

# “Dissertation in Language Sciences”

s1107496

January 20, 2015

## Abstract

Particle swarm optimization (PSO) has been proposed as a means of modeling changes in human behaviour in a social context. PSO has also been shown to be an effective optimization method for non-differentiable problems with continuous search spaces, and has seen widespread use as such. In this paper, the language game presented in Rohde et al.’s “Communicating with Cost-based Implicature: a Game-Theoretic Approach to Ambiguity” is modeled as a particle swarm optimization task, with the aim of creating a system which captures the results seen in Rohde et al.’s experiments. Two such possible systems are presented, one of which is successful in modeling these results. Further to this, PSO’s suitability for use in language games with highly dynamic objective functions, how we might frame PSO’s behaviour from game-theoretic and psycholinguistic standpoints, and predictions made by the successful PSO-based model are also discussed.

## 1 Introduction

Computational modeling is a powerful tool which has been used extensively in the field of linguistics. By building computational models of linguistic phenomena, it is possible to create simulations which explore the experimentally infeasible or impossible. Computational models can also be used to extrapolate from data collected in the lab, the results of which can be used to drive further studies in potentially fruitful directions. Further to this, computational models uniquely allow for direct inspection of their parameters and processes, which can shed light on aspects of human cognition. Of the multitudinous uses of modeling in the recent linguistics literature, notable examples include the use of iterated learning to demonstrate the spontaneous emergence of syntax from holistic languages (Kirby & Hurford, 2002), the use of incremental probabilistic Earley parsing to model garden-path sentence comprehension (Hale, 2001), the use of Bayesian statistics to model pragmatic reference disambiguation (Frank & Goodman, 2012), Modeling to gain insight to linguistics - other successful models - Elman - Frank and Goodman - Kirby - “This paper considers how modeling...”?

Reference - what is reference? Frege sense/reference (RTM) paradigm? - Speakers/comprehenders - Cost/ambiguity (balance cost with failure) - Coordination to reduce cost - how do we capture this? - analysis through game theory - Jaeger

In “Communicating with Cost-based Implicature: a Game-Theoretic Approach to Ambiguity” (2012), Rohde et al. present an iterated language game in which participants aim

to indicate an object to their partner via use of one of several possible referring terms. Participants gain points upon successful communication, but must spend points in order to communicate. Each referent has a corresponding unambiguous form that players may choose to send to their partners, alternatively, players may send an ambiguous form with a different cost that could potentially indicate a number of referents. The studies run by Rohde et al. demonstrated the likelihood of a pair of participants to successfully coordinate their use of an ambiguous reference to be partially contingent on the relative costs of the unambiguous and ambiguous references.

Particle swarm optimization, as originally introduced by Kennedy & Eberhart (1995), can serve not only as a general optimization method, but also as a means of modeling human social behaviour, especially in the context of collaborative problem solving (Kennedy, 1997). Modeling the language games used in Rohde et al. as particle swarm optimization tasks, if possible, would allow a more exhaustive exploration of the effects of various form costs on reference coordination, offering a comprehensive picture of the circumstances under which ambiguous form entrainment is possible and likely. Further to this, an accurate particle-swarm-based model of human coordination in a discourse setting might shed light on more fundamental aspects of human cognition and social interaction as represented via the model’s parameters. A successful model could also suggest particle swarm optimization’s suitability for the modeling of other linguistic phenomena.

Conversely, if particle swarm optimization is not a viable means of modeling the results observed in Rohde et al., this could suggest a fundamental difference between human linguistic behaviours and other social behaviours as successfully modeled via PSO. A negative finding might also suggest that specialized models, such as Liu et al.’s “human behaviour-based particle swarm optimization” (2014), are required for the modeling of more complex human social interactions.

To address these possibilities, this paper models Rohde et al.’s experiments as a particle swarm optimization task, utilizing a mixed strategy search space to represent form production and comprehension. Two model variants are presented. In the first, particles are restricted to a unit interval search space, resulting in a model which fails to capture the effects seen in Rohde et al. The second model, which utilizes a technique suggested in Engelbrecht 2005, p. 252 to allow for an unrestricted search space, exhibits responses to form costs in keeping proportionally with the results found by Rohde et al..

