

# **Keen: Kotlin Genetic Algorithms Framework**

*Thesis for the degrees of  
Civil Engineer in Computer Science  
and  
Master of Science in Computing*



**Ignacio Slater Muñoz**

*Departamento de Ciencias de la Computación*

*Facultad de Ciencias Físicas y Matemáticas*

*Universidad de Chile.*

*Thesis Author*

**Nancy Hitschfeld. PhD.**

*Departamento de Ciencias de la Computación*

*Facultad de Ciencias Físicas y Matemáticas*

*Universidad de Chile*

*Thesis Supervisor*

**Alexandre Bergel. PhD.**

*Relational AI*

*Switzerland*

*Second Supervisor*

Santiago, Chile

2023-12-04

v0.5.2310-2

## **Abstract**

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Motivation . . . . .	1
1.2	Hypothesis and Research Questions . . . . .	2
1.2.1	Hypothesis . . . . .	2
1.2.2	Research Questions . . . . .	2
1.3	Objectives . . . . .	2
1.3.1	Main Objective . . . . .	2
1.3.2	Specific Objectives . . . . .	2
1.4	Methodology . . . . .	3
1.5	Thesis Structure . . . . .	3
<b>2</b>	<b>Theoretical Framework</b>	<b>5</b>
2.1	Evolutionary Algorithms . . . . .	5
2.2	Genetic Algorithms . . . . .	6
2.2.1	Representation and Evaluation . . . . .	6
2.2.2	Initialization . . . . .	8
2.2.3	Selection . . . . .	9
2.2.4	Variation . . . . .	9
2.2.5	Termination . . . . .	13
2.3	Genetic Programming . . . . .	14
2.3.1	Representation and Evaluation . . . . .	15
2.3.2	Initialization . . . . .	20
2.3.3	Selection . . . . .	23
2.3.4	Variation . . . . .	23
2.3.5	The Generalization Problem in GP . . . . .	28
<b>3</b>	<b>State of the Art: Relevant Work</b>	<b>29</b>
3.1	The One Max Problem . . . . .	29
3.2	Agile Artificial Intelligence in Pharo . . . . .	30
3.3	Distributed Evolutionary Algorithms in <i>Python</i> (DEAP) . . . . .	31
3.4	Jenetics: Java Genetic Algorithm Library . . . . .	33
3.5	ECJ: A Java-based Evolutionary Computation Research System . . . . .	35
3.6	GeneticSharp: Comprehensive Overview of a .NET Genetic Algorithm Library . . . . .	39
3.7	Other Libraries . . . . .	42
<b>4</b>	<b>The Keen Framework</b>	<b>43</b>
4.1	Introduction . . . . .	43
4.2	Architecture . . . . .	43
4.3	Genetic Algorithms . . . . .	45
4.3.1	Genetic Material . . . . .	45
4.3.2	Ranker . . . . .	47
4.3.3	Evolution Engine . . . . .	47
4.3.4	Evolution Listeners . . . . .	49
4.4	Genetic Operators . . . . .	50

4.4.1	Selection . . . . .	50
4.4.2	Mutation . . . . .	52
4.4.3	Crossover . . . . .	55
4.5	Genetic Programming . . . . .	61
4.5.1	Primitive Set . . . . .	61
4.5.2	Genetic Operators . . . . .	65
4.6	Extensibility . . . . .	68
4.6.1	Structured Extensibility Through Interfaces . . . . .	68
4.6.2	Factory Method Pattern in Genetic Construction . . . . .	68
4.6.3	Sustaining Future Evolution . . . . .	69
4.7	Conclusion . . . . .	69
<b>5</b>	<b>Case Study: Real Function Optimization</b>	<b>71</b>
5.1	Introduction . . . . .	71
5.2	Problem Description . . . . .	71
5.3	Solution . . . . .	72
5.3.1	Chromosome Structure . . . . .	72
5.3.2	Variation Operators . . . . .	72
5.3.3	Implementation . . . . .	72
5.4	Results . . . . .	74
5.4.1	Random Selector . . . . .	75
5.4.2	Tournament Selector . . . . .	76
5.4.3	Roulette Selector . . . . .	77
5.5	Conclusion . . . . .	78
<b>6</b>	<b>Case Study: Knapsack Problem</b>	<b>79</b>
6.1	Introduction . . . . .	79
6.2	Problem Description . . . . .	79
6.2.1	0-1 Knapsack Problem . . . . .	79
6.2.2	Unbounded Knapsack Problem . . . . .	80
6.3	Solution Strategies for Knapsack Problems . . . . .	80
6.3.1	0-1 Knapsack Problem . . . . .	80
6.4	Results . . . . .	81
6.5	Conclusion . . . . .	81
<b>7</b>	<b>Case Study: Crash Reproduction</b>	<b>83</b>
7.1	Introduction . . . . .	83
7.2	Problem Description . . . . .	83
7.3	Solution . . . . .	84
7.4	Results . . . . .	85
7.5	Conclusion . . . . .	85
<b>8</b>	<b>Conclusions</b>	<b>87</b>
8.1	Summary . . . . .	87
8.2	Contributions . . . . .	88
8.3	Future Work . . . . .	88
8.3.1	Genetic Algorithms . . . . .	88
8.3.2	Genetic Programming . . . . .	89
8.3.3	Multi-Objective Evolution . . . . .	90
8.3.4	Neuroevolution . . . . .	91
8.3.5	Coevolution . . . . .	91
8.3.6	Evolution Strategy . . . . .	92
8.3.7	Other evolutionary algorithms . . . . .	93
8.3.8	Crash Reproduction . . . . .	93

<b>A Glossary</b>	<b>95</b>
<b>B Test Functions for Optimization</b>	<b>99</b>
B.1 Ackley Function . . . . .	99
B.2 Beale Function . . . . .	100
B.3 Booth Function . . . . .	100
B.4 Bukin Function N.6 . . . . .	100
B.5 The Cross-in-Tray Function . . . . .	101
B.6 Easom Function . . . . .	102
B.7 Eggholder Function . . . . .	102
B.8 Goldstein-Price Function . . . . .	103
B.9 Himmelblau's Function . . . . .	104
B.10 Hölder Table Function . . . . .	105
B.11 Lévi Function N.13 . . . . .	105
B.12 Matyas Function . . . . .	106
B.13 McCormick Function . . . . .	107
B.14 Rastrigin Function . . . . .	107
B.15 The Rosenbrock Function . . . . .	109
B.16 Schaffer Function N.2 . . . . .	109
B.17 Schaffer Function N.4 . . . . .	110
B.18 The Sphere Function . . . . .	112
B.19 Styblinski-Tang function . . . . .	112
B.20 Three-Hump Camel Function . . . . .	113
<b>C Additional Listings</b>	<b>115</b>
<b>D Keen Open Source License: BSD 2-Clause License</b>	<b>121</b>

# Chapter 1

## Introduction

### 1.1 Motivation

The inception of artificial intelligence (AI) as we perceive it today is deeply rooted in the mid-20th century when Alan Turing proposed the intriguing question, “Can machines think?” [1]. This seminal query catalyzed a remarkable evolution in the field, leading to the emergence of a diverse spectrum of AI applications that persistently extend and advance.

Among the early endeavors to emulate evolution artificially, Nills Aall Barricelli’s initiative in 1954 stands out. Barricelli construed evolution as a purely statistical process, setting a novel precedent for the field of evolutionary computation (EC) [3]. This groundbreaking exploration established the parameters for an evolutionary system, which necessitated the system’s components to exhibit capabilities of reproduction and mutation, thereby facilitating evolution through natural selection or the survival of the fittest.

Expanding upon these foundational principles, John Holland’s influential book, *Adaptation in Natural and Artificial Systems* [6], unveiled genetic algorithms (GA). This innovative methodology for addressing optimization problems drew inspiration from natural selection and genetic mechanisms. Holland’s work solidified the field of evolutionary algorithms (EA), igniting the development of a plethora of techniques and algorithms, including but not limited to, genetic algorithms, genetic programming, and evolutionary strategies.

Despite the proven efficacy of evolutionary algorithms in resolving intricate optimization problems ranging from scheduling and data mining to machine learning and beyond, their execution can be challenging, frequently demanding extensive, repetitive coding. As the uptake of AI continues to soar, the requirement for streamlined, intuitive libraries and frameworks that empower efficient experimentation and application of these potent algorithms grows concomitantly.

The *Kotlin* programming language emerges as a promising platform for developing such a framework [60]. *Kotlin* distinguishes itself with its lucid syntax, static typing, and flawless interoperability with Java, making it an attractive choice for AI and EC enthusiasts [49, 50, 61]. The significance of *Kotlin* has been formally acknowledged in *Android* app development [48], and it continues to gain popularity among developers, with over 60% of *Android* developers now using it [47]. Furthermore, according to a *Github* study, *Kotlin* emerged as the ninth fastest-growing language worldwide from 2021 to 2023, overtaking *Python* by 0.4% [67].

The allure of *Kotlin* is further amplified by its multiplatform feature, which permits code sharing across a wide range of platforms, including JVM, *JavaScript*, *Android*, *iOS*, and native desktop applications. This cross-platform compatibility promotes enhanced productivity and consistency in software development.

*Kotlin* also supports functional programming (FP) and coroutines, which simplify the development of asynchronous and concurrent applications. This functionality proves particularly beneficial in the realm of evolutionary algorithms that often necessitate the execution of numerous concurrent processes. Moreover, *Kotlin*’s capabilities align well with the design of *Domain Specific Languages* (DSLs) [68], which can be used to concisely and intuitively define the problem domain.

Given these compelling features, this thesis aims to exploit the prowess of *Kotlin* to architect and implement a pioneering genetic algorithms framework, thereby augmenting the ever-growing application of *Kotlin* within the AI community. We expect this

framework to stimulate research and applications in the field, hastening the creation of novel algorithms and techniques.

## 1.2 Hypothesis and Research Questions

### 1.2.1 Hypothesis

The *Kotlin* programming language, known for its expressive syntax, simplicity, and cross-platform capabilities, presents an opportunity for enhancing the field of evolutionary algorithms. We hypothesize that the development of an efficient and user-friendly genetic algorithms framework in *Kotlin* could significantly streamline the implementation process of these algorithms, thereby catalyzing innovation and expediting research in this domain.

### 1.2.2 Research Questions

To assess the validity of our hypothesis and further investigate the potential benefits of the proposed *Kotlin*-based framework, we outline the following pivotal research questions:

1. What constitute the *fundamental requirements* for a versatile, efficient, and user-friendly genetic algorithm framework?
2. How can the *unique language features* of *Kotlin* be leveraged to fulfill these requirements and facilitate the development and use of genetic algorithms?
3. How does the performance and *ease of use* of a *Kotlin*-based genetic algorithms framework compare to that of existing solutions, particularly in terms of syntactical complexity?
4. In what ways can the proposed *Kotlin*-based framework be applied to *address complex real-world problems*, thereby demonstrating its practical viability and potential to contribute to advancements in AI and evolutionary computation?
5. What potential opportunities exist for the *enhancement and expansion* of the proposed *Kotlin*-based genetic algorithm framework in the future?

These research questions serve as guideposts for our exploration into the feasibility and potential impact of using *Kotlin* to develop a novel genetic algorithms framework. Each question targets a specific facet of the project, offering a holistic understanding of the technical intricacies, practical applications, and future prospects of our proposed framework.

## 1.3 Objectives

The primary objectives of this thesis are set forth to establish a solid foundation for the design, implementation, and evaluation of a novel, efficient, and user-friendly genetic algorithms framework using the *Kotlin* programming language.

### 1.3.1 Main Objective

The main objective of this study is to develop a *Kotlin*-based genetic algorithms framework that is versatile, efficient, and user-friendly. The framework should be able to support classical GAs, as seen in section 2.2 on page 6, and be easily extensible to support future enhancements and expansions such as new algorithms, like GP, and genetic operators.

### 1.3.2 Specific Objectives

Building upon the primary objectives, the following specific objectives have been established to guide this research:

1. **Language Feature Utilization:** Examine and utilize *Kotlin*'s unique language features, such as its expressive syntax, static typing, and its ease of developing DSLs, to design and implement a genetic algorithms framework. This exploration should enhance the design and implementation of the genetic algorithms framework, ensuring it is efficient, robust, and user-friendly.
2. **Framework Efficiency:** Design the genetic algorithms framework to optimize computational efficiency. This should involve algorithmic improvements, effective use of *Kotlin*'s language features, and careful resource management.

3. **Comparative Study:** Conduct comparative studies between the *Kotlin*-based genetic algorithms framework and existing solutions, particularly in terms of syntactical complexity, and ease of use. This should provide a benchmark to evaluate the advantages and potential areas of improvement for our framework.
4. **Real-world Applications:** Apply the proposed *Kotlin*-based framework to address complex real-world problems. This will not only demonstrate the practical viability of the framework but also its potential to contribute to advancements in AI and EC.
5. **Future Enhancements and Expansion:** Anticipate and propose potential enhancements and extensions for the *Kotlin*-based genetic algorithms framework. Discuss its adaptability to incorporate future advancements in AI and EC, ensuring its sustainability and continued relevance.

These specific objectives aim to add more granularity to our research focus, ensuring a comprehensive and detailed exploration of the possibilities offered by *Kotlin* for the development of a novel genetic algorithms framework.

## 1.4 Methodology

Our research methodology employs a multi-faceted approach comprised of distinct phases. This approach maintains a balance between theoretical exploration and practical application, with a consistent emphasis on rigorous analysis and evaluation.

1. **Literature Review and Language Exploration:** We initiate our research with an exhaustive study and analysis of the current literature pertaining to genetic algorithm frameworks, thereby discerning their respective merits and shortcomings. This analysis will serve as a cornerstone for defining the quintessential attributes for a versatile, efficient, and user-friendly genetic algorithms framework. Concurrently, we will conduct an in-depth investigation of the unique language features offered by *Kotlin*, discerning their potential contributions to the design and development of the proposed framework.
2. **Framework Design and Development:** Informed by the comprehensive literature review and language exploration, we will transition to the design and implementation phase of the *Kotlin*-based genetic algorithms framework. The design strategy will involve the thoughtful employment of *Kotlin*'s unique features such as its expressive syntax, static typing, and ease of crafting DSLs, thereby ensuring that the resulting framework is efficient, robust, and user-friendly.
3. **Application to Real-world Challenges:** Subsequently, we will utilize the developed *Kotlin*-based framework to address complex real-world problems. This stage will function as a testing ground to demonstrate the practical applicability, potential contributions, and viability of the framework in fostering advancements in the realms of AI and evolutionary computation.
4. **Future Adaptability and Expansion:** In the final phase of our research, we will deliberate upon prospective enhancements and extensions for the *Kotlin*-based genetic algorithms framework. Moreover, this phase will entail an assessment of the framework's adaptability and scalability to accommodate future advancements in AI and evolutionary computation, thereby assuring its long-term sustainability and relevance.

Throughout the implementation of our research methodology, we will uphold an analytical and critical perspective to discern the significance of our findings and their implications. We acknowledge and uphold the ethical principles of transparency and honesty in reporting our results and documenting the progress of our research.

## 1.5 Thesis Structure

1. **Background (chapter 2 on page 5):** This chapter offers a necessary foundation for understanding the central concepts utilized in this thesis. It encompasses a theoretical analysis of the relevant algorithms, setting the stage for the discussions and investigations that follow.
2. **State of the Art (chapter 3 on page 29):** Here, we present a comprehensive review of the current advancements in the domain of evolutionary computation. We particularly emphasize and examine the prevailing frameworks in this field, facilitating a comparative context for our research.
3. **Framework Design and Implementation (chapter 4 on page 43):** This chapter elucidates the architecture of our framework, detailing the design considerations and the subsequent implementation.

4. **Function Optimization Case Study (chapter 5 on page 71):** In this section, we put our framework to the test by applying it to a real-world function optimization problem. We utilize 20 classical benchmark functions to assess the framework's performance, with the functions elaborately displayed in appendix B on page 99.
5. **Knapsack Problem Case Study (chapter 6 on page 79):** This chapter features another practical application of the framework, in this case, tackling the Knapsack Problem. We test the framework against two different variants of this problem: the 0-1 Knapsack Problem and the Unbounded Knapsack Problem, evaluating its adaptability and robustness.
6. **Crash Reproduction Problem Case Study (chapter 7 on page 83):** In our third case study, we explore the application of the framework to a crash reproduction problem, showcasing its utility in a distinctly different domain.
7. **Conclusion and Future Work (chapter 8 on page 87):** In the concluding chapter, we summarize the breadth of work accomplished in this thesis and propose potential avenues for future investigations and enhancements based on our findings.

This structure facilitates a progressive narrative of our research, allowing for a coherent understanding of our objectives, methodology, and outcomes.

## Chapter 2

# Theoretical Framework

The objective of this chapter is to provide the reader with the theoretical background necessary to understand the rest of the document.

## 2.1 Evolutionary Algorithms

In the field of computational intelligence, *evolutionary algorithms* (EA) [26] are a family of algorithms inspired by the process of natural selection. They are part of the larger field of *evolutionary computation*,<sup>1</sup> which is a subfield of *metaheuristics*.<sup>2</sup>

EAs are algorithms that perform optimization or learning tasks by evolving solutions to a given problem via emergent intelligence [10]. These tasks may range from function optimization to machine learning or game AI development. EAs have three main characteristics:

- **POPULATION-BASED:** These algorithms work with a **population of solutions**, allowing them to explore the search space in *parallel*.
- **FITNESS-ORIENTED:** The solutions in the population are evaluated using a *fitness function*, which is a **problem-dependent function** that assigns a value to each solution based on its quality. The goal of the algorithm is to find the solution with the highest<sup>3</sup> fitness.
- **VARIATION-DRIVEN:** The candidate solutions are modified using *genetic operators*, such as mutation, crossover, and selection, to create new solutions. These operators are usually based on the biological processes of *mutation* and *recombination*.

**Remark.** “Parallel” in this context means that the algorithm is exploring multiple points in the search space in the same generation (iteration). It should not be confused with parallel computing, which is a technique used to speed up the execution of the algorithm by running it on multiple processors.

Even though EAs are parallel in nature, they can be run on a single processor. Nevertheless, this parallel nature makes them a good candidate for parallel computing; in this thesis we will explore the use of parallel computing to speed up the execution of certain stages of the algorithm.

While these principles serve as the foundation for most EAs, it’s important to note that some variants may prioritize some principles over others, or introduce new principles. This diversity allows EAs to be adapted to a wide range of problems and scenarios.

<sup>1</sup>See definition A.7 on page 96

<sup>2</sup>See definition A.14 on page 97

<sup>3</sup>In some cases, the goal is to minimize the fitness function, in which case the algorithm will aim to find the solution with the lowest fitness.

## 2.2 Genetic Algorithms

Genetic Algorithms (GA)<sup>4</sup> [6, 7, 26, 29] are a type of EA where a **population** of **individuals**<sup>5</sup> representing candidate solutions to an optimization problem evolves towards better solutions. Each individual is defined by its location in the search space, known as its **genotype**<sup>6</sup>, and its fitness value, computed by a **fitness function**. At a high level, GA is an automatic method for problem-solving, starting from a *high-level statement* of the desired outcome, without needing the user to predefine the solution's form or structure.

The classical GA operates as follows:

Listing 2.1– Genetic algorithm

```
var population = initialize population // Creates a random population of individuals
population.forEach { it -> evaluate it } // Calculates the fitness of each individual
while (termination condition is not met) {
    ← // Could be a pre-defined number of generations, a desired fitness level, etc.
    val survivors = select survivors from population
    ← // Selects individuals for the next generation
    val parents = select parents from population
    ← // Selects a subset of individuals as parents
    val offspring = alter parents
    ← // Applies genetic operators to parents to create new individuals
    offspring.forEach { it -> evaluate it }
    ← // Calculates the fitness of each new individual
    population = survivors + offspring // Creates the next generation
}
return fittest from population // Returns the most fit individual
```

The exact implementation of each of these steps depends on the specific problem at hand. Factors such as the problem's complexity, the representation of individuals, or even the computational resources available, can greatly influence the choice of methods used for initialization, selection, alteration, and survivor selection.

### 2.2.1 Representation and Evaluation

A pivotal component of a Genetic Algorithm (GA) is the representation of individuals, which encodes potential solutions into a form manipulable by the GA. This representation delineates the algorithm's search space and is a prime determinant of its performance.

The predominant representation is a matrix of genes<sup>7</sup> termed the **genotype**<sup>8</sup>, wherein each column is denoted as a **chromosome**<sup>9</sup>.

<sup>4</sup>Also known as Simple Genetic Algorithms (SGA) [26], or Traditional Genetic Algorithms (TGA) [29].

<sup>5</sup>See definition A.12 on page 96.

<sup>6</sup>See definition A.11 on page 96

<sup>7</sup>See definition A.8 on page 96

<sup>8</sup>See [69]

<sup>9</sup>See definition A.4 on page 95

**Definition 2.1** (Cardinality of the search space). *The cardinality of the search space is the number of different individuals that can be represented by the encoding.*

Formally, given a vector of alphabets  $(\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n)$ , and a representation  $\mathbf{G}$  with  $n$  chromosomes of lengths  $(m_1, m_2, \dots, m_n)$  where each chromosome is encoded using the alphabet  $\mathcal{A}_i$ , the cardinality of the search space  $S$  is defined as:

$$|S| = \prod_{i=1}^n |\mathcal{A}_i|^{m_i} \quad (2.1)$$

Note that this definition assumes that the chromosomes are independent, which may not be the same for all evolutionary algorithms.

**Remark.** In the original publication of the GA [6], the genotype was known as the **environment** ( $E$ ) and the search space was defined as a class  $\mathcal{E}$  of all possible environments.

To illustrate this concept, consider the following problem: given a binary string of length  $n$ , find the string that has the most ones, this is known as the **One Max problem** (OMP) [14]<sup>10</sup> (refer to section 3.1 on page 29 for a more detailed description of the problem). In this case, we can use a single column matrix  $\mathbf{G}$  to represent the individual, where each gene  $g_i \in \mathcal{A}$  represents the  $i$ -th bit of the string, where  $\mathcal{A} = \{0, 1\}$  is the alphabet containing the two possible values of a bit.

Then,

$$|S_{\text{OMP}}| = \prod_{i=1}^1 |\mathcal{A}|^n = 2^n$$

Knowing this, we can conclude that an exhaustive search of the search space would require evaluating  $2^n$  individuals, and thus the algorithm would have a time complexity of  $\mathcal{O}(2^n)$ .

This is a very simple example, but we can see how a naive search algorithm would have a very high time complexity. This would be of the utmost importance in a real world problem, where the search space would be much larger.

With a representation defined, we can now define an evaluation method for the individuals, which is done using a **fitness function**.

**Definition 2.2** (Fitness function). *A fitness function is a function  $\phi : S \rightarrow \mathbb{R}^n$ , where  $S$  is the search space and  $n$  is the number of objectives of the optimization problem, that takes a genotype as input and returns a vector of real numbers representing how close the individual is to the global optimum of each objective.*

*The fitness function is usually defined by the user of the algorithm, and it is problem dependent.*

**Definition 2.3** (Batch fitness function). *A batch fitness function  $\Phi : \mathbb{P} \rightarrow \mathbb{R}^{m \times n}$  is a function that maps a population to a matrix of real numbers, where  $m$  is the number of individuals in the population and  $n$  is the number of objectives of the optimization problem.*

The one max problem is a maximization problem with a single objective, so the fitness function would be defined as follows:

$$\phi_{\mathbf{G}} = \sum_{i=1}^n g_i \quad (2.2)$$

Having laid the groundwork by illustrating the pivotal role of representation in GAs, we've delineated how the representation forms the backbone of the algorithm's search space and in turn determines its performance. Key terminologies, including the genotype, chromosomes, cardinality of the search space, and the crucial role of the fitness function, have been elucidated with references and examples. With this foundational knowledge, in the forthcoming section, we will delve deeper into the

<sup>10</sup>Also **Ones Counting** problem [69], or **Max Ones** problem [54].

initialization phase of the algorithm, setting the stage for how the GA seeds its initial population and embarks on the quest for optimal solutions.

### 2.2.2 Initialization

In the initialization phase of a genetic algorithm, the foundational stage is set for the algorithm's evolutionary journey. Here, we define and setup the population of individuals to be employed in the search process. Whether informed by prior knowledge or randomly generated, each individual's evaluation is pivotal to steer the algorithm's quest for optimal solutions.

As we delve further, stages such as selection (section 2.2.3 on the facing page), crossover (section 2.2.4.1 on page 10), and mutation (section 2.2.4.2 on page 11) build upon this foundational phase.

A GA operates on a group of individuals termed a ***population***. Determining the population's size and its initialization is crucial. Typically, this initialization process leans on randomness, but it can also be informed by insights about the problem at hand [20].

Upon initializing the population, each individual undergoes evaluation to receive a ***fitness value***. This step is imperative to glean insights about the problem, thereby directing the search towards enhanced solutions.

Using the ***One Max*** problem as a backdrop, where there's an absence of prior problem knowledge, the initialization encompasses a blind search of the search space, rendering it random. For each population member, a random binary string of length  $n$  is generated.

Let's assume that we have a population of size 4, and that the length of the binary strings is  $n = 4$ .

The initialization process could generate the following individuals:<sup>11</sup>

Generation 0		
Individual	Binary string	Fitness
$I_1$	1100	2
$I_2$	0001	1
$I_3$	0000	0
$I_4$	0100	1

Table 2.1: Population of individuals in generation 0

	Fitness	Individual
Best	2	$I_1$
Worst	0	$I_3$
Average		1
Standard deviation		0.817

Table 2.2: Fitness of the individuals in generation 0

Emphasis should be placed on the aggregation functions outlined in table 2.2. The ***average fitness*** is straightforward, offering an overall view of the population's performance; a higher<sup>12</sup> average suggests better overall performance. On the other hand, the ***standard deviation*** gauges the dispersion of fitness values around this average. A low standard deviation indicates a homogeneous population, which might hint at ***premature convergence*** due to limited search space exploration. Conversely, a high standard deviation signifies a diverse population, potentially suggesting ***excessive exploration*** without honing in on promising areas.

The initialization phase of a genetic algorithm sets the foundation by establishing the population of individuals for the search. This population might be randomized or informed by prior domain knowledge. Each member of this population undergoes

<sup>11</sup>Since the nature of genetic algorithms is stochastic, the initialization process could generate different individuals each time the algorithm is run. For this example, we selected a specific set of individuals in a way that makes it easier to get a grasp of the algorithm.

<sup>12</sup>Remember, a "higher fitness" refers to proximity to the optimal solution, not necessarily the greatest numerical value.

a fitness evaluation, directing the search towards optimal outcomes. Using our OMP example, we commenced with a four-individual population with binary strings of length  $n = 4$ , assessing their fitness. This foundational step paves the way for the ensuing phases of selection (section 2.2.3), crossover (section 2.2.4.1 on the next page), and mutation (section 2.2.4.2 on page 11).

### 2.2.3 Selection

Following the initialization phase, a genetic algorithm (GA) progresses through its main evolutionary cycle. Key to this cycle is the process of selection, which mirrors natural selection by favoring individuals with higher fitness for reproduction.

Given a population  $P$  comprising  $N$  individuals, each identified by their fitness value  $\phi_i$  (where  $i \in \{1, \dots, N\}$ ), the survival rate  $\sigma$  helps control elitism.<sup>13</sup> This rate determines the portion of the population that proceeds to the next generation without change. Specifically,  $\lfloor \sigma N \rfloor$  individuals continue, while the remaining  $\lceil (1 - \sigma)N \rceil$  are replaced by their descendants.<sup>14</sup>

**Definition 2.4** (Selection Operator). *A tool for choosing specific members from a population, formally expressed as:*

$$\Sigma : \mathbb{P} \times \mathbb{N} \times \dots \rightarrow \mathbb{P}; (P, n, \dots) \mapsto \Sigma(P, n, \dots)$$

Parameters include:

- $\mathbb{P}$ : Set of possible populations.
- $\mathbb{N}$ : Set of natural numbers.
- $P$ : A given population.
- $n$ : Number of selections from  $P$ .
- $\Sigma(P, n)$ : Resultant population from the selections.

This selection operator usually integrates randomness, bringing some unpredictability to the process. As an exemplification, let's consider the *roulette wheel* selection operator.<sup>15</sup> Here, individuals are assigned selection probabilities according to their fitness, as expressed by:

$$\rho_{\Sigma}(i) = \frac{\phi_i}{\sum_{j=1}^N \phi_j} \quad (2.3)$$

Individual	Fitness	Selection Probability
$I_1$	2	50%
$I_2$	1	25%
$I_3$	0	0%
$I_4$	1	25%

Table 2.3: Probabilities of selection for our sample population.

Following this, individuals are probabilistically selected for survival. For instance, if  $I_2$  persists, the remaining  $I_1$ ,  $I_3$ , and  $I_4$  yield their places to the offspring.

Concluding this introduction to selection in GAs, the subsequent sections will investigate this process further and introduce the variation operators crucial for offspring generation.

### 2.2.4 Variation

Variation is the process of creating new individuals from existing ones in the pursuit of exploring the solution space. This is crucial in a Genetic Algorithm (GA) to avoid premature convergence to sub-optimal solutions. In a GA, variation is achieved by applying **variation operators** to the individuals in the population. The most common variation operators are **crossover** and **mutation**, which will be explored in this section.

<sup>13</sup>For more on elitism, refer to definition A.5 on page 95.

<sup>14</sup>Using both the *floor* and *ceiling* functions ensures a consistent total population of  $N$ .

<sup>15</sup>Refer to section 4.4.1.2 on page 51 for details.

**Definition 2.5** (Variation operator). A variation operator is used to create new individuals from existing ones. Formally, it is a variadic function represented as

$$\varphi : \mathbb{P} \times \mathbb{R} \times \dots \rightarrow \mathbb{P}; (P, \rho_\varphi, \dots) \mapsto \varphi(P, \rho_\varphi, \dots)$$

where:

- $\mathbb{P}$  is the set of all possible populations,
- $\mathbb{R}$  is the set of real numbers,
- $P$  is the population to be varied,
- $\rho_\varphi$  is the probability of applying the operator to an individual in the population.

