Guarantees of Progress for Geometric Semantic Genetic Programming

Tomasz P. Pawlak
tpawlak@cs.put.poznan.pl

Krzysztof Krawiec
krawiec@cs.put.poznan.pl

Institute of Computing Science
Poznan University of Technology

Parallel Problem Solving from Nature XIII, 2014

Outline

Motivation

- What are the limitations of GSGP
  - What can we expect?
  - What will never happen?
- Which metrics are better for GSGP operators?

Preliminary definitions

Definition

Semantics $s \in S$ is a tuple of n elements corresponding to inputs.

Hence:

- semantics is description of program behavior $\rightarrow s(p)$
- $S \equiv D^n$, where D is the codomain (type) of output values produced by the programs [in the considered programming language]

Assume:

- each program $p \in P$ has semantics $s(p)$
- there may exist semantics in S without counterpart in program set P

Preliminary definitions

Definition

Semantics $s \in S$ is a tuple of n elements corresponding to inputs.

Hence:

- semantics is description of program behavior $\rightarrow s(p)$
- $S \equiv D^n$, where D is the codomain (type) of output values produced by the programs [in the considered programming language]

Assume:

- each program $p \in P$ has semantics $s(p)$
- there may exist semantics in S without counterpart in program set P

Preliminary definitions

Definition

Semantics $s \in S$ is a tuple of n elements corresponding to inputs.

Hence:

- semantics is description of program behavior $\rightarrow s(p)$
- $S \equiv D^n$, where D is the codomain (type) of output values produced by the programs [in the considered programming language]

Assume:

- each program $p \in P$ has semantics $s(p)$
- there may exist semantics in S without counterpart in program set P

Preliminary definitions

Definition

Target semantics $t \in S$ is semantics representing the desirable behavior a program.

Definition

Programming task is a task with objective to create program $p^* : s(p^*) = t$.

Definition

Fitness function $f(p) = d(t, s(p))$, where $d(\cdot,\cdot)$ is a metric.

Hence:

- $f(\cdot)$ measures divergence from target t
- $f(p^*) = 0$

Preliminary definitions

Definition

Target semantics $t \in S$ is semantics representing the desirable behavior a program.

Definition

Programming task is a task with objective to create program $p^* : s(p^*) = t$.

Definition

Fitness function $f(p) = d(t, s(p))$, where $d(\cdot, \cdot)$ is a metric.

Hence:

- $f(\cdot)$ measures divergence from target t
- $f(p^*) = 0$

Preliminary definitions

Definition

Target semantics $t \in S$ is semantics representing the desirable behavior a program.

Definition

Programming task is a task with objective to create program $p^* : s(p^*) = t$.

Definition

Fitness function $f(p) = d(t, s(p))$, where $d(\cdot, \cdot)$ is a metric.

Hence:

- $f(\cdot)$ measures divergence from target t
- $f(p^*) = 0$

Preliminary definitions

Definition

Target semantics $t \in S$ is semantics representing the desirable behavior a program.

Definition

Programming task is a task with objective to create program $p^* : s(p^*) = t$.

Definition

Fitness function $f(p) = d(t, s(p))$, where $d(\cdot, \cdot)$ is a metric.

Hence:

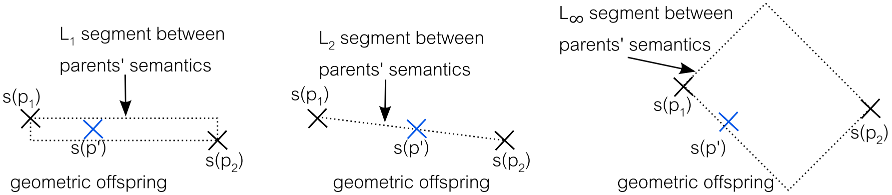
- $f(\cdot)$ measures divergence from target t
- $f(p^*) = 0$

Geometric Semantic Genetic Programming

[A. Moraglio, K. Krawiec, C. Johnson, 2012]

Definition

Geometric crossover is binary operator that produces all offspring in the *d*-metric segment connecting semantics of its parents



Implications

- Fitness landscape is the graph of fitness function when plotted for the solutions arranged according to the neighborhood structure induced by a search operator.
- **Key observation:** In GSGP, that spatial arrangement is consistent with the adopted metric.

