ISSN 2224-087X. Електроніка та інформаційні технології. 2024. Випуск 28. С. 38–45
Electronics and information technologies. 2024. Issue 28. P. 38–45

UDC 004.9 DOI: https://doi.org/10.30970/eli.28.4

THE SPECIATION IN GENETIC ALGORITHMS FOR PRESERVING POPULATION DIVERSITY AND OPTIMIZATION OF FUNCTIONS WITH SUBOPTIMAL SOLUTIONS

V. Pretsel, R. Shuvar

Department of System Design,
Ivan Franko National University of Lviv
50 Drahomanova St., UA-79005 Lviv, Ukraine
vitalii.pretsel@lnu.edu.ua , *roman.shuvar@lnu.edu.ua*

Background. Genetic algorithms are used in various tasks and show advantages compared to other optimization methods, which may not always be applicable. However, there are cases when they can't provide the best possible solution. One of them is the premature convergence of the population to a single solution. As population diversity decreases, the search domain becomes limited, and potential solutions may be overlooked. This scenario is particularly common in multimodal functions where multiple local optima exist. To address this, a class of techniques known as niching techniques has been developed. These methods preserve population diversity and prevent premature convergence to suboptimal solutions.

Materials and Methods. In this work we investigate the method of speciation and how it helps to find solutions for given tasks. For this, several experiments were conducted, in which the highest value of the function was found in the given interval. It was compared how results of the optimization differed if we used speciation and didn't. To assess the diversity of speciation, the values of the average fitness of the population and the standard deviation of the values of the individuals in the population were compared. We also evaluated how speciation helps with optimization for tasks with suboptimal solutions, comparing how many successful solutions were obtained in experiments with and without speciation.

Results and Discussion. The results show that the speciation method preserves population diversity and improves optimization outcomes for multimodal functions. In the experiments where speciation was applied, the population maintained a higher level of diversity, as indicated by a larger standard deviation in population individuals' values. It resulted in increasing of the number of successful solutions in tasks with multiple local optima.

Conclusion. Speciation effectively preserves population diversity and helps to avoid premature convergence in genetic algorithms. This leads to better optimization results, especially in tasks with multiple local optima. This highlights the importance of diversity-preserving techniques, such as speciation, in addressing the limitations of genetic algorithms, especially in complex optimization tasks.

Keywords: genetic algorithms, optimization, niching techniques, speciation

Introduction

Genetic algorithms (GA) have a wide spread of use in optimization tasks and competes with other optimization techniques, sometimes showing better results in terms of efficiency and also is capable to be used in specific cases where other methods are not applicable. That said, GA also have a range of problems that come from unique features of algorithm itself. That forces us to find configuration appropriate for specific tasks or even invent new modifications and techniques.

One of common problems with using GA is converging to single solution, that makes population through evolution process less diverse. This behavior is called a genetic drift. Less diverse population have negative impact on GA [1]. Although we can often receive solutions of high quality, it also could be a crucial problem for tasks with suboptimal solutions, that leads GA to not being able to find best solution. Diverse population also provides following advantages: global exploration of the search space, facilitating crossover, diverse set of solutions for decision making, robustness.

To maintain population diversity many solutions were proposed. They include preserving diversity through balancing exploitation and exploration via parameter tuning and careful designing of selection mechanisms. Other techniques control diversity explicitly with mechanisms embedded to evolutionary algorithms, such as eliminating duplicates and niching techniques [2], that maintain population diversity based on the distance between the population members.

While niching techniques already were investigated in other works [3-5], in this work we will cover speciation technique, a method for promoting diversity by grouping similar individuals into subpopulations or species, and compare how results improved based on desirable solutions of defined optimization tasks and diversity of population. To measure the dissimilarity between individuals, we will employ the Hamming distance, a well-established metric for quantifying differences in binary representations.

Materials and methods.

Algorithm of speciation.

The speciation technique involves dividing a population into subgroups (species) based on genetic similarity. To implement it few additional steps are embedded into the genetic algorithm. After speciation is performed, we proceed with all remaining steps of genetic algorithm such as selection, mutation and crossover.

Determining similarity metric. A similarity (or distance) metric is selected to measure how genetically different two individuals are. Most common metrics include:

- Hamming distance: used for binary-encoded genomes. It counts the number of differing bits;
- Euclidean distance: used for real-valued vectors. It measures the straight-line distance between two individuals in multi-dimensional space.

Depending on the genome encoding, other metrics such as Manhattan distance or cosine similarity can be used.

