sinha, M. K. Reddy, P. Bhattacharyya, P. Pandey and L. Kashyap, Hindi Word Sense Disambiguation. Proceedings of International Symposium on Machine Translation, Natural Language Processing and Translation Support System, Delhi India, 2003
- [14] M. Lesk. 1986, Automatic sense disambiguation using machine readable dictionaries: how to tell a pine cone from an ice cream cone. Proceedings of the 5th annual international conference on Systems documentation, Toronto, Ontario, Canada, 1986.
- [15] D. Walker and R. Amsler, The Use of Machine Readable Dictionaries in Sublanguage Analysis. In Analyzing Language in Restricted Domains, Grishman and Kittredge (eds), LEA Press, pp. 69-83, 1986.
- [16] D. Yarowsky, Word sense disambiguation using statistical models of Roget's categories trained on large corpora. Proceedings of the 14th International Conference on Computational Linguistics (COLING), Nantes, France, 454-460, 1992.
- [17] D. Yarowsky, 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 (ACL), Las Cruces, U.S.A., 88-95, 1994.
- [18] D. Yarowsky, Unsupervised word sense disambiguation rivaling supervised methods. Proceedings of the 33rd Annual Meeting of the Association for Computational Linguistics (ACL), Cambridge, MA, 189-196, 1995.
- [19] E. Agirre & R. German, Word sense disambiguation using conceptual density. Proceedings of the 16th International Conference on Computational Linguistics (COLING), Copenhagen, Denmark, 1996
- [20] Ng, Hwee T. & Hian B. Lee, Integrating multiple knowledge sources to disambiguate word senses: An exemplar-based approach. Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics (ACL), Santa Cruz, U.S.A., 40-47, 1997
- [21] Ng, Hwee T., Exemplar-based word sense disambiguation: Some recent improvements. Proceedings of the 2nd Conference on Empirical Methods in Natural Language Processing (EMNLP), Providence, U.S.A., 208-213.
- [22] D. Lin, Using syntactic dependency as local context to resolve word sense ambiguity. Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics (ACL), Madrid, 64-71, 1997.
- [23] R. Mihalcea, D. I. Moldovan, An automatic method for generating sense tagged corpora. Proceedings of the sixteenth national conference on Artificial intelligence

and the eleventh Innovative applications of artificial intelligence conference innovative applications of artificial intelligence. Orlando, Florida, United States, 1999.

[24] P. Resnik, 1999, Semantic Similarity in a Taxonomy: An Information-Based Measure and its Application to Problems of Ambiguity in Natural Language. *Journal of Artificial Intelligence Research*, 1999.

[25] E. Agirre, J. Atserias, L. Padr, G. Rigau, Combining Supervised and Unsupervised Lexical Knowledge Methods for Word Sense Disambiguation Computers and the Humanities. Special Double Issue on SensEval. Eds. Martha Palmer and Adam Kilgarriff. 34:1,2, 2000.

[26] R. Mihalcea and D. Moldovan, An Iterative Approach to Word Sense Disambiguation. *Proceedings of Florida Artificial Intelligence Research Society Conference (FLAIRS 2000)*, [pg.219-223] Orlando, FL, May 2000.

[27] A. Eneko, A. Olatz, H. Eduard, M. David, Enriching WordNet concepts with topic signatures. *Proceedings of the NAACL workshop on WordNet and Other lexical Resources: Applications, Extensions and Customizations*. Pittsburg, 2001.

[28] M. Diab and P. Resnik, An Unsupervised Method for Word Sense Tagging Using Parallel Corpora. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, Philadelphia, Pennsylvania, July 2002.

[29] Jean Veronis., HyperLex: Lexical cartography for information retrieval. *Computer Speech & Language*, 18(3):223-252, 2004.

[30] R. Mihalcea and E. Faruque, SenseLearner: Minimally Supervised Word Sense Disambiguation for All Words in Open Text. *Proceedings of ACL/SIGLEX Senseval-3*, Barcelona, Spain, July 2004.