The additional arguments depend on the specific implementation of the variation operator. The role of these arguments will be clarified in section 4.4 on page 50.

#### 2.2.4.1 Crossover

In genetic algorithms, a prominent variation operator is the **crossover**. This operator mirrors the genetic recombination seen in nature.<sup>16</sup> It facilitates the exchange of genetic material between two individuals, spawning a new generation.

**Definition 2.6** (Crossover operator). The crossover operator recombines genetic material from existing individuals to create new ones. Formally, it is represented as:

$$X : \mathbb{P} \times \mathbb{R} \times \dots \rightarrow \mathbb{P}; (P, \rho_X, \dots) \mapsto X(P, \rho_X, \dots)$$

with the following parameters:

- $\mathbb{P}$  – the set of all possible populations,
- $\mathbb{R}$  – the set of real numbers,
- $P$  – the population under variation,
- $\rho_X$  – probability of applying the operator to an individual.

In our analysis, we employ a condensed version of the **single-point crossover** operator.<sup>17</sup> This operator picks the first half of the genes from two parental figures and births two new offspring by swapping these chosen genes.

Take, for instance, two parent individuals,  $I_1 = 1100$  and  $I_2 = 0001$ . Utilizing the single-point crossover operator, we select the first half of the genes: 11 from  $I_1$  and 00 from  $I_2$ . This exchange produces the offspring  $O_1 = 1101$  and  $O_2 = 0000$ , demonstrated in fig. 2.1 on the next page.

Following another iteration of the single-point crossover operator, we can generate a result as shown in table 2.4, leading to a new population  $\mathbf{O} = \{(0000, 0), (1101, 3), (0101, 2)\}$ .

Generation 0 → Generation 1			
I	$\Phi_I$	O	$\Phi_O$
$\begin{bmatrix} 1100 \\ 0001 \end{bmatrix}$	$\begin{bmatrix} 2 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0000 \\ 1101 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 3 \end{bmatrix}$
$\begin{bmatrix} 0001 \\ 0100 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0101 \\ \cdot \end{bmatrix}$	$\begin{bmatrix} 2 \\ \cdot \end{bmatrix}$

Table 2.4: Illustration of the single-point crossover operation. In this procedure, two parent individuals are selected and a cut point is chosen. Each offspring is then formed by combining the genes from the parents: one gets the genes from the first part of the first parent and the second part of the second parent, while the other gets the genes from the first part of the second parent and the second part of the first parent. Here, · represents a “discarded” value (since according to the survival rate, only three offspring need to be produced).

<sup>16</sup>Also known as **crossing-over** in [6].

<sup>17</sup>Refer to section 4.4.3.2 on page 56.



Figure 2.1: Single-point crossover

If we now use these offspring as-is to create the next generation, we would obtain the population shown in table 2.5:

Generation 1		
Individual	Binary String	Fitness
$I_2$	0001	1
$O_1$	0000	0
$O_2$	1101	3
$O_3$	0101	2

Table 2.5: Population after applying the single-point crossover operator. Note that  $I_2$  is the survivor of the previous generation picked in section 2.2.3 on page 9.

	Fitness	Individual
Best	3	$O_2$
Worst	0	$O_1$
Average		1.25
Standard deviation		1.291

Table 2.6: Fitness of the population after applying the single-point crossover operator. “Best” refers to the individual with the highest fitness, and “Worst” refers to the individual with the lowest fitness

By the metrics in table 2.6, the population’s average fitness rises from 1 to 1.25, with the fittest individual’s score jumping from 2 to 3. This uplift highlights the crossover operator’s role in directing the search towards enhanced solutions.

While the crossover operation elevates the average population fitness, incorporating a **mutation** operator can further boost genetic diversity<sup>18</sup> and sidestep early convergence to inferior solutions. The subsequent section delves into this operation.

#### 2.2.4.2 Mutation

While the crossover operator effectively recombines existing genetic material, it is limited by the genetic diversity already present in the population. This can sometimes lead to premature convergence, especially for complex problems characterized by nu-

<sup>18</sup>See definition A.10 on page 96.

merous local optima, like the *Rastrigin function*.<sup>19</sup>

To combat this and infuse fresh *diversity*, the *mutation operator* is employed. It introduces small, probabilistic changes to the genetic makeup of individuals.

**Definition 2.7** (Mutation operator). *The mutation operator introduces variations in an individual's genetic material based on a predefined probability, resulting in a new population. Formally, a mutation operator can be represented as:*

$$M : \mathbb{P} \times [0, 1] \times \dots \rightarrow \mathbb{P}; (P, \mu, \dots) \mapsto M(P, \mu, \dots)$$

Where:

- $\mathbb{P}$  denotes the set of all possible populations.
- $\mathbb{R}$  signifies the set of real numbers.
- $P$  stands for the current population.
- $\mu$  indicates the mutation rate – the chance of an individual undergoing mutation.

Additional parameters vary depending on the specific mutation operator in play.

For example, the “One Max” problem might employ a *bit-flip* mutation.<sup>20</sup> This mutation scans every gene in an individual, flipping it based on a predetermined probability.

Let's say we apply this mutation with a rate of  $\mu = 1$  to the population post-crossover (as discussed in section 2.2.4.1 on page 10). The resulting mutated population,  $O'$ , is illustrated in table 2.7.

Generation 0 → Generation 1			
I	$\Phi_I$	$O'$	$\Phi_O$
$\begin{bmatrix} 0000 \\ 1101 \\ 0101 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 1111 \\ 0010 \\ 1010 \end{bmatrix}$	$\begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}$

Table 2.7: Illustration of the *bit-flip* mutation operator applied to the population resulting from the crossover operation in section 2.2.4.1 on page 10.

By analyzing the resulting offspring, as highlighted in table 2.8, we can observe the mutation's role in enhancing diversity. For instance, no member of the original population had a 1 in the third position, rendering crossover incapable of producing such a gene configuration. However, post-mutation, three individuals exhibit this trait.

Generation 1		
Individual	Binary String	Fitness
$I_2$	0001	1
$O'_1$	1111	4
$O'_2$	0010	1
$O'_3$	1010	2

Table 2.8: Population after applying the *bit-flip* mutation operator to the population resulting from the crossover operation in section 2.2.4.1 on page 10.

<sup>19</sup>Refer to appendix B.14 on page 107.

<sup>20</sup>For more details, refer to section 4.4.2.2 on page 53.

	Fitness	Individual
Best	4	$O'_1$
Worst	0	$(I_2, O'_2)$
Average	2	
Standard deviation	1.414	

Table 2.9: Fitness of the population after applying the *bit-flip* mutation operator to the population resulting from the crossover operation in section 2.2.4.1 on page 10.

In summary, the mutation operator serves a pivotal role in genetic algorithms. It rejuvenates population diversity, mitigating early convergence to suboptimal solutions. By fostering exploration of the search landscape, it allows potentially superior traits to surface. Nonetheless, the mutation rate's calibration is a balancing act: excessively high rates might destabilize favorable traits, while overly conservative rates might inadequately deter premature convergence. The specific mutation operator and its associated rate significantly influence the genetic algorithm's exploratory efficiency.

Having concluded the variation phase, we can transition to the subsequent stage of the genetic algorithm.

### 2.2.5 Termination

The genetic algorithm assesses the termination criteria after generating each new population. If met, the algorithm ends and outputs the best-found individual. Otherwise, the generational cycle repeats.

Suppose our termination criterion is the identification of an individual with all ones, represented as 1111. This would mean the fitness function,  $\phi_G$ , attains the value 4.

During our exploration, after applying the variation operators, we indeed discovered the individual 1111. This satisfies our termination criterion, prompting the algorithm to conclude its operation.

Yet, as illustrated in table 2.10, the algorithm has not scoured the entire search space. Instead, the genetic algorithm prioritizes a fitness-guided exploration over a complete one. Still, witnessing the increasing fitness of individuals over generations signals a trend towards an optimal or near-optimal solution.

	00	01	10	11
00				
01				
10				
11				

	00	01	10	11
00				
01				
10				
11				

Table 2.10: A map of the search space explored by the genetic algorithm. Dark gray cells denote the candidates the algorithm reviewed. The position of each individual corresponds to its binary representation, using the row for the first two bits and the column for the last two.

In small search arenas, the distinction between a genetic algorithm and random searches might appear negligible. However, as the search space expands, which we delve into later in this thesis, the difference becomes substantial.

Despite their stochastic nature, genetic algorithms do not always promise optimal outcomes. Their efficacy hinges on various components, including the fitness function, representation, variation operators, and selection techniques. This thesis delves deeper into these elements and analyzes their performance across diverse problems.

In essence, the termination phase is pivotal in dictating the final outcome of the genetic algorithm. By setting a clear termination condition, like achieving the highest fitness score, the algorithm efficiently navigates the search terrain. Although it might not exhaustively search, it employs a fitness-centric strategy to steer closer to optimal solutions. It's vital to acknowledge the intrinsic limitations of genetic algorithms. Yet, when configured right, their capability to overshadow random searches, especially in vast spaces, is undeniable. We'll further explore this subject in the subsequent segments of this thesis.

## 2.3 Genetic Programming

**Genetic Programming** (GP) [7, 9, 21, 26] has been carved out as a specialized subfield of EAs which emphasizes evolving a collection of computer programs tailored to address specific problems. While GP can be viewed as a natural evolution of GAs, it distinguishes itself by its primary goal: GAs fine-tunes parameters to optimize a given function, whereas GP is geared towards program induction.<sup>21</sup>

At their core, both GP and GA operate on similar foundations. They use a population-based strategy, employ a fitness function to assess individual performance, and apply genetic operators to produce new individuals. Yet, the representation of individuals in GP, along with its unique genetic operators, differentiates it from GA.

**Remark.** While GP is primarily a fitness-centric search in the set of computer programs, its fundamental nature is optimization, echoing the core of GA.

Within the population of GP, every individual represents a computer program, built from a library of primitives, distinguished as **functions** and **terminals**. One way to represent this is as a composite pattern where functions are internal nodes, and terminals are the leaves. This structure is commonly referred to as an abstract syntax tree (AST). For a visual representation, refer to fig. 2.2.

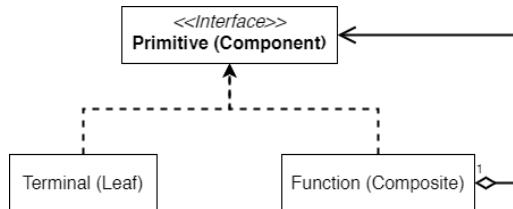


Figure 2.2: Representation of GP individuals as a composite structure.

The GP algorithm, delineated in listing 2.2, bears structural resemblance to the GA algorithm. The key differentiator rests in the specific genetic operations tailored for GP.

Listing 2.2– Structural blueprint of the Genetic Programming algorithm, underscoring its architectural alignment with the Genetic Algorithm.

```

var population = recursively construct random programs
population.forEach { it -> execute it then evaluate it }
while (termination criteria is not met) {
    val survivors = select survivors from population
    val parents = select parents from population
    val offspring = alter parents
    offspring.forEach { it -> execute it then evaluate it }
    population = survivors + offspring
}
return fittest from population
  
```

In the following sections, we will delve into the key components of the GP algorithm, clarifying these concepts with practical examples.

<sup>21</sup>A more detailed definition of program induction can be found in definition A.20 on page 97.

### 2.3.1 Representation and Evaluation

#### 2.3.1.1 Representation

As with GAs, the representation of the individuals is one of the most important aspects of GP. The representation is the encoding of potential solutions to the problem into a form that can be manipulated,<sup>22</sup> executed and evaluated by the algorithm.

Various methods exist for program representation, such as utilizing an abstract syntax tree, a linear sequence of instructions, a stack of instructions, or a combination of these approaches. However, the most classical representation is the ***tree representation***, where the program is represented as a composite data structure like the one shown in the introduction to this section.

Let's illustrate this with an example problem: given a set of  $n$  points in the plane, find the curve that best fits the points. This is a very common problem in statistics, and it is known as the ***symbolic regression*** problem [7]. In this example, our goal is to use symbolic regression to approximate the function

$$f(x) = 5x^3 - 2x^2 + \sin(x) - 7; x \in [-1, 1] \quad (2.4)$$

using this function, we can generate a set of points that lies on the curve as shown in fig. 2.3 and table 2.11 on the following page.

The next step in preparing our GP setup is to define the primitive set, which includes the functions and terminals that the algorithm can use to construct candidate solutions. In this case, we will use the following set of functions and terminals:

- **Functions:** 1. + (Addition) 2. – (Subtraction) 3. × (Multiplication) 4. / (Division) 5. sin (Sine) 6. cos (Cosine) 7. pow (Power)
- **Terminals:** 1.  $x$  (The variable  $x$ ) 2.  $\{c \mid c \in [1, 7] \wedge c \in \mathbb{Z}\}$  (An ephemeral constant)<sup>23</sup>

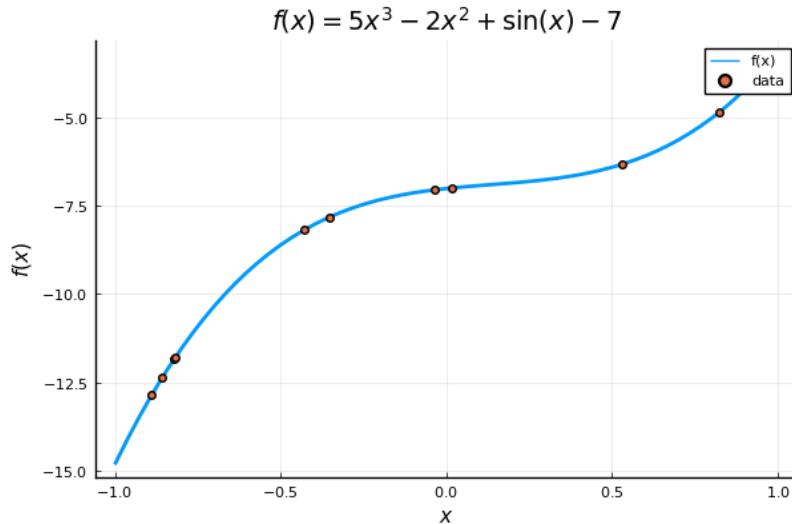


Figure 2.3: A set of points generated from the function  $5x^3 - 2x^2 + \sin(x) - 7$

Using this set of functions and terminals, we can represent the program as a tree, as shown in fig. 2.4 on the following page.

Note that this definition arises the possibility of having a program that has an infinite number of nodes, as the tree can grow indefinitely. This leads the search to be unsuccessful since the probability of finding a solution is close to zero, for example, the probability of finding a solution on the initial population would be:  $\lim_{x \rightarrow \infty} \frac{1}{x} = 0$ . To avoid this issue of potentially infinite trees, we typically impose certain ***constraints*** on the generation of the trees. The most common constraints are the ***maximum height*** of the tree and the ***maximum number of nodes*** in the tree.

<sup>22</sup>For example, by applying genetic operators.

<sup>23</sup>See definition A.6 on page 95.

$x$	$y$
0.889 160	-12.872 629
0.856 103	-12.358 361
0.821 295	-11.851 004
0.818 193	-11.807 452
0.429 859	-8.183 442
0.352 328	-7.812 033
0.035 776	-7.038 557
0.017 450	-6.983 134
0.529 010	-6.314 804
0.821 101	-4.848 557

Table 2.11: A set of points generated from the function  $5x^3 - 2x^2 + \sin(x) - 7$

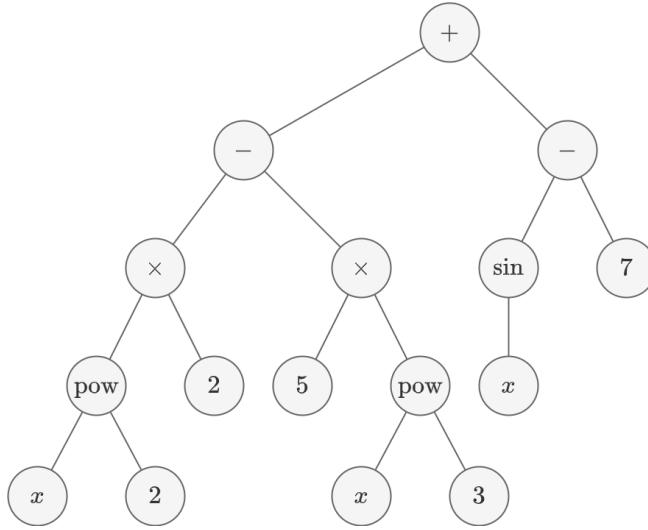


Figure 2.4: A possible tree representation of the program  $5x^3 - 2x^2 + \sin(x) - 7$

### 2.3.1.2 Search space

Using this representation, we can define the **genotype** of the individuals to contain only one **chromosome** which is composed of a single **gene**<sup>24</sup> that is the tree representation of the program. Recalling the definition of cardinality presented in definition 2.1 on page 7, we can see that the cardinality of the search space will be the number of possible trees that can be generated using the primitive set and the maximum height of the tree.

**Lemma 2.1.** Let  $\mathbb{T}_H$  be the set of all possible **labeled trees** of height  $H$ , with  $H \in \mathbb{N}$ . Given the sets  $\mathcal{T}$  and  $\mathcal{F}$  corresponding to the possible labels of terminal nodes (nodes that do not have children) and the possible labels of internal nodes (nodes that have children) respectively, the number of trees in  $\mathbb{T}_H$  is given by the following recurrence relation:

$$|\mathbb{T}_H(\mathcal{T}, \mathcal{F})| = \begin{cases} |\mathcal{T}| & \text{if } H = 0 \\ \sum_{f \in \mathcal{F}} |\mathbb{T}_{H-1}(\mathcal{T}, \mathcal{F})|^{A(f)} & \text{if } H > 0 \end{cases} \quad (2.5)$$

where  $A(f)$  is the arity of the node  $f$ .

<sup>24</sup>Although the most common representation is to have a single gene referencing the root of the tree, several variations that use multi-gene chromosomes have been proposed, such as Koza's *Automatically Defined Functions* [9], Angeline and Pollack's *Genetic Library Builder* [8, 10], or Rosca and Ballard's *Adaptive Representation* [11].

**Proof.** For the proof, we will use induction on the height of the tree. For the sake of brevity, we will use the notation  $|\mathbb{T}_H(\mathcal{T}, \mathcal{F})| = |\mathbb{T}_H|$ .

**Base case:**  $H = 0$  If the height of the tree is 0, then the tree is composed of a single node, which is a terminal node. Thus, the number of possible trees is equal to the number of possible terminal nodes, which arises:

$$|\mathbb{T}_0| = |\mathcal{T}|$$

**Base case:**  $H = 1$  If the height of the tree is 1, then the tree is composed of a root node, which is an internal node, and a set of children, which are terminal nodes. Suppose that the root node has the label  $f \in \mathcal{F}$  and arity  $A(f)$ . Since the children are terminal nodes, each child can have any of the labels in  $\mathcal{T}$ . Thus, the number of trees rooted at  $f$  is equal to the number of possible combinations of  $A(f)$  elements (with repetition and order) from the set  $\mathcal{T}$ , this is:

$$\prod_{i=1}^{A(f)} |\mathcal{T}| = |\mathcal{T}|^{A(f)}$$

Since the root node can have any of the labels in  $\mathcal{F}$ , the number of possible trees of height 1 is equal to:

$$|\mathbb{T}_1| = \sum_{f \in \mathcal{F}} |\mathcal{T}|^{A(f)}$$

**Inductive step:**  $H > 1$  Suppose the statement holds true for  $H = h$ . We aim to prove that the statement also holds true for  $H = h + 1$ .

Since a terminal node cannot have children,<sup>a</sup> each tree of height  $h + 1$  has a root with one of the labels from the set  $\mathcal{F}$ , and the remaining  $h$  layers are fully formed subtrees of height  $h$ .

For a given node label  $f \in \mathcal{F}$  with arity  $A(f)$ , each child is the root of a subtree of height  $h$ . Given our inductive assumption, there are  $|\mathbb{T}_h|$  possible such subtrees.

Since all subtrees are independent, the number of possible trees with the root  $f$  is  $|\mathbb{T}_h|^{A(f)}$ , which is the product of  $|\mathbb{T}_h|$  over the arity of  $f$ .

We can sum this quantity over all  $f \in \mathcal{F}$  to get the total number of possible trees of height  $h + 1$ :

$$|\mathbb{T}_{h+1}| = \sum_{f \in \mathcal{F}} |\mathbb{T}_h|^{A(f)}$$

□

<sup>a</sup>This could also be interpreted as a terminal node having an arity of 0, or that all terminal nodes are leaves.

**Lemma 2.2.** Let  $\mathbb{T}_{\leq H}$  be the set of all possible **labeled trees** of height  $h \leq H$ , with  $H \in \mathbb{N}$  and  $h \in \mathbb{N}$ . Given the sets  $\mathcal{T}$  and  $\mathcal{F}$  corresponding to the possible labels of terminal nodes and the possible labels of internal nodes respectively, the number of trees in  $\mathbb{T}_{\leq H}$  is given by the following recurrence relation:

$$|\mathbb{T}_{\leq H}(\mathcal{T}, \mathcal{F})| = \begin{cases} |\mathcal{T}| & \text{if } H = 0 \\ |\mathbb{T}_H(\mathcal{T}, \mathcal{F})| + |\mathbb{T}_{\leq H-1}(\mathcal{T}, \mathcal{F})| & \text{if } H > 0 \end{cases} \quad (2.6)$$

Where  $\mathbb{T}_H$  is the set of all possible trees of height  $H$ .

**Proof.** For the sake of simplicity, we will use the notation  $|\mathbb{T}_{\leq H}| = |\mathbb{T}_{\leq H}(\mathcal{T}, \mathcal{F})|$  and  $|\mathbb{T}_H| = |\mathbb{T}_H(\mathcal{T}, \mathcal{F})|$ .

The set  $\mathbb{T}_{\leq H}$  can be partitioned into two disjoint sets: the set of all possible trees of height  $H$  and the set of all possible trees of height  $h < H$ . Thus we have:

$$|\mathbb{T}_{\leq H}| = |\mathbb{T}_H| + |\mathbb{T}_{\leq H-1}|$$

□

**Theorem 2.1.** Let  $\mathbb{T}_{\leq H}$  be the set of all possible **labeled trees** of height  $h \leq H$ , with  $H \in \mathbb{N}$  and  $h \in \mathbb{N}$ . Given the sets  $\mathcal{T}$  and  $\mathcal{F}$  corresponding to the possible labels of terminal nodes and the possible labels of internal nodes respectively, the number of trees in  $\mathbb{T}_{\leq H}$  is given by the following recurrence relation:

$$|\mathbb{T}_{\leq H}(\mathcal{T}, \mathcal{F})| = \begin{cases} |\mathcal{T}| & \text{if } H = 0 \\ \left( \sum_{h=0}^{H-1} \sum_{f \in \mathcal{F}} |\mathbb{T}_h(\mathcal{T}, \mathcal{F})|^{A(f)} \right) + |\mathcal{T}| & \text{if } H > 0 \end{cases} \quad (2.7)$$

Where  $\mathbb{T}_H$  is the set of all possible trees of height  $H$  and  $A(f)$  is the arity of the node  $f$ .

**Proof.** From lemma 2.2 on the preceding page we know the number of trees of height  $H$  or less is given by:

$$|\mathbb{T}_{\leq H}| = |\mathbb{T}_H| + |\mathbb{T}_{\leq H-1}|$$

Then, by applying lemma 2.1 on page 16, we get:

$$|\mathbb{T}_{\leq H}| = \left( \sum_{f \in \mathcal{F}} |\mathbb{T}_{H-1}|^{A(f)} \right) + |\mathbb{T}_{\leq H-1}|$$

By unrolling the recurrence relation, we get:

$$\begin{aligned} |\mathbb{T}_{\leq H}| &= \left( \sum_{f \in \mathcal{F}} |\mathbb{T}_{H-1}|^{A(f)} \right) + |\mathbb{T}_{\leq H-1}| \\ &= \left( \sum_{f \in \mathcal{F}} |\mathbb{T}_{H-1}|^{A(f)} \right) + \left( \sum_{f \in \mathcal{F}} |\mathbb{T}_{H-2}|^{A(f)} \right) + \dots + \left( \sum_{f \in \mathcal{F}} |\mathbb{T}_1|^{A(f)} \right) + \left( \sum_{f \in \mathcal{F}} |\mathbb{T}_0|^{A(f)} \right) + |\mathbb{T}_0| \\ &= \left( \sum_{h=0}^{H-1} \sum_{f \in \mathcal{F}} |\mathbb{T}_h(\mathcal{T}, \mathcal{F})|^{A(f)} \right) + |\mathcal{T}| \end{aligned}$$

□

Given a program viewed as a **labeled tree**, we can determine the cardinality of the genetic programming algorithm's search space for a specific maximum height  $H$ . Considering the sets  $\mathcal{T} = \{x, c\}$  and  $\mathcal{F} = \{+, -, \times, /, \text{sin}, \text{cos}, \text{exp}, \text{log}\}$ :

$$A(f) = \begin{cases} 2 & \text{for } f \in \{+, -, \times, /, \text{pow}\} \\ 1 & \text{for } f \in \{\text{sin}, \text{cos}\} \end{cases}$$

For an AST with a target program height of 4, we can set 5 as the maximum height for the programs in our search space, offering some flexibility to the generated programs. Given that  $c$  has 7 potential values,  $|\mathcal{T}| = 8$ , the search space's cardinality can be calculated accordingly.

$$\begin{aligned} |\mathcal{S}| &= |\mathbb{T}_{\leq 5}(\mathcal{T}, \mathcal{F})| = \left( \sum_{h=0}^4 \sum_{f \in \mathcal{F}} |\mathbb{T}_h(\mathcal{T}, \mathcal{F})|^{A(f)} \right) + 8 \\ &\approx 9.531\,142 \times 10^{37} \\ &\approx 10^{38} \end{aligned}$$

Thus, the search space of the genetic programming algorithm is of the order of  $10^{38}$  programs.<sup>25</sup> It should be easy to see that the size of the search space make it unfeasible to perform an exhaustive search.<sup>26</sup>

This is why we need to use a heuristic search algorithm, such as genetic programming. In fig. 2.5 we can see how the number of trees of height less or equal to  $h$  rapidly increases as  $h$  increases.



Figure 2.5: Total number of trees of height less or equal to  $h$  for  $h \in \{0, \dots, 5\}$  and  $\mathcal{T} = \{x, c\}$  and  $\mathcal{F} = \{+, -, \times, /, \sin, \cos, \text{pow}\}$ . Note that the Y axis is in logarithmic scale.

### 2.3.1.3 Evaluation

To determine a program's fitness, various methods exist, but a prevalent choice is the **mean squared error** (MSE) between the points and the program.

**Definition 2.8** (Mean Squared Error). *Given a vector of  $n$  predictions derived from  $n$  data points across all variables, where  $y_i$  represents the  $i$ -th observed value and  $\hat{y}_i$  is the  $i$ -th prediction, the MSE of the predictor is a function  $\text{MSE} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  defined by:*

$$\text{MSE}(\mathbf{y}, \hat{\mathbf{y}}) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.8)$$

The MSE is a benchmark for assessing an estimator's quality, especially in *machine learning* contexts.<sup>27</sup>

For our purposes, we employ the MSE to assess a program's fitness. Given a program  $P$  and two point sets,  $\mathbf{x}$  and  $\mathbf{y}$ , as detailed in table 2.11 on page 16, and presuming  $P[\mathbf{x}]$  as the points produced by evaluating  $P$  on  $\mathbf{x}$ , with  $P(x)$  being the evaluation result for point  $x$ , the fitness of  $P$  can be described as:

$$\phi_P = \text{MSE}(\mathbf{y}, P[\mathbf{x}]) = \frac{1}{n} \sum_{i=1}^n (y_i - P(x_i))^2 \quad (2.9)$$

In summation, when gauging a program's fitness in GP, error measures play a pivotal role, with MSE emerging as a primary standard. By contrasting a program's predictions with actual data, the MSE offers a valuable assessment of prediction accuracy. Its inherent simplicity effectively establishes a quality standard for future enhancements. Such an organized methodology guarantees that GP systematically traverses the solution domain, persistently optimizing candidate solutions.

<sup>25</sup>This value was computed using the script shown in listing C.1 on page 115.

<sup>26</sup>If we approximate the time to evaluate a program to be 1 nanosecond, it would take approximately 3 sextillion years to evaluate all the programs in the search space.

<sup>27</sup>While we favor the mean squared error for its straightforward nature and prominence in literature, other error measures like the *mean absolute error* (MAE) or the *cross-entropy* (CE) loss functions can also be considered.

### 2.3.2 Initialization

As with other evolutionary algorithms, the algorithm starts by generating a population of random individuals.

There are many ways to generate random individuals, but a very common (and simple) method is the *grow method* [7]. In this method, a maximum height is defined, and the algorithm then creates a random tree with a given minimum height and a maximum height.

**Remark.** A tree with a height of 0 is a tree with only one node, the root.

The grow method then proceeds to recursively generate the trees with random nodes until a terminal node is selected or the maximum height is reached. This method is shown in listing 2.3.

Listing 2.3– The grow method for generating random trees.

```
// Given:
val lo: N // Minimum height
val hi: N // Maximum height
require(lo ≤ hi)
val n = random.int(lo..hi) // Random height
fun grow(ts: Set<Terminal>, fs: Set<Function>, depth: N) {
    require(ts != ∅ && fs != ∅)
    val c = ∅
    if (
        depth == n || (depth ≥ lo && random.double(0..1) ≤ ts.size / (ts.size + fs.size))
    ) {
        // d = n ∨ (d ≥ l ∧ random() < |T| / (|T| + |F|))
        return random node from ts
    } else {
        val f = random node from fs
        for (i in 1..arity(f)) {
            c += grow(ts, fs, d + 1)
        }
        return a tree with f as root and c as children
    }
}
```

With this method, the algorithm can generate trees where the size of the longest path from the root to a leaf is a number  $n \in [l, h]$ .