Defining speciation threshold. Speciation threshold δ_t determines the maximum distance between two individuals that allows them to belong to the same species. The threshold is a critical parameter that controls how similar two individuals must be to be placed in the same species. A smaller threshold leads to more species (higher diversity), while a larger threshold may result in fewer species and faster convergence.

Assigning individuals to species. Each individual is assigned to a species based on their genetic similarity to the others. For each individual, we compare it with the picked species' representative. If the genetic distance between the individual and the representative of any existing species is less than the speciation threshold δ_t, it is added to that species. If no species meets the threshold, the individual forms a new species and becomes the representative for that species.

At the end of this process, each individual is grouped into a species where all members are genetically similar within the bounds of the threshold.

Fitness Sharing. For each individual, its fitness score is adjusted based on the number of individuals in its species. The adjusted fitness is computed by dividing the individual's raw fitness by a factor that reflects how crowded its species is. This adjustment reduces the reproductive advantage of individuals in large species, promoting the survival of individuals in smaller or less populated species.

Experiment configuration.

We will perform a series of experiments to see how applying speciation changes the optimization results and its process at all. First we will perform test on function with one optimal solution to compare how GA with and without speciation will behave. Second we will have function with few suboptimal solutions. One solution will be in short range of values, which should complicate the task and population diversity will play a crucial role in finding best solution.

For GA we will use classic model with selection, mutation and crossover. Our chromosomes will be represented as bit strings in Gray code, so we will have following operators: tournament selection with 3 individuals per group, uniform crossover, bit flip mutation. The other configurations are: population size 50, mutation rate 0.05, crossover rate 0.5. We also save 2 best individuals in each generation.

For speciation we calculate distance as Hamming distance between individual and species mascot, which is randomly picked individual from a species. If distance is less than specified threshold $\delta_t = 0.3$ individual belongs to same species as mascot. After all individuals in population are distributed between different species, we calculate their new fitness values as their original fitness divided by number of individuals in species they belong to. Thus, individuals which are more unique will have higher chance to survive and as a result we should receive new generation with higher diversity.

To be able to perform comparisons we will use mean and best fitness values in population and standard deviation of individual values.

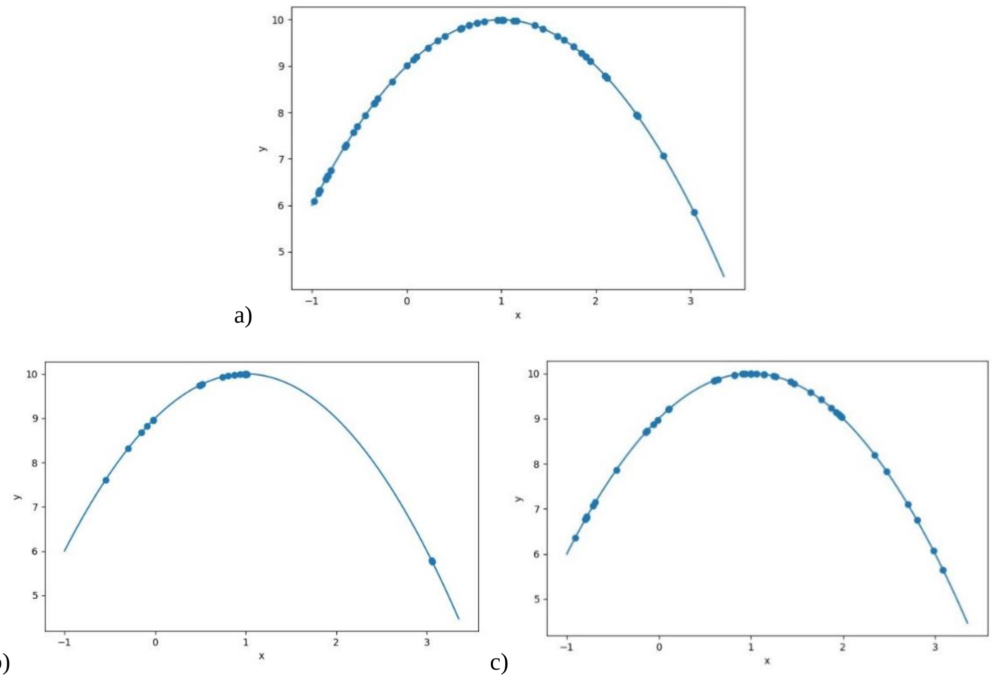
Results and discussion.

We were able to receive desired results. Although we think that results may differ if we test these methods on other tasks and applying another configuration may also change the outcome, testing those methods with different parameters or operators didn’t affect the main conclusion.