[31] I. Bhattacharya, L. Getoor, Y. Bengio, 2004, Unsupervised sense disambiguation using bilingual probabilistic models. In *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*, Spain, 2004.

[32] Y. K. Lee, Hwee T. Ng & Tee K. Chia, Supervised word sense disambiguation with support vector machines and multiple knowledge sources. *Proceedings of Senseval-3: Third International Workshop on the Evaluation of Systems for the Semantic Analysis of Text*, Barcelona, Spain, 137-140.

[33] R. Mihalcea, Large vocabulary unsupervised word sense disambiguation with graph-based algorithms for sequence data labeling. *Proceedings of the Joint Human Language Technology and Empirical Methods in Natural Language Processing Conference (HLT/EMNLP)*, Vancouver, Canada, 411-418, 2005.

[34] M. Carpuat, D. Wu, Evaluating the Word Sense Disambiguation Performance of Statistical Machine Translation.

[35] M. Carpuat and D. Wu, Word sense disambiguation vs. statistical machine translation. *Proceedings of the 43rd Annual Meeting of the Association for*

Computational Linguistics (ACL'05), pages 387-394, Ann Arbor, Michigan, June 2005.

[36] R. Navigli, P. Velardi, Structural Semantic Interconnections: A Knowledge-Based Approach to Word Sense Disambiguation. IEEE Transactions On Pattern Analysis and Machine Intelligence, July 2005.

[37] M Ciaramita, Y. Altun, Broad-Coverage Sense Disambiguation and Information Extraction with a Supersense Sequence Tagger. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2006).

[38] Y. Seng Chan, HweeTou Ng, Estimating class priors in domain adaptation for word sense disambiguation. Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the ACL, Sydney, 2006.

[39] R. Navigli, Meaningful clustering of senses helps boost word sense disambiguation performance. Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the ACL, Sydney, 2006.

[40] R. Navigli, Ensemble methods for unsupervised WSD. Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the ACL, Sydney, 2006.

[41] Z. Lu, Haifeng Wang, Jianmin Yao, Ting Liu, Sheng Li, "An equivalent pseudoword solution to Chinese word sense disambiguation", in Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the ACL, Sydney, 2006.

[42] Y. S. Chan, H. Tou Ng and D. Chiang, Word Sense Disambiguation Improves Statistical Machine Translation. Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics, Prague, 2007.

[43] E. Agirre and O. L. deLacalle, Supervised domain adaption for wsd. In EACL '09: Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics, pages 42–50, Morristown, NJ, USA. Association for Computational Linguistics.

[44] E. Agirre and D. Martinez, The effect of bias on an automatically-built word sense corpus. In Proceedings of the 4rd International Conference on Languages Resources and Evaluations (LREC). 2004

[45] E. Agirre, O. L. deLacalle, C. Fellbaum, A. Marchetti, A. Toral, and P. Vossen, Semeval-2010 task 17: all-words word sense disambiguation on a specific domain. DEW '09: Proceedings of the Workshop on Semantic Evaluations: Recent Achievements and Future Directions, pages 123–128, Morristown, NJ, USA. Association for Computational Linguistics.

[46] E. Agirre, O. L. De Lacalle, and A. Soroa, Knowledge-based wsd on specific domains: Performing better than generic supervised wsd. In Proceedings of IJCAI. 2009

