Predict the Success or Failure of an Evolutionary Algorithm

Gopinath Chennupati, Conor Ryan, R. Muhammad Atif Azad

Gopinath.Chennupati@ul.ie, Conor.Ryan@ul.ie, Atif.Azad@ul.ie

Introduction

- Grammatical Evolution (GE) [3] is an evolutionary algorithm (EA) that evolves computer programs in any language through CFGs represented in BNF grammars.
- Since EAs are computationally demanding, sensible utilisation of computational resources, especially when problems scale up, is essential.

Aim

This paper proposes to identify and cull the evolutionary runs that are unlikely to produce solutions of acceptable quality.

- Use Ant Colony Optimization (ACO) [1] to predict the failure of a GE run.
- Analyse and hand-tune the ACO produced predictive models.
- Then, use the predictive model to terminate potentially poor runs very early (10 generations).

Conclusion

We improved the solution quality of GE runs using a completely novel prediction approach. This also has significantly reduced the time spent on executing GE. We will focus on further improving the prediction through rapid retraining.

References

[1] M. Medland and F. E. B. Otero. A study of different quality evaluation functions in the cant-miner(pb)
classification algorithm. In Genetic and Evolutionary Computation Conference, pages 49–56, 2012.

[2] M. Keijzer. Improving symbolic regression with interval arithmetic and linear scaling. In C. Ryan et at., Genetic Programming, volume 2610 of LNCS, pages70–82. Springer Berlin, Heidelberg, 2003.

[3] C. Ryan, J. J. Collins, and M. O'Neill. Grammatical evolution: Evolving programs for an arbitrary language. In W. Banzhaf et al., Proceedings of the First European Workshop on Genetic Programming, Vol 1391 of Springer LNCS, pp. 83-95, 1998.

Acknowledgements

This work has been supported by the Science Foundation of Ireland (SFI). Grant No. I O/IN. I/I303l.

The Run Prediction Model (RPM)

- The Run Prediction Model (RPM) + GE armed with a rule based model that predicts the quality of each run.
- *Figure 1* presents the block diagram of RPM+GE.
- The predictive model notes the changes in 4 parameters over first 10 generations: best fitness (BFC), average fitness (AFC), average actual length (AALC), average effective length (AELC).
- We subjectively select a different threshold for each problem as a measure of acceptable quality.
- The predictive model then judges whether a GE run can cross that threshold; hence, a binary classification of each run.

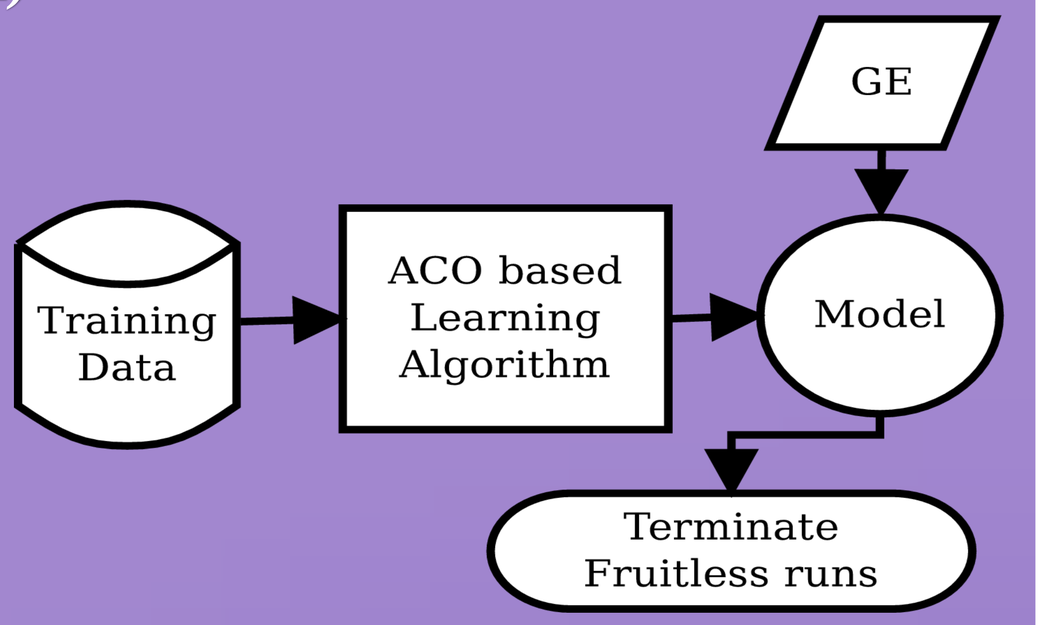

Figure 1: Block diagram of the Run Prediction Model (RPM) applied GE.