Investigation of population diversity without and with speciation. In this experiment we will try to find maximum value for $f(x) = -x^2 + 2x + 9$ on range $[-1;\ 2.095]$ with step 0.001. In this experiment we will run 100 generations for each test. Same initially generated random population will be used for finding solution both for method without and with speciation.

At first let’s take a look in Figure 1 to see how individuals are distributed before and after optimization process. As we can see initially individuals are evenly distributed, while after optimization we see difference between methods. In first case where we didn’t use speciation most of individuals converged to one point, which is the individual with highest fitness value. In second case where speciation was applied individuals are distributed in entire range more evenly. This is also can be seen on population fitness dynamic throughout evolving: mean fitness was closer to best fitness if speciation wasn’t applied (Figure 2).

Gathering standard deviation of individuals values from population at the generation with best found fitness and performing 100 experiment gives us 0.57 mean standard deviation for GA without speciation and 0.96 mean standard deviation for GA with speciation. This confirms that speciation helps to maintain diversity in population.



a) b) c)

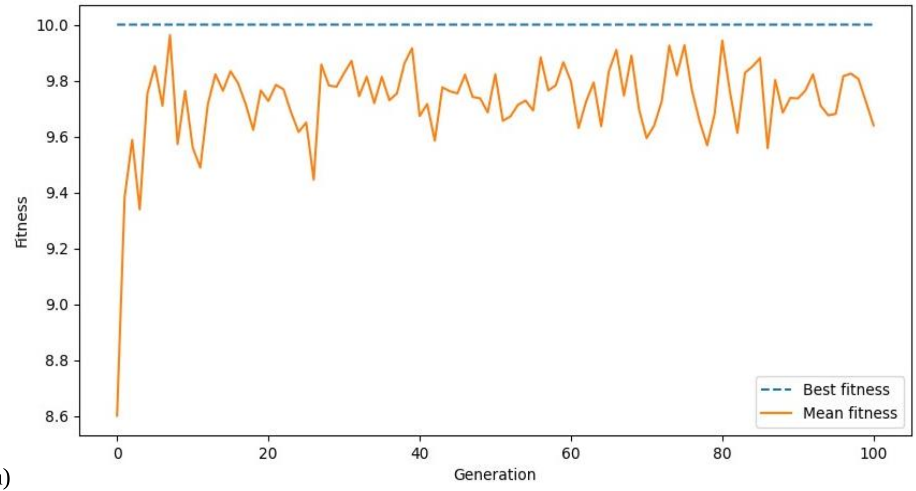
Fig. 1. Population depicted on investigated function with one optimal solution: a) initial population b) last generation for GA without speciation c) last generation for GA with speciation

***Applying speciation to find optimal solution in multimodal function*.** In this experiment function will have few local maximum values:

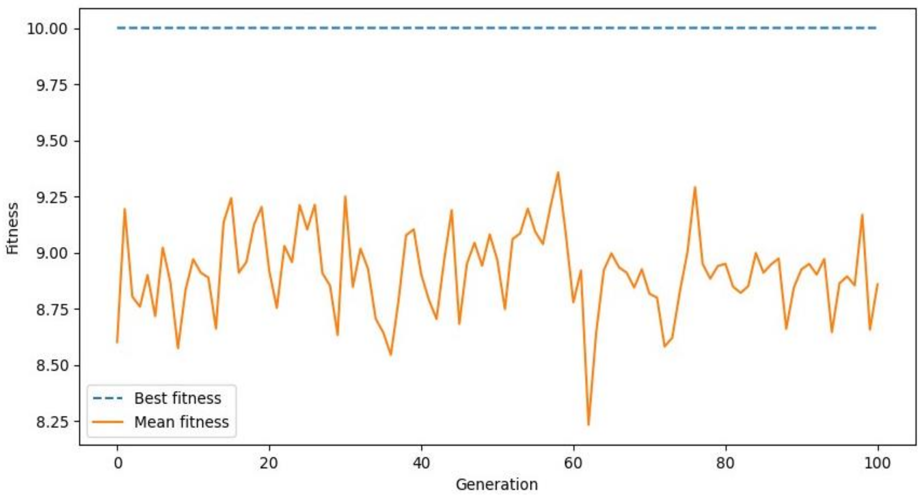
$$f(x) = \begin{cases} 1.5, & 15 \le x \le 15.01 \\ \dfrac{\left(sin(x) * x + \dfrac{(x ** 2 - 10 * x)}{5} + 10\right)}{10}, & x < 15 \; and \; x > 15.01 \end{cases}$$

on range $[0; 32.767]$. In this experiment we will run 400 generations for each test. Same as before same initially generated population will be used for finding solution.