Consequences:

- Exactly one optimum – at the target
- For any program p, elevation on the fitness landscape at $s(p)$ is the same as its distance to the target semantics in S

Implications

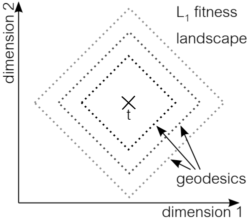
- Fitness landscape is the graph of fitness function when plotted for the solutions arranged according to the neighborhood structure induced by a search operator.
- **Key observation:** In GSGP, that spatial arrangement is consistent with the adopted metric.

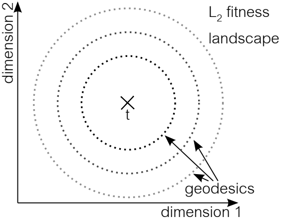
Consequences:

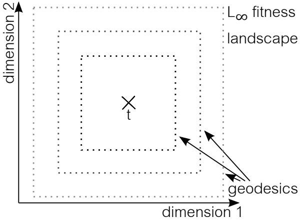
- Exactly one optimum – at the target
- For any program p, elevation on the fitness landscape at $s(p)$ is the same as its distance to the target semantics in S
  - $f(p) = f(t, s(p))$

Shape of fitness landscape

Fitness landscape is d-metric cone with the target in apex







Weak guarantee of progress

Definition

An operator has *weak guarantee of progress (WGP)* if all the produced offspring is not worse than the worst of its parents

Hence:

- There is no guarantee that the operator having WGP produce a *strictly better* solution

Weak guarantee of progress

Definition

An operator has *weak guarantee of progress (WGP)* if all the produced offspring is not worse than the worst of its parents

Hence:

-
-
- There is no guarantee that the operator having WGP produce a *strictly better* solution

Incomplete guarantee of progress

Definition

An operator has *incomplete guarantee of progress (IGP)* if for every pair of parents, there exists a produced offspring that is not worse than the best of its parents

Strong guarantee of progress

Definition

An operator has *strong guarantee of progress (SGP)* if all the produced offspring is not worse than the best of its parents

Hence:

- The worst fitness in the next population must be not worse than the best fitness of individuals chosen for recombination in current population
- Operator having SGP has also IGP and WGP

Strong guarantee of progress

Definition

An operator has *strong guarantee of progress (SGP)* if all the produced offspring is not worse than the best of its parents

Hence:

- The worst fitness in the next population must be not worse than the best fitness of individuals chosen for recombination in current population
- Operator having SGP has also IGP and WGP

Properties of GSGP crossover

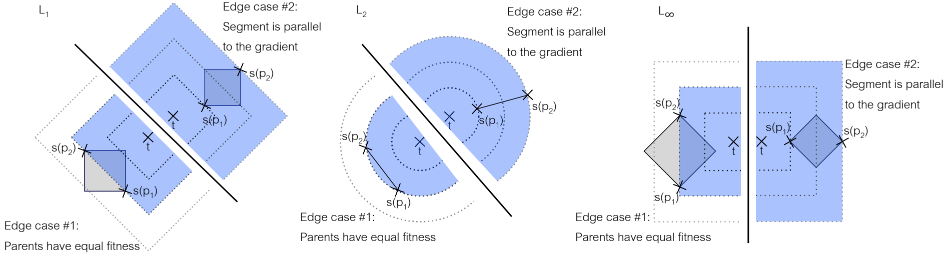
Metric	WGP	IGP	SGP
L_1[1]	×	✓	×
L_2	✓	✓	×
L_∞	×	✓	×

[1] Also applies to Hamming metric

'Visual' proofs for WGP

Consider two parents $p_1, p_2 : s(p_1) \neq s(p_2)$ and the segment connecting their semantics. There are two edge cases:

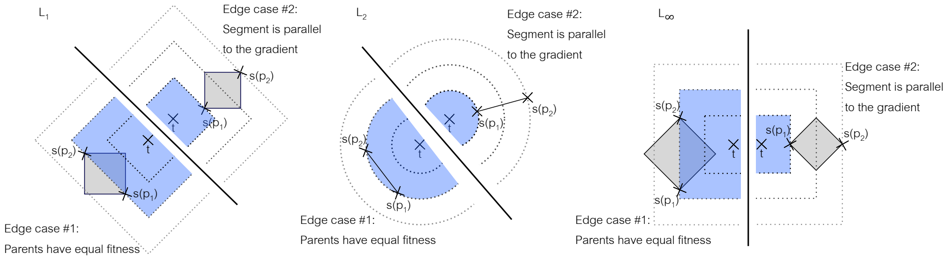
Metric	Case 1: $f(p_1) = f(p_2)$	Case 2: $\nabla f \parallel \overline{s(p_1)s(p_2)}$	Conclusion
L_1	Part of segment has fitness $> f(p_1)$	Entire segment has fitness $< f(p_2)$	No WGP
L_2	Entire segment has fitness $\leq f(p_1)$	Entire segment has fitness $\leq f(p_2)$	WGP
L_∞	Part of segment has fitness $> f(p_1)$	Entire segment has fitness $\leq f(p_2)$	No WGP



‘Visual’ proofs for IGP

Consider two parents $p_1, p_2 : s(p_1) \neq s(p_2)$ and the segment connecting their semantics. There are two edge cases:

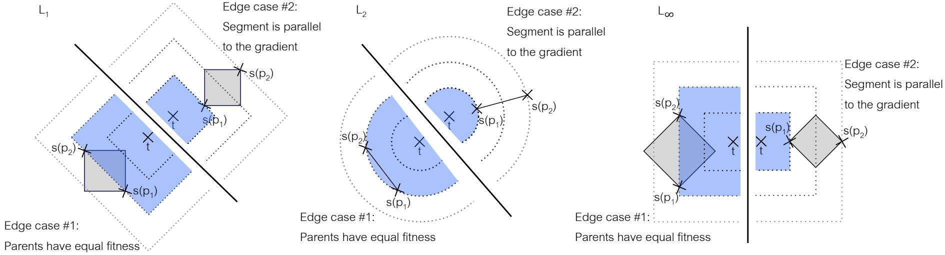
Metric	Case 1: $f(p_1) = f(p_2)$	Case 2: $\nabla f \parallel \overline{s(p_1)s(p_2)}$	Concl.
L_1	Part of segment has fitness $\leq f(p_1)$	Exactly one point has fitness $= f(p_1)$	IGP
L_2	Entire segment has fitness $\leq f(p_1)$	Exactly one point has fitness $= f(p_1)$	IGP
L_∞	Part of segment has fitness $\leq f(p_1)$	Exactly one point has fitness $= f(p_1)$	IGP



'Visual' proofs for SGP

Consider two parents $p_1, p_2 : s(p_1) \neq s(p_2)$ and the segment connecting their semantics. There are two edge cases:

Metric	Case 1: $f(p_1) = f(p_2)$	Case 2: $\nabla f \parallel \overline{s(p_1)s(p_2)}$	Conclusion
L_1	Part of segment has fitness $> f(p_1)$	Entire segment has fitness $\geq f(p_1)$	No SGP
L_2	Entire segment has fitness $\leq f(p_1)$	Entire segment has fitness $\geq f(p_1)$	No SGP
L_∞	Part of segment has fitness $> f(p_1)$	Entire segment has fitness $\geq f(p_1)$	No SGP



How to design a crossover with SGP?

- Always choose the best offspring candidate in the segment between parents
- Or...

How to design a crossover with SGP?