- [47] S.Banerjee and T.Pedersen, An adapted lesk algorithm for word sense disambiguation using wordnet. In CILCling '02: Proceedings of the Third International Conference on Computational Linguistics and Intelligent Text Processing, pages 136–145, London, UK. Springer-Verlag.
- [48] M.Ciaramita, Y.Altun, Broad-Coverage Sense Disambiguation and Information Extraction with a Supersense Sequence Tagger Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP 2006), pages 594–602, Sydney, July 2006, 2006, Association for Computational Linguistics.
- [49] Lovins, J.B. 1968. Development of a stemming algorithm, Mechanical Translations and Computational Linguistics Vol.11 Nos 1 and 2, pp. 22-31.
- [50] Porter M.F. 1980. An algorithm for suffix stripping, Program; 14, 130-137
- [51] Porter M.F. 2006. Stemming algorithms for various European languages, Available at [URL] <http://snowball.tartarus.org/texts/stemmersoverview.html>
- [52] Goldsmith John A. 2001. Unsupervised Learning of the morphology of a Natural Language, Computational Linguistics, 27(2): 153-198
- [53] Goldsmith John A. 2006. An algorithm for the unsupervised Learning of morphology, Natural Language Engineering, 12(4): 353-371
- [54] Creutz, Mathis, and Krista Lagus. 2005. Unsupervised morpheme segmentation and morphology induction from text corpora using Morfessor 1.0., Technical Report A81, Publications in Computer and Information Science, Helsinki University of Technology.
- [55] Creutz, Mathis, and Krista Lagus. 2007. Unsupervised models for morpheme segmentation and morphology learning, Association for Computing Machinery Transactions on Speech and Language Processing, 4(1):1-34
- [56] Grzegorz Chrupala 2006. Simple data-driven context-sensitive lemmatization, Chrupala, Grzegorz (2006) Simple data-driven context-sensitive lemmatization.
- [57] Ramanathan, Ananthakrishnan, and Durgesh D. Rao, 2003. A Lightweight Stemmer for Hindi. , Workshop on Computational Linguistics for South-Asian Languages, EACL
- [58] Cormen, Thomas H. and Stein, Clifford and Rivest, Ronald L. and Leiserson, Charles E. 2001. Introduction to Algorithms, 2nd Edition, ISBN:0070131511, McGraw-Hill Higher Education
- [59] JolPlisson, Nada Lavrac, Dunja Mladenic 2008. A Rule based Approach to Word Lemmatization, Proceedings of 7th International Multiconference Information Society, IS 2004, Institute Jozef Stefan, Ljubljana, pp.83-86, 2008

- [60] Snigdha Paul, Nisheeth Joshi and ItiMathur 2013. Development of a Hindi Lemmatizer, CoRR, DBLP:journals/corr/abs/1305.6211 2013
- [61] Pushpak Bhattacharyya, AnkitBahuguna, LavitaTalukdar, and BornaliPhukan. . 2014 Facilitating Multi-Lingual Sense Annotation: Human Mediated Lemmatizer. In Global WordNet Conference, 2014
- [62] Tarek El-Shishtawy and Fatma El-Ghannam 2012. An Accurate Arabic Root-Based Lemmatizer for Information Retrieval Purposes, IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 1, No 3, January 2012 ISSN (Online): 1694-0814.
- [63] DabreRaj,AmberkarArchana and Bhattacharyya Pushpak 2012. Morphology Analyser for Affix Stacking Languages: a case study in Marathi, COLING 2012, Mumbai, India, 10-14 Dec, 2012.
- [64] Massimo Melucci and Nicola Orio 2003. A novel method of Stemmer Generation Based on Hidden Markov Models, CIKM '03, New Orleans, Louisiana, USA.
- [65] LauriKarttunen 1983. KIMMO: A General Morphological Processor , Texas Linguistic Forum, 22 (1983), 163186.
- [66]Majumder, Prasenjit, MandarMitra, Swapan K. Parui, GobindaKole, PabitraMitra, and KalyankumarDatta. 2007. YASS: Yet another suffix stripper, Association for Computing Machinery Transactions on Information Systems, 25(4):18-38..
- [69] Bart Jongejan and Hercules Dalianis 2009. Automatic training of lemmatization rules that handle morphological changes in pre-, in- and suffixes alike, Proceedings of the 47th Annual Meeting of the ACL and the 4th IJCNLP of the AFNLP, pages 145153, Suntec, Singapore, 2-7 August 2009
- [70] A. Chatterjee, S. R. Joshi, M. M. Khapra and P. Bhattacharyya 2010. Introduction to Tools for IndoWordnet and Word Sense Disambiguation, The 3rd IndoWordnet Workshop, Eighth International Conference on Natural Language Processing (ICON 2010), IIT Kharagpur, India
- [71] T. Petrolito and F. Bond, A survey of WordNet annotated corpora, Proceedings of the 7th Global WordNet Conference (GWC 2014), 2014.
- [72] S. Landes, C. Leacock and R. I. Teng, Building semantic concordances, WordNet: An electronic lexical database, 1998: 199-216.
- [73] H. Tou Ng and H. B. Lee. Integrating multiple knowledge sources to disambiguate word sense: An exemplar-based approach. In Proceedings of the 34th annual meeting on Association for Computational Linguistics, pages 40–47, 1996.
- [74] F. Bond, T. Baldwin, R. Fothergill & K. Uchimoto, Japanese SemCor: A sense-tagged corpus of Japanese, Proceedings of the 6th International Conference of the Global WordNet Association (GWC), 2012.