Now that the algorithm can generate random trees, it can generate a random population of trees by generating a random tree for each individual in the population. Assuming a population size of  $p = 4$ , a maximum height of  $h = 3$ , a minimum height of  $l = 1$ , and the primitives set defined in the previous section, the algorithm could generate the population:  $\mathbf{P} = \{\mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3, \mathbf{I}_4\} = \left\{ \frac{3}{\sin(2)} \times 5^3, 7 - (5 + \sin(x)), 7 + 2, 5x^2 \right\}$ , as shown in fig. 2.6 on the facing page.

The next step is to calculate the fitness of each individual in the population. If we recall, the fitness function is the MSE between the expected output and the actual output of the individual.



Figure 2.6: A population of random trees.

$$\begin{aligned}
 \text{MSE}(\mathbf{y}, \mathbf{I}_1[\mathbf{x}]) &= \frac{1}{10} \sum_{i=1}^{10} (\mathbf{y}_i - \mathbf{I}_1(\mathbf{x}_i))^2 = \frac{1}{10} \sum_{i=1}^{10} \left( \mathbf{y}_i - \frac{3}{\sin(2)} \cdot 5^3 \right)^2 \\
 &\approx 177\,596.851\,131 \\
 \text{MSE}(\mathbf{y}, \mathbf{I}_2[\mathbf{x}]) &= \frac{1}{10} \sum_{i=1}^{10} (\mathbf{y}_i - \mathbf{I}_2(\mathbf{x}_i))^2 = \frac{1}{10} \sum_{i=1}^{10} (\mathbf{y}_i - 7 - (5 + \sin(x_i)))^2 \\
 &\approx 137.398\,836 \\
 \text{MSE}(\mathbf{y}, \mathbf{I}_3[\mathbf{x}]) &= \frac{1}{10} \sum_{i=1}^{10} (\mathbf{y}_i - \mathbf{I}_3(\mathbf{x}_i))^2 = \frac{1}{10} \sum_{i=1}^{10} (\mathbf{y}_i - (7 + 2))^2 \\
 &\approx 331.924\,267 \\
 \text{MSE}(\mathbf{y}, \mathbf{I}_4[\mathbf{x}]) &= \frac{1}{10} \sum_{i=1}^{10} (\mathbf{y}_i - \mathbf{I}_4(\mathbf{x}_i))^2 = \frac{1}{10} \sum_{i=1}^{10} (\mathbf{y}_i - 5x^2)^2 \\
 &\approx 138.079\,865
 \end{aligned}$$

With this, we can assign a fitness to each individual in the population as shown in table 2.12 on the next page. A summary of the population's fitness is shown in table 2.13 on the following page.

We can observe that the worst individual has an error significantly larger than the best individual. This is to be expected, as the MSE is a measure of the error that escalates exponentially with the difference between the expected and actual output.

Generation 0			
Individual	Program	Height	Fitness
$I_1(x)$	$\frac{3}{\sin(2)} \cdot 5^3$	3	177 596.851 131
$I_2(x)$	$7 - (5 + \sin(x))$	3	137.398 835
$I_3(x)$	$7 + 2$	1	331.924 267
$I_4(x)$	$5x^2$	2	138.079 865

Table 2.12: Initial population of the genetic programming algorithm

	Fitness	Individual
Best	137.398 835	$I_2(x)$
Worst	177 596.851 131	$I_1(x)$
Average	44 551.063 525	
Standard deviation	76 814.062 197	

Table 2.13: Fitness of the individuals in generation 0

A graphical representation of the population is shown in fig. 2.7. It is clear from the figure that the worst individual is  $I_1$ .

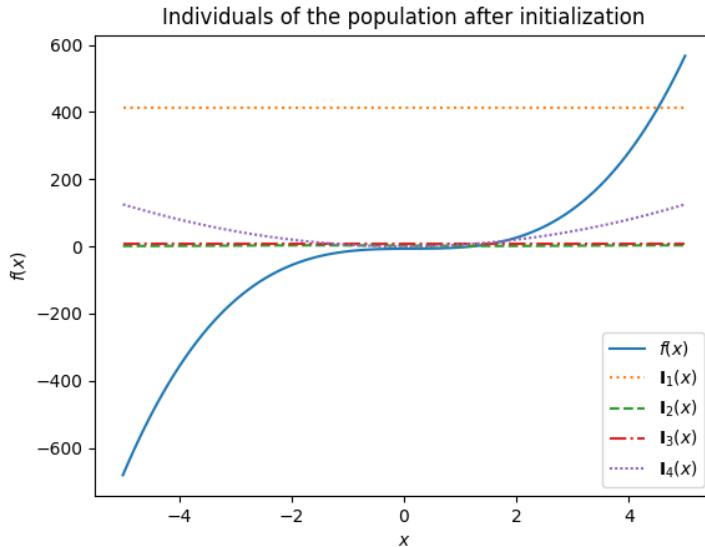


Figure 2.7: Graphical representation of the population in generation 0 compared to the expected output.

In summary, the initialization phase plays a pivotal role in setting the foundational landscape for the Genetic Programming process. Through randomized generation methodologies, like the grow method, the algorithm nurtures a diverse pool of candidate solutions. This diverse initial population enhances the algorithm's probability of exploring different regions of the solution space, thereby increasing the odds of locating an optimal solution. Yet, these generated programs are merely starting points. Their evaluation, determined through metrics such as the MSE, provides an empirical compass to guide the evolution process. Once the fitness landscape has been charted out, the algorithm is primed for the next pivotal phase: selection. In the upcoming section, we delve into the mechanics of selection, a process crucial for ensuring the survival and propagation of the fittest candidates.

### 2.3.3 Selection

The selection process in GP is similar to that of GAs. However, unique modifications have been proposed for the standard process in GP, which consider the semantics, or the functional behavior, of the programs being evolved [31]. This work will not delve into this topic, as it is beyond the scope of this document.

For the problem under study, we use a selection process similar to the one used in the *One Max* problem (a simple optimization problem where the objective is to maximize the number of ones in a binary string). The equation for calculating the selection probability will diverge from eq. (2.3) on page 9, as it presumes the fittest individual is the one with the highest fitness value. In contrast, the symbolic regression problem, which is our focus, considers the individual with the lowest fitness value as the fittest.

For this particular case of the symbolic regression problem, we adjust our approach to selection. We introduce a *corrected fitness function*,  $\phi'$ , defined as:

$$\phi'(I) = \left( \sum \Phi_{\mathbf{P}} \right) - \phi_I \quad (2.10)$$

Here,  $\phi_I$  signifies the fitness of individual  $I$ , and  $\Phi_{\mathbf{P}}$  represents the *batch fitness function* defined in definition 2.3 on page 7 applied to the population  $\mathbf{P}$ . We then define the selection probability for an individual  $\mathbf{P}_i$  as:

$$p_i = \frac{\phi'(\mathbf{P}_i)}{\sum_{j=1}^N \phi'(\mathbf{P}_j)} \quad (2.11)$$

In this equation,  $N$  stands for the size of the population.

With this methodology, we calculate the selection probabilities for the population as illustrated in table 2.14. The outcome shows that the individual with the highest error has a considerably low probability of being selected, while the other individuals have roughly equal chances.

Individual	Fitness	Selection probability
$\mathbf{I}_1(x)$	607 402 968	0.113 615%
$\mathbf{I}_2(x)$	178 066.855 263	33.307 633%
$\mathbf{I}_3(x)$	177 872.329 832	33.271 246%
$\mathbf{I}_4(x)$	178 066.174 234	33.307 505%

Table 2.14: Selection probabilities for the symbolic regression problem.

Assuming a *survival rate* of 50%, let's consider that the selection process favors  $I_2$  and  $I_3$  as survivors due to their higher selection probabilities (as shown in the previous table). In this scenario,  $I_1$  and  $I_4$  are identified as the ones to be replaced by the offspring in the next generation. A comparison between the survivors and the target function is shown in fig. 2.8 on the next page.

Selection remains a cornerstone of evolutionary algorithms, ensuring that individuals possessing beneficial traits have a greater likelihood of transferring those traits to subsequent generations. While nuances distinguish the selection processes of Genetic Programming (GP) and Genetic Algorithms (GAs), the underlying principle remains: to direct the search towards more promising regions of the solution space. As we progress further, it's essential to recognize the interplay between selection and other operators, understanding that their combined effects mold the trajectory of the evolutionary process. The insights gleaned from this understanding will aid in fine-tuning algorithms and achieving better results in future endeavors.

### 2.3.4 Variation

Analogous to GA, the individuals in a GP population undergo a variation process. For the symbolic regression problem under consideration, we introduce variation using two specific operators: *crossover* and *mutation*.

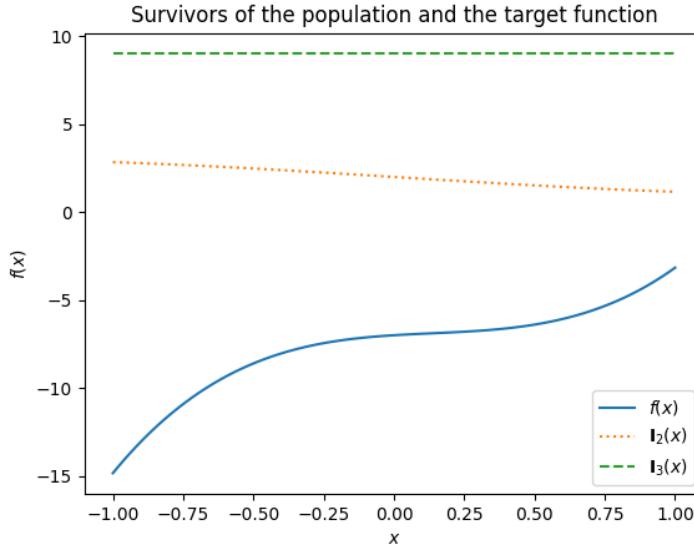


Figure 2.8: Comparison between the survivors and the target function.

#### 2.3.4.1 Crossover

Variation operators in GP must maintain the syntactic correctness of the programs or individuals. For crossover, this implies that the resultant offspring must be syntactically correct programs.

The crossover operator used in GAs, described in section 2.2.4.1 on page 10, is not typically suitable for GP, as it does not guarantee the syntactic correctness of the resulting individuals. Though there could be instances where the crossover operator used in GAs is applicable to GP, as depicted in chapter 7 on page 83, these are not common scenarios.

The choice of operator in GP depends on the representation of the individuals.

For tree-based GP, the fundamental crossover operator is the **subtree crossover**, referenced in section 4.5.2.1 on page 65.

This operator selects a random node from each parent and exchanges the subtrees rooted at these nodes. Usually, a constraint similar to the one used for generating the initial population is applied to this operator to prevent the creation of overly large trees.

Assuming we select two individuals,  $\mathbf{I}_1$  and  $\mathbf{I}_2$ , from the population, the subtree crossover operator chooses a random node from each individual, say  $\clubsuit = 7$  from  $\mathbf{I}_1$  and  $\diamond = x^2$  from  $\mathbf{I}_2$ . The subtrees rooted at these nodes are then interchanged, as shown below:

$$\begin{aligned} \chi(\mathbf{I}_2, \mathbf{I}_4) &= \chi(\clubsuit - (5 + \sin(x)), 5\diamond) \\ &= (\diamond - (5 + \sin(x)), 5\clubsuit) \\ \Leftrightarrow (\mathbf{O}_1, \mathbf{O}_2) &= (x^2 - (5 + \sin(x)), 5 \cdot 7) \end{aligned}$$

where  $\chi$  signifies the subtree crossover operator between two individuals. The crossover of  $\mathbf{I}_1$  and  $\mathbf{I}_2$  is depicted in fig. 2.9 on the next page.

Following the application of the subtree crossover operator, the fitness of the individuals in the population is evaluated as shown in table 2.15 on the facing page. A summary of the population's fitness is given in table 2.16 on the next page.



Figure 2.9: Crossing over of  $I_2 = 7 - (5 + \sin(x))$  and  $I_4 = 5x^2$ , producing  $O_1 = x^2 - (5 + \sin(x))$  and  $O_2 = 5 \cdot 7$ .

Generation 1		
Individual	Program	Fitness
$I_2(x)$	$7 - (5 + \sin(x))$	137.398 835
$I_3(x)$	$7 + 2$	331.924 267
$O_1(x)$	$5 \cdot 7$	1944.288 127
$O_2(x)$	$x^2 - (5 + \sin(x))$	33.740 766

Table 2.15: Population after applying the subtree crossover operator.

	Fitness	Individual
Best	33.740 766	$O_2(x)$
Worst	1944.288 127	$O_1(x)$
Average	611.837 999	
Standard deviation	776.701 997	

Table 2.16: Fitness summary of the population after applying the subtree crossover operator.

A notable improvement in the population's fitness is observed after the application of the subtree crossover operator. Comparing the results from table 2.13 on page 22, we find that the average fitness (or error) has dropped from 44 551.063 525 to 611.837 999, equating to an improvement of approximately 98.627%.

$$\frac{\bar{\Phi}_i - \bar{\Phi}_X}{\bar{\Phi}_i} = \frac{44\,551.063\,525 - 611.837\,999}{44\,551.063\,525} \approx 98.627\%$$

where  $\bar{\Phi}_i$  is the average fitness of the population after initialization and  $\bar{\Phi}_X$  is the average fitness of the population after applying the subtree crossover operator.

A reduction in the population's fitness standard deviation, from 76 814.062 197 to 776.701 997, is also seen. This decrease of around 98.989% indicates that the population's diversity has reduced.

$$\frac{\sigma_i - \sigma_X}{\sigma_i} = \frac{76\,814.062\,197 - 776.701\,997}{76\,814.062\,197} \approx 98.989\%$$

where  $\sigma_i$  is the standard deviation of the fitness of the population after initialization and  $\sigma_X$  is the standard deviation of the fitness of the population after applying the subtree crossover operator.

In the following section, we discuss the mutation operator, which can be used to introduce diversity into the population, thereby preventing premature convergence.

#### 2.3.4.2 Mutation

Mutation, another crucial genetic operator, introduces new genetic material into the population, thus maintaining genetic diversity and preventing premature convergence to suboptimal solutions. In the context of GP, the mutation operator modifies a program in the population, while ensuring the resultant individual's syntactic correctness.

In tree-based GP, a common form of mutation is the ***point mutation*** [21, 69], which selects a random node from an individual and replaces it with a random primitive with the same arity. This operator is similar to the bit-flip mutation operator used on section 2.2.4.2 on page 11. As with bit-flip mutation, point mutation can also be applied with a certain probability to each node in an individual, meaning that more than one node can be mutated in a single individual.

Suppose we mutate all the individuals in the population resulting from the crossover operation in section 2.3.4.1 on page 24, and that exactly one node is mutated in each individual. Let the selected nodes be  $\clubsuit = \sin$  in  $I_2$ ,  $\spadesuit = 7$  in  $I_3$ ,  $\heartsuit = 5$  in  $O_1$ , and  $\diamondsuit = 2$  in  $O_2$ . Then, a possible result of applying the point mutation operator can be:

$$M \left( \begin{bmatrix} I_2 \\ I_3 \\ O_1 \\ O_2 \end{bmatrix} \right) = M \left( \begin{bmatrix} 7 - (5 + \clubsuit(x)) \\ \spadesuit + 2 \\ \heartsuit \cdot 7 \\ x^{\diamondsuit} - (5 + \sin(x)) \end{bmatrix} \right) = \left( \begin{bmatrix} 7 - (5 + \clubsuit'(x)) \\ \spadesuit' + 2 \\ \heartsuit' \cdot 7 \\ x^{\diamondsuit'} - (5 + \sin(x)) \end{bmatrix} \right) = \left( \begin{bmatrix} 7 - (5 + \cos(x)) \\ 6 + 2 \\ 6 \cdot 7 \\ x^3 - (5 + \sin(x)) \end{bmatrix} \right)$$

Here,  $\clubsuit'$ ,  $\spadesuit'$ ,  $\heartsuit'$ , and  $\diamondsuit'$  are random primitives with the same arity as  $\clubsuit$ ,  $\spadesuit$ ,  $\heartsuit$ , and  $\diamondsuit$ , respectively. That being,  $\clubsuit' = \cos$ ,  $\spadesuit' = 6$ ,  $\heartsuit' = 6$ , and  $\diamondsuit' = 3$ .

The fitness of the individuals in the population after applying the subtree mutation operator is then evaluated. The results of this fitness evaluation are shown in ?? on page ???. A summary of the fitness of the population is presented in ?? on page ??.

Generation 1		
Individual	Program	Fitness
$M(I_2(x))$	$7 - (5 + \cos(x))$	112.297 411
$M(I_3(x))$	$6 + 2$	296.910 273
$M(O_1(x))$	$6 \cdot 7$	2 609.386 088
$M(O_2(x))$	$x^3 - (5 + \sin(x))$	15.295 863

Table 2.17: Population after applying the node mutation operator.

	Fitness	Individual
Best	15.295 863	$M(\mathbf{O}_2(x))$
Worst	2 609.386 088	$M(\mathbf{O}_1(x))$
Average	758.472 409	
Standard deviation	1 073.402 832	

Table 2.18: Fitness summary of the population after applying the node mutation operator.

Just like the crossover operator, mutation can also significantly influence the fitness and diversity of the population. By generating new structures in the population, mutation can help prevent stagnation and maintain diversity, thus avoiding premature convergence to suboptimal solutions.

With this, we can conclude that after at the end of the generation, the population shows an improvement of 98.298% in fitness, and 98.603% in standard deviation.

$$\frac{\bar{\Phi}_i - \bar{\Phi}_M}{\bar{\Phi}_i} = \frac{(44\,551.063\,525 - 758.472\,409)}{44\,551.063\,525} \approx 98.298\%$$

$$\frac{\sigma_i - \sigma_M}{\sigma_i} = \frac{(76\,814.062\,197 - 1\,073.402\,832)}{76\,814.062\,197} \approx 98.603\%$$

fig. 2.10 shows the population after applying the node mutation operator. Two points should be clear from this figure. First, all individuals of the population are closer to the target function than the individuals in the initial population. Second, the new fittest individual is  $\mathbf{O}_2$ , this individual has a shape similar to the target function.

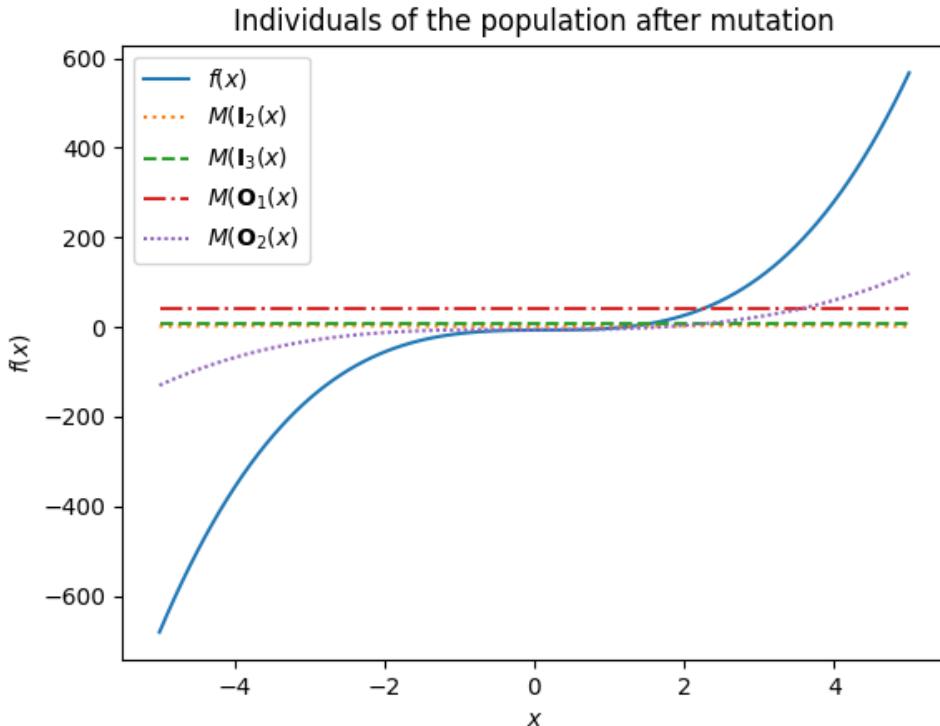


Figure 2.10: Population after applying the node mutation operator

In summary, mutation plays a pivotal role in the genetic programming process. It injects fresh genetic material into the population, thereby promoting diversity and ensuring the continual exploration of the search space. As demonstrated, mutation not only alters the structure of individuals but can also lead to significant enhancements in fitness. By tactfully combining crossover and mutation operations, genetic programming can efficiently traverse the vast landscape of possible solutions. This underscores the importance of maintaining a delicate balance between exploration (introducing new genetic structures) and exploitation (refining existing successful solutions).

### 2.3.5 The Generalization Problem in GP

#### 2.3.5.1 Understanding Generalization

Generalization in Genetic Programming (GP) relates to a program's proficiency in handling unseen data, not solely the data it evolved with. A truly generalized program goes beyond rote memorization of training data. It discerns underlying patterns, equipping it to make precise predictions or decisions on new, unencountered instances.

#### 2.3.5.2 Significance of Generalization in GP

Generalization holds a paramount position in GP. The overarching objective of evolving programs is to address real-world challenges. A program's utility diminishes if it can't generalize, limiting its effectiveness to familiar data. Particularly in predictive modeling, the emphasis isn't on a model's ability to recall past data, but on its precision with upcoming, unpredictable data. Therefore, the act of generalizing surpasses basic accuracy measures on training datasets, signaling a model's resilience and practical relevance.

#### 2.3.5.3 The Overfitting Dilemma

Overfitting stands as an adversary to generalization. A program succumbs to overfitting when it molds excessively to training data, picking up on its noise and irregularities rather than its broad trends. Such overfitted programs may boast of stellar accuracy on training datasets but falter with unseen data. Within the GP context, visualizing overfitting can be akin to observing a program that's evolved into a convoluted form, brimming with redundant branches or logic catering to the peculiarities of training data.

#### 2.3.5.4 Addressing the Generalization Problem

The academic landscape has brought forth numerous techniques to counter the generalization conundrum in GP, some such works include:

- Chen et al.'s approach hinged on feature selection tailored for high-dimensional symbolic regression [33].
- Kushchuk advocated the adoption of multiple fitness functions, a measure aimed at bolstering generalization [15].
- Enhancing generalization through the Rademacher distribution was a strategy put forth by Chen et al. [38].

While this thesis will sidestep an in-depth exploration of the generalization issue in GP, it's pivotal to acknowledge its profound implications in GP and its standing as a fervently pursued research frontier.

## Chapter 3

# State of the Art: Relevant Work

This chapter delves into the current state of the art in evolutionary computation frameworks. While evolutionary algorithms themselves aren't new, their applications keep evolving<sup>1</sup> with technological advancements in programming languages and tools.

We'll spotlight five major frameworks: *Agile Artificial Intelligence in Pharo*, *DEAP*, *Jenetics*, *ECJ*, and *GeneticSharp*. Each section provides an overview of the respective framework, complete with code samples, showcasing their syntax and distinct features. This exploration aims to give readers a broad understanding of the array of tools in this domain.

Moreover, this review sets the context for our forthcoming contribution: an evolutionary computation framework crafted in Kotlin. To relate, listing C.2 on page 115 features the *Keen* framework's approach to the *OneMax* problem, akin to the examples herein.

While we strive for objectivity, we recognize that personal biases may color our perceptions. We don't position ourselves as experts in the covered frameworks or languages, and there's a possibility that the provided implementations could be optimized further. We invite readers to delve into these frameworks firsthand and shape their own conclusions.

For transparency, it's worth noting that this review is impartial. We haven't received any compensation from, nor are we affiliated with, any of the framework developers aside from Bergel's *Agile Artificial Intelligence in Pharo* book. This chapter reflects our individual insights and experiences.

### 3.1 The One Max Problem

For the purpose of illustrating the use of the different frameworks, we will use the *One Max problem* introduced on section 2.2.1 on page 6.

The *One Max problem* is a classic and straightforward optimization problem often used as a benchmark in the study of evolutionary algorithms and other heuristic search methods. It serves as a deceptively simple yet effective test of an algorithm's optimization ability.

Given a binary string of length  $n$ , the *One Max problem* is to find a binary string such that the sum of its bits (counting the number of ones) is maximized. In formal terms, if we denote the binary string as  $x = (x_1, x_2, \dots, x_n)$  where each  $x_i \in \{0, 1\}$  for all  $i = 1, 2, \dots, n$ , the *fitness function*  $\phi(x)$  to be maximized can be expressed as:

$$\phi(x) = \sum_{i=1}^n x_i \tag{3.1}$$

The function  $\phi(x)$  counts the number of ones in the string  $x$ . The maximum possible value of  $\phi(x)$  is  $n$ , which is achieved when all bits in the string are one. The *One Max problem* is an instance of a unimodal problem since there's only one local maximum which is also a global maximum.

---

<sup>1</sup>No pun intended.

It's important to note that despite its simplicity, the *One Max problem* does provide a non-trivial task for many search algorithms. For a binary string of length  $n$ , there are  $2^n$  possible solutions. For larger  $n$ , an exhaustive search of the solution space is not feasible, hence the need for efficient optimization algorithms.

Due to its characteristics, the *One Max problem* is often used to evaluate the performance of optimization algorithms especially genetic and evolutionary algorithms. It's particularly well-suited for genetic algorithms as the operations of crossover and mutation can directly change the number of ones in a binary string, thus impacting the fitness of a potential solution.

Despite the straightforward objective function, the *One Max problem* is invaluable in the study of heuristic search methods due to its accessibility, simplicity and the vastness of its search space, which permits the analysis and comparison of the performance of different optimization techniques.

## 3.2 Agile Artificial Intelligence in Pharo

Alexandre Bergel's “*Agile Artificial Intelligence in Pharo*” [36] delivers a comprehensive exploration of genetic algorithms, encapsulating their theory and application within the *Pharo* programming environment. This language, dynamic and reflective, finds its roots in *Smalltalk*. The source code of the framework is licensed under a *freeware license* and is available on *GitHub*.<sup>2</sup>

Bergel's tome not only introduces readers to genetic algorithms but also to neural networks and the concept of neuroevolution. The framework's design and the reasoning behind it are meticulously detailed, offering readers insights into its three main components: **genetic operators**, **selection operators**, and the **evolution engine**. Accompanying this detailed framework are test cases. These serve dual roles: exemplifying the framework's application and guiding enthusiasts in developing their own robust and extensible frameworks.

The book also provides utilities for the visualization of genetic algorithms and neural networks. These are implemented using *Roassal*, a visualization engine for *Pharo*. The engine is capable of rendering a variety of graphs, including directed and undirected graphs, trees, and more relevant to this document, a summary of the evolution fitness over time.

To showcase the framework's utility, the OMP problem resolution via a genetic algorithm is provided:

Listing 3.1– A simple genetic algorithm using Bergel's framework.

```

1 engine := GAEngine new.
2 engine random: (Random seed: 11).
3 engine populationSize: 20.
4 engine numberOfGenes: 20.
5 engine createGeneBlock: [ :rand :index :ind |
6   (0 to: 1) atRandom: rand ].
7 engine fitnessBlock: [ :ind | ind count: [ :each | each = 1 ] ].
8 engine endIfFitnessIsAbove: 20.
9 engine run.
10 engine trace: 'Target fitness reached at generation '.
11 engine traceCr: engine logs last generationNumber.
12 engine trace: 'Best individual is: '.
13 engine traceCr: engine logs last fittestIndividual genes .
14 engine trace: 'with fitness: '.
15 engine traceCr: engine logs last bestFitness .

```

Breaking it down:

- 1 A new genetic algorithm engine is created.
- 2 The random seed is set to 11.
- 3-4 The population and number of genes are both set to 20.
- 5-6 The gene creation block is outlined for initial population generation, returning a random digit either 0 or 1.

<sup>2</sup><https://github.com/Apress/agile-ai-in-pharo/tree/master>

- 7 The fitness block is then detailed, here counting the number of 1s in an individual.
  - 8 The algorithm's termination condition is set, halting if an individual's fitness surpasses 20.
  - 9 The algorithm is then executed.
- 10-18 The algorithm's logs are then printed to the console. Note that the engine keeps a log of the evolution process.

While Bergel's framework, as presented in the book, stands as a clear and potent genetic algorithm tool for *Pharo*, its primary intent is instructional. As such, it may not rival the robustness of seasoned frameworks like DEAP or *Jenetics*.

### 3.3 Distributed Evolutionary Algorithms in Python (DEAP)

**Distributed Evolutionary Algorithms in Python** (DEAP) [52] is a powerful evolutionary computation framework designed for rapid prototyping and validation of concepts. It stands apart from many other evolutionary computation libraries due to its significant modularity and versatility, which enables the construction of a broad range of evolutionary algorithms, genetic algorithms, and even hybrid algorithms. DEAP is open-source and available under the *GNU Lesser General Public License v3.0* (LGPL-3.0) [57].

DEAP is structured around two primary components: the *Creators* and the *Toolbox*. The *Creators* module facilitates the generation of new classes integral to the genetic algorithm, such as individuals and populations. In contrast, the *Toolbox* serves as a comprehensive repository for various operators necessary in evolutionary algorithms, including the evaluation, selection, mutation, and crossover functions.