- Always choose the best offspring candidate in the segment between parents

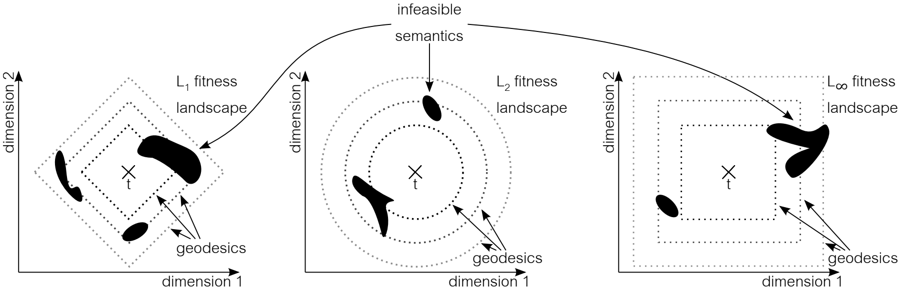
- Go outside the segment!
- Extrapolate
- See [**?**] for example

One more thing: The true shape of fitness landscape

- The presented landscapes stretch across semantic space S
- However the space being searched is program space!
- By excluding from S infeasible semantics in the given programming language the fitness landscape may feature holes

One more thing: The true shape of fitness landscape

- The presented landscapes stretch across semantic space S
- However the space being searched is program space!
- By excluding from S infeasible semantics in the given programming language the fitness landscape may feature holes



Conclusions

- Defined types of guarantees of progress for GSGP
- Guarantees verified for GSGP crossover
- Constructive observations for designing new search operators

Bibliography

Analytical proof for L_2, WGP and crossover

Let p_1, p_2 be parents, p' be their offspring, l be dimensionality of search space, $t = (t_i)_{i=1..l}$, $s(p_1) = (s_{1i})_{i=1..l}$, $s(p_2) = (s_{2i})_{i=1..l}$, $s(p') = \alpha s(p_1) + (1-\alpha)s(p_2) = (\alpha s_{1i} + (1-\alpha)s_{2i})_{i=1..l}$, i.e., the semantics of the offspring is linear combination of semantics of its parents. Hence the fitness of offspring is

$$f(p') = d(t, s(p')) = \sqrt{\sum_{i=1..l} (\alpha s_{1i} + (1-\alpha)s_{2i} - t_i)^2}$$

Then

$$\frac{\partial f(p')}{\partial \alpha} = \frac{\sum_{i=1}^{l} (\alpha s_{1i} + (1-\alpha)s_{2i} - t_i)(s_{1i} - s_{2i})}{\sqrt{\sum_{i=1}^{l} (\alpha s_{1i} + (1-\alpha)s_{2i} - t_i)^2}}$$

$$\frac{\partial^2 f(p')}{\partial \alpha^2} = \frac{\sum_{i=1}^{l} (s_{1i} - s_{2i})^2}{\sqrt{\sum_{i=1}^{l} (\alpha s_{1i} + (1-\alpha)s_{2i} - t_i)^2}} - \frac{(\sum_{i=1}^{l} (\alpha s_{1i} + (1-\alpha)s_{2i} - t_i)(s_{1i} - s_{2i}))^2}{(\sum_{i=1}^{l} (\alpha s_{1i} + (1-\alpha)s_{2i} - t_i)^2)^{3/2}}$$

Let

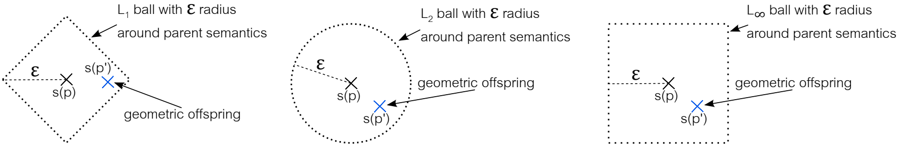
$$MAX \equiv \begin{cases} \frac{\partial f(p')}{\partial \alpha} = 0 \\ \frac{\partial^2 f(p')}{\partial \alpha^2} < 0 \end{cases} \qquad MIN \equiv \begin{cases} \frac{\partial f(p')}{\partial \alpha} = 0 \\ \frac{\partial^2 f(p')}{\partial \alpha^2} > 0 \end{cases}$$