Results

Problem	Class		Threshold
	No	Yes	(BF <)
$f_1 = (1 + x)^3$	590	410	0.7
$f_2 = x^4 - x^3 - y^2 - y$	606	394	0.65
$f_3 = x^3 - y^3 - y - x$	848	152	0.7
$f_4 = x^y$	518	482	1

Table 1: Experimental problems and the classification of the data sets.

- *Table 1* presents the experimental problems and the data sets used.
- Data sets are classified based on a predefined success threshold.
- Defined a separate success threshold for each and every problem as the solution producing ability of GE is unique on each problem.
- Greatly improved the solution quality rather the computational effort is high for data set preparation.

- Run Prediction Model (RPM) follows the traditional design of GE except for a small change at the run execution level.
- *Figure 1* presents the block diagram of RPM.
- Model takes training data prepared from 1000 GE runs that contains 4 attributes best fitness change (BFC), average fitness change (AFC), average actual length change (AALC), average effective length change (AELC).
- This is a binary classification problem as each data point is classified based on a separate threshold value for each problem.
- ACO algorithm predicts a rule based model that takes GE as input and terminates the poor quality runs.

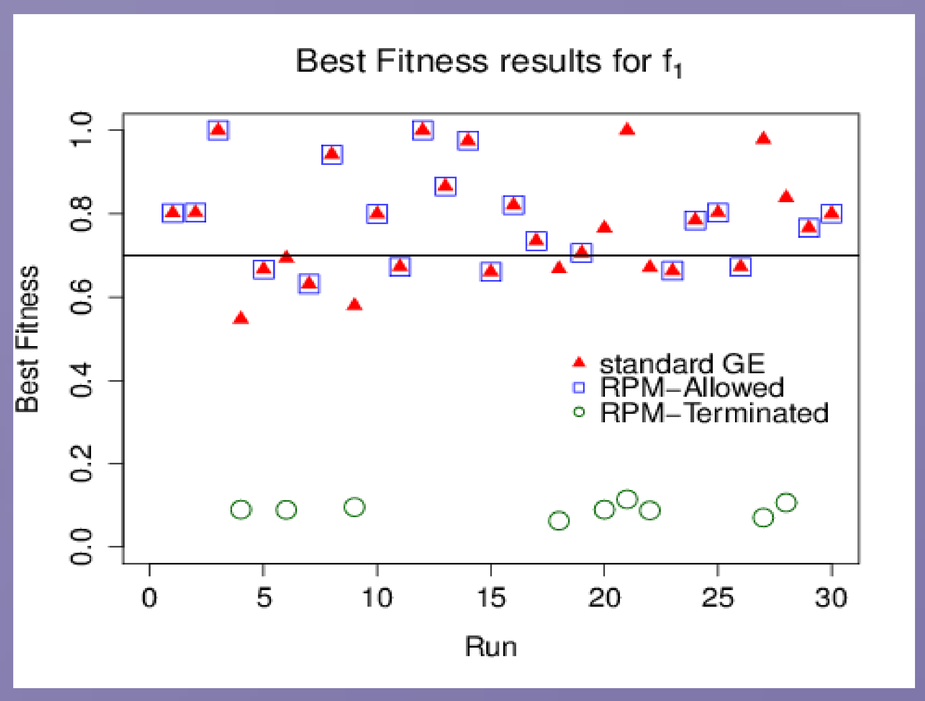

Figure 2: Comparison of end of run fitness results both in standard and RPM+GE.

Problem	GE	RPM+GE	
	Fruitful Runs	Fruitful Runs	Terminated Runs
f_1	18	15	9
f_2	7	3	23
f_3	6	5	9
f_4	5	12	13

Table 2: Success rate of both the approaches.

- *Table 2* compares the standard GE results with RPM applied GE across 30 independent runs.
- The results also contain the information regarding the number of discontinued runs.
- Wilcoxon Signed Rank tests show a significant improvement in the number of successful runs at $\alpha = 0.05$.
- RPM+GE system exerts certain amount of error in predicting the performance of runs.
- Ex:- f_3 contains the small amount of improvement in the solution.

BDS Group, Computer Science and Information Systems (CSIS) Department, University of Limerick, Ireland.