The following illustrates the use of DEAP to solve the OMP problem using a genetic algorithm:

Listing 3.2– A simple genetic algorithm using DEAP.

```

1 creator.create("FitnessMax", base.Fitness, weights=(1.0,))
2 creator.create("Individual", list, fitness=creator.FitnessMax)
3 toolbox = base.Toolbox()
4 toolbox.register("attr_bool", random.randint, 0, 1)
5 toolbox.register("individual", tools.initRepeat, creator.Individual, toolbox.attr_bool, n=
   ↵ 20)
6 toolbox.register("population", tools.initRepeat, list, toolbox.individual)
7 toolbox.register("evaluate", lambda i: (sum(i),))
8 toolbox.register("mate", tools.cxTwoPoint)
9 toolbox.register("mutate", tools.mutFlipBit, indpb=0.05)
10 toolbox.register("select", tools.selTournament, tournsize=3)
11 TARGET_FITNESS = 20
12 if __name__ == "__main__":
13     random.seed(11)
14     pop = toolbox.population(n=20)
15     hof = tools.HallOfFame(1)
16     stats = tools.Statistics(lambda i: i.fitness.values)
17     stats.register("max", max)
18     gen = 0
19     while True:
20         gen += 1
21         offspring = algorithms.varAnd(pop, toolbox, cxpb=0.5, mutpb=0.2)
22         fits = toolbox.map(toolbox.evaluate, offspring)
23         for fit, ind in zip(fits, offspring):
24             ind.fitness.values = fit
25         pop = toolbox.select(offspring, k=len(pop))
26         hof.update(pop)
27         record = stats.compile(pop)

```

```

28     if record['max'][0] >= TARGET_FITNESS:
29         break
30     print(f"Target fitness reached at generation {gen}.")
31     print(f"Best individual is: {''.join(map(str, hof[0]))}")
32     print(f"with fitness: {hof[0].fitness.values[0]}")

```

When executed, the program outputs the following:

```

Target fitness reached at generation 12.
Best individual is: 11111111111111111111
with fitness: 20.0

```

This example demonstrates the use of DEAP to implement a simple genetic algorithm, with the added functionality to stop evolution once a target fitness is achieved.

Here is a breakdown of the code:

- 1-2 DEAP uses a meta-factory to *dynamically* create user-defined classes. Here, we create two classes: `FitnessMax` for the fitness (single-objective, to be maximized) and `Individual` to represent an individual in the population. The second argument to `create` is the base class, meaning that `FitnessMax` is a subclass of `base.Fitness` and `Individual` is a subclass of `list`.
- 3 The `Toolbox` is created, it is used to store most of the functions and arguments required to perform the genetic algorithm.
- 4-5 We register two functions in the toolbox: `attr_bool` which generates a random binary number (0 or 1), and `individual` which creates a new individual using the `attr_bool` function. The individual consists of 20 binary numbers (genes).
- 6-7 We then register two more functions: `population` which creates a population of individuals, and `evaluate` which evaluates an individual's fitness as the sum of its genes.
- 8-10 We register three more functions in the toolbox: `mate` for performing *two-point crossover*, `mutate` with a *bit-flip mutator* with a 5% probability, and `select` for performing tournament selection with a tournament size of 3.
- 11 We set the target fitness to 20 (since we aim to maximize the sum of the genes, the maximum possible fitness is 20)
- 12-32 In the main section of the code, we seed the random number generator to ensure reproducibility, create a population of 20 individuals, and set up the *Hall of Fame* and a statistics object to keep track of the maximum fitness in the population.
- 19-29 A while loop is used to run the genetic algorithm until the target fitness is reached. Inside the loop:
  - 21-22 We apply crossover and mutation to the population using the `varAnd` function,<sup>3</sup> and then evaluate the fitness of the offspring
  - 23-24 We assign the newly computed fitness values to the individuals.
  - 25 We replace the old population with the selected individuals from the offspring.
  - 26-27 We update the Hall of Fame and compile the statistics.
  - 28-29 We check if the maximum fitness has reached the target fitness. If it has, we break the loop and the algorithm stops.
- 30-32 Finally, we print the results.

This example illustrates the key aspects of using DEAP: creating custom classes, setting up a toolbox, defining and registering functions, and manually controlling the loop of the genetic algorithm.

DEAP provides robust support for multi-objective algorithms and parallelization, which are common requirements in complex optimization problems. The library also includes a set of benchmark functions and examples to help users understand the various algorithms' behavior and performance.

<sup>3</sup>This function simply applies the variation operators to the population.

DEAP is widely regarded as one of the most comprehensive and cutting-edge genetic algorithm frameworks currently available, encompassing a wide range of algorithms, including GAs, GP, Evolutionary Strategies (ES), and more. It boasts a wealth of documentation and a robust community of users and contributors. However, certain facets of DEAP may pose challenges for users, as reflected in the preceding code example.

- **Toolbox Convention:** DEAP's verbose code and toolbox usage can confuse those new to evolutionary algorithms.
- **Flexibility vs. Usability:** DEAP offers flexibility but its dependence on code injection complicates static analysis and IDE integration.
- **Production Risks:** DEAP's code injection can introduce vulnerabilities if not handled properly.
- **Python's Typing:** The dynamic typing in Python, while flexible, can lead to runtime errors and harder-to-understand code.
- **Dependency on NumPy:** DEAP requires the NumPy library, introducing an additional dependency to manage.
- **Library Mastery:** Extending DEAP's capabilities demands a deep understanding of its architecture and often manual algorithm modifications.
- **Abbreviated Variables:** DEAP's use of abbreviated variable names, coupled with Python's keyword arguments, can hinder code interpretation without documentation.

Despite this, DEAP's flexibility and modularity have made it a popular choice among researchers and practitioners in the field of evolutionary computation.

### 3.4 Jenetics: Java Genetic Algorithm Library

*Jenetics* [69] is a robust EC library developed in *Java*. It is a comprehensive framework that provides a *wide range* of genetic algorithms and operators. The library is open-source and available under the *MIT* license [19].

*Jenetics* structure revolves around two main concepts: **Phenotype**<sup>4</sup> and **Engine**. The **Phenotype** is a representation of a single candidate solution to a given optimization problem. Meanwhile, the **Engine** is the core of the library, which controls the evolution process. It is responsible for the initialization of the population, the execution of the evolution, and the termination of the evolution once the termination condition is met (this uses a very similar approach to the one we presented in section 3.2 on page 30).

One thing that sets *Jenetics* apart from other libraries is the use of the *Java's Stream API* [66] to make a seamless integration between it and the **EvolutionStream** class used by the **Engine**. This means that the evolution has access to all the features provided by the Stream API, such as parallelization, filtering, and mapping.

Below, we demonstrate the use of *Jenetics* to solve the OMP with a genetic algorithm:

Listing 3.3– Solution to the OMP using *Jenetics*

```

1 public class OneMax {
2     public static void main(String[] args) {
3         RandomRegistry.random(new Random(11));
4         final var engine =
5             Engine.builder((Genotype<BitGene> gt) ->
6                 gt.chromosome().as(BitChromosome.class).bitCount(),
7                 BitChromosome.of(20, 0.5))
8                 .maximizing()
9                 .populationSize(20)
10                .alterers(new Mutator<>(0.05), new SinglePointCrossover<>(0.5))
11                .build();
12
13     Phenotype<BitGene, Integer> best = engine.stream()
```

<sup>4</sup>See definition A.18 on page 97

```

13     .limit(Limits.byFitnessThreshold(19))
14     .collect(EvolutionResult.toBestPhenotype()));
15     System.out.println("Target fitness reached at generation: " + best.generation());
16     System.out.println("Best individual is: " + best.genotype());
17     System.out.println("with fitness: " + best.fitness());
18   }
19 }
```

When executed, the program outputs the following:

```

Target fitness reached at generation: 40
Best individual is: [00001111|11111111|11111111]
with fitness: 20
```

This example demonstrates *Jenetics*' usage for implementing a straightforward genetic algorithm, with the added ability to stop evolution once steady fitness is achieved.

Here is an explanation of the code:

- 1-2 First, a class containing a `main` method is established. This forms the launching point of the program.
- 3 The `RandomRegistry` is then utilized to preset the random seed, a step that ensures consistency and replicability of program outputs.
- 4-11 The core configuration of the evolutionary engine occurs next, tailored specifically to the problem at hand.
  - 5-7 An `Engine` is instantiated using the `Engine.builder` method, requiring two arguments: a `Function` and a `Genotype`. The `Function` assesses the fitness of a `Genotype` by tallying the quantity of 1s in the `Genotype`. A `Genotype` is created via the `BitChromosome.of` method, which mandates the length of the chromosome and the probability of a 1 appearing within the chromosome. Here, `BitChromosome` is a distinct version of the `Chromosome` interface, used to depict a `Genotype` composed of `BitGenes`.
  - 8 The `maximizing` method signals that the objective is to boost fitness to its maximum possible value.
  - 9 The `populationSize` method sets the size of the evolutionary population.
  - 10 The `alterers` method sets the genetic operations to be used in the evolution. Here, the `Mutator` and `SinglePointCrossover` are selected. The `Mutator` behaves similarly to a *bit-flip mutation* operator in this context.
  - 11 Finally, the `build` method assembles the functional `Engine`.
- 12-14 Execution of the evolution process follows.
  - 12 The `stream` method generates an `EvolutionStream` from the `Engine`, allowing for control over the evolution process.
  - 13 The `limit` method sets the termination condition. The evolution halts once a `Phenotype` possessing a fitness **exceeding** 19 is located.
  - 14 Lastly, the `collect` method gathers the evolution's output, in this case, capturing the optimal `Phenotype` discovered.
- 15-17 The final stage involves printing the evolution results to the console for review and analysis.

This example showcases the key aspects of using *Jenetics*: creating genotypes, setting up an engine, and controlling the evolution process.

*Jenetics* provides extensive support for multi-objective algorithms and parallel execution, which are common necessities in intricate optimization problems.

Recognized as one of the most comprehensive and cutting-edge genetic algorithm libraries today, *Jenetics* receives wide-ranging acclaim. However, certain aspects of *Jenetics* could be challenging for its users.

- **Java Limitations:** *Jenetics* is hindered by Java's verbosity and complexity. As a Java library, *Jenetics* unavoidably incorporates some of Java's limitations, notably its verbosity and complexity.
- **Documentation Gaps:** The library's documentation is not exhaustive and sometimes outdated. The library's documentation could be more comprehensive. Currently, some examples are outdated or have redundant implementations.
- **Seed Replicability:** The RandomRegistry's claim on replicability conflicts with actual outcomes, indicating a documentation deficiency. The insufficient documentation is most evident in the use of the RandomRegistry to set the random seed, as seen in example line 3. The documentation claims this ensures the replicability of the program's output, but practical execution proves otherwise. More complete code samples and concrete use cases would ameliorate this issue.
- **Flexibility vs. Usability:** *Jenetics* offers adaptability but leans on builders and factories, sometimes obscuring the code. There is a delicate balance between the framework's flexibility and its usability. *Jenetics'* notable adaptability leans heavily on builders and factories, which can occasionally obfuscate the code.
- **Verbose Type Inference:** Java's type inference limitations can lead to verbose code, especially with classes like Phenotype. Java's lack of robust type inference can lead to verbosity, particularly evident when using classes like Phenotype requiring multiple type parameters.
- **Parallelization Approach:** *Jenetics* uses the Java Stream API for parallelization, a non-standard approach in evolutionary computation that may be tricky for novices. *Jenetics* relies on the Java Stream API for parallelization to maintain efficiency. However, this approach is not typical in the field of evolutionary computation, making it challenging for beginners. Furthermore, the use of the Stream API can result in difficult-to-debug code due to the execution mainly handled by Java's standard library.
- **Default Parallelization:** The library's default parallel operation can hamper performance on smaller tasks due to overhead. By default, *Jenetics* operates in parallel. This can detrimentally affect program performance when working on small populations or problems with fewer variables due to parallelization's overhead.
- **Complex Problem Handling:** Addressing intricate problems in *Jenetics* introduces multiple complexities in the code. Complex problem-solving with *Jenetics* can quickly complicate the code, introducing codecs, proxies, and the Problem interface.
- **Non-standard Features:** *Jenetics* relies heavily on features outside Java's standard library, posing challenges for beginners. *Jenetics* extensively employs features not found in Java's standard library, particularly immutable data, thereby steepening the learning curve for beginners.
- **Java Version Compatibility:** Compatibility with only newer Java versions restricts its usability, especially for many Android devices. *Jenetics* is compatible with newer versions of Java, potentially problematic for users unable to update their Java version. This issue is especially relevant for most Android devices, which do not support newer Java versions.
- **Visualization Gap:** The library lacks a comprehensive mechanism for visualizing the evolutionary process. There's a lack of a robust mechanism for visualizing the evolution process, a feature commonly found in other evolutionary computation libraries.
- **Type Safety:** *Jenetics* sometimes requires explicit type casting, hinting at a lack of full type safety. *Jenetics* isn't entirely type-safe, necessitating explicit casts via the as method.

Despite these shortcomings, *Jenetics* is a powerful and flexible library that can be used to solve a wide variety of optimization problems. The library is actively maintained, and it is one of the most popular evolutionary computation libraries available today.

### 3.5 ECJ: A Java-based Evolutionary Computation Research System

The *Evolutionary Computation in Java* (ECJ) [54] is a powerful and flexible framework for conducting research and experiments in the field of evolutionary computation. ECJ has been under active development for more than two decades, and during this period, it has grown into one of the most comprehensive open-source<sup>5</sup> libraries for evolutionary computation.

---

<sup>5</sup>No license is provided with the ECJ source code.

What sets ECJ apart from other libraries is its ability to handle a broad range of evolutionary computation paradigms. It offers support for genetic algorithms, genetic programming, multi-objective optimization, co-evolution, and many more. It also comes with built-in functionality for distributed computing, allowing researchers to harness the power of large-scale computing clusters for their experiments.

The ECJ library is designed with flexibility and extensibility in mind. It employs a design pattern that uses parameter files for configuration, thus enabling researchers to customize the algorithms according to their specific needs without having to modify the source code. The library is highly modular, and its components can be easily replaced or extended.

The following code sample demonstrates how to solve the *OneMax* problem using ECJ:<sup>6</sup>

Listing 3.4– ECJ’s solution to the OMP

```

1 public class OneMax extends Problem implements SimpleProblemForm {
2     @Override
3     public void evaluate(final EvolutionState state,
4                         final Individual ind,
5                         final int subpopulation,
6                         final int threads) {
7         if (ind.evaluated) {
8             return;
9         }
10        if (!(ind instanceof BitVectorIndividual)) {
11            state.output.fatal("Individual must be a BitVectorIndividual")
12        }
13        int sum = 0;
14        // This is necessary to safe-cast ind to BitVectorIndividual
15        BitVectorIndividual ind2 = null;
16        if (ind instanceof BitVectorIndividual) {
17            ind2 = (BitVectorIndividual) ind;
18        }
19        if (ind2 != null) {
20            for (int x = 0; x < ind2.genome.length; x++) {
21                sum += (ind2.genome[x] ? 1 : 0);
22            }
23
24            if (!(ind2.fitness instanceof SimpleFitness)) {
25                state.output.fatal("Fitness must be a SimpleFitness")
26            }
27            ((SimpleFitness) ind2.fitness)
28                .setFitness(state,
29                    /* fitness */ sum / (double) ind2.genome.length,
30                    /* isIdeal */ sum == ind2.genome.length);
31            ind2.evaluated = true;
32        }
33    }
34 }
```

An explanation of the code:

- 1 The class `OneMax` extends `Problem` and implements `SimpleProblemForm`, which is ECJ’s interface for problems that can be solved with generational evolutionary algorithms.
- 2-6 The `evaluate` method is overridden from the `Problem` class. It is responsible for evaluating the fitness of an `Individual` (a candidate solution in the population). The arguments include the current state of the evolution, the individual to be

<sup>6</sup>We couldn’t find a way to set the random seed in ECJ, so the results are not reproducible.

evaluated, the subpopulation to which the individual belongs, and the number of the thread executing this method (useful in a multithreaded setting).

- 7-9 The method checks if the individual has already been evaluated. If it has, then the method immediately returns to avoid unnecessary computation.
- 10-12 If the individual is not an instance of `BitVectorIndividual` (which represents individuals whose genome consists of bits), a fatal error is reported. The OneMax problem assumes a bit string representation of individuals.
- 13-18 Variable `sum` is initialized to keep track of the total number of 1s in the individual's bit string representation. The individual is then safe-casted to `BitVectorIndividual`. This is necessary because the `ind` argument is of type `Individual`, but we need to work with its `BitVectorIndividual` specifics (such as its `genome` property).
- 20-22 If the casting is successful, a loop iterates through the genome (bit string) of the individual, incrementing `sum` for every bit set to 1.
- 24-26 If the fitness of the individual is not an instance of `SimpleFitness` (which represents a single floating-point fitness value), a fatal error is reported.
- 27-30 The individual's fitness is then set to the proportion of bits set to 1 in the bit string (`sum / (double) ind2.genome.length`). If all bits are set to 1 (i.e., `sum` equals the length of the genome), then the individual is marked as ideal.
- 31 Finally, `ind2.evaluated` is set to `true`, indicating that the individual has been evaluated.

To run the OneMax problem, we need to create a parameter file that specifies the problem, the evolutionary algorithm, and the parameters of the algorithm. The following is a sample parameter file for the OneMax problem:

Listing 3.5– Configuration file for the OMP

```

1 breedthreads = 1
2 evalthreads = 1
3 seed.0 = 11
4 eval.problem = ec.app.onemax.OneMax
5 pop.subpop.0.species.ind = ec.vector.BooleanVectorIndividual
6 pop.subpop.0.species.genome-size = 20
7 pop.subpop.0.species.fitness = ec.FitnessSimple
8 pop.subpop.0.size = 20
9 pop.subpop.0.species.pipe = ec.vector.breed.VectorMutationPipeline
10 pop.subpop.0.species.pipe.source.0 = ec.vector.breed.VectorCrossoverPipeline
11 pop.subpop.0.species.pipe.source.0.prob = 0.5
12 pop.subpop.0.species.pipe.source.0.source.0 = ec.select.TournamentSelection
13 pop.subpop.0.species.pipe.source.0.source.1 = ec.select.TournamentSelection
14 eval.problem.terminateFitness = 20

```

This is a parameter file for ECJ, which specifies various configuration settings for an evolutionary computation run.

- 1-2 `breedthreads` and `evalthreads` determine the number of threads to be used for breeding and evaluation operations, respectively. Both are set to 1, which means a single-threaded operation.
- 3 `seed.0` sets the initial seed for the random number generator. Different seeds will lead to different runs, even with identical configuration parameters.
- 4 `eval.problem` sets the problem to be solved, in this case, `ec.app.onemax.OneMax`, which refers to the OneMax problem class implemented earlier.
- 5-6 `pop.subpop.0.species.ind` sets the type of individual in the population. `ec.vector.BooleanVectorIndividual` means that each individual will be a Boolean vector, which is appropriate for the OneMax problem. `pop.subpop.0.species.genome-size` sets the length of the Boolean vector to 20.
- 7-8 `pop.subpop.0.species.fitness` sets the type of fitness to be used. `ec.FitnessSimple` represents a single floating-point fitness value. `pop.subpop.0.size` sets the population size to 20.

- 9 `pop.subpop.0.species.pipe` defines the pipeline for breeding new individuals. `ec.vector.breed.VectorMutationPipeline` means a mutation operation will be applied.
- 10 `pop.subpop.0.species.pipe.source.0` sets the source of individuals for the mutation operation to be a crossover operation (`ec.vector.breed.VectorCrossoverPipeline`), with a crossover probability of 0.5.
- 12-13 `pop.subpop.0.species.pipe.source.0.source.0` and `pop.subpop.0.species.pipe.source.0.source.1` set the source of individuals for the crossover operation to be tournament selection (`ec.select.TournamentSelection`).
- 14 `eval.problem.terminateFitness` sets the fitness level at which the evolutionary run should terminate. When an individual reaches a fitness of 20, the run will stop, implying that a perfect solution (all bits set to 1) has been found for the OneMax problem.

To run the OneMax problem, we need to execute the following command:

Listing 3.6– Running the OMP with ECJ

```
java -cp .
ec.Evolve -file .\ec\app\onemax\one_max.properties
```

This command will run the OneMax problem with the parameter file `one_max.properties` in the package `ec.app.onemax`. According to the official ECJ documentation, this command should be able to solve the OneMax problem, however, we were unable to get it to work. The following is the output of the command:

Listing 3.7– Output of the command in listing 3.6

```
# Code omitted for brevity
Exception in thread "main" ec.util.ParamClassLoadException:
No class name provided.
PARAMETER: state
        at ec.util.ParameterDatabase.getInstanceForParameter(ParameterDatabase.java:493)
        at ec.Evolve.initialize(Evolve.java:479)
        at ec.Evolve.initialize(Evolve.java:412)
        at ec.Evolve.main(Evolve.java:758)
```

The error message indicates that the parameter `state` is missing. This parameter is required by ECJ, but is not specified in the parameter file. We tried to add the parameter to the file, but the error persisted. We also tried to run the OneMax problem with the parameter file provided by the ECJ documentation, but the error persisted. We were unable to find a solution to this problem.

While ECJ stands as one of the most comprehensive and sophisticated Evolutionary Computation (EC) frameworks available, it presents users with a daunting learning curve. We've identified several areas of concern that contribute to this complexity:

- **Verbose and obscure code:** The framework's verbosity hinders code comprehension. To illustrate, the `eval` method in the `OneMax` class spans 31 lines, yet only 3 of these lines (20-22) actually pertain to the OneMax problem itself—the rest are framework-related.
- **Complex configuration:** Setting up the framework involves numerous layers of boilerplate code, necessitating a deep understanding of the underlying concepts and terminology. Lines such as `pop.subpop.0.species.pipe.source.0.source.0` can be particularly opaque to newcomers.
- **Outdated documentation and implementation:** The `README.md` file has not been updated in 4 years, and the `OneMax` class equivalent was last refreshed 5 years ago. Even the most recent version (27) of the framework dates back 4 years, rendering much of the documentation and implementation obsolete.
- **Legacy Java syntax:** ECJ's reliance on older Java syntax, specifically the lack of generics, results in frequent casting and `instanceof` checks. This not only makes the code more brittle but also complicates understanding.

- **Exclusion from popular package managers:** The absence of ECJ from widely-used package managers like *Maven* or *JitPack* imposes additional burdens on integration with other projects, as manual download and compilation become necessary.
- **Limited IDE compatibility:** The design of ECJ emphasizes command line compilation and execution, undermining the utility of modern Integrated Development Environments (IDEs) and complicating debugging efforts.
- **Inconvenience for integration:** Rather than being crafted for use as a library, the framework is designed with the expectation of direct source code inclusion. This requirement to import source code into users' own projects increases the difficulty of integration.
- **Generic name complicates research:** The term “ECJ” is sufficiently common to cause confusion when seeking information on the framework. Queries such as “ECJ” or “ECJ Java” yield results concerning the *European Court of Justice* or the *Eclipse Compiler for Java*, respectively, rather than the intended framework.
- **Limited community and resources:** As of 2023, ECJ lacks a robust community of developers or users. The dearth of online resources like tutorials or Stack Overflow discussions compounds the challenges for new users seeking to leverage the framework effectively.
- **Non-utilization of modern Java features:** Beyond the lack of generics, the framework fails to harness modern Java features like lambda expressions, lists, or Optional types. These features, by enabling more concise and readable code, can significantly enhance error prevention.

In conclusion, the *Evolutionary Computation in Java* (ECJ) framework, while extensive and capable, poses considerable barriers to entry for users due to its complexity, outdated documentation, and heavy reliance on legacy *Java* syntax. Furthermore, despite its versatility in supporting various evolutionary computation paradigms and distributed computing, the steep learning curve, absence from popular package managers, and limited compatibility with modern IDEs make it challenging for newcomers and experienced developers alike. The verbose code, intricate configuration requirements, and limited use of modern *Java* features also contribute to the difficulty in understanding and integrating the ECJ into projects. Our exploration further exposed issues with the execution of sample problems, highlighting the need for updated documentation and maintenance. However, ECJ's modular design, extensibility, and breadth of capabilities maintain its status as a valuable tool for research in the field of evolutionary computation. Future efforts could focus on modernizing the framework, improving documentation, and fostering an active user community to mitigate the identified issues and enhance its usability.

### 3.6 GeneticSharp: Comprehensive Overview of a .NET Genetic Algorithm Library

*GeneticSharp* is a high-performance, extensible, cross-platform C# Genetic Algorithm library tailored for multi-threading. This open-source library, designed to streamline the integration of genetic algorithms into applications, is licensed under the *MIT License* [19].

*GeneticSharp* revolves around three foundational concepts: *Chromosome*, *Population*, and *GeneticAlgorithm*:

- *Chromosome* serves as the foundation for various chromosome types, responsible for gene storage and genetic operator management.
- *Population* maintains a collection of chromosomes undergoing evolution.
- *GeneticAlgorithm* acts as the algorithm's primary driver, encompassing its main loop.

Distinctively, *GeneticSharp* capitalizes on parallel computing, accelerating fitness evaluation, crossover, and mutation. Furthermore, its extension-oriented design simplifies the addition of new operators or further customizations to the genetic algorithm.

Consider the following example, which demonstrates solving the OMP using *GeneticSharp*:

Listing 3.8– Solving OMP with *GeneticSharp*

```

1 public sealed class BinaryChromosome : BinaryChromosomeBase {
2     public BinaryChromosome(int length) : base(length) {
3         CreateGenes();

```

```

4
5     }
6     public override IChromosome CreateNew() {
7         return new BinaryChromosome(Length);
8     }
9 }
10 public class OneMaxFitness : IFitness {
11     public double Evaluate(IChromosome chromosome) {
12         return chromosome.GetGenes().Count(gene => (int)gene.Value == 1);
13     }
14 }
15 public class ReproducibleRandom : RandomizationBase {
16     private static readonly object GlobalLock = new();
17     private static readonly ThreadLocal<Random?> ThreadRandom = new(NewRandom);
18     private static Random? Instance => ThreadRandom.Value;
19     private static Random NewRandom() {
20         lock (GlobalLock) {
21             return new Random(11);
22         }
23     }
24     public override int GetInt(int min, int max) {
25         Debug.Assert(Instance != null, nameof(Instance) + " != null");
26         return Instance.Next(min, max);
27     }
28     public override float GetFloat() {
29         Debug.Assert(Instance != null, "ReproducibleRandom.Instance != null");
30         return (float)Instance.NextDouble();
31     }
32     public override double GetDouble() {
33         Debug.Assert(Instance != null, "ReproducibleRandom.Instance != null");
34         return Instance.NextDouble();
35     }
36 const int chromosomeLength = 20;
37 const int populationSize = 20;
38 RandomizationProvider.Current = new ReproducibleRandom();
39 var selection = new EliteSelection();
40 var crossover = new UniformCrossover();
41 var mutation = new FlipBitMutation();
42 var fitness = new OneMaxFitness();
43 var chromosome = new BinaryChromosome(chromosomeLength);
44 var population = new Population(populationSize, populationSize, chromosome);
45 var ga = new GeneticAlgorithm(population, fitness, selection, crossover, mutation) {
46     Termination = new FitnessThresholdTermination(chromosomeLength)
47 };
48 ga.Start();
49 Console.WriteLine($"Target fitness reached at generation: {ga.GenerationsNumber}");
50 var bestChromosome = ga.BestChromosome as BinaryChromosome;
51 Console.WriteLine($"Best individual is: {bestChromosome}");
52 Console.WriteLine($"with fitness: {bestChromosome?.Fitness}");

```

Dissecting the code reveals the following:

- 1-8 The `BinaryChromosome` class is sealed<sup>7</sup> to inhibit subclassing and encapsulates a chromosome with binary genes. Sealing is an advisable practice here due to the constructor's invocation of the virtual `CreateGenes` method, thus eliminating potential ambiguities from subclass method overrides.
- 2-4 The constructor, accepting a length parameter, invokes the base constructor. Given that the base may call a virtual method, sealing thwarts unforeseen behaviors from subclasses unknowingly overriding it.
- 4-7 `CreateNew` method proffers a fresh `BinaryChromosome` instance retaining the original length.
- 9-13 `OneMaxFitness` class quantifies a chromosome's fitness.
- 10-12 `Evaluate` method tallies genes equating to 1 (equivalent to binary `true`), serving as the OMP's fitness metric.
- 14-35 The `ReproducibleRandom` class furnishes reproducibly random numbers, critical for consistent genetic algorithm outcomes across iterations.
- 15-17 A global lock alongside a thread-local `Random` instance (`ThreadRandom`) ensures synchronized seeding across threads for consistent randomness.
- 18-22 `NewRandom` method spawns a `Random` object seeded with 11, bolstered by the `lock` keyword for thread-safety.
- 23-34 Methods `GetInt`, `GetFloat`, and `GetDouble` yield values from the `Random` instance.
- 36-37 Constants pertinent to the genetic algorithm, such as chromosome length and population size, are delineated.
- 38 Randomization defaults to the `ReproducibleRandom` class.
- 39-44 Configuration lines dictate the genetic algorithm's components, encompassing selection (elite), crossover (uniform), mutation (flip bit), and the fitness function (`OneMaxFitness`).
- 45-47 The genesis population and genetic algorithm instance are instantiated.
- 46 The algorithm's cessation is predicated on a chromosome attaining peak fitness, tantamount to its length.
- 48 The genetic algorithm is set in motion.
- 49-52 Outputs include the generation count to attain target fitness, the paramount chromosome, and its corresponding fitness.

Despite its many advantages, *GeneticSharp* also has certain limitations:

- **Overwhelming verbosity:** The library's detailed nature can overshadow its functionality, especially in contrast to succinct frameworks like Jenetics.
- **Externalized fitness function:** The architecture demands that fitness functions be isolated in distinct classes. This design choice might make trivial problems seem more convoluted than when using, for example, higher-order functions.
- **Sparse documentation:** Compared to competitors like DEAP, the library lags in both instructional content and practical examples. Novices might find this shortage hampers their initial engagement with the library.
- **Parallelization trade-offs:** Incorporating parallel functionality amplifies the complexity of defining custom genetic operators. Although parallel processing boosts execution speed, developers need to grapple with the intricacies it introduces.
- **Advanced C# intricacies:** Features unique to advanced C# such as the `sealed` attribute can disorient beginners. New users might be deterred by such advanced constructs, especially without prior experience in C#.
- **Developer-centric focus:** *GeneticSharp* caters predominantly to application creation, potentially sidelining academic researchers. Scholars deeply entrenched in genetic algorithm studies might find the library's emphasis misaligned with their interests.
- **Budding community support:** The fledgling nature of the *GeneticSharp* community in 2023 could impede access to resources and prompt support. This nascent state might affect user engagement, library updates, and peer assistance.
- **Conventional algorithm scope:** The library's focus remains anchored to traditional genetic algorithms, omitting newer evolutionary algorithm derivations. Practitioners might find the lack of support for paradigms like genetic programming limiting.

