





All the other edges in the graph, including links between adjectives and adverbs, or links across different parts-of-speech, are drawn using the *lesk* measure. The results indicate that the right combination of similarity metrics can lead to a performance competing with the state-of-the-art in unsupervised word sense disambiguation.

### 3 WSD using the Lesk Algorithm

The Lesk algorithm [10] uses dictionary definitions (gloss) to disambiguate a polysemous word in a sentence context. The major objective of his idea is to count the number of words that are shared between two glosses. The more overlapping the words, the more related the senses are.

To disambiguate a word, the gloss of each of its senses is compared to the glosses of every other word in a phrase. A word is assigned to the sense whose gloss shares the largest number of words in common with the glosses of the other words. Figure 1 shows the graphic representation of the Lesk Algorithm.

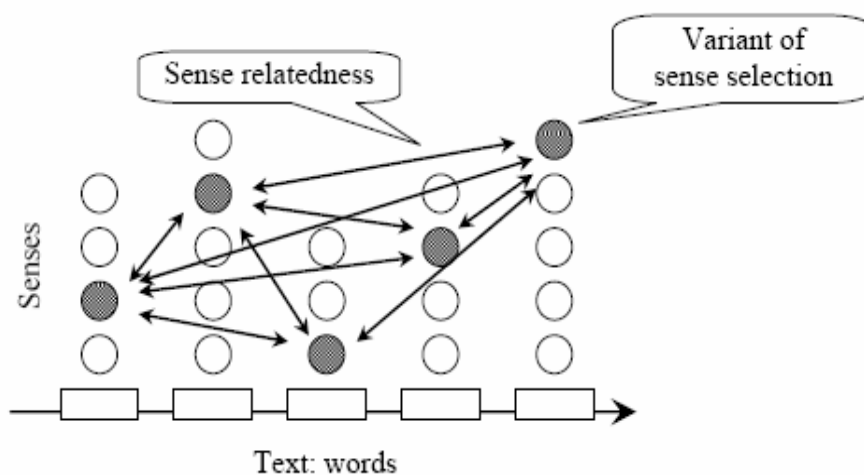


Figure 1. Graphic Representation of the Lesk Algorithm

For example: In performing disambiguation for the "pine cone" phrasal, according to the Oxford Advanced Learner's Dictionary, the word "pine" has two senses:

- sense 1: kind of evergreen tree with needle-shaped leaves,
- sense 2: waste away through sorrow or illness.

The word "cone" has three senses:

- sense 1: solid body which narrows to a point,
- sense 2: something of this shape whether solid or hollow,
- sense 3: fruit of a certain evergreen tree.

















9. Leacock and M. Chodorow. Combining local context and WordNet sense similarity for word sense identification. In *WordNet, An Electronic Lexical Database*. The MIT Press, 1998.
10. M. Lesk. Automatic sense disambiguation using machine readable dictionaries: How to tell a pine cone from a ice cream cone. In *Proceedings of SIGDOC '86*, 1986.
11. Lin. An information-theoretic definition of similarity. In *Proceedings of the 15th International Conference on Machine Learning*, Madison, WI, 1998.
12. M. Palmer, C. Fellbaum, S. Cotton, L. Delfs, and H. Dang. English tasks: all-words and verb lexical sample. In *Proceedings of ACL/SIGLEX Senseval-2*, Toulouse, France, 2001.
13. M. Palmer and M. Light. Introduction to the special issue on semantic tagging. *Natural Language Engineering*, 5(2): i-iv.
14. S. Patwardhan, S. Banerjee, and T. Pedersen. Using Measures of Semantic Relatedness for Word Sense Disambiguation. In *Proceedings of the Fourth International Conference on Intelligent Text Processing and Computational Linguistics*, pp. 241–257, Mexico City, Mexico, February, 2003.
15. S. Pradhan, E. Loper, D. Dligach, and M. Palmer. SemEval-2007 task 17: English lexical sample, SRL and all words. In *Proceedings of the 4th international workshop on semantic evaluations (SemEval-2007)*, pp. 87–92, Prague, Czech Republic, 2007.
16. P. Resnik. Using information content to evaluate semantic similarity. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, Montreal, Canada, 1995.
17. Sebtı and A. Barfroush. A new word sense similarity measure in wordnet. In *Proceedings of the IEEE International Multiconference on Computer Science and Information Technology*, October, 2008.
18. R. Sinha and R. Mihalcea. Unsupervised Graph-based Word Sense Disambiguation Using Measures of Word Semantic Similarity, In *Proceedings of the IEEE International Conference on Semantic Computing (ICSC 2007)*, Irvine, CA, September 2007.
19. B. Snyder and M. Palmer. The English all-words task. In *Proceedings of ACL/SIGLEX Senseval-3*, Barcelona, Spain, July 2004.
20. Senseval. *Evaluation Exercises for the Semantic Analysis of Text* <http://www.senseval.org>