Guarantees of Progress for Geometric Semantic Genetic Programming

Tomasz P. Pawlak, Krzysztof Krawiec*

1 Introduction

In this study, we succintly rephrase the basic theory on geometric semantic genetic programming [1], and propose definitions and discuss properties of guarantees of progress for search operators. In the second part of the paper we show how these definitions apply to the geometric operators.

Definition 1. *Semantics* $s(p) \in S$ of a program p is a n-tuple of values obtained by applying p to a given set of n inputs. We assume that S is a metric space endowed with metric d.

Since S incorporates all possible semantics , we call it *semantic space*. We identify a semantics with a description of program behavior, which, for particular program p, we denote as $s(p)$. Thus there is a projection $P \to S$, such that $\forall p \in P, \exists s(p) \in S$ and $\forall p_1, p_2 \in P : p_1 = p_2 \Rightarrow s(p_1) = s(p_2)$, however in general there may exist a semantics in S having no corresponding program in P, and distinct programs may have equal semanticsl.

We distinguish in S a *target* semantics $t \in S$, which represents the desirable behavior of a program. The target implicitly defines a *programming task*, the objective of which is to construct a program p^* such that $s(p^*) = t$. To assess programs, we define a *fitness function* $f(p) = d(t, s(p))$, where d is a metric, i.e., f measures the divergence of program's semantics $s(p)$ from the target semantics t, and for the optimal program $f(p^*) = 0$ by definition.

An n-ary *search operator* is a function $o : P \times \ldots \times P \to P$. The arguments of a search operator are called parents of the resulting program. After [1], we define geometric recombination operators under metric d:

Definition 2. *Geometric crossover* is a binary ($n = 2$) search operator such that all offspring it produces are in the d-metric segment between semantics of its parents.

Definition 3. *Geometric ϵ-mutation* is an unary ($n = 1$) search operator such that all offspring it produces is are the d-ball of radius ϵ centered in the semantics of the parent.

*The authors are with Institute of Computing Science, Poznan University of Technology, e-mails: {tpawlak, krawiec}@cs.put.poznan.pl. Work supported by Polish National Science Centre, T. Pawlak grant no. DEC-2012/07/N/ST6/03066 and K. Krawiec grant no. DEC-2011/01/B/ST6/07318.

2 Guarantees of progress

Let $I = I(o, p_1, p_2, \ldots, p_n)$ denote the set of all *potential offspring* that can be produced by applying a search operator o to the parents $p_1, p_2, \ldots, p_n$ (due to o's internal stochasticity, there is typically more than one such offspring).

Definition 4. An operator has *weak guarantee of progress* if all the produced offspring are not worse the worst of its parents, i.e., $\forall p \in I, \forall i \in [1, n] : f(p) \geqslant f(p_i)$.

In other words, an operator having weak guarantee of progress must not deteriorate fitness. If it is the only search operator, the worst fitness of the next population cannot be worse than the worst in the current one. However, such operator does not guarantee producing a better solution.

Definition 5. An operator is *potentially progressive* if for every set of parents, there exists a potential offspring that is not worse than the best of its parents, i.e., $\exists p \in I : f(p) \geqslant \max_{i \in [1,n]} f(p_i)$.

In practice, an operator should be considered potentially progressive if it has a non-zero probability of improvement w.r.t. the parents.

Definition 6. An operator has *strong guarantee of progress* if all the produced offspring is not worse than the best of its parents, i.e., $\forall p \in I : f(p) \geqslant \max_{i \in [1,n]} f(p_i)$

Strong guarantee of progress implies weak guarantee of progress and potential progress. Note however that potential progress does not imply weak guarantee of progress.

We formulate and prove the following theorems:

Theorem 1. *Geometric crossover under L_1 metric is potentially progressive, and under L_2 metric is potentially progressive and has weak guarantees of progress.*

Theorem 2. *Geometric ϵ-mutation under L_1 and L_2 metrics is potentially progressive.*

In the workshop presentation, we will present the above formal framework, prove Theorems 1 and 2, and discuss the practical implications of and nuances related to these observations.

References

[1] Alberto Moraglio, Krzysztof Krawiec, and Colin G. Johnson. Geometric semantic genetic programming. In *PPSN XII*, volume 7491 of *LNCS*, pages 21–31. Springer, 2012.