---

<sup>7</sup>Details on definition A.21 on page 97.

In summary, *GeneticSharp* offers a dynamic, parallelized platform for crafting genetic algorithms in C#. Its customizability and developer-centric approach make it a compelling choice for those aiming to harness genetic algorithms within their software projects. Yet, users must be cognizant of its more intricate API, the nuances introduced by parallelization, and potential challenges arising from a smaller community and limited documentation. Ultimately, while *GeneticSharp* solidifies its position as a valuable tool within the .NET arena, its optimal utilization depends on aligning its capabilities with the user's specific requirements and familiarity with genetic algorithms.

### 3.7 Other Libraries

In the preceding sections, we detailed a select set of premier frameworks within EC. However, the scope of our research extended beyond this list to a wider array of libraries. This comprehensive survey was instrumental in shaping our proposed *Kotlin*-based EC framework. Through understanding the strengths and peculiarities of each library, we aimed to incorporate their best features while sidestepping prevalent pitfalls.

Below, we provide a concise overview of some additional frameworks:

- **EvolvingObjects** (EO) [55]: An object-oriented C++ framework. It offers various evolutionary algorithms and operators, including GAs, ES, and Particle Swarm Optimization (PSO). Notably, certain aspects, like GP, are not currently supported.
- **Inspyred** [58]: A Python-focused framework inspired by De Jong [18]. It clearly distinguishes between algorithmic computations and problem-specific ones. While it supports a range of evolutionary algorithms, GP is not among them.
- **Pyevolve**: A Python-based framework that provides a gamut of evolutionary algorithms, including GAs and GP. However, it is currently inactive.
- **PGAPack** [63]: Crafted in C, it is a parallel genetic algorithm library. It offers compatibility with *Fortran* and C++ and a myriad of parallel architectures, it also features a *Python* interface [64].
- **pagmo** [44]: A C++ library that prioritizes parallel optimization. It delivers a cohesive interface for various optimization algorithms and provides a *Python* interface [46].
- **easy\_ga** [53]: Developed in *Rust*, this framework simplifies GA prototyping, primarily focusing on traditional GAs.
- **genevo** [56]: Another *Rust* framework, it majorly supports classic GAs.
- **Evolutionary Computation Framework** (ECF) [40]: ECF is a C++ framework that provides a comprehensive set of evolutionary algorithms. Its depth rivals that of DEAP and ECJ. Nonetheless, it faces challenges like limited documentation and a challenging learning curve.

The breadth and capabilities of these libraries highlight the vast potential within the domain of genetic algorithms. Our exhaustive exploration empowers us to develop a *Kotlin*-based framework that amalgamates the strengths of each while pioneering unique features.

## Chapter 4

# The Keen Framework

### 4.1 Introduction

Within the expansive landscape of *evolutionary computation* (EC), a multitude of methodologies and algorithms have emerged, each demonstrating significant success in addressing complex optimization problems. Nevertheless, the practical application of these techniques often presents its own challenges, most notably, the extensive repetitive coding required and the absence of an intuitive, user-friendly platform. To address these challenges, this chapter introduces *Keen*, an innovative genetic algorithms framework developed using the *Kotlin* programming language.

As a comprehensive and adaptable framework, Keen embarks on an ambitious journey to promote the application and acceptance of evolutionary algorithms across various domains. This well-structured framework is currently designed to simplify the implementation process of *genetic algorithms* (GAs) and *genetic programming* (GP), with an emphasis on allowing extension to other evolutionary algorithms, thereby accelerating the pace of research and development in the EC field.

This chapter provides an in-depth examination of the *Keen framework*, illuminating its architecture, fundamental components, and the array of functionalities it offers. We explore the range of genetic operators within Keen, including various selection, mutation, and crossover techniques. A comprehensive account of the GAs and GP methods within *Keen's* repertoire will be provided, underscoring their distinctive attributes and potential use-cases.

Beyond its existing features, the extensibility of the Keen framework will also be examined, demonstrating its readiness to incorporate novel algorithms and techniques. Keen is designed with a view to facilitate easy expansion and modification of its components, encouraging innovation and continuous progression, in tandem with advancements in the EC field.

The introduction of *Keen* marks a valuable contribution to both the *Kotlin* and AI communities, offering a robust, user-friendly platform for the development, testing, and deployment of evolutionary algorithms. We envision that this pioneering framework will encourage broader application of evolutionary algorithms, thus pushing the boundaries of artificial intelligence research and application.

### 4.2 Architecture

*Keen*, our sophisticated and flexible evolutionary computation library, organizes its architectural blueprint into five distinguishable modules: *Evolution*, *Genetic Material*, *Genetic Operators*, *Utility*, and *Programs*.

The *Utility* module serves as the library's toolbox, providing various auxiliary classes and functions that the remaining modules can utilize. As its content is primarily supplementary, it won't be thoroughly examined within this document.

The *Genetic Material* module encapsulates the foundational elements defining the characteristics of our evolving entities or Individuals. Its core constituents are the Individual and Genotype classes, accompanied by the encompassing Chromosome and Gene hierarchies.

Within the ***Genetic Operators*** module, we define the different genetic operators that the *Evolution Engine* leverages to stimulate the evolutionary process. This module accommodates classes embodying crossover, mutation, and selection operators, systematically grouped into corresponding categories: **Crossover**, **Mutation**, and **Selection**.

The ***Evolution*** module houses classes that construct the engine of evolution itself, along with essential classes promoting the smooth progression of the evolution process.

Finally, the nascent ***Programs*** module showcases genetic programming paradigms through a collection of programs developed under this concept. However, as it is in its preliminary stages, this document will not cover it in detail.

A majority of the *Keen* classes are parameterized by the type of data they handle and the gene type composing the population of individuals, like an integer or an integer gene. This design choice lends *Keen* a wide range of flexibility, enabling users to tailor the genetic material for the evolution engine. Moreover, thanks to *Kotlin*'s proficient type inference, we can generally omit type parameters, thereby boosting code legibility.

*Keen* upholds the principle of immutability, rendering most of its classes immutable. This provides an assurance of robustness and predictability, as objects remain unaltered post-creation, enhancing parallelization potential by allowing shared memory use without the need for synchronization mechanisms.

Furthermore, *Keen* offers factory methods for most classes, fostering an intuitive and straightforward programming experience. It employs language-oriented programming techniques to form a domain-specific language (DSL), streamlining object creation with more readable code. This approach is evident in libraries such as *Kotlinx.html* [43], *Kotlin Telegram Bot* [42], and *Kotest* [59]. While we promote factory methods as the primary mode of object creation, the design caters to manual object creation, granting users greater control over the process when needed.

Here's an illustrative DSL example, where we create a Genotype for the *Room Scheduling* problem. Here, each Chromosome represents a meeting, composed of a single integer gene indicating the assigned room. The full code snippet can be found in listing C.3 on page 116.

Listing 4.1– Usage of the *Keen* Domain Specific Configuration Language (KDSCL) to create a Genotype for the *Room Scheduling* problem.

```
genotype { // Create a genotype
    repeat(meetings.size) { // Create a chromosome for each meeting
        chromosomeOf { // Create a chromosome
            ints { // Containing an integer gene
                size = 1
                ranges += meetings.indices.first..meetings.indices.last
            }
        }
    }
}
```

The approach we see here closely mirrors the use of configuration files in ECJ. However, as our DSL is defined entirely within *Kotlin*, it offers significant advantages. These include type safety, robust Integrated Development Environment (IDE) support, and the ability to leverage the full capabilities of the language in defining our genotypes.

In the given example, we employed the standard `repeat` function from the *Kotlin* library to create a number of chromosomes equal to the number of meetings. It is essential to note that this function is not part of our DSL but a standard feature provided by the language.

Furthermore, we claim that our approach enhances the readability of the code. Compared to using ECJ's configuration files, the DSL presents a more intelligible and cleaner interface. By defining the configurations in the same language and even within the same files as the rest of the code, we significantly improve code readability. This consolidation offers a more streamlined, unified development environment that facilitates easier understanding and manipulation of the codebase.

In the following sections, we will delve deeper into the pivotal classes within these modules and examine their functions and significance.

## 4.3 Genetic Algorithms

As we delve deeper into the intricate workings of the framework, it becomes imperative to explore its foundational components. While the theoretical underpinnings of Genetic Algorithms (GAs) have been exhaustively discussed in section 2.2 on page 6, this section aims to shed light on the actual implementation and design specifics of the GAs within our framework. Our exploration will be segmented into three primary components: the representation and structuring of the *genetic material*, the core mechanics of the *evolution engine*, and the crucial role of the *evolution listeners*. Each subsection will provide a granular understanding of how these elements function and interconnect, ensuring that the GAs are optimally tailored to meet the design goals of the framework.

### 4.3.1 Genetic Material

The genetic material stands at the heart of GAs, representing the problem domain's primary aspect. Consequently, it's essential to structure this material efficiently and flexibly.

Our approach is “gene-centric”. Rather than relying on indirect representations, the genetic material is depicted as a collection of genes. Such a representation bolsters the flexibility of the API design. For instance, the genetic material can manifest as a single gene, a collection of genes (referred to as a chromosome), or even as multiple collections of genes (genotype).

Furthermore, the genetic material is immutable. This immutability ensures thread safety and adheres to the “copy-on-write” principle, both of which are indispensable for crafting robust GAs.

**Gene** The gene stands as the fundamental unit of genetic material within our framework, representing the minutest unit that can undergo manipulation.

Fundamentally, each gene must be capable of: 1. *mutation*, 2. *reproduction*, and 3. *information storage*. To promote extensibility, genes are designed to handle minimal operations, passing most tasks to operators (refer to section 4.4 on page 50). This approach ensures the framework's adaptability across diverse genetic materials, including binary, integer, real, and custom types.

Thus, we suggest the following interface for gene implementation:

Listing 4.2– Gene interface

```
interface Gene<T, G> where G : Gene<T, G> {
    val value: T
    fun mutate(): G
    fun duplicateWithValue(value: T): G
}
```

**Chromosome** A chromosome is defined as an iterable collection of genes. Its primary role is to offer a streamlined way to manipulate a group of genes. It's crucial to keep the chromosome's responsibilities minimal to foster adaptability across diverse genetic materials.

To this end, we suggest the following interface for each chromosome:

Listing 4.3– Chromosome interface

```
interface Chromosome<T, G> : Collection<G> where G : Gene<T, G> {
    val genes: List<G>
    fun duplicateWithGenes(genes: List<G>): Chromosome<T, G>
    interface Factory<T, G> where G : Gene<T, G> {
        fun make(): Chromosome<T, G>
    }
}
```

Incorporated within is the Factory interface, dedicated to the creation of new chromosomes. This is harnessed by the framework whenever there's a need to generate new chromosomes.

**Genotype** Similar to the chromosome, the genotype serves as an iterable collection of chromosomes. Though they share analogous responsibilities, the genotype operates on a broader spectrum. While many frameworks tend to merge the definitions of chromosomes and genotypes, distinguishing between them is paramount to our framework's adaptability. This distinction facilitates the simultaneous manipulation of multiple chromosomes, paving the way for more intricate genetic materials.

Utilizing parametric polymorphism, we negate the necessity for a hierarchical type relationship between genotypes and chromosomes. Consequently, the genotype can be represented by a concrete class, as demonstrated below:

Listing 4.4– Genotype class

```
data class Genotype<T, G>(val chromosomes: List<Chromosome<T, G>>) :
    Collection<Chromosome<T, G>> where G : Gene<T, G> {

    class Factory<DNA, G : Gene<DNA, G>> {
        fun make(): Genotype<DNA, G> { ... }
    }
}
```

Embedded within is the `Factory` class, designated for the generation of new genotypes.

Note this implementation does not depend on the `Chromosome` type, thus allowing for the use of any chromosome implementation that adheres to the `Gene` type. For example, one could have:

Listing 4.5– Genotype class with custom chromosome

```
genotype {
    chromosome0f {
        intImpl1 { ... }
    }
    chromosome0f {
        intImpl2 { ... }
    }
}
```

where `intImpl1` and `intImpl2` are custom implementations of the `Chromosome` interface. Thus allowing more flexibility in the design of the genetic material.

**Individual** The individual epitomizes the expression of the genetic material. Serving as the evaluated manifestation of the genotype, it embodies the candidate solution. Fundamentally, the individual hinges on:

1. A *fitness* value, and
2. The associated *genotype*.

In representing individual entities within a population, it's imperative that a individual can be compared with its peers. In light of this, the structure for the individual is delineated as:

Listing 4.6– Individual class

```
data class Individual<T, G>(
    val genotype: Genotype<T, G>,
    val fitness: Double = Double.NaN
) where G : Gene<T, G> { ... }
```

### 4.3.2 Ranker

The **ranker** is a component of the **evolution engine** responsible for ranking the population. The main purpose of this component is to compare individuals and determine their relative fitness. The ranker is a crucial component of the evolutionary process, as it is the primary source of information for the selection process.

A tempting interface for the rankers is depicted in listing 4.7.

Listing 4.7– Ranker interface

```
interface IndividualRanker<T, G> where G : Gene<T, G> {
    operator fun invoke(first: Individual<T, G>, second: Individual<T, G>): Int
    fun fitnessTransform(fitness: List<Double>)
}
```

Here, the **ranker** acts as a function that receives two individuals and returns an integer representing their relative fitness.

The **ranker** is also responsible for transforming the fitness values of the population in case that it is necessary. For example, the *roulette wheel* selection assigns a higher probability of being selected to individuals with higher fitness values, this is not suitable for minimization problems, hence the fitness values must be transformed.

The framework currently provides two rankers: the *Fitness Max Ranker*, that ranks individuals in descending order of fitness, and the *Fitness Min Ranker*, that ranks individuals in ascending order of fitness.

### 4.3.3 Evolution Engine

The **evolution engine** serves as the framework's nucleus, orchestrating the evolution of the population. It encapsulates all functions essential for the evolutionary process. This engine integrates the fundamental structure of the evolutionary algorithms, invoking relevant functions from associated entities like the **genetic material** and **genetic operators**.

Specifically, the engine manages:

1. Population initialization;
2. Population evaluation;
3. Selection of individuals for the forthcoming generation;
4. Genetic operator application on chosen individuals;
5. Evaluation of the newly formed individuals;
6. Supplanting the old population with the new;
7. Iteration from step 3 until meeting termination criteria.

Furthermore, the engine offers hooks for both pre and post-processing steps related to the population and individuals. These hooks facilitate user-driven customization of the evolution trajectory. Although this feature remains outside this work's scope, its availability merits acknowledgment.

The engine's architecture is depicted in listing 4.8.

Listing 4.8– Evolution engine structure

```
1 class EvolutionEngine<T, G>(
2     populationConfig: PopulationConfig<T, G>,
3     selectionConfig: SelectionConfig<T, G>,
4     alterationConfig: AlterationConfig<T, G>,
5     evolutionConfig: EvolutionConfig<T, G>,
6 ) : Evolver<T, G> where G : Gene<T, G> {
```

```

8     private var state: EvolutionState<T, G> = EvolutionState.empty(ranker)
9
10    override fun evolve(): EvolutionState<T, G> {
11        do {
12            state = iterateGeneration(state)
13        } while (limits.none { it(state) })
14        return state
15    }
16
17    fun iterateGeneration(state: EvolutionState<T, G>): EvolutionState<T, G> {
18        val interceptedStart = interceptor.before(state)
19        val initialPopulation = startEvolution(interceptedStart)
20        val evaluatedPopulation = evaluatePopulation(initialPopulation)
21        val parents = selectParents(evaluatedPopulation)
22        val survivors = selectSurvivors(evaluatedPopulation)
23        val offspring = alterOffspring(parents)
24        val nextPopulation = survivors.copy(population = survivors.population + offspring.
25            → population)
25        val nextGeneration = evaluatePopulation(nextPopulation)
26        val interceptedEnd = interceptor.after(nextGeneration)
27        return interceptedEnd.copy(generation = interceptedEnd.generation + 1)
28    }
29
30    ...
31
32    class Factory<T, G>(...) where G : Gene<T, G> {
33        ...
34        fun make() = EvolutionEngine(
35            populationConfig = PopulationConfig(genotypeFactory, populationSize),
36            selectionConfig = SelectionConfig(survivalRate, parentSelector,
37                → survivorSelector),
38            alterationConfig = AlterationConfig(alterers),
39            evolutionConfig = EvolutionConfig(
40                limits, ranker, listeners, evaluator.creator(fitnessFunction), interceptor
41            )
42        )
43    }
44 }
```

One of the key design decisions of the engine is to make it store as little information as possible. This decision is motivated by the fact that the engine is the most complex component of the framework, and thus, it should be as lightweight as possible to reduce its reasons for change and increase its maintainability.

The only information stored by the engine is its current state, which is composed of the current generation, the population, and the ranker used to evaluate the population. This state is passed as a parameter to all the functions that need it, and the result of these functions is a new state.

Instead of depending on the engine to maintain information about the evolution process, the framework relies on the ***evolution listeners*** to store information about the evolution process.

The supplementary **Factory** class streamlines the engine configuration process. Although its utilization is recommended for engine instantiation, it remains optional, ensuring flexibility for users.

Finally, the **Evolver** interface is designed to accommodate varying engine implementations, with the **EvolutionEngine** class being the exclusive existing implementation to date.

#### 4.3.4 Evolution Listeners

The *Keen* framework introduces an `EvolutionListener` interface designed with an *observer pattern* to offer users insights into the evolution process in real-time. This interface provides hooks for various events in the evolution cycle, ranging from its commencement and culmination to the beginning and conclusion of each phase, including generation, evaluation, and selection, among others.

The primary objective of the `EvolutionListener` is to empower users to oversee the evolution process, leveraging this data for diverse purposes. Users can employ this interface to conduct a detailed analysis of the evolution process, generate comprehensive reports based on its results, or persistently store the best-performing individual of each generation.

To facilitate these functions, the interface features the following methods:

Listing 4.9– Evolution Listener Interface (some variables and functions omitted for brevity)

```
interface EvolutionListener<T, G> where G : Gene<T, G> {
    fun onGenerationStarted(state: EvolutionState<T, G>) = Unit
    fun onGenerationEnded(state: EvolutionState<T, G>) = Unit
    fun onInitializationStarted(state: EvolutionState<T, G>) = Unit
    fun onInitializationEnded(state: EvolutionState<T, G>) = Unit
    fun onEvaluationStarted(state: EvolutionState<T, G>) = Unit
    fun onEvaluationEnded(state: EvolutionState<T, G>) = Unit
    fun onParentSelectionStarted(state: EvolutionState<T, G>) = Unit
    fun onParentSelectionEnded(state: EvolutionState<T, G>) = Unit
    fun onSurvivorSelectionStarted(state: EvolutionState<T, G>) = Unit
    fun onSurvivorSelectionEnded(state: EvolutionState<T, G>) = Unit
    fun onAlterationStarted(state: EvolutionState<T, G>) = Unit
    fun onAlterationEnded(state: EvolutionState<T, G>) = Unit
    fun onEvolutionStarted(state: EvolutionState<T, G>) = Unit
    fun onEvolutionEnded(state: EvolutionState<T, G>) = Unit
}
```

Within the *Keen* framework, several concrete implementations of this interface are available:

- `EvolutionPlotter`: Visualizes the best, worst, and average fitness throughout the generations in a plot format.
- `EvolutionPrinter`: Outputs essential evolution data every  $n$  generations, detailing aspects such as fitness metrics, the top-performing individual, and the progression of generations.
- `EvolutionSummary`: Prints a comprehensive summary post-evolution, encapsulating fitness statistics, leading individuals, generation count, time metrics for each evolution phase, and the evolution's total duration.

**Remark.** The *Evolution plotter* uses *Lets-Plot Skia Frontend* to render the plots, which is a multi-platform plotting library for Kotlin, meaning that the plots can be rendered in any platform supported by Kotlin Multi-Platform.

The implementation of this interface is a novel feature of the *Keen* framework, and it is not present in the frameworks that were used as a reference for this work.

##### 4.3.4.1 Records

The framework harnesses records to chronicle information pertaining to the evolution process.

**Generation Record** The `GenerationRecord` class embodies data of an individual generation, encapsulating each phase of the evolutionary journey within a generation record.

**Evolution Record** The `EvolutionRecord` class, while centered on the entire evolutionary process, primarily delegates data storage responsibilities to the `GenerationRecord` class.

## 4.4 Genetic Operators

Genetic operators are the driving force behind the evolution process in any genetic algorithm or genetic programming paradigm. These operators stimulate change in the population and guide the search towards optimal solutions. This section illuminates the fundamental genetic operators incorporated within *Keen*, categorized into *Selection*, *Mutation*, and *Crossover*.

*Selection* operators guide the evolutionary process by determining which individuals from the population will contribute to the next generation. These mechanisms often favor individuals that exhibit superior fitness, thereby promoting the inheritance of advantageous traits.

*Mutation* operators introduce randomness into the population by modifying individual genes or gene sequences. This mechanism ensures genetic diversity, aiding the population to explore a broader search space and escape local optima.

*Crossover*, also known as recombination, is another crucial genetic operator. It combines the genetic material from a set of parent individuals to generate offspring, simulating the biological process of sexual reproduction.

Each of these categories encompasses several techniques, providing *Keen* with a versatile repertoire of genetic operators. The subsequent subsections provide a comprehensive discussion on each of these operators, elucidating their working principles and application contexts.

### 4.4.1 Selection

Building upon the concepts presented in definition 2.4 on page 9, the **selector operator** is mathematically represented as  $\Sigma : \mathbb{P} \times \mathbb{N} \times \dots \rightarrow \mathbb{P}$ . At its core, the selector operator processes a given population  $P$ , selects  $n$  individuals based on specific criteria, and outputs a subset of the initial population. To ensure adaptability, the design challenge centers around establishing a versatile interface to cater to diverse selection strategies.

A promising approach consists of introducing an interface encompassing a method that requires the population and desired number of individuals as inputs. We can incorporate any supplementary parameters as properties of a specific selector object. This conceptualization can be defined as:

Listing 4.10– Selector interface

```
interface Selector<T, G> where G : Gene<T, G> {
    fun invoke(
        state: EvolutionState<T, G>, outputSize: Int
    ): EvolutionState<T, G>
}
```

In this model, the `Selector` interface gets parameterized according to the *value* and *gene* types. The incorporated `invoke` method explicitly mentions the state and output size parameters, which are essential for the selection process. Thanks to Kotlin's capacity for operator overloading, this framework allows direct function invocation through the selector object, depicted as `selector(...)`.<sup>1</sup>

#### 4.4.1.1 Random Selector

The `RandomSelector` class acts as a selection strategy, assigning equal selection probabilities to every individual within a population. Its primary utility is to serve as a baseline for comparison with other selection methods.

<sup>1</sup>Syntactic sugar for `selector.invoke(...)`.

**Definition 4.1** (Random Selector). *The **random selector** operates under a uniform approach, ensuring every individual in the population receives the same selection probability. Formally, the RandomSelector is defined as:*

$$\Sigma_{\text{random}} : \mathbb{P} \times \mathbb{N} \rightarrow \mathbb{P}; (P, n) \mapsto \Sigma_{\text{random}}(P, n) \quad (4.1)$$

*Each individual is then selected based on a uniform probability distribution:*

$$\rho_i = \frac{1}{|P|} \quad (4.2)$$

Although the RandomSelector is not frequently mentioned in literature, its straightforward nature provides a valuable touchstone for assessment.

#### 4.4.1.2 Roulette Wheel Selector

The RouletteWheelSelector, commonly known as “fitness proportionate selection”, is a widely-used selection method in GAs. It allocates each individual in a population a segment of a roulette wheel, proportional to its fitness. Thus, individuals with higher fitness have a greater chance of selection.

**Definition 4.2** (Roulette Wheel Selector). *The **roulette wheel selector** ensures selection chances proportional to an individual's fitness. It's defined as:*

$$\Sigma_{\text{roulette}} : \mathbb{P} \times \mathbb{N} \times \{0, 1\}; (P, n, b) \mapsto \Sigma_{\text{roulette}}(P, n, b) \quad (4.3)$$

where  $P$  represents the population,  $n$  is the individuals' count to select, and  $b$  is a boolean for preceding selection with a sort. Let  $\phi_i$  be the fitness of the  $i^{\text{th}}$  individual. Suppose we have a function  $t : \mathbb{R}^n \rightarrow \mathbb{R}^n$  that transforms the fitness into a value that assigns greater probabilities to the fittest individuals according to the ranking strategy.

The transformed fitness, meaning the fitness adjusted to the current ranking strategy, is defined as:  $\Phi' = t(\Phi)$ . Consider  $\Phi'_{\min} = \min \{\Phi_i, 0\}$  the minimum fitness in the population, and  $\phi'_i = \Phi'_i - \min \{\Phi'_{\min}, 0\}$  its adjusted fitness. The selection probability  $p_i$  for the  $i^{\text{th}}$  individual is:

$$\rho_i = \frac{\phi'_i}{\sum_{j=1}^{|P|} \phi'_j} \quad (4.4)$$

Selection is based on a random number between 0 and 1, selecting  $i$  if the summed probabilities up to  $i$  surpass this number.

**Remark.** The fitness adjustment,  $\phi'_i$ , ensures non-negative probabilities. However, it makes the least fit individual unselectable, with a zero probability.

This approach, while ensuring a selection chance for almost every individual, can face challenges. If few individuals have significantly higher fitness, they might dominate, reducing diversity over time. A more in depth discussion on this topic can be found in section 5.5 on page 78.

#### 4.4.1.3 Tournament Selector

The **Tournament Selector** is a selection operator that employs the widely-recognized tournament selection strategy found in evolutionary computation. At its core, this strategy involves randomly selecting a subset of individuals from the population and evaluating them based on their fitness. From this subset, the individual with the highest fitness is deemed the winner and becomes eligible for participation in the formation of the subsequent generation. The “tournament” process is iteratively carried out until the required number of eligible individuals is chosen.

**Definition 4.3** (Tournament Selector). *The tournament selector embodies the tournament selection approach and is formally defined as:*

$$\Sigma_{\text{tournament}} : \mathbb{P} \times \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{P}; (P, n, k) \mapsto \Sigma_{\text{tournament}}(P, n, k) \quad (4.5)$$

Herein,  $P$  represents the population,  $n$  stands for the count of individuals to be deemed eligible, and  $k$  pertains to the tournament or sample size.

The methodological sequence for selection is as follows:

1. A random subset of  $k$  individuals is culled from the population.
2. From this subset, the individual bearing the most superior fitness score is singled out.
3. Steps 1 and 2 are reiterated until the set quota of eligible individuals is reached.

**Remark.** The selection dynamic is profoundly influenced by the tournament size,  $k$ . Opting for a larger  $k$  leans towards elitism, predominantly favoring individuals with superior fitness scores. In contrast, a reduced  $k$  champions diversity, helping to deter the premature convergence of the genetic algorithm.

For an exhaustive delve into the intricacies and theoretical frameworks surrounding the tournament selection modality, the reader is directed to Bickle et al.'s seminal paper [12].

#### 4.4.2 Mutation

Mutation stands as a pivotal operator in the realm of evolutionary computation, serving as the primary catalyst for genetic variation. This operation sporadically alters parts of a solution, commonly referred to as a chromosome or individual. Such alterations ensure that the genetic material remains diverse and rich—qualities vital for producing robust and adaptable solutions. Even though mutation is often linked with genetic algorithms, its importance is felt throughout various evolutionary computation methods, from genetic programming to evolutionary strategies.

The methodology of mutation differs depending on the solution space's nature and the problem's unique demands. This section will explore the myriad mutation strategies, shedding light on their distinctive traits and use-cases.

From our previous discussion in section 2.2.4.2 on page 11, a mutation operator can be articulated as a function  $M : \mathbb{P} \times [0, 1] \times \dots \rightarrow \mathbb{P}$ . This operator selects individuals from a population based on a probability,  $\mu$ , and then alters a portion of its genetic composition.

The approach adopted by Keen mirrors that of selection operators. The proposed interface is as follows:

Listing 4.11– Mutator interface (Note: For clarity, certain parts of the code have been omitted.)

```
Mutator<T, G> where G : Gene<T, G> {
    val individualRate: Double
    val chromosomeRate: Double
    operator fun invoke(
        state: EvolutionState<T, G>, outputSize: Int
    ): EvolutionState<T, G>
    fun mutateIndividual(individual: Individual<T, G>): Individual<T, G>
    fun mutateChromosome(chromosome: Chromosome<T, G>): Chromosome<T, G>
}
```

This design contemplates mutations at different stages of the solution blueprint, focusing primarily on the chromosome level. It provides default `invoke` and `mutateIndividual` methods (omitted from the code for the sake of brevity), directing the mutation task to the `mutateChromosome` method. This design choice facilitates ease for users: they only need to implement the `mutateChromosome` method when devising a new mutator, but still retain the freedom to override other methods if necessary.

#### 4.4.2.1 Random Mutator

The `RandomMutator` represents a fundamental mutation operator that introduces random alterations to the genes within a chromosome. Its versatility lies in its capacity to function with any gene type, making it a universally adaptable mutation mechanism.

To underpin this mutator's generic nature, each gene is equipped with a `mutate` function. This function, when invoked, produces a randomly mutated gene of the same type as the original. By default, the function generates a new gene mirroring the initial gene's value. This design choice provides users with the flexibility to determine if a specific mutation operator should employ this function. As we'll observe, certain mutation operators opt not to utilize it.

**Definition 4.4** (Random Mutator). *The **random mutator** serves as a rudimentary mutation operator, undertaking random alterations to the genes of a given chromosome. Formally, it is articulated as:*

$$M_{\text{random}} : \mathbb{P} \times [0, 1] \times [0, 1] \times [0, 1] \rightarrow \mathbb{P}; (P, \mu, \rho_c, \rho_g) \mapsto M_{\text{random}}(P, \mu, \rho_c, \rho_g) \quad (4.6)$$

Here,  $P$  symbolizes the population,  $\mu$  represents the likelihood of an individual undergoing mutation,  $\rho_c$  is the probability of a chromosome being selected for mutation, and  $\rho_g$  is the probability of a gene being selected for mutation.