In Figure 3 we can see that evolving process is similar to previous example, but in this experiment finding best solution required more generations for GA with speciation, GA without speciation wasn't able to find best solution before it converged to one solution.

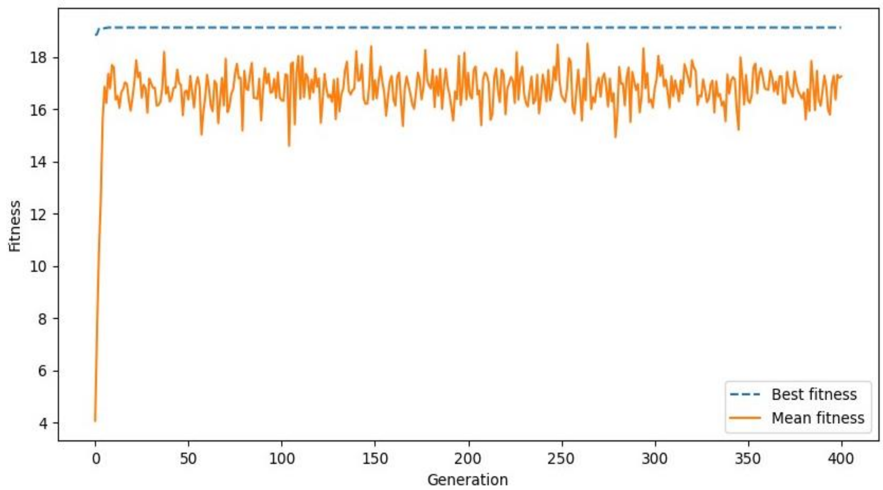


a)

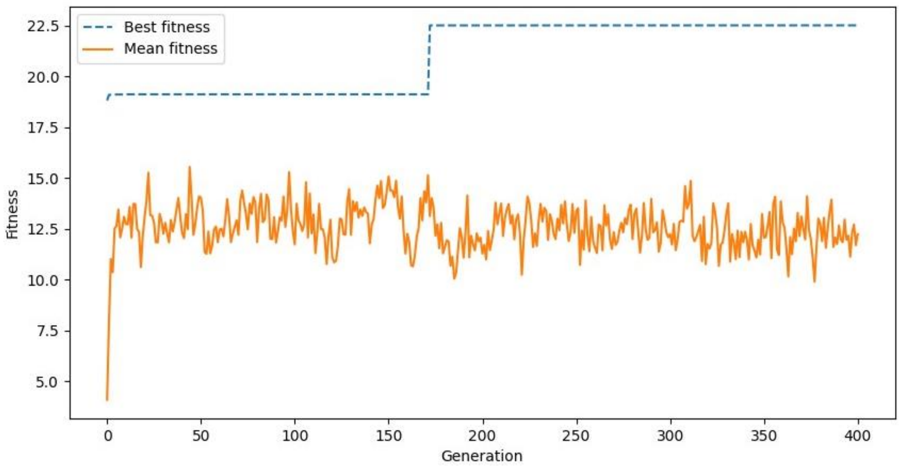


b)

Fig. 2. Best and mean fitness values: a) GA without speciation, b) GA with speciation



a)



b)

Fig. 3. Best and mean fitness values throughout evolving for multimodal function: a) GA without speciation b) GA with speciation

Performing 100 experiments for each we were able to achieve best result 6 times for GA without speciation and 88 times for GA with speciation. Individuals standard deviation for GA without speciation was 4.13 and for GA with speciation 6.32. As we can see results gradually improved after applying speciation.

Conclusion.

In this paper, we described how to use speciation technique with genetic algorithms for optimization task and compared how results have improved after applying speciation. We used the mean and best fitness values and standard deviation for the individuals' values in population to compare population diversity and how successful were optimization results.

Experiments results were according to our expectations. We achieved higher population diversity with speciation, and also were able to improve probability of finding best solution for function with suboptimal solutions gradually. These results confirm effectiveness of speciation in resolving optimization problem for multimodal functions.

Also we tend to think that for some tasks speciation may not be the best technique, and results may differ depending on other GA configurations, but general approach has shown its advantages and may be used in pair with other optimization techniques to improve results.

Compliance with ethical standards.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author contributions.