## 2 Background

### 2.1 Particle swarm optimization

#### 2.1.1 Origin and concept

Particle swarm optimization was first formulated by Kennedy & Eberhart (1995) in an attempt to model human social behaviour. The initial model was based off of Heppner & Grenander (1990)’s work in modeling bird flocking and roosting behaviour in two-dimensional space, refined to allow for an arbitrary number of dimensions and for particles to share the same or arbitrarily close positions within said multidimensional space. Kennedy & Eberhart demonstrated that their new optimization method was suitable for use not only in modeling

human social behaviour, but also for the general-purpose optimization of non-linear continuous problems. Specifically, particle swarm optimization was found to be successful both in optimizing the weights of an artificial neural network and in the Schaffer *f6* problem, a standard benchmark for general-purpose optimization methods (Davis, 1991).

The concept behind the particle swarm optimization algorithm is relatively simple. Potential solutions to some problem with  $n$  dimensions are represented as particles existing within an  $n$ -dimensional search space. Each particle has both a position within this space and some velocity vector. Each particle also keeps track of the best position it has been in as evaluated by the given objective function, which is known as its “personal” best position; a “global” best position representing the best position found by all particles across a swarm or group is also maintained (Chong & Zak, 2013). A particle swarm optimization task is run iteratively, until some stopping criterion is reached (Solnon, 2010, p. 80). Every iteration begins by updating the velocities of all particles. In doing so, a particle is acted upon by two forces: the attraction of the particle to its personal previous best known position, as governed by a “cognitive” parameter, and its attraction to the best known position within its group, as governed by a “social” parameter (Chong & Zak, 2013). The particle also maintains some momentum from its previous velocity. Following this, particle positions are updated in accordance with their velocities, with new positions being evaluated via the given objective function and best-known personal and global positions updated where appropriate. Initial particle velocities are assigned randomly; in doing so, exploration of the search space is encouraged, reducing the likelihood of the swarm as a whole becoming caught in a local extremum or local extrema of the objective function (Yang, 2014, p. 32; Solnon, 2010, p. 78, 81).

As noted by Yang (2014, p. 99), particle swarm optimization’s applicability to a large domain of problems paired with its straightforward conceptual groundings and implementation have spurred on its widespread use in a number of fields, as well as the development of numerous refinements and derivatives of the original algorithm. This paper uses the well-known inertial variant of particle swarm optimization, which offers a “noticeable improvement” in speedy convergence on good solutions as compared to standard PSO (Yang, 2014, p. 101). Further to this variant, for the purposes of this paper each particle’s position is updated with respect to the best-found solution within a predefined neighborhood or group of particles, as opposed to the global best-found solution, as presented in Solnon (2010, p. 79). When these groups do not intersect, this is simply equivalent to running a number of independent particle swarm optimization tasks equal to the number of groups.

### 2.1.2 Previous work

- Noto et al. (2013) “Agent-based Social Simulation Model for Analyzing Human Behaviors using Particle Swarm Optimization” (“Agent-based Social Simulation Model for Analyzing Human Behaviors using Particle Swarm Optimization”)
- Cheng et al. (2008) “A modified Particle Swarm Optimization-based human behavior modeling for emergency evacuation simulation system”
- Franken & Engelbrecht (2005) “Particle swarm optimization approaches to coevolve strategies for the iterated prisoner’s dilemma”

### 2.1.3 Parameters and equations

In the formulation of PSO employed in this paper, a particle  $i$  with position  $x_i$  has a velocity  $v_i$  at time  $t$  such that

$$v_i^t = \theta(t)v_i^{t-1} + \alpha\epsilon_1(x_{N(i)}^* - x_i^{t-1}) + \beta\epsilon_2(x_i^* - x_i^{t-1})$$