#### 4.4.2.2 Bit Flip Mutator

The `BitFlipMutator` stands out as one of the quintessential mutation operators specifically tailored for binary chromosomes. As the name suggests, its operation revolves around inverting individual bits within a chromosome, providing a simple yet effective means of introducing variability in the genetic representation.

At its core, the `BitFlipMutator` iterates over each bit of a binary chromosome. With a predetermined mutation probability  $\mu$ , it decides whether to flip a given bit from 0 to 1 or vice versa. Such a straightforward approach ensures that the fundamental structure of the chromosome remains intact, while still allowing for diverse offspring.

**Definition 4.5** (Bit Flip Mutator). *The **bit flip mutator** is a mutation operator tailored for binary genetic representations. It can be mathematically formulated as:*

$$M_{0/1} : \mathbb{P} \times [0, 1] \times [0, 1] \rightarrow \mathbb{P}; (P, \mu, \rho_c, \rho_g) \mapsto M_{0/1}(P, \mu, \rho_c, \rho_g) \quad (4.7)$$

In this equation,  $P$  symbolizes the population, and  $\mu$  is the predefined probability determining the likelihood of a given individual undergoing mutation,  $\rho_c$  is the probability of a chromosome being selected for mutation, and  $\rho_g$  is the probability of a gene being selected for mutation.

Note that this operator is similar to applying a `RandomMutator` to a binary chromosome with a probability of  $\mu$ . However, when a given gene is selected for mutation, the `BitFlipMutator` always flips the gene's value, whereas the `RandomMutator` randomly assigns a new value to the gene. Take for example a binary chromosome with the following genetic representation: 01101001. If the `BitFlipMutator` selects the fourth gene for mutation (denoted by the red color), it will always produce the following result: 01110001. On the other hand, the `RandomMutator` will randomly assign a new value to the gene, which might result in the following: 01110001 or 01101001.

#### 4.4.2.3 Swap Mutator

The `SwapMutator` is a mutation operator designed to enhance population diversity by interchanging the positions of two genes in a chromosome. Unlike conventional mutation techniques that modify the intrinsic value of a gene, the `SwapMutator` focuses on the positional rearrangement of genes within a chromosome.

**Definition 4.6** (Swap Mutator). *The swap mutator, denoted by*

$$M_{\text{swap}} : \mathbb{P} \times [0, 1] \times [0, 1] \times [0, 1] \rightarrow \mathbb{P}; (P, \mu, \rho_c, \rho_g) \mapsto M_{\text{swap}}(P, \mu, \rho_c, \rho_g) \quad (4.8)$$

*is an operator that, with given probabilities, chooses a chromosome from a population and swaps the positions of two genes within it. In the equation above,  $P$  represents the population,  $\mu$  denotes the probability of mutation,  $\rho_c$  quantifies the likelihood of chromosome selection, and  $\rho_g$  determines the gene selection probability.*

*Its versatility lies in its applicability to any chromosome, regardless of its specific genetic representation.*

The mechanism of the SwapMutator is elucidated below:

Listing 4.12– Schematic representation of the SwapMutator. This representation simplifies the mutator's workings, particularly considering the assumption of chromosome mutability. In contrast, Keen adopts an immutable approach.

```
var genesToSwap = []
for (gene in chromosome) {
    if (random.nextDouble() < ρ_g) {
        genesToSwap += gene
    }
}
for (gene in genesToSwap) {
    chromosome.swap(gene.index, random.nextInt(0..chromosome.size))
}
```

**Remark.** *This swap mutator implementation permits the possibility of selecting the same gene multiple times, leading to potential multiple swaps involving the same gene. While increasing the population's diversity, this approach also reflects Keen's design choice. Implementers seeking to negate this behavior can design a tailored mutator.*

One of the prominent use-cases of the swap mutator is in combinatorial optimization tasks, where gene positioning holds significance. An exemplar problem is the *Travelling Salesman Problem* (TSP), aiming to ascertain the shortest possible route that visits each city precisely once, subsequently returning to the starting city. Given that both the order of visiting and the distances between cities matter, the swap mutator becomes invaluable in exploring potential solutions without altering the cities' identity. Comprehensive studies elucidating the swap mutator's efficacy in TSP are beyond the scope of this thesis. However, interested readers can refer to Bergel's work [36] and Abdoun et al. [28].

#### 4.4.2.4 Inversion Mutator

The class `InversionMutator` embodies the inversion mutation operator. This operator, upon activation, selects a subset of genes from a chromosome and inverts their sequence. Analogous to the `SwapMutator`, it primarily targets the gene ordering.

**Definition 4.7** (Inversion Mutator). *The inversion mutator works by selecting a random subset of genes from a chromosome and subsequently inverting their sequence. Formally, the operator can be represented as:*

$$M_{\text{inv}} : \mathbb{P} \times [0, 1] \times [0, 1] \times [0, 1] \rightarrow \mathbb{P}; (P, \mu, \rho_c, \rho_{ib}) \mapsto M_{\text{inv}}(P, \mu, \rho_c, \rho_{ib}) \quad (4.9)$$

Where:

- $P$  represents the population.
- $\mu$  denotes the mutation probability.
- $\rho_c$  is the chromosome-wise mutation rate.
- $\rho_{ib}$  is the inversion boundary probability. This probability is used to determine the starting and ending index of the inversion.

The operator's function can be outlined as follows:

Listing 4.13– Pseudo-code depicting the inversion mutation process. For illustrative purposes, we consider the chromosome as mutable, even if it isn't in the *Keen* framework.

```

var start = 0
var end = chromosome.size - 1
for (i in chromosome.indices) {
    if (random.double(0..1) < ρib) {
        start = i
        break
    }
}
for (i in chromosome.indices) {
    if (random.double(0..1) > ρib) {
        end = i
        break
    }
}
chromosome.invert(start, end)

```

Two pivotal steps are present in this operation: 1. The determination of the starting index, and 2. The identification of the ending index.

The starting index is discerned by sequentially scanning the chromosome and selecting a gene based on the gene-wise mutation rate. The ending index is determined in a similar manner, initiating from the previously established start index. This methodology ensures only one iteration over the chromosome, optimizing the time complexity of the operation. The inversion is subsequently applied to the selected gene subset.

Much like the SwapMutator, the inversion mutator exhibits prominence in permutation-centric problems, with the TSP being a notable example. Abdoun et al. provide a comprehensive analysis of the inversion operator's efficacy, especially in the context of the TSP [28].

The content is clear, but there are some minor improvements that can be made to further enhance clarity and maintain formality:

**Remark.** A vital consideration must be made when interpreting the inversion operator within the *Keen* framework. Specifically, the gene-wise mutation rate,  $\rho_{ib}$ , signifies the probability of selecting a particular gene for the starting or ending index. It should not be misconstrued as a metric indicating the percentage of genes selected for inversion.

#### 4.4.3 Crossover

Crossover, also termed recombination, serves as an indispensable operator in a plethora of evolutionary computation algorithms. Its primary function is to generate novel candidate solutions by amalgamating attributes from multiple input solutions. This operation, inspired by genetic mechanisms observed in biological reproduction, is pivotal for striking a balance between exploration and exploitation within an algorithm's search space. By fostering a diverse set of solution structures and retaining beneficial traits, crossover steers the evolutionary process closer to optimal or near-optimal solutions.

Within the *Keen* framework, a diverse array of crossover techniques has been conceptualized and developed, each epitomizing a distinctive recombination approach. This section endeavors to furnish an exhaustive overview of these methodologies, elucidating their mechanics, applications, and subtle intricacies.

What distinguishes *Keen* is its adaptable approach to crossover. Many frameworks typically limit a crossover to two parents, but *Keen* offers users the flexibility to determine the number of parents they wish to involve. Research supports this broader view of crossover, with studies advocating for multi-parent crossover approaches, including those by Tsutsui and Yamamura [13], Elsayed et al.[27], and Arram and Ayob[34].

#### 4.4.3.1 Combine Crossover

One of the straightforward strategies for executing crossover in genetic algorithms is to utilize a function that transforms numerical inputs into numerical outputs. Termed the *combine crossover*, this approach processes multiple inputs through a designated function to generate outputs. In *Keen*, this is realized through the `CombineCrossover` class.

**Definition 4.8 (*n*-Combine Crossover).** *The  $n$ -combine crossover operator ingests a set of  $n$  numerical inputs and employs a specific function to them, producing combined offspring. More formally, a combine crossover is represented as:*

$$X_{n\text{-comb}} : \mathbb{P}_{\mathbb{R}} \times (\mathbb{R}^n \rightarrow \mathbb{R}) \times [0, 1] \times [0, 1] \times [0, 1] \rightarrow \mathbb{P}_{\mathbb{R}}; (P, f, \rho_i, \rho_c, \rho_g) \mapsto X_{n\text{-comb}}(P, f, \rho_i, \rho_c, \rho_g) \quad (4.10)$$

with:

- $P$  denoting a population of numerical entities.
- $f$  representing a combiner function that processes a set of numerical inputs to yield a singular numerical outcome.
- $\rho_i$  symbolizing the likelihood of the crossover's application on an individual of the population.
- $\rho_c$  symbolizing the likelihood of the function's application on a chromosome.
- $\rho_g$  indicating the probability of the function influencing a gene.

**Remark.** *The essence of combine crossover is rooted in the multi-parent crossover paradigm, where the number of input parents matches the arity of the employed combiner function.*

Significantly, the *Combine crossover* operator in *Keen* is architected as an **open class**, empowering developers to extend it and craft bespoke combiner operators. A case in point is the `AverageCrossover` operator, which computes the arithmetic mean of the inputs.

#### 4.4.3.2 Single-Point Crossover

Single-point crossover is one of the most foundational and widely utilized crossover methods in the domain of EC. Rooted in its simplicity and efficiency, this method selects a random “crossover point” on the input chromosomes and exchanges the subsequences of genes situated after this point between the two input chromosomes, producing two outputs.

**Definition 4.9 (Single-Point Crossover).** *The single-point crossover operator takes a pair of input chromosomes and randomly selects a crossover point on them. It then exchanges the subsequences of genes situated after this point between the two input chromosomes to produce two outputs. More formally, a single-point crossover is represented as:*

$$X_{sp} : \mathbb{P} \times [0, 1] \times [0, 1] \rightarrow \mathbb{P}; (P, \rho_i, \rho_c) \mapsto X_{sp}(P, \rho_i, \rho_c) \quad (4.11)$$

where:

- $P$  represents a population of binary chromosomes.
- $\rho_i$  symbolizes the likelihood of applying the crossover to an individual.
- $\rho_c$  symbolizes the likelihood of applying the crossover to a chromosome.

**Remark.** *The simplicity of single-point crossover makes it a popular choice in many EC applications. However, it might not always ensure adequate exploration of the search space, especially in problems where gene positions have strong interactions. In these scenarios, permutation crossover methods like ordered crossover and partially mapped crossover might be more effective.*

For better clarity on the single-point crossover operation, see the graphical representation in fig. 2.1 on page 11, which showcases the random selection of a crossover point and the subsequent exchange of gene subsequences.

Note: There exists a generalized version of single-point crossover named *multi-point crossover*. This method selects multiple crossover points on the input chromosomes and exchanges the subsequences of genes situated after these points between the two input chromosomes to produce two outputs. However, this method is currently not supported in *Keen*.

#### 4.4.3.3 Ordered Crossover (OX)

Ordered Crossover [36], commonly referred to as OX, is a specialized permutation crossover. In the realm of genetic algorithms, it is particularly useful for problems where the solution representation is a sequence or order of items. The technique was developed to preserve order and avoid duplicate genes. Its effectiveness is evident in applications like the TSP.

**Definition 4.10** (Ordered Crossover). *The ordered crossover operator takes two parent chromosomes. Through a series of steps, it ensures that the offspring chromosomes inherit the order of sequences or genes from both parents. Formally, the ordered crossover is represented as:*

$$X_{\text{ox}} : \mathbb{P} \times [0, 1] \times [0, 1] \rightarrow \mathbb{P}; (P, \rho_i, \rho_c) \mapsto X_{\text{ox}}(P, \rho_i, \rho_c) \quad (4.12)$$

where:

- $P$  denotes a population of ordered chromosomes.
- $\rho_i$  is the probability of applying the crossover to an individual.
- $\rho_c$  represents the chance of employing the ordered crossover on a chromosome.

There are multiple implementations of the ordered crossover operator. The following displays the variant used in the *Keen* framework:

Listing 4.14– OX algorithm as implemented in the *Keen* framework.

```
(index1, index2) = random.indices in input1
crossSection = input1[index1..index2]
for (i in 0..input1.size) {
    if (i < index1 or i >= index2) {
        if (input2[i] not in crossSection) {
            output1[i] = input2[i]
        }
        if (input1[i] not in crossSection) {
            output2[i] = input1[i]
        }
    }
}
return output1, output2
```

In this approach, a random section from the first parent chromosome is copied to an offspring chromosome in the same position. Subsequently, genes from the second parent chromosome, not in the copied section, fill the offspring in the same order. The second offspring chromosome uses the second parent chromosome for the copied section and the first parent for the remaining genes. For a visual elucidation of this process, refer to fig. 4.1.

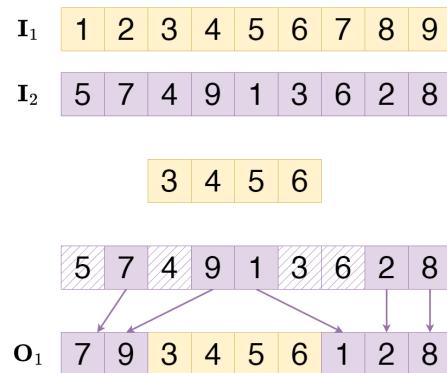


Figure 4.1: Ordered Crossover (OX) operator.

#### 4.4.3.4 Partially Mapped Crossover (PMX)

Partially Mapped Crossover (PMX) is another robust permutation crossover technique highly regarded in the world of genetic algorithms. Introduced by Goldberg and Lingle [5], it primarily caters to problems that involve ordering or sequencing, much like the TSP. PMX has been conceived to ensure output inherit a blend of both input chromosomes without disturbing the position-based information vital for many sequencing problems.

**Definition 4.11** (Partially Mapped Crossover). *The partially mapped crossover (PMX) operates by choosing a random substring of one input chromosome and then mapping the position of these genes on the other input chromosome. This procedure ensures output obtain genes from both inputs, maintaining the original order. Formally, PMX can be articulated as:*

$$X_{\text{pmx}} : \mathbb{P} \times [0, 1] \times [0, 1] \rightarrow \mathbb{P}; (P, \rho_i, \rho_c) \mapsto X_{\text{pmx}}(P, \rho_i, \rho_c) \quad (4.13)$$

where:

- $P$  denotes a population of sequenced chromosomes.
- $\rho_i$  stands for the likelihood of the crossover being applied to an individual.
- $\rho_c$  signifies the probability of triggering the PMX operation on a chromosome.

The implementation of PMX in the *Keen* framework operates as follows:

Listing 4.15– PMX algorithm as described in the *Keen* framework.

```
(output1, output2) = input1, input2
(lo, hi) = random.indices in input1
crossSection1 = input1[lo..hi]
crossSection2 = input2[lo..hi]
for (i in 0..input1.size) {
    if (i < lo or i >= hi) {
        while (output1[i] in crossSection2) {
            index = crossSection2.indexOf(output1[i])
            output1[i] = crossSection1[index]
        }
        while (output2[i] in crossSection1) {
            index = crossSection1.indexOf(output2[i])
            output2[i] = crossSection2[index]
        }
    }
}
return output1, output2
```

The essence of the above implementation is to select a random section from the first input chromosome and copy it to the output chromosome in the same position. Then, the genes from the second input chromosome, not in the copied section, fill the output in the same position as the first input. Finally, the remaining genes are mapped from the second input chromosome to the output chromosome, ensuring that the output chromosome is a valid permutation.

Consider the following example:

**Initial Chromosomes** Let's begin with two input chromosomes:

- $I_1 = 1\ 2345\ 6789$
- $I_2 = 5\ 7491\ 3628$ .

**Section Selection** Next, a random section from the first input chromosome is selected. For instance, consider the section  $I_1[3\dots 6] = 3456$ . This section is replaced with the equivalent section from the second input chromosome,  $I_2[3\dots 6] = 4913$ .

The resulting intermediate output looks like `_ _ 4913 _ _ _`.

**Gene Transfer** The next step involves copying genes. Any genes from the first input chromosome not present in the copied section are transferred to the output chromosome, keeping their original positions. This action results in the intermediate output: `_ 2 4913 7 _`.

**Gene Mapping** The final step is gene mapping. Here, remaining genes from the second input chromosome are mapped onto the output chromosome. The aim is to form a valid permutation. Genes already present in the output chromosome get replaced by corresponding genes from the second input chromosome. This process might require repeated iterations if the mapped gene is already in the output chromosome. The mapping is continued until all genes find their places.

To clarify with an example:

1. Start with the gene “1” from the first position of  $I_1$ . Locate its position in the output and replace it with gene “5” from  $I_2$ , producing the intermediate sequence `5 249137 _`.
2. Take gene “9” from the last position of  $I_1$ . Finding its position in the output reveals another “9”. Thus, refer back to  $I_1$  and take the gene “4”. Since it maps to “3”, which is already in the output, turn to gene “3” in  $I_1$ . This maps to “6” from  $I_2$ , which is finally placed in the output, resulting in  $O_1 = 52491376$ .

**Visualization** For a clearer visual representation of this entire process, refer to fig. 4.2.

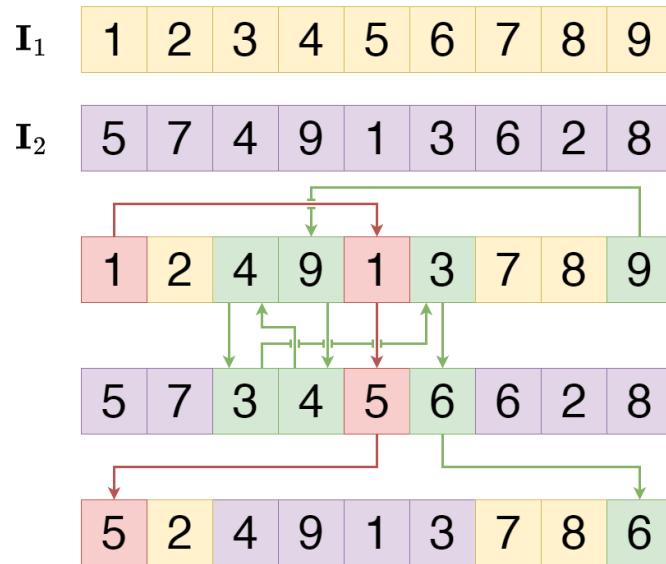


Figure 4.2: Partially Mapped Crossover (PMX) operator.

#### 4.4.3.5 Position Based Crossover (PBX)

Position Based Crossover, commonly known as PBX, is a unique permutation crossover strategy. In genetic algorithms, this technique is invaluable when dealing with problems where the solution is represented as a sequence or order of items. It's designed to preserve the absolute positions of certain genes while allowing for flexibility in others. This leads to its efficacy in tasks like job scheduling.

**Definition 4.12** (Position Based Crossover). *The position based crossover operator works with two input chromosomes. It strategically copies certain positions from one input, while the rest of the genes are filled in from the other input without duplicating any genes. Formally, the position based crossover is illustrated as:*

$$X_{\text{pbx}} : \mathbb{P} \times [0, 1] \times [0, 1] \rightarrow \mathbb{P}; (P, \rho_i, \rho_c) \mapsto X_{\text{pbx}}(P, \rho_i, \rho_c) \quad (4.14)$$

where:

- $P$  signifies a population of ordered chromosomes.
- $\rho_i$  denotes the probability of utilizing the crossover for an individual.
- $\rho_c$  signifies the likelihood of adopting the position based crossover for a chromosome.

The way PBX is implemented in the *Keen* framework is as follows:

Listing 4.16– PBX algorithm as executed in the *Keen* framework.

```
val positions = random.indices in input1
val output = input1
for (i in positions) {
    output[i] = input1[i]
}
var j = 0
for (i in 0..input2.size) {
    if (input2[i] not in output) {
        output[positions[j++]] = input2[i]
        if (j >= positions.size) {
            break
        }
    }
}
return output
```

In the PBX method, select positions from the first input chromosome are directly copied to the output. The remaining genes from the second input chromosome are then filled in the order they appear, skipping any that are already present in the output. A visual representation of this process can be gleaned from fig. 4.3.

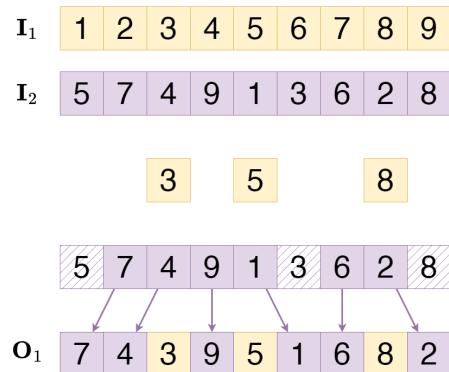


Figure 4.3: Position Based Crossover (PBX) operation.

## 4.5 Genetic Programming

While the principles of Genetic Programming (GP) have been extensively discussed in section 2.3 on page 14, its application within the *Keen* framework offers a unique perspective. This section sheds light on the preliminary implementation of GP in the Keen framework, emphasizing its prototypal nature. As the Keen framework continues its evolution, GP is anticipated to be a focal component, promising a refined synergy between theoretical foundations and innovative enhancements. Although it is still in the early stages, the Keen framework's GP implementation showcases the platform's adaptability and ambition to provide a comprehensive toolset for evolutionary computation.

### 4.5.1 Primitive Set

The quintessential component of any GP implementation is the set of *primitives* that define the language of the GP system. The *Keen* framework's GP implementation is no exception, and its primitives are defined in the `prog` package.

The `prog` package exposes three main components: ***reducible*** expressions, ***program*** trees, and ***generator*** functions.

#### 4.5.1.1 Reducible Expressions

Within the *Keen* framework, “reducible expressions” serve as the building blocks for creating more complex mathematical or logical constructs. Primitives in *Keen* are captured using the `Reducible` interface. This interface outlines a single method, `invoke`, which accepts a list of arguments and yields a single value. Drawing parallels, the `invoke` method resembles a function call, and the `Reducible` interface stands as a function signature.

The architecture of the proposed `Reducible` interface can be appreciated in listing 4.17.

Listing 4.17– The `Reducible` interface

```
interface Reducible<T> : Node<Reducible<T>>, SelfReferential<Reducible<T>> {
    override val contents: Reducible<T> get() = this
    operator fun invoke(environment: Environment<T>, args: List<T>): T
    operator fun <T> Reducible<T>.invoke(environment: Environment<T>, vararg args: T): T =
        invoke(environment, args.toList())
}
```

An important aspect of the `Reducible` interface is the ***environment***; a data structure that stores the values of the variables used in the expression.

It's essential to underline that the `Reducible` interface is an extension of `Node<Reducible<T>>`. This inherently implies that any reducible expression is also a node within an expression tree. Such an arrangement forms a cornerstone of *Keen*'s Genetic Programming (GP) approach as it lays the foundation for generating expressions of arbitrary intricacy.

Venturing forward, *Keen* subdivides the `Reducible` into two primitive subsets: terminals and functions. The former, terminals, embody reducible expressions that act as leaf nodes in an expression tree. Conversely, functions serve as reducible expressions that denote intermediate nodes in said tree.

Listing 4.18– Prototype of the Terminal interface

```
interface Terminal<T> : Reducible<T>, Leaf<Reducible<T>> {
    fun create(): Terminal<T>
    override val arity: Int
        get() = 0
}
```

Listing 4.19– Prototype of the Fun open class

```
open class Fun<T>(
    private val name: String,
    override val arity: Int,
    private val body: (List<T>) -> T
) : Reducible<T>, Intermediate<Reducible<T>> {
    override fun invoke(args: List<T>) = body(args)
}
```

Note that both the Terminal interface and the Fun class are parameterized by a type T. This type represents the return type of the invoke method, and is used to enforce type safety when creating expressions.

A crucial aspect of our design strategy involves the employment of an open class for the Fun class, a decision we consider more fitting than resorting to an interface. Given that functions symbolize expressions with an innate arbitrary behavior, architecting an interface and extending it for each unique function emerges as an impractical venture. In contrast, an open class stands out as a versatile choice, facilitating the crafting of generic functions and paving the way for class extensions to enhance the behavior of the function. This structural framework empowers users to directly harness the Fun class for function creation or extend it for more intricate functions, as exemplified in listing C.4 on page 118.

#### 4.5.1.2 Program trees

The Reducible interface and its subsets—terminals and functions—are foundational components that aid in the construction of program trees in *Keen*. The framework represents these program trees as immutable tree structures, a design decision rooted in advantages such as thread safety, predictability, and reduced chances of accidental data modification, in particular, making program trees immutable one can assure that the arity of a node is always correct.

In essence, a program tree is characterized by a reducible expression situated at its root, which in turn references other programs as its descendants. This design is vividly portrayed in the blueprint provided in listing 4.20.

Listing 4.20– Prototype of the Program class

```
class Program<V>(
    val reducible: Reducible<V>,
    override val children: List<Program<V>> = emptyList()
) : Tree<Reducible<V>, Program<V>>, Copyable<Program<V>> {
    // ...
    operator fun invoke(vararg args: V): V = reducible(children.map { it(*args) })
}
```

The Program class employs a type V as a parameter, symbolizing the return type of the invoke method intrinsic to the Reducible interface. Such an approach ensures type-safety during the construction of program trees. Notably, the invoke function adopts a recursive mechanism. It calls upon the invoke method of the root reducible expression, subsequently passing the results derived from the recursive invocations of its children.

Forming immutable trees, particularly in recursive architectures, can be a challenging endeavor. To counteract this complexity, the Tree interface introduces a method to instantiate trees by conducting a depth-first traversal across a list of nodes. The underlying algorithm, designed to transform a list of nodes into trees, is presented in listing 4.21.

Listing 4.21– Algorithm to create a tree from a list of nodes

```
val stack = []
nodes.reversed().forEach { node ->
    val children = stack.take(node.arity)
    stack.removeAll(children)
```

```

    val node = createNode(node.value, children)
    stack += node
}
return stack.first()

```

This algorithm takes a list of nodes and creates a tree by traversing the list in reverse order. For each node, it takes the required number of children from the top of the stack, removes them from the stack, and creates a new node with the given value and children. Finally, it returns the top of the stack, which is the root of the tree.

#### 4.5.1.3 Generation Methods

Every individual within the population should consistently be a valid program throughout the evolution process. To guarantee the syntactic correctness of all newly created or varied individuals, the GP algorithm must generate an initial population of syntactically accurate individuals. The designated algorithm for producing this initial population is termed the *generation method* or *generator function*.

Koza's influential work [7] introduced two pivotal algorithms for generating the initial population: the *full* and *grow* methods.

The *full* method creates a tree with a specified maximum depth  $D$ , ensuring that all terminal nodes are located at depth  $D$ .

Contrastingly, the *grow* method crafts a tree with a maximum depth of  $D$  and a minimum depth of  $d$ , positioning all terminal nodes at depths ranging between  $d$  and  $D$ . This method was elaborated upon in section 2.3.2 on page 20.

It's noteworthy that the *full* method can be viewed as a specific instance of the *grow* method when  $d = D$ . Given this perspective, we can further conceptualize the *grow* method as a “depth-conditional tree generation” strategy. The detailed algorithm for this is presented in listing 4.22.

Listing 4.22– Depth-conditional tree generation

```

fun generateRecursive(
    intermediates: List<Intermediate>,
    leafs: List<Leaf>,
    depth: Int,
    maxHeight: Int,
    condition: (maxHeight: Int, depth: Int) -> Boolean,
    leafFactory: (Leaf) -> Tree,
    intermediateFactory: (Intermediate, List<Tree>) -> Tree,
) {
    val children = []
    val node = if (condition(maxHeight, depth)) {
        leafFactory(random node from leafs)
    } else {
        val intermediate = random node from intermediates
        repeat intermediate.arity times {
            children += generateRecursive(
                intermediates,
                leafs,
                depth + 1,
                maxHeight,
                condition,
                leafFactory,
                intermediateFactory
            )
        }
    }
    intermediateFactory(intermediate, children)
}

```

```

    }
    return node
}

```

The `generateRecursive` function accepts a list of intermediate nodes, a list of leaf nodes, a maximum height, a current depth, and a condition function. The condition function receives the maximum height and the current depth of a node, returning a boolean value. If the condition is met (true), a leaf node is formed; if not (false), an intermediate node is created. This function is invoked recursively to form the descendants of the current node, halting the recursion when the condition is true or when the maximum height is attained. Additionally, the `generateRecursive` function takes in two other functions dedicated to the creation of both leaf and intermediate nodes, enabling its adaptability for any tree type, including program trees.

A salient feature of this algorithm is its bottom-up tree generation approach, commencing with terminal nodes and culminating with the root node. This strategy facilitates the production of immutable trees, obviating the necessity for crafting temporary trees.

The `generateRecursive` function underpins the `generate` function, which simply invokes the recursive function, setting a maximum height chosen randomly between  $d$  and  $D$ .