- [75] P. Vossen, A. Görög, R. Izquierdo, & A. van den Bosch, DutchSemCor: Targeting the ideal sense-tagged corpus, LREC, 2012.
- [76] P. Vossen, I. Maks, R. Segers, & H. VanderVliet, Integrating Lexical Units, Synsets, and Ontology in the Cornetto Database, Proceedings on the sixth international Conference on Language Resources and Evaluation (LREC), 2008.
- [77] S. Koeva, S. Lesseva and M. Todorova, Bulgarian sense tagged corpus. 5th SALT MIL Workshop on Minority Languages, Genoa, Italy, May 23rd 2006
- [78] S. Koeva, S. Mihov, & T. Tinchev, "Bulgarian Wordnet—Structure and Validation." Romanian Journal of Information Science and Technology 7, 2004.
- [79] S. Urooj, S. Shams, S. Hussain, F. Adeeba, Sense Tagged CLE Urdu Digest Corpus, Proc. Conf. on Language and Technology, Karachi, 2014.
- [80] Y. Wu, P. Jin, Y. Zhang & S. Yu, "A Chinese corpus with word sense annotation" Computer Processing of Oriental Languages. Beyond the Orient: The Research Challenges Ahead, Springer, Berlin Heidelberg, 2006.
- [81] Mitesh M. Khapra et.al. All words domain adapted WSD: Finding a middle ground between supervision and unsupervision (2010), In Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL)
- [82] Sivanandanam SN, Sumathi S, Deepa SN 2009. Introduction to Neural Networks Using Matlab6.0 , 9th Edition, ISBN:100070591121, Tata McGraw-Hill
- [83] Statistical Summaries, Ethnologue. 2012
- [84] Languages in Descending Order of Strength — India, States and Union Territories – 1991 Census, *Census Data Online*.
- [85] [http://www.isical.ac.in/~rc\\_bangla/bangla.html](http://www.isical.ac.in/~rc_bangla/bangla.html) as seen on 06.01.2015
- [86] Navarro, Gonzalo (2001). "A guided tour to approximate string matching", ACM Computing Surveys 33 (1): 31–88
- [87] [http://tdil-dc.in/index.php?option=com\\_vertical&parentid=60&lang=en](http://tdil-dc.in/index.php?option=com_vertical&parentid=60&lang=en) as seen on 08.06.2015
- [88] Frakes, W. "Stemming Algorithms." In Frakes, W. and R. Baeza-Yates, (ed.) Information Retrieval: Data Structures and Algorithms. Englewood Cliffs, NJ: Prentice-Hall, 1992.
- [89] . A. Roy, S. Sarkar, and B. S. Purkayastha , Knowledge Based Approaches to Nepali Word Sense Disambiguation, IJNLIC, Vol. 3, No.3, June 2014, Pg. 51-63