

# Convergence of symbolic regression using metaheuristics

Ben Cardoen  
ben.cardoen@student.uantwerpen.be

April 29, 2017

## 1 Abstract

Symbolic regression (SR) fits a symbolic expression to a set of expected values. Amongst its advantages over other techniques is the ability for a practitioner to interpret the resulting expression, determine important features by their usage in the expression, and an insight into the behavior of the resulting model (e.g. continuity, derivatives, extrema). SR combines a discrete combinatoric problem (combining base functions) with a continuous optimization problem (selecting and mutating constants). One of the main algorithms used in SR is Genetic Programming (GP). The convergence characteristics of SR using GP is still an open issue. The continuous aspect of the problem has traditionally been an issue in GP based symbolic regression. This paper will study convergence of a GP-SR implementation on selected use cases known for bad convergence. We introduce modifications to the classical mutation and crossover operators and observe their effects on convergence. The constant optimization problem is studied using a two phase approach. We apply a variation on constant folding in the GP algorithm and evaluate its effects. The hybridization of GP with 3 metaheuristics (Differential Evolution, Artificial Bee Colony, Particle Swarm Optimization) is studied.

- 2 Introduction**
- 3 Design**
- 4 Experiments**
- 5 Related Work**
- 6 Conclusion**
- 7 Future work**