Using this method, the grow and full techniques can be illustrated as:

Listing 4.23– Grow and full methods

```

fun <T> generateProgramGrowing(
    terminals: List<Terminal<T>>,
    functions: List<Fun<T>>,
    min: Int,
    max: Int,
): Program<T> {
    val condition = { h: Int, d: Int ->
        d == h ||
        d >= min && random.double() < terminals.size / (terminals.size + functions.size)
    }
    return generateProgram(functions, terminals, min, max, condition)
}

fun <T> generateProgramFull(
    terminals: List<Terminal<T>>,
    functions: List<Fun<T>>,
    min: Int,
    max: Int,
): Program<T> {
    val condition = { height: Int, depth: Int ->
        depth == height
    }
    return generateProgram(functions, terminals, min, max, condition)
}

fun <T> generateProgram(
    functions: List<Fun<T>>,
    terminals: List<Terminal<T>>,
    min: Int,
    max: Int,
    condition: (Int, Int) -> Boolean,
) = Tree.generate(terminals, functions, min, max, condition, ::Program) { f, c ->

```

```
    Program(f, c)
}
```

The `generateProgramGrowing` function obtains a list of terminal nodes, a collection of intermediate nodes, a minimum and maximum height, and returns a program tree. Its condition function is articulated based on the following criteria:

- If the present depth mirrors the maximum height, the condition holds true.
- Should the current depth equal or surpass the minimum height and a random value between 0 and 1 is less than the proportion between the quantity of terminal nodes and the number of intermediate nodes, the condition is validated.
- In all other scenarios, the condition remains unfulfilled.

The `generateProgramFull` function mirrors the aforementioned one, with the distinction that its condition is only validated when the current depth matches the maximum height.

Koza also suggested that population members could be derived using a *ramped half-and-half* technique. Here, the initial population is formed utilizing both the *full* and *grow* strategies, with the method for each individual being selected at random. To implement this, *Keen* offers the `generateProgramWith` function, receiving a list of generation strategies, a compilation of terminals and functions, and a minimum and maximum height. This function randomly chooses a generation method to produce the program tree. By defining the function this way, integrating new generation techniques into the library becomes seamless.

Crucially, with this approach, the *ramped half-and-half* technique simply becomes a specialized version of the `generateProgramWith` function, wherein the list of generation strategies is confined to the *full* and *grow* methodologies.

### 4.5.2 Genetic Operators

The preceding sections introduced genetic operators within the realm of genetic algorithms. Yet, when we transition to Genetic Programming (GP), these operators assume a different character. In GP, the evolutionary landscape is punctuated by program trees or structures, making the operators' responsibilities more multifaceted. They must not only bring about genetic variations but also uphold the structural, syntactic, and semantic sanctity of the programs they modify.

The forthcoming subsections shine a light on the bespoke genetic operators devised for GP in the *Keen* framework, elucidating their roles, mechanisms, and significance.

**Remark.** *This operators are still in a prototypal stage, and their implementation is prone to change in future versions of the framework.*

#### 4.5.2.1 Subtree Crossover

In the realm of Genetic Programming (GP), the subtree crossover operator emerges as a pivotal tool to foster genetic diversity and probe the vast solution space. This operator is adept at melding genetic information from two parent program trees, paving the way for outputs that may exhibit enhanced fitness values.

**Definition 4.13** (Subtree Crossover). *The subtree crossover is a genetic operator designed for tree-structured chromosomes. For two given trees, it randomly chooses a node or subtree in each. The subtrees anchored at these points are then interchanged, birthing two new outputs. The operator can be mathematically expressed as:*

$$X_{\text{subtree}} : \mathbb{P} \times \{0, 1\} \times [0, 1] \times [0, 1] \times [0, 1] \rightarrow \mathbb{P}; (P, e, \rho_i, \rho_c, \rho_g) \mapsto X_{\text{subtree}}(P, e, \rho_i, \rho_c, \rho_g) \quad (4.15)$$

Here's a brief rundown of the parameters:

- $P$ : A population of program trees.
- $e$ : A flag indicating if an individual can participate in crossover more than once.
- $\rho_i$ : Probability of an individual undergoing crossover.
- $\rho_c$ : Chance of initiating the subtree crossover.
- $\rho_g$ : Likelihood of selecting a gene for the crossover process.

**Remark.** Keen's rendition of the subtree crossover currently equips all tree nodes with equal selection probabilities. This strategy may evolve in subsequent versions to encompass a more sophisticated selection mechanism.

To comprehend the subtree crossover's operation within *Keen*, let's walk through the code snippet below:

Listing 4.24– Illustration of Subtree Crossover in *Keen*

```
val (node1, node2) = (random node from input1, random node from input2)
val slices = (gene1.dna.searchSubtree(node1), gene2.dna.searchSubtree(node2))
val newTree1 = gene1.dna.replaceSubtree(slices.first, node2)
val newTree2 = gene2.dna.replaceSubtree(slices.second, node1)
```

Delve deeper into this operator by consulting section 2.3.4.2 on page 26, where a vivid example of the subtree crossover comes alive.

*Keen* is judiciously engineered to address potential challenges posed by the subtree crossover, such as generating unmanageably large trees or producing trees devoid of semantic significance for a given problem. Such astute handling ensures the subtree crossover's unwavering contribution to steering the evolutionary journey towards excellence.

In summation, *Keen*'s GP module masterfully executes the subtree crossover, striking a harmonious balance between bolstering genetic diversity, upholding program integrity, and channeling the evolutionary exploration towards promising horizons.

#### 4.5.2.2 Point Mutation

In Genetic Programming (GP), the capability to mutate program trees is paramount for infusing genetic diversity within a population. Among various mutation operators, Point Mutation stands out. Its specialty lies in introducing delicate perturbations to a program's structure, all the while upholding its syntactic and semantic consistency.

**Definition 4.14** (Point Mutation). *The point mutation operator intricately selects a random node from a program tree, then searches for another node boasting the same arity (i. e., number of child nodes). The initial node is subsequently replaced by the latter, thus producing a variation of the original program while preserving the overarching tree structure.*  
*Mathematically, the operation can be delineated as:*

$$M_{point} : \mathbb{P} \times [0, 1] \times [0, 1] \rightarrow \mathbb{P}; (P, \mu_c, \mu_g) \mapsto M_{point}(P, \mu_c, \mu_g) \quad (4.16)$$

*Explicating the parameters:*

- $P$ : Represents a population of program trees.
- $\mu_c$ : Denotes the probability of a chromosome experiencing mutation.
- $\mu_g$ : Signifies the likelihood of a gene's selection for mutation.

The *Keen* framework houses a lucid implementation of the Point Mutation operator, encapsulated in a succinct four-step methodology:

Listing 4.25– Exemplification of Point Mutation in *Keen*

```
val original = random node from program tree
val replacements = nodes with identical arity as original
val replacement = random node from replacements
val mutated = program tree substituting original with replacement
```

The Point Mutation's cornerstone, especially in *Keen*, is its knack for retaining the architectural coherence of program trees amidst the introduction of genetic variations.

**Remark.** The emphasis on matching arity during node replacement is non-trivial. This meticulous approach ensures the resultant mutated tree mirrors the structure of its predecessor. Such precision mitigates potential runtime anomalies or semantic ambiguities, both of which could impede the evolutionary trajectory.

While Point Mutation might seem simplistic at a cursory glance, its potency shouldn't be underestimated. Its strategy of fine-tuned alterations forms a judicious equilibrium between retaining the essence of high-performing solutions and embarking on novel exploratory ventures.

A visual representation of the Point Mutation operator's workings can be observed in fig. 4.4.

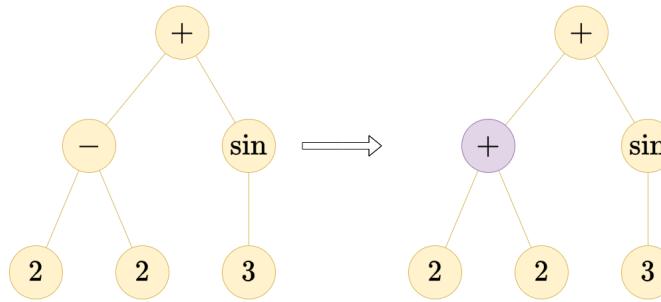


Figure 4.4: A graphical elucidation of the Point Mutation operator's effect on a program tree. A random node is selected and substituted with another of identical arity, resulting in a structurally-consistent variant of the original tree.

#### 4.5.2.3 Subtree Mutation

One key technique among mutation operators in GP is the Subtree Mutation. It serves to adjust the genetic structure of program trees.

**Definition 4.15** (Subtree Mutation). *The subtree mutation operator works by selecting a node from a program tree. It then replaces the subtree originating from this node with a new subtree. This process keeps the general shape of the program while introducing fresh genetic elements, which might enhance the program's effectiveness.*

*To paint a mathematical portrait of this operation:*

$$M_{\text{subtree}} : \mathbb{P} \times [0, 1] \rightarrow \mathbb{P}; (P, \mu_{\text{st}}) \mapsto M_{\text{subtree}}(P, \mu_{\text{st}}) \quad (4.17)$$

*Breaking down the parameters:*

- $P$ : Embodies a population of program trees.
- $\mu_{\text{st}}$ : Represents the probability of a subtree undergoing mutation.

Here's a more simplified version:

In the *Keen* framework, the Subtree Mutation process happens like this:

##### Listing 4.26– Explaining Subtree Mutation in Keen

```
val targetNode = choose random node from program tree
val newSubtree = create a new subtree
val mutatedTree = replace targetNode's subtree in program tree with newSubtree
```

In essence, the Subtree Mutation in *Keen* adds new genetic components without heavily changing the main structure of the tree.

**Remark.** The Subtree Mutation is great at making significant genetic changes. However, it's important to make sure the new subtrees fit well with the problem you're solving. Without this check, you might end up with trees that don't make sense or are too complex.

*Keen* handles this by using generation methods mentioned in section 4.5.1.3 on page 63.

The beauty of the Subtree Mutation method lies in its balance. It keeps what works while trying out new genetic possibilities.

For a visual representation of the Subtree Mutation process, see fig. 4.5.

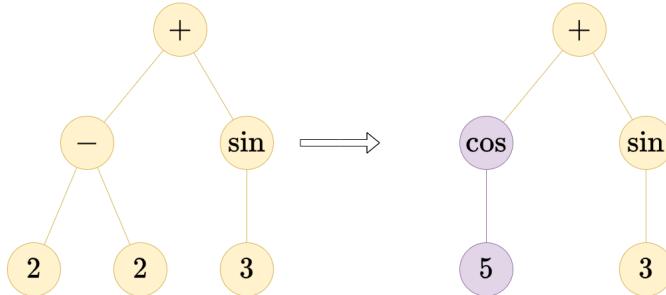


Figure 4.5: A visual guide to the Subtree Mutation process. The operator picks a node and replaces its attached subtree with a new one, giving a changed but still recognizable program tree.

## 4.6 Extensibility

*Keen* is architected with adaptability at its core, catering to a wide array of genetic problems and computational scenarios. This adaptability is principally achieved through the combination of intuitive interfaces and the integration of the Factory Method Pattern.

### 4.6.1 Structured Extensibility Through Interfaces

*Keen* employs well-defined interfaces to encapsulate essential functionalities. These interfaces act as templates, enabling developers to easily extend and modify the framework. The process is straightforward: to enhance or customize capabilities, one simply implements the provided interfaces.

### 4.6.2 Factory Method Pattern in Genetic Construction

The Factory Method Pattern plays a pivotal role in *Keen*, facilitating the dynamic creation of genetic material. By decoupling the framework from specific implementations, this pattern provides the flexibility to introduce varied genetic construction techniques seamlessly.

Consider the `Chromosome.Factory` interface and the `Chromosome.AbstractFactory`. These provide a standardized mechanism for generating chromosome objects, ensuring easy adaptability to diverse problem domains such as the Knapsack problem (refer to chapter 6 on page 79).

For illustration, consider the `SimpleChromosome`:

Listing 4.27– Implementation of the `SimpleChromosome` using the `Chromosome.Factory` interface.

```
class SimpleChromosome(override val genes: List<SimpleGene>) :
    Chromosome<Int, SimpleGene> {
    // ... Other implementations ...
}

class Factory(override var size: Int, private val geneFactory: () -> SimpleGene) :
```

```

    Chromosome.AbstractFactory<Int, SimpleGene>() {
        override fun make() = SimpleChromosome((0 until size).map { geneFactory() })
    }
}

```

Using the `SimpleChromosome` is straightforward:

Listing 4.28– Utilization of the `SimpleChromosome` within an evolutionary engine setup.

```

data class SimpleGene(val dna: Int) : Gene<Int, SimpleGene> {
    // ... Implementation details for SimpleGene ...
}

fun main() {
    val engine = engine(
        ::fitnessFn,
        genotype {
            chromosome {
                SimpleChromosome.Factory(10) { SimpleGene((0..100).random()) }
                    // Random genes between 0 and 100
            }
        }
    ) {
        // ... Other configurations ...
    }
    val result = engine.evolve()
    // ... Use the result ...
}

```

### 4.6.3 Sustaining Future Evolution

Designed with a forward-looking perspective, *Keen* leverages interfaces and the Factory Method Pattern to ensure its sustained evolution. This positions *Keen* advantageously, ensuring it remains relevant and adaptable to emerging challenges in genetic algorithms and computations.

## 4.7 Conclusion

Throughout this chapter, we've journeyed through the conceptual foundations and architectural intricacies of the *Keen* framework. As we've seen, *Keen* is not just another computational tool, but a flexible and extensible platform poised to tackle the myriad challenges associated with genetic algorithms and related computational techniques. Its design philosophy, emphasizing modularity and extensibility, promises to make it a valuable asset for both usage and research.

But, understanding a framework's theory and architecture is just the beginning. The true power and versatility of *Keen* are best experienced in action. In the subsequent chapters, we will shift our focus from the abstract to the concrete. We'll explore how to harness *Keen*'s capabilities for real-world problems, diving deep into its usage patterns and practical extensions.

As we move forward, remember the essence of *Keen*: a platform built for evolution, by evolution.



## Chapter 5

# Case Study: Real Function Optimization

### 5.1 Introduction

Real function optimization is a prevalent task in numerous fields, from *data science* and *machine learning* (ML) to *operations research* and *engineering*. It presents a common class of challenges that can be effectively addressed using GAs.

This chapter embarks on a practical exploration of *Keen*. We aim to illustrate the robust capabilities of *Keen* by employing it to solve various classic optimization problems, the details of which are provided in appendix B on page 99.

Please note that this chapter doesn't intend to delve into the formulation of these optimization problems; the discussion is primarily centered around how *Keen* can be leveraged as a solution tool. For a comprehensive understanding of *Keen*, its design, and its features, refer to chapter 4 on page 43.

The ensuing sections will navigate through the problem descriptions, the solutions employed using *Keen*, the corresponding results, and the consequent analysis. In the following sections, we will systematically explore and demonstrate the efficiency and versatility of *Keen* in tackling a range of optimization problems.

### 5.2 Problem Description

The primary task in this chapter revolves around the optimization of several real functions. These functions, which can be found in appendix B on page 99, are selected based on their complexity and their widespread use in the field of optimization. We aim to find the **global minimum** for each of these functions, a problem known as global optimization.

In the context of this case study, the search space is defined as the set of all possible solutions to the optimization problem. Since we are dealing with real-valued functions, the search space will be a subset of the set of real numbers,  $\mathbb{R}$ . The exact subset that forms the search space can vary depending on the function we are optimizing. For example, the *Cross-in-Tray* function has a search space of  $[-10, 10]^2$ , whereas the *Easom* function has a search space of  $[-100, 100]^2$ .

Given the nature of our functions and their respective search spaces, it is vital to understand some fundamental concepts related to sets, cardinalities, and functions. These concepts will not only help us better comprehend the vastness and complexity of the search space but will also lay the groundwork for the use of *Keen* in this context.

The following definitions, theorems, and corollaries will elucidate the necessary background.

**Definition 5.1** (Set cardinality inequality). *For any two sets A and B, if there exists an injective function  $f : A \rightarrow B$ , then  $|A| \leq |B|$ .*

**Theorem 5.1** (Schröder–Bernstein theorem). *Given two sets A and B, if there exist two injective functions  $f : A \rightarrow B$  and  $g : B \rightarrow A$ , then there exists a bijective function  $h : A \rightarrow B$ .*

The proof for theorem 5.1 on the preceding page is attributed to J. König [[konigTheorieEnsembles1906](#)].

**Corollary 5.1.1.** *For any two sets A and B, if  $|A| \leq |B|$  and  $|B| \leq |A|$ , then  $|A| = |B|$ .*

**Proof.** The conditions  $|A| \leq |B|$  and  $|B| \leq |A|$  imply that there exist injective functions  $f : A \rightarrow B$  and  $g : B \rightarrow A$ , respectively. By the Schröder–Bernstein Theorem (theorem 5.1 on the previous page), the existence of these injective functions guarantees a bijective function  $h : A \rightarrow B$ .

A bijective function is one that is both injective and surjective. This means that every element of  $A$  is mapped to a unique element in  $B$  and every element of  $B$  is the image of some element in  $A$ .

Therefore, there can't be more elements in  $A$  than in  $B$  (or vice versa), as this would contradict the surjectivity or the injectivity of the function  $h$ . Thus, by the definition of set cardinality (definition 5.1 on the preceding page), we conclude that  $|A| = |B|$ .  $\square$

**Theorem 5.2.** *Let  $(x, y) \in \mathbb{R}^2$  with  $x \neq y$ . Then, the cardinality of the interval  $[x, y]$ , denoted as  $|[x, y]|$ , is the same as the cardinality of  $\mathbb{R}$ . Thus,  $|[x, y]|$  is uncountable.*

**Proof.** Let us define a function  $f : [x, y] \rightarrow \mathbb{R}$  defined as  $f(z) = \tan\left(\frac{\pi(z-x)}{y-x} - \frac{\pi}{2}\right)$  which maps the interval  $[x, y]$  bijectively onto  $\mathbb{R}$ .

This means that there is an injective function  $f : [x, y] \rightarrow \mathbb{R}$  and an injective function  $g : \mathbb{R} \rightarrow [x, y]$ .

By the definition of cardinality inequality (definition 5.1 on the previous page), we conclude that  $|[x, y]| \leq |\mathbb{R}|$  and  $|\mathbb{R}| \leq |[x, y]|$ .

By theorem 5.1 on the preceding page, we conclude that  $|[x, y]| = |\mathbb{R}|$ .  $\square$

These results make it clear that, in most cases, the search space for our real function optimization problems will have an uncountable number of potential solutions. This vastness and complexity underline the necessity for robust and efficient optimization methods like those provided by Keen.

## 5.3 Solution

To solve this problem, we aim to optimize certain functions using a Genetic Algorithm (GA). This approach will leverage a double-valued chromosome where each gene represents the  $x_i$  values. The fitness function encapsulates the function we aim to optimize. Let's break down the process:

### 5.3.1 Chromosome Structure

Each gene in the chromosome will be a double-valued gene, with a specified range of valid values depending on the optimization function. For the Bukin function N.6 (appendix B.4 on page 100), valid values range between  $[-15, -5]$  for the first gene, and  $[-3, 3]$  for the second.

### 5.3.2 Variation Operators

Regardless of the specific function we're optimizing, we will consistently employ the following variation operators:

- Recombination: *average crossover* (see section 4.4.3.1 on page 56)
- Mutation: *random mutation* (see section 4.4.2.1 on page 53)

For selection, we compare the performance of *random selector*, *tournament selector*, and *roulette selector* (see section 4.4.1.1 on page 50, section 4.4.1.3 on page 51, and section 4.4.1.2 on page 51).

### 5.3.3 Implementation

For illustration purposes, we'll employ the Bukin function N.6. This function's definition is as follows:

Listing 5.1– Implementation of the Bukin function N.6 in the context of *Keen*.

```
fun bukinN6(genotype: Genotype<Double, DoubleGene>) = genotype.flatten().let { (x, y) ->
    100 * (y - 0.01 * x.pow(2) + 1).pow(2) + 0.01 * (x + 10).pow(2)
}
```

Here, the function receives a genotype and computes the fitness value. The genotype gets flattened into a pair of doubles, representing  $x$  and  $y$  values. We then calculate the fitness using the Bukin function N.6.

To reuse the evolution engine creation across different optimization functions, we'll abstract its definition:

Listing 5.2– Function that creates an evolution engine for a given function to optimize.

```
fun createEngine(
    fitnessFunc: (Genotype<Double, DoubleGene>) -> Double,
    vararg ranges: DoubleRange,
) = engine(
    fitnessFunc,
    genotype {
        chromosome {
            doubles {
                this.size = 2
                this.ranges += ranges
            }
        }
    }
) {
    populationSize = 500
    optimizer = FitnessMinimizer()
    alterers = listOf(
        RandomMutator(0.06),
        AverageCrossover(0.3)
    )
    limits = listOf(SteadyGenerations(50))
}
```

The preceding code snippet defines a function that receives a fitness function and a list of ranges of valid values for each gene.<sup>1</sup> The function then creates an evolution engine that uses the received fitness function and a genotype with two double-valued genes. The ranges of valid values for each gene are defined by the received list of ranges.

The constructed evolution engine uses the provided fitness function and a genotype with two double-valued genes, defined by the input list of valid value ranges.

The engine's configuration encompasses:

- A population size of 500 individuals.
- A *fitness minimizer* optimizer.
- Alterers like *random mutator* with a mutation rate of 0.06 and *average crossover* with a crossover rate of 0.3. These values, chosen after some experimentation, may not necessarily be the best for this problem.
- A *steady generations* limit of 50 generations, meaning that the evolution will stop after 50 generations without any improvement.

<sup>1</sup>The `vararg` keyword allows us to pass a variable number of arguments to a function. The arguments are then accessible inside the function as an array. This is equivalent to using `*args` in Python, or `Object... args` in Java.

With the evolution engine set up, you can now define a function to run the Genetic Algorithm (GA) for a designated optimization function:

Listing 5.3– Function that runs the GA for a given function to optimize.

```
fun main() {  
    val engine = createEngine(::bukinN6, -15.0..-5.0, -3.0..3.0)  
    val result = engine.evolve()  
    println("Result: $result")  
}
```

This function initializes an evolution engine for the Bukin function N.6 and then executes the GA. Post-evolution, the resultant output is printed.

## 5.4 Results

This section presents the optimization results of 20 classical functions using three different selectors: Random, Tournament, and Roulette. The tables detail the time spent in each phase of the evolutionary process, as well as the outcomes from the optimization procedures.

Each experiment was run 4 times, dropping the first iteration to avoid JVM warmup effects. The results presented are the average of the remaining 3 iterations.

### 5.4.1 Random Selector

Function	Selection	Total
Ackley ms	0.22476124444444412	1.1540964333333334 s
Beale ms	0.12055760000000007	1.0416774 s
Booth ms	0.12132855555555554	1.0239565333333334 s
Bukin N.6 ms	0.12815200000000032	0.9967993333333333 s
Cross-in-tray ms	0.12550486666666665	1.033344 s
Easom ms	0.1243147333333315	1.0729220333333334 s
Egg holder ms	0.1178918222222216	0.9910291999999999 s
Goldstein-Price ms	0.1326573999999999	0.9626545333333333 s
Himmelblau ms	0.1347557777777776	1.0103663999999999 s
Holder table ms	0.1274720666666666	1.0294960666666666 s
Levi ms	0.11814286666666658	1.0163812333333333 s
Matyas ms	0.1249572444444447	1.0444310666666665 s
McCormick ms	0.1204162444444456	0.9626910666666667 s
Rastrigin ms	0.1215248222222223	1.0027034333333333 s
Rosenbrock ms	0.1310438888888885	1.1014783333333333 s
Schaffer N.2 ms	0.11766186666666661	1.0315578666666665 s
Schaffer N.4 ms	0.1170894222222206	1.043442666666667 s
Styblinski-Tang ms	0.1236970222222217	1.0122185666666665 s
Sphere ms	0.1254926888888888	1.0411549333333334 s
Three-hump camel ms	0.1298584888888882	1.0756696333333333 s
Average	0.1293640311111111	1.0324035166666667

Table 5.1: Time spent in each phase of the evolution process with a Random Selector.

Function	Generations	Fittest	Error
Ackley	500.0	(-0.13394883451482661, 0.10179317814216116)	3.943728441167162
Beale	500.0	(0.07823776218960378, 0.11955346519294004)	13.42706286558303
Booth	500.0	(-2.1850854919936555, -0.0199912893706677)	254.06273318409762
Bukin N.6	500.0	(-10.80479985228797, -0.2151680192633476)	125.27612990733026
Cross-in-tray	500.0	(1.5328646464195066, -0.16240602749885066)	0.1554508011743374
Easom	500.0	(17.787362052439622, 19.41394522290803)	1.0
Egg holder	500.0	(-33.73042665354185, 66.13578937508247)	898.8723913876921
Goldstein-Price	500.0	(0.13236613718684337, 0.13254815514598114)	1084.6647893162292
Himmelblau	500.0	(1.1508780026445862, -0.8755824864740784)	129.3648232523963
Holder table	500.0	(0.5196152859095905, 0.5585089785935319)	18.68846675464214
Levi	500.0	(-0.42322048197184964, 0.12505388160608266)	9.131079684541946
Matyas	500.0	(-0.0695391039455038, 0.49496763044637054)	0.8715086285046144
McCormick	500.0	(1.1064923222718617, 0.918660628954033)	4.594145921707965
Rastrigin	500.0	(0.4106889277736023, -0.2661147890304942)	7.975168917614923
Rosenbrock	500.0	(-0.02408663355305204, -0.07775339200797114)	12.246229264836906
Schaffer N.2	500.0	(-2.728183302602101, 4.711117086548835)	0.4658580287979734
Schaffer N.4	500.0	(-18.217116384258492, 12.447576482522038)	0.49559481620297513
Styblinski-Tang	500.0	(-0.40060119125064036, -0.26453235448743)	24.213652024305162
Sphere	500.0	(0.02831845956873839, 0.957533075380755)	2.9049027371863825
Three-hump camel	500.0	(-0.04249951382806109, -0.07015804752693917)	0.7408489164696809
Average	500.0	—	129.65472824252404

Table 5.2: Results of the optimization process with a Random Selector.

### 5.4.2 Tournament Selector

Function	Selection	Total
Ackley ms	0.1722495107632091	0.3283717666666666 s
Beale ms	0.08278604882909812	0.3957252666666667 s
Booth ms	0.05777837582625109	0.34720913333333336 s
Bukin N.6 ms	0.055187279151943464	0.3456149333333333 s
Cross-in-tray ms	0.05710198598130846	0.30195846666666665 s
Easom ms	0.05908053927315363	0.2919613666666667 s
Egg holder ms	0.05792807017543866	0.4810475 s
Goldstein-Price ms	0.05974518686296707	0.2916315 s
Himmelblau ms	0.05722125256673511	0.37256146666666673 s
Holder table ms	0.05875948553054657	0.3632463666666667 s
Levi ms	0.07155889164598853	0.3182966 s
Matyas ms	0.06000960302457464	0.6332080666666666 s
McCormick ms	0.0552641059602649	0.2722128666666666 s
Rastrigin ms	0.05857562500000002	0.27712773333333335 s
Rosenbrock ms	0.053423104265402824	0.5166973333333333 s
Schaffer N.2 ms	0.057860843373494036	0.2961707666666667 s
Schaffer N.4 ms	0.0553688748685595	0.30088173333333335 s
Styblinski-Tang ms	0.05994333333333314	0.2566216 s
Sphere ms	0.061840923694779164	0.5332526666666667 s
Three-hump camel ms	0.08042557681485646	0.6974308666666666 s
Average	0.06660543084709522	0.3810614

Table 5.3: Time spent in each phase of the evolution process with a Tournament Selector.

Function	Generations	Fittest	Error
Ackley	113.55555555555554	(2.714557171913187E-10, 8.338312829559817E-10)	2.5008522660148933E-9
Beale	223.0	(2.9756742560022995, 0.4930758899822217)	2.443501094058878E-4
Booth	117.666666666666667	(1.0, 3.0)	0.0
Bukin N.6	125.777777777777779	(-9.999920778058002, -1.4585757385130478E-5)	6.633869150552885E-10
Cross-in-tray	95.11111111111113	(-1.349406597635503, -1.3493998085716579)	1.8706670414185094E-6
Easom	94.777777777777777	(3.1415926539169874, 3.141592653689459)	0.0
Egg holder	202.666666666666666	(149.46897626232973, 421.6346724107953)	32.705629617656314
Goldstein-Price	98.1111111111111	(2.0276310494922466E-9, -0.9999999992683337)	7.71234927772942E-14
Himmelblau	108.22222222222221	(3.1948094467768304, 0.7172911576785322)	0.0
Holder table	172.77777777777778	(-2.6844266114600797, 9.665720249842868)	5.754373661659429E-5
Levi	134.33333333333334	(1.0, 1.0)	1.3497838043956716E-31
Matyas	293.88888888888889	(-3.9676961981915714E-20, -4.041678369961017E-20)	1.9286074504938834E-40
McCormick	83.88888888888889	(-0.36998949140111453, -1.5470317233766142)	0.0874615740477706
Rastrigin	88.88888888888889	(-1.8951401788104123E-11, 6.84897169857076E-11)	0.0
Rosenbrock	187.55555555555554	(0.9726960167315691, 0.9464774700495662)	0.0010178373288836763
Schaffer N.2	92.22222222222223	(-2.000697449624541E-8, -4.434497313163789E-9)	0.0
Schaffer N.4	105.666666666666667	(-0.0069024657494170345, 0.49954418513095655)	0.29258768921744405
Styblinski-Tang	90.0	(-2.9035340261657274, -2.9035340275650756)	39.16634140754285
Sphere	276.66666666666667	(7.549263436563391E-9, 1.4287115935043707E-8)	7.833391805280381E-16
Three-hump camel	394.88888888888889	(2.0969751991550495E-10, -2.0932706061327837E-10)	2.6360565768184573E-19
Average	154.9833333333333	—	3.6126670946735318

Table 5.4: Results of the optimization process with a Tournament Selector.

### 5.4.3 Roulette Selector

Function	Selection	Total
Ackley ms	0.17056164444444416	1.0807452 s
Beale ms	0.15100291111111053	1.0999544333333333 s
Booth ms	0.13605653333333298	1.0591936333333332 s
Bukin N.6 ms	0.14861524444444424	1.1450094000000002 s
Cross-in-tray ms	0.1362319333333331	1.0848857333333333 s
Easom ms	0.20113288888888917	1.0845391999999998 s
Egg holder ms	0.1496153777777776	1.202205 s
Goldstein-Price ms	0.14033877777777787	1.0740913 s
Himmelblau ms	0.1547297111111113	1.1500744666666667 s
Holder table ms	0.13423242222222231	1.0950766333333333 s
Levi ms	0.14451297777777808	1.0559435666666666 s
Matyas ms	0.15002555555555597	1.1181167 s
McCormick ms	0.13963766666666672	1.0885313666666667 s
Rastrigin ms	0.15620186666666616	1.1324616333333333 s
Rosenbrock ms	0.16064142222222247	1.0342551666666668 s
Schaffer N.2 ms	0.1498355555555554	1.0825938333333334 s
Schaffer N.4 ms	0.14426979999999978	1.0626771999999998 s
Styblinski-Tang ms	0.150212288888889	1.1550819666666667 s
Sphere ms	0.14043651111111116	1.0473681666666665 s
Three-hump camel ms	0.13449742222222233	1.0629636 s
Average	0.1496394255555555	1.09578841

Table 5.5: Time spent in each phase of the evolution process with a Roulette Selector.

Function	Generations	Fittest	Error
Ackley	500.0	(1.0488541629733732, 0.8957445383336756)	6.839668782944339
Beale	500.0	(0.3081505517495053, -0.17484593986456695)	10.591099547320836
Booth	500.0	(0.9228224735100191, 0.11791814539701513)	58.19770979283484
Bukin N.6	500.0	(-10.672766796668993, 0.45742007821650027)	15.66706099917838
Cross-in-tray	500.0	(0.552029931564722, 0.4361050357937735)	0.242671640561385
Easom	500.0	(3.1327776189248646, 3.1245350565779546)	0.02929707999400148
Egg holder	500.0	(96.87188731616186, 93.90325777549764)	925.2665967126995
Goldstein-Price	500.0	(0.15323069574941106, -0.12414246714202803)	1960.71891810742
Himmelblau	500.0	(1.23650088121654, 1.099883428818093)	245.76742587320123
Holder table	500.0	(2.662324487653543, -9.255820683804545)	1.8446389200179343
Levi	500.0	(0.2608301637487597, 0.03800459994074693)	17.770157848017817
Matyas	500.0	(-0.9471069479231572, -0.19434283656504645)	1.0775276439400463
McCormick	500.0	(1.8138819876395145, -0.09418393411948285)	4.3745193757530565
Rastrigin	500.0	(0.220516897122612, 1.3071386843051218)	17.28431729816618
Rosenbrock	500.0	(-0.09011204721660471, -0.14523942448830085)	12.91514333522403
Schaffer N.2	500.0	(-6.823689518504916, 3.796492519126597)	0.4026344237361048
Schaffer N.4	500.0	(5.0805888642370105, -2.525877929661206)	0.5373324019814119
Styblinski-Tang	500.0	(-2.1216420971940946, -2.4435406767197017)	27.991493859541265
Sphere	500.0	(-0.3796898245556603, 0.1057702219749745)	0.9625662569996849
Three-hump camel	500.0	(0.8603201794580596, -0.02766857565810721)	1.1328783825638105
Average	500.0	—	165.4806829141048

Table 5.6: Results of the optimization process with a Roulette Selector.

## 5.5 Conclusion

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

# Chapter 6

## Case Study: Knapsack Problem

### 6.1 Introduction

The knapsack problem stands as a classic optimization dilemma, posing the task of determining the optimal combination of items with given weights and values while adhering to a weight constraint. Its practical relevance spans diverse fields, making it an essential study in the realm of optimization.

This chapter delves into the two prominent variants of the knapsack problem: the **unbounded** and the **0-1** knapsack. The former allows an infinite number of each item, whereas the latter restricts each item to a singular instance. Both versions are analyzed in depth, elucidating their inherent complexities and solutions.

### 6.2 Problem Description

The **Knapsack Problem** is a canonical combinatorial optimization problem. It involves selecting items, each characterized by a specific weight and value, to maximize the total value while adhering to a weight constraint.

The goal is to ascertain the optimal subset of items (maximize profit) for the knapsack, ensuring the total weight remains under the knapsack's capacity.

Two primary variants of the Knapsack Problem exist: **0-1 Knapsack** ( $K_{0/1}$ ) [4, 36, 39] and **Unbounded Knapsack** ( $K_*$ ) [14, 36].

#### 6.2.1 0-1 Knapsack Problem

In this variant, each item can be taken once or disregarded. This constraint can be expressed as  $x_i \in \{0, 1\}$ . The mathematical formulation is:

$$\begin{aligned} & \text{maximize} && \sum_{i=1}^n v_i x_i \\ & \text{subject to} && \sum_{i=1}^n w_i x_i \leq W \\ & && x_i \in \{0, 1\} \quad \forall i \in \{1, 2, \dots, n\} \end{aligned} \tag{6.1}$$

Analyzing the problem's search space, we note that the potential combinations of selected items equate to a size of  $2^n$ . Given that the **Subset Sum Problem**, to which this problem is reducible, is NP-hard,  $K_{0/1}$  inherits this complexity.

**Example:** Consider 8 items  $(w, v) = \{(11, 1), (21, 11), (31, 21), (33, 23), (43, 33), (53, 43), (55, 45), (65, 55)\}$ . For a knapsack with a capacity  $W = 110$ , the optimal selection consists of items 1, 3, 4, 5, and 6. This results in a cumulative value of 159 and a total weight of 109.

## 6.2.2 Unbounded Knapsack Problem

Here, there's no limit to the number of instances of each item that can be selected.

Formally, the problem can be expressed as:

$$\begin{aligned} & \text{maximize} && \sum_{i=1}^n v_i x_i \\ & \text{subject to} && \sum_{i=1}^n w_i x_i \leq W \\ & && x_i \in \mathbb{N} \quad \forall i \in \{1, 2, \dots, n\} \end{aligned} \tag{6.2}$$

It's easy to see that the search space is infinite, since there's no limit on the number of each item that can be selected. The problem is still NP-hard.

To evaluate the performance of the solution, we are going to use the benchmarks from [14].

## 6.3 Solution Strategies for Knapsack Problems

Given two distinct knapsack problem variants, our approach encompasses unique solutions tailored to each. We employ the *Keen* framework's built-in Genetic Algorithm (GA) capabilities for the K<sub>0/1</sub> problem and augment the *Keen* framework for the K<sub>\*</sub> problem.

### 6.3.1 0-1 Knapsack Problem

To solve the 0-1 Knapsack problem:

1. **Problem Representation:** We represent each item's inclusion in the knapsack as a binary string. This string is facilitated by the `BoolGene` class, while the collection of these genes is managed by the `BoolChromosome` class, essentially acting as a wrapper for the `BoolGene`.
2. **Item Representation:** Each item is captured as a `Pair` of integers, with the first element indicating the item's value and the second indicating its weight.
3. **Fitness Function Definition:** The core objective of the fitness function is to maximize the total value, ensuring the aggregated weight does not breach the knapsack's weight limit. Should the weight exceed the allowed capacity, a proportional penalty is applied to the fitness score.

The fitness function for this problem, implemented in Kotlin, is as follows:

Listing 6.1 – Fitness Function for the 0/1 Knapsack Problem via Keen

```
fun fitnessFn(genotype: Genotype<Boolean, BoolGene>): Double {
    val profit = (genotype.flatten() zip items).sumOf { (isInBag, item) ->
        if (isInBag) item.first else 0
    }
    val weight = (genotype.flatten() zip items).sumOf { (isInBag, item) ->
        if (isInBag) item.second else 0
    }
    val penalty = if (weight > MAX_WEIGHT) abs(weight - MAX_WEIGHT) else 0
    return (profit - penalty).toDouble()
}
```

## 6.4 Results

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

## 6.5 Conclusion

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

## Chapter 7

# Case Study: Crash Reproduction

### 7.1 Introduction

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

### 7.2 Problem Description

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between

this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

### 7.3 Solution

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the

selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

## 7.4 Results

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

## 7.5 Conclusion

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

# Chapter 8

## Conclusions

### 8.1 Summary

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

## 8.2 Contributions

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

## 8.3 Future Work

### 8.3.1 Genetic Algorithms

#### 8.3.1.1 Genetic Operators

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between

this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

### 8.3.2 Genetic Programming

#### 8.3.2.1 Primitive Set

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

### 8.3.2.2 Genetic Operators

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

### 8.3.3 Multi-Objective Evolution

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

### 8.3.4 Neuroevolution

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

### 8.3.5 Coevolution

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

### 8.3.6 Evolution Strategy

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between

this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

### **8.3.7 Other evolutionary algorithms**

#### **8.3.7.1 Ant Colony Optimization**

#### **8.3.7.2 Artificial Immune Systems**

#### **8.3.7.3 Differential Evolution**

#### **8.3.7.4 Estimation of Distribution Algorithms**

#### **8.3.7.5 Evolutionary Programming**

#### **8.3.7.6 Grammatical Evolution**

#### **8.3.7.7 Particle Swarm Optimization**

### **8.3.8 Crash Reproduction**

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

After this fourth paragraph, we start a new paragraph sequence. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

This is the second paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

And after the second paragraph follows the third paragraph. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.



# Appendix A

## Glossary

In this appendix we present a glossary of terms used throughout this document that may be unfamiliar to the reader and are not defined in the main text.

### A

**Definition A.1** (Alteration). *See section 2.2.4 on page 9*

**Definition A.2** (Alterer). *See definition 2.5 on page 10.*

**Definition A.3** (Arity). *The number of arguments a function takes, or the number of children a node has in a tree.*

### C

**Definition A.4** (Chromosome). *Representation of a single column of genetic information of a candidate solution to a given optimization problem.*

*Formally, a chromosome is a vector  $c = (g_1, g_2, \dots, g_n)$ , where  $g_i$  is a gene (see definition A.8 on the next page).*

### E

**Definition A.5** (Elitism). *In evolutionary computation, elitism refers to a strategy wherein the top-performing individuals from a population are given a higher probability of being selected for the next generation. While this doesn't guarantee their direct transfer, it increases the likelihood of retaining the best solutions in subsequent generations, ensuring that the quality of the population does not degrade.*

**Definition A.6** (Ephemeral Constant). *A constant that is randomly generated at the start of the program and remains constant throughout the execution of the program.*

*This is used to represent constant values in the program.*

**Definition A.7** (Evolutionary computation). *Family of algorithms for global optimization inspired by the process of natural selection.*

*This typically involves processes mimicking natural selection, mutation, recombination, and survival of the fittest.*

*The solutions to a problem are encoded as a set of “individuals” in a “population”.*

*Over multiple generations, these individuals are selected and modified (via genetic operators like crossover and mutation) in order to find better solutions.*

## G

**Definition A.8** (Gene). *Representation of a single component of a candidate solution to a given optimization problem. Formally, for a multi-dimensional function  $f$ , a gene is an element  $g$  in the domain of  $f$ .*

**Definition A.9** (Generation). *Number of iterations performed by an evolutionary algorithm.*

**Definition A.10** (Genetic Diversity). *In evolutionary computation, diversity refers to the degree of variation or difference in the genetic representation of individuals within a population. Maintaining diversity is essential for several reasons:*

1. **Exploration vs. Exploitation:** A diverse population can explore various regions of the solution space, ensuring that the algorithm doesn't focus solely on one area. This balance between exploration (searching new areas) and exploitation (optimizing known good areas) is fundamental in evolutionary computation.
2. **Preventing Premature Convergence:** Without adequate diversity, the population may converge too quickly to a suboptimal solution, known as a local optimum. By maintaining diversity, the population has a better chance of discovering the global optimum.
3. **Adaptability:** A diverse population can better adapt to changing environments or requirements, making it more resilient against dynamic problems.

Diversity can be measured in several ways, including genetic distance metrics, fitness-based measures, or phenotypic variance. To maintain or introduce diversity, techniques like mutation, crossover variations, immigration, and diversity preservation strategies (like fitness sharing or crowding) are employed.

**Definition A.11** (Genotype). *Representation of the full genetic information of a candidate solution to a given optimization problem.*

*Formally, a genotype is a matrix  $\mathbf{G} = (\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n)$ , where  $\mathbf{c}_i$  is a chromosome (see definition A.4 on the preceding page).*

## I

**Definition A.12** (Individual). *A candidate solution to a given optimization problem.*

*Formally, an individual is a pair  $(\mathbf{G}, \mathbf{f})$ , where  $\mathbf{G}$  is the genotype (see definition A.11) and  $\mathbf{f}$  is the fitness value of the individual.*

## K

**Definition A.13** (Kronecker delta function). *Kronecker delta function is a function of two variables, usually integers, which is 1 if they are equal, and 0 otherwise. Formally:*

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad (\text{A.1})$$

**M**

**Definition A.14** (Metaheuristics). *Problem-independent algorithmic method that yields a sufficiently good solution within reasonable time for an optimization problem, especially for complex problems where an exact solution is not crucial.*

**Definition A.15** (Multimodal function). *A multimodal function is a function that has multiple local maxima and/or multiple local minima within a given domain or interval. In other words, it has several peaks and troughs. The term "multi" in "multimodal" signifies the presence of multiple modes or peaks in the function.*

*Mathematically, if  $f$  is a multimodal function in an interval  $[a, b]$ , then there exist multiple points  $c_1, c_2, \dots, c_n$  in  $[a, b]$  such that  $f$  has local maxima and/or minima at these points.*

*Multimodal functions can be more challenging for optimization algorithms because the presence of multiple local optima can trap optimization techniques in suboptimal solutions, preventing them from finding the global optimum. This complexity often necessitates more sophisticated or stochastic optimization techniques, like simulated annealing or genetic algorithms, to effectively explore the search space.*

**Definition A.16** (Mutator). *See definition 2.7 on page 12.*

**P**

**Definition A.17** (Parameter optimization). *Optimization problem where the solution is a set of parameters that optimize a given function.*

**Definition A.18** (Phenotype). *Same as definition A.12 on the preceding page.*

**Definition A.19** (Population). *Set of candidate solutions to a given optimization problem.*

**Definition A.20** (Program induction). *Inference of an algorithm or program featuring recursive calls or repetition control structures, starting from information that is known to be incomplete, called the evidence, such as positive and negative I/O examples or clausal constraints.*

**S**

**Definition A.21** (Sealed class). *A sealed class or interface is a type of class or interface that restricts its subclassing or implementation to a limited set of classes, usually defined within the same module or file. This ensures tighter control over inheritance, allowing developers to dictate where and how a class or interface can be extended or implemented.*

**Definition A.22** (Search space). *Set of all candidate solutions to a given optimization problem.*

**Definition A.23** (Selector). *See definition 2.4 on page 9.*

**U**

**Definition A.24** (Unimodal function). *A unimodal function is a function that, within a given domain or interval, has only one local maximum and one local minimum, which can also be the global maximum and minimum respectively. In other words, the function increases to a certain point and then decreases, or vice versa. This single peak or trough is the “uni” in “unimodal”.*

*Mathematically speaking, if  $f$  is a unimodal function in the interval  $[a, b]$ , then there exists some point  $c$  in  $[a, b]$  such that:*

1.  $f(x)$  is increasing on  $[a, c]$  and decreasing on  $[c, b]$ , or
2.  $f(x)$  is decreasing on  $[a, c]$  and increasing on  $[c, b]$ .

**V**

**Definition A.25** (Variadic function). *Function that accepts a variable number of arguments.*

## Appendix B

# Test Functions for Optimization

This appendix contains the test functions used in the numerical experiments in chapter 5 on page 71.

### B.1 Ackley Function

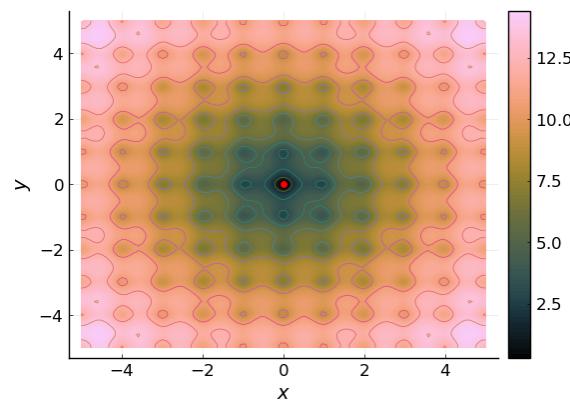
In the field of optimization, especially within evolutionary algorithms and swarm intelligence, the **Ackley function** serves as a prevalent benchmark function. Named after *David H. Ackley*, who introduced it during his research,<sup>1</sup> this function is particularly challenging for optimization algorithms due to its property of possessing a large number of local minima, despite having a single global minimum. This trait can often cause such algorithms to become entrapped in local minima.

**Definition B.1** (Ackley Function). *The Ackley Function, denoted as  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , is mathematically expressed as:*

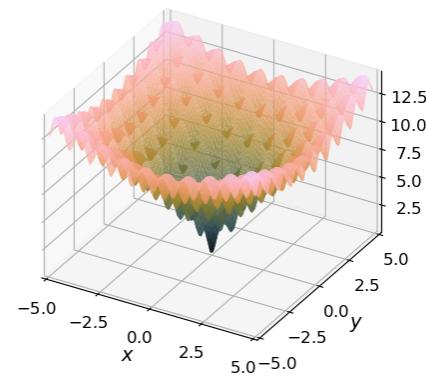
$$f(x, y) = -20 e^{-0.2\sqrt{0.5(x^2+y^2)}} - e^{0.5 [\cos(2\pi x)+\cos(2\pi y)]} + e + 20 \quad (\text{B.1})$$

*This function is evaluated within the range  $x, y \in [-5, 5]$ .*

The global minimum of the Ackley function is located at  $f(0, 0) = 0$ . Visualizations of the Ackley function are depicted as a contour plot and a surface plot in fig. B.1.



(a) Contour plot of the Ackley Function



(b) Surface plot of the Ackley Function

Figure B.1: Illustrations of the Ackley Function with the global minimum indicated by a red dot.

<sup>1</sup>Ackley, D. H. (1987) "A connectionist machine for genetic hillclimbing", Kluwer Academic Publishers, Boston MA.

## B.2 Beale Function

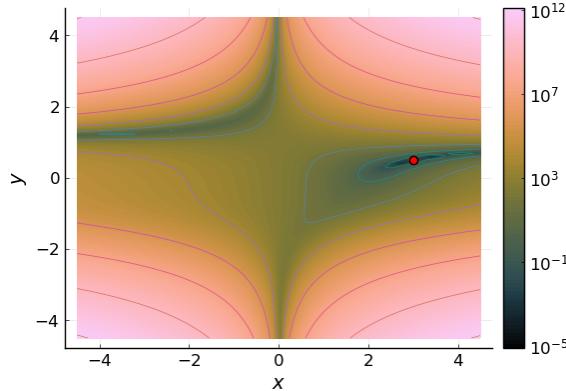
Unveiled by Beale in 1958<sup>2</sup>, the Beale function is recognized for its multimodal characteristics and sharp peaks that define the domain's corners.

**Definition B.2** (Beale Function). *The **Beale Function**, symbolized as  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , is expressed by the equation:*

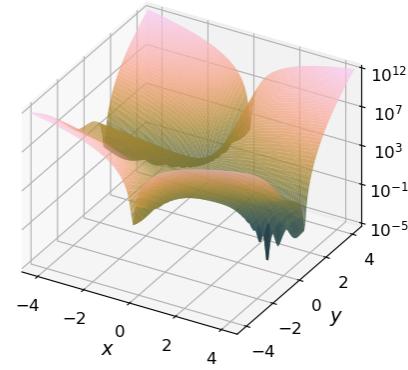
$$f(x, y) = (1.5 - x + xy)^2 + (2.25 - x + xy^2)^2 + (2.625 - x + xy^3)^2 \quad (\text{B.2})$$

*It's evaluated within the domain  $x, y \in [-4.5, 4.5]$ .*

The global minimum of the Beale function is found at  $f(3, 0.5) = 0$ . Both contour and surface plots of the Beale function are depicted in fig. B.2.



(a) Contour plot of the Beale Function



(b) Surface plot of the Beale Function

Figure B.2: Visual representation of the Beale Function

## B.3 Booth Function

The Booth function is a quadratic test problem used in the optimization field, specifically tailored for algorithms that handle two-dimensional search spaces.

**Definition B.3** (Booth Function). *The **Booth function**,  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , is defined as:*

$$f(x, y) = (x + 2y - 7)^2 + (2x + y - 5)^2 \quad (\text{B.3})$$

*where:*

- $x, y \in \mathbb{R}$  represent the decision variables.

The global minimum of the Booth function is  $f(1, 3) = 0$ . The function is usually constrained to the square  $[-10, 10]^2$ . A contour plot and a surface plot of the Booth function are illustrated in fig. B.3 on the facing page.

## B.4 Bukin Function N.6

The **Bukin function N.6**, a two-dimensional benchmark problem, is renowned for its inherent complexity and frequent utilization in assessing optimization algorithms. Notable for a sharply defined, deep valley, the Bukin function N.6 presents distinct hurdles for optimization techniques owing to its abrupt discontinuity and non-differentiability at  $x = 0$ .

<sup>2</sup>Beale, E. M. (1958). "On an Iterative Method for Finding a Local Minimum of a Function of More than One Variable".

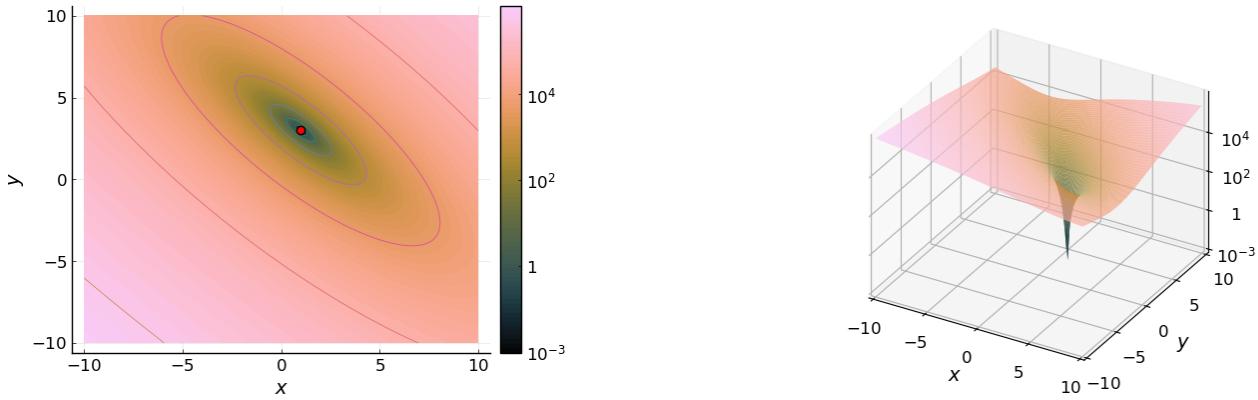


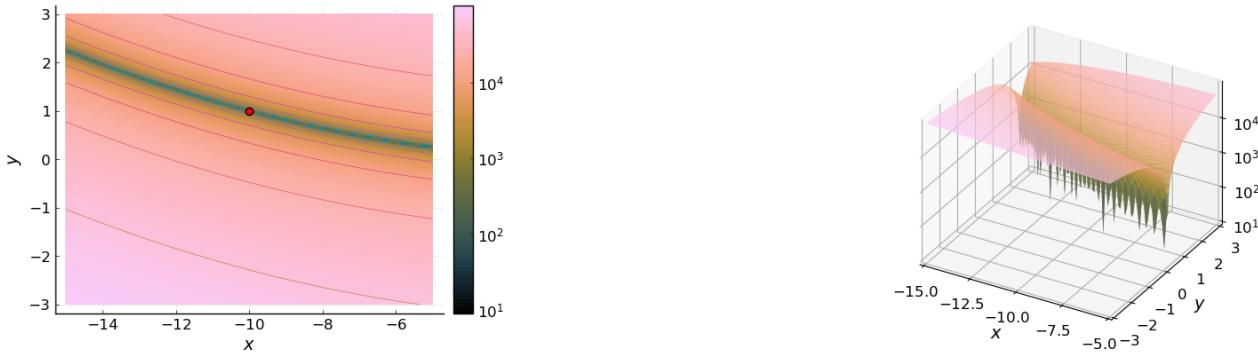
Figure B.3: Booth Function

**Definition B.4** (Bukin Function N.6). *The **Bukin function N.6**, given by  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , is mathematically represented as:*

$$f(x, y) = 100\sqrt{|y - 0.01x^2|} + 0.01|x + 10| \quad (\text{B.4})$$

*The decision variables,  $x, y \in \mathbb{R}$ , typically have the prescribed domains:  $-15 \leq x \leq -5$  and  $-3 \leq y \leq 3$  respectively.*

The function attains its global minimum at  $(x, y) = (-10, 1)$ , yielding a value of zero. The Bukin function N.6, due to its unique characteristics, offers a striking visualization. A profound ridge, extending diagonally across the domain, forms a distinguishing feature. The contour and surface plots illustrating the Bukin function N.6 are presented in fig. B.4.



(a) Contour plot of Bukin function N.6

(b) Surface plot of Bukin function N.6

Figure B.4: Contour and Surface Plots of Bukin Function N.6

## B.5 The Cross-in-Tray Function

The Cross-in-Tray function, known for its utility in testing optimization algorithms, poses a challenge due to its numerous local minima and four identical global minima, thereby making it an effective benchmark for evaluating an algorithm's capability to escape local optima and locate global optima.

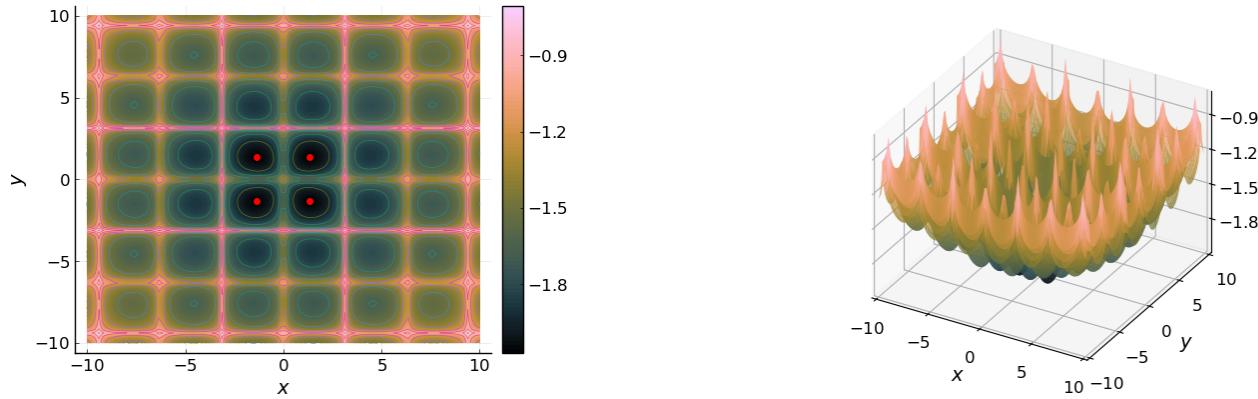
**Definition B.5** (Cross-in-Tray Function). *The Cross-in-Tray Function, denoted as a mapping from  $\mathbb{R}^2 \rightarrow \mathbb{R}$ , is formally defined as:*

$$f(x, y) = -0.0001 \left[ \left| \sin(x) \sin(y) \exp \left( \left| 100 - \frac{\sqrt{x^2 + y^2}}{\pi} \right| \right) \right| + 1 \right]^{0.1} \quad (\text{B.5})$$

This function is usually evaluated over the domain  $-10 \leq x, y \leq 10$ .

This function exhibits four identical global minima located at  $(\pm 1.34941, \pm 1.34941)$ , with each of these points having the function value of  $f(x, y) = -2.06261$ .

In fig. B.5, contour and surface plots are presented to provide a visual understanding of the function's intricate topology.



(a) Contour plot showcasing the complex landscape of the Cross-in-Tray function. The red dot denotes the location of the global minimum.

(b) Surface plot providing a 3D representation of the Cross-in-Tray function, enhancing the visualization of its global and local minima.

Figure B.5: Contour and surface plots illustrating the topological complexity of the Cross-in-Tray function

## B.6 Easom Function

The **Easom function** is a well-known unimodal benchmark function employed in the evaluation of optimization algorithms. It gains its name from Charles Easom and is distinctively recognized by its 'needle'-like global minimum. This feature presents a demanding task for optimization algorithms due to its confined optimal search space.

**Definition B.6** (Easom Function). *The Easom function, defined for  $f : [-100, 100]^2 \rightarrow \mathbb{R}$ , is formally expressed as:*

$$f(x, y) = -\cos(x) \cos(y) \exp(-((x - \pi)^2 + (y - \pi)^2)) \quad (\text{B.6})$$

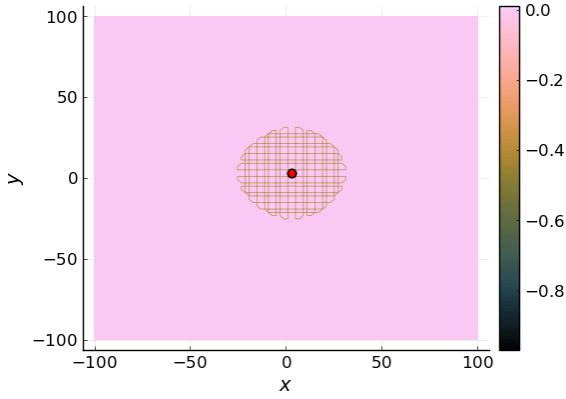
Here,  $x$  and  $y$  constitute the decision variables.

The global minimum of the Easom function is situated at the coordinates  $f(\pi, \pi) = -1$ .

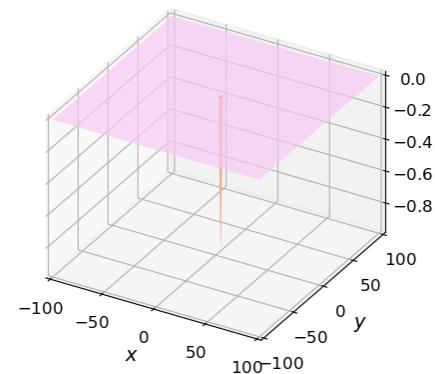
The intricate structure and narrow optimal space of the Easom function are clearly revealed in its contour and