Since MAX has no solution, $f(p')$ has no maximum and since MIN has exactly one solution $\alpha^* = -(\sum_{i=1}^{l} s_{1i}s_{2i} - s_{2i}^2 - (s_{1i} - s_{2i})t_i) / \sum_{i=1}^{l} (s_{1i} - s_{2i})^2$, there is one minimum of $f(p')$ at α^*. Thus the line through points of parents' semantics is split by point $s^* = \alpha^* s(p_1) + (1-\alpha^*)s(p_2)$ into two monotonously increasing parts w.r.t. $f(p')$. By definition of geometric crossover $\alpha \in [0,1]$, since it guarantees that $s(p)$ is in L_2 segment $\overline{s(p_1)s(p_2)}$. Two cases occur: (i) $0 \le \{\alpha, \alpha^*\} \le 1$ or (ii) $\alpha^* < 0 \le \alpha \vee \alpha \le 1 < \alpha^*$. In the former one s^* lies in the segment $\overline{s(p_1)s(p_2)}$, thus due to monotonicity there is an induced order between parents' and offspring's fitness: $f(s^*) \le f(p) \le f(p_1)$ or $f(s^*) \le f(p') \le f(p_2)$ depending on which ray $s(p')$ belongs to, i.e., $s(p') \in \overrightarrow{s^* s(p_1)}$ or $s(p') \in \overrightarrow{s^* s(p_2)}$, respectively. For (ii) the order is $f(s^*) \le f(p_1) \le f(p') \le f(p_2)$ or $f(p_1) \ge f(p') \ge f(p_2) \ge f(s^*)$ depending on the relation between parents.

Geometric Semantic Mutation

Definition

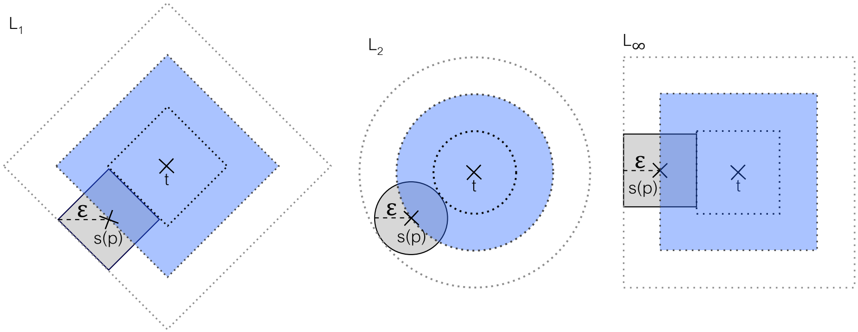
Geometric ϵ-mutation is unary operator that produces all offspring in the d-metric ball of ϵ radius centered in the parent semantics.



'Visual' proofs for WGP and SGP and mutation

Consider a parent p and a ball centered in its semantics.

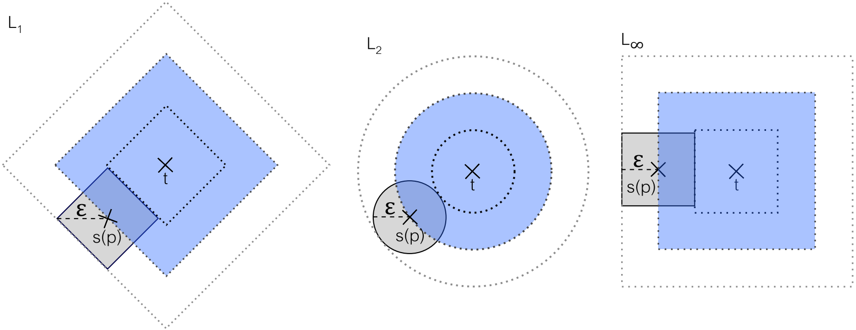
Metric		Conclusion
L_1	Part of ball has fitness $> f(p)$	No WGP, no SGP
L_2	Part of ball has fitness $> f(p)$	No WGP, no SGP
L_∞	Part of ball has fitness $> f(p)$	No WGP, no SGP



'Visual' proofs for IGP and mutation

Consider a parent p and a ball centered in its semantics.

Metric		Conclusion
L_1	Part of ball has fitness $\leq f(p)$	IGP
L_2	Part of ball has fitness $\leq f(p)$	IGP
L_∞	Part of ball has fitness $\leq f(p)$	IGP



How to create mutation having SGP?

- Always choose the best offspring candidate in the ball centered in parent
- Or...

How to create mutation having SGP?

- Always choose the best offspring candidate in the ball centered in parent
- Create offspring in the ball centered in target with radius equal to parent fitness $f(p)$