, where  $\theta$  is the inertial scheduling function,  $\alpha$  is the cognitive component,  $\beta$  is the social component,  $x_{N(i)}^*$  is the global best position known for  $i$ 's neighborhood or group  $N(i)$ ,  $x_i^*$  is  $i$ 's personal best known position, and  $\epsilon_1$  and  $\epsilon_2$  are randomly chosen values within  $(0.0, 1.0]$ . Although  $\theta$  can take many forms including, most commonly, a constant function (Yang, 2014, p. 101), for the purposes of this paper  $\theta$  is defined with respect to a base inertia  $\tau$  and inertial dampening factor  $\sigma$  such that

$$\theta(t) = \frac{\tau}{\sigma^t}$$

Finally, the position  $x$  of  $i$  at time  $t$  is defined as

$$x_i^t = x_i^{t-1} + v_i^t$$

## 2.2 Rohde et al. 2012

### 2.2.1 More detailed overview of method

### 2.2.2 More detailed overview of results

### 2.2.3 Game-theoretic model

## 3 Methods

### 3.1 Modeling language game as PSO task

#### 3.1.1 Mixed-strategy search space

#### 3.1.2 Particles are agents

#### 3.1.3 Criteria for “ambiguous coordination” and “unambiguous coordination”

#### 3.1.4 Model with particles restricted to search space vs. not

### 3.2 Optimizing parameters for PSO with PSO

#### 3.2.1 Evaluating error

## 4 Results

### 4.1 Results of meta-optimization (PSO parameters)

### 4.2 Results for model with search space restriction

- Rates at approximately correct levels, but reacts incorrectly to cost adjustments

### 4.3 Results for model without

- Ratios of rates track well onto data from experiments, but rates much lower

## 5 Discussion

## 6 Conclusion

## References

- Chong, E., & Zak, S. (2013). Particle swarm optimization. In *An introduction to optimization* (chap. Global search algorithms). Hoboken, New Jersey: John Wiley & Sons.
- Davis, L. (1991). *Handbook of genetic algorithms*. New York: Van Nostrand Reinhold.
- Engelbrecht, A. P. (2005). *Fundamentals of computational swarm intelligence*. John Wiley & Sons Ltd., England.
- Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, 336(6084), 998-998.
- Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of the second meeting of the north american chapter of the association for computational linguistics on language technologies* (p. 1-8). Stroudsburg, PA, USA: Association for Computational Linguistics. doi: 10.3115/1073336.1073357
- Heppner, F., & Grenander, U. (1990, December). A stochastic nonlinear model for coordinated bird flocks. In S. Krasner (Ed.), *The ubiquity of chaos* (p. 233-). Washington, D.C: American Association for the Advancement of Science.
- Kennedy, J. (1997, April). The particle swarm: social adaptation of knowledge. In *IEEE international conference on evolutionary computation, 1997.* (p. 303-308). doi: 10.1109/ICEC.1997.592326
- Kennedy, J., & Eberhart, R. (1995, November). Particle swarm optimization. In *IEEE international conference on neural networks, 1995. Proceedings.* (Vol. 4, p. 1942-1948). doi: 10.1109/ICNN.1995.488968
- Kirby, S., & Hurford, J. R. (2002). The emergence of linguistic structure: An overview of the iterated learning model. In A. Cangelosi & D. Parisi (Eds.), *Simulating the evolution of language* (p. 121-147). Springer London. doi: 10.1007/978-1-4471-0663-0\_6
- Liu, H., Xu, G., Ding, G., & Sun, Y. (2014, April). Human behavior-based particle swarm optimization. *The Scientific World Journal*. doi: 10.1155/2014/194706
- Rohde, H., Seyfarth, S., Clark, B., Jäger, G., & Kaufmann, S. (2012, September). Communicating with cost-based implicature: A game-theoretic approach to ambiguity. In *The 16th workshop on the semantics and pragmatics of dialogue*. Paris.

- Solnon, C. (2010). *Ant colony optimization and constraint programming*. London Hoboken, NJ: Wiley Online Library.
- Yang, X.-S. (2014). Chapter 7 - Particle swarm optimization. In *Nature-inspired optimization algorithms* (p. 99-110). Oxford: Elsevier. doi: 10.1016/B978-0-12-416743-8.00007-5