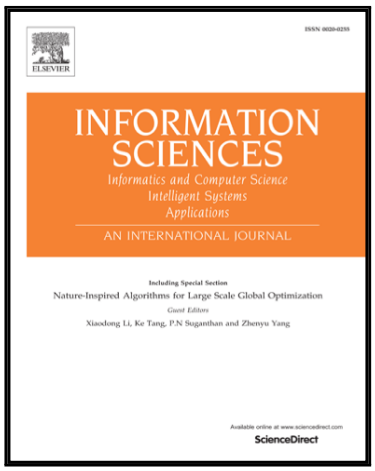


# A Hybrid Genetic-Ant Colony Optimization Algorithm for the Word Sense Disambiguation Problem

Wojdan Alsaeedan, Mohamed El Bachir Menai, Saad Al-Ahmadi

PII: S0020-0255(17)30130-5
DOI: 10.1016/j.ins.2017.07.002
Reference: INS 12968

To appear in: *Information Sciences*

Received date: 9 January 2017
Revised date: 23 June 2017
Accepted date: 1 July 2017

Please cite this article as: Wojdan Alsaeedan, Mohamed El Bachir Menai, Saad Al-Ahmadi, A Hybrid Genetic-Ant Colony Optimization Algorithm for the Word Sense Disambiguation Problem, *Information Sciences* (2017), doi: 10.1016/j.ins.2017.07.002

# A Hybrid Genetic-Ant Colony Optimization Algorithm for the Word Sense Disambiguation Problem

Wojdan Alsaeedan[a], Mohamed El Bachir Menai[a,*], Saad Al-Ahmadi[a]

[a]*Department of Computer Science, College of Computer and Information Sciences, King Saud University*

**Abstract**

Word sense disambiguation (WSD) is a natural language processing problem that occurs at the semantic level. It consists of determining the sense of a polysemous word that is suitable in a particular context. WSD has been addressed using several approaches, including metaheuristic algorithms. We propose hybrid algorithms for WSD that consist of a self-adaptive genetic algorithm (SAGA) and variants of ant colony optimization (ACO) algorithms: max-min ant system (MMAS) and ant colony system (ACS). SAGA is used to automatically tune the parameters of MMAS and ACS. The ACO algorithms are adapted based on a combination of semantic relatedness between sequences of senses corresponding to the context words and semantic relatedness between the sense of a target word and the sense of a context word. We evaluated the performance of the two ACO algorithms (MMASWSD and ACSWSD) and their hybridization with SAGA (GMMASWSD and GACSWSD) on fine-grained and coarse-grained corpora, and compared them with the best-performing algorithms. The empirical results indicate that GMMASWSD outperformed the other variants and all of the rival algorithms on the fine-grained corpora. However, GMMASWSD did not achieve the best performance on the coarse-grained corpus, even though its performance was close to that of the best algorithm.

*Keywords:* Natural language understanding, Word sense disambiguation, Ant colony optimization, Genetic algorithms.

## 1. Introduction

Word sense disambiguation (WSD) addresses one of the main characteristics of natural language applications (e.g., machine translation, question answering, and document classification)—the problem of ambiguous words (words that have more than one meaning). Thus, WSD is one of the main research directions in the field of natural language processing. The goal of methods to solve the WSD problem is to assign a proper meaning to a polysemous word in a context by selecting the appropriate meaning from an inventory of word meanings. For example, the word "partner" can have two very different meanings, for example, "within a marriage relationship" versus "within a law or accounting firm". The meaning or sense inventories contain words, their senses, and extra information about the words. The two main types of inventories used in methods intended to perform WSD are structured inventories (e.g., thesauri and ontologies) and unstructured inventories (e.g., corpora). One of the more common inventories used in WSD is WordNet (WN)[1]. In WN, words are arranged in groups of synonyms (named synsets). Each synset contains a gloss, which is the textual definition, and there are lexical and semantic relations between a pair of synsets. Following the original work on WN, some additions have been made to enhance the relations between pairs of synsets. In eXtended WordNet (XWN)[2], the words in the gloss of a synset are annotated with their senses. Then, relations are created between the synset and the senses of these words. Another work, called WordNet++ (WNPP), produced relations between noun synsets in WN based on Wikipedia [1]. Magnini and Cavaglia developed the WordNet Domain

*Corresponding author. P.O. Box 51178, Riyadh 11543, Saudi Arabia
*Email address:* menai@ksu.edu.sa (Mohamed El Bachir Menai)

[1]https://wordnet.princeton.edu

[2]http://www.hlt.utdallas.edu/ xwn/about.html