Conceptualization, [V. Pretsel]; methodology, [V. Pretsel]; validation, [R. Shuvar]; formal analysis, [V. Pretsel].; investigation, [V. Pretsel]; resources, [V. Pretsel, R. Shuvar]; data curation, [V. Pretsel]; writing – original draft preparation, [V. Pretsel]; writing – review and editing, [R. Shuvar]; visualization, [V. Pretsel] supervision, [R. Shuvar]; project administration, [R. Shuvar]; funding acquisition, [V. Pretsel, R. Shuvar].

All authors have read and agreed to the published version of the manuscript.

References

[1] *Sudholt, D. (2019)* "The Benefits of Population Diversity in Evolutionary Algorithms: A survey of rigorous runtime analyses". *Natural computing series, 359–404*. https://doi.org/10.1007/978-3-030-29414-4_8

[2] *Mahfoud, S.W. (1996)* "Niching methods for genetic algorithms". *University of Illinois at Urbana-Champaign* https://dl.acm.org/citation.cfm?id=240028

[3] *Goldberg, D.E. & Richardson, J. (1987)* "Genetic algorithms with sharing for multimodal function optimization". *International Conference on Genetic Algorithms, 41–49*. http://dblp.uni-trier.de/db/conf/icga/icga1987.html#GoldbergR87

[4] *Li, J.-P., Balazs, M. E., Parks, G. T. & Clarkson, P. J (2002)* "A species conserving genetic algorithm for multimodal function optimization". *Evolutionary Computation, 10(3), 207–234.* https://doi.org/10.1162/106365602760234081

[5] *Della Cioppa, A., Marcelli, A. & Napoli, P. (2011)* "Speciation in evolutionary algorithms". *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation (GECCO '11), 1053–1060.* https://doi.org/10.1145/2001576.2001719

ВИДОУТВОРЕННЯ В ГЕНЕТИЧНИХ АЛГОРИТМАХ ДЛЯ ЗБЕРЕЖЕННЯ РІЗНОМАНІТНОСТІ ПОПУЛЯЦІЇ ТА ОПТИМІЗАЦІЇ ФУНКЦІЙ ІЗ СУБОПТИМАЛЬНИМИ РОЗВ'ЯЗКАМИ

В. Прецель, Р. Шувар

Львівський національний університет імені Івана Франка,
вул. Драгоманова 50, 79005 Львів, Україна
vitalii.pretsel@lnu.edu.ua , *roman.shuvar@lnu.edu.ua*

Вступ. Генетичні алгоритми використовуються в різних задачах і демонструють переваги в порівнянні з іншими методами оптимізації, які не завжди можуть бути застосовані. Однак існує ряд недоліків, через які вони можуть не завжди забезпечити найкраще можливе рішення. Одним з них є передчасне згортання популяції до єдиного рішення. У міру зменшення різноманітності популяції область пошуку стає обмеженою, і потенційні рішення можуть бути пропущені. Цей сценарій особливо поширений у мультимодальних функціях, де існує кілька локальних оптимумів. Щоб вирішити цю проблему, було розроблено клас методів, відомих як методи ніші. Ці методи зберігають різноманітність популяції та запобігають передчасній конвергенції до неоптимальних рішень.

Матеріали та методи. У цій роботі досліджується метод видоутворення та те, як він допомагає знаходити рішення для заданих завдань. Для цього було проведено кілька дослідів, у яких знаходилось найбільше значення функції на заданому проміжку. Було порівняно, як відрізняються результати оптимізації з та без застосування видоутворення. Для оцінки різноманітності видоутворення порівнювались значення середньої пристосованості популяції та стандартного відхилення значень особин популяції. Також було оцінено, як видоутворення допомагає в задачах оптимізації з субоптимальними рішеннями, порівнявши скільки успішних розв'язків було отримано в експериментах з та без видоутворення.

Результати. Результати показують, що видоутворення забезпечує збереження різноманітності популяції та покращує результати оптимізації неунімодальних функцій. В експериментах, де застосовувалося видоутворення, популяція зберігала різноманітності, на що вказує більше стандартне відхилення у значеннях особин популяції. Це призвело до збільшення кількості успішних розв'язків у задачах із декількома локальними екстремумами.

Висновки. Видоутворення ефективно зберігає різноманітність популяції та допомагає уникнути передчасного згортання в генетичних алгоритмах. Це призводить до кращих результатів оптимізації, особливо в задачах із кількома локальними екстремумами. Це підкреслює важливість методів збереження різноманітності популяції, таких як видоутворення, у вирішенні недоліків генетичних алгоритмів, особливо в складних задачах оптимізації.

Ключові слова: генетичні алгоритми, оптимізація, методи нішування, видоутворення

The article was received by the editorial office on 18.09.2024.
Accepted for publication on 01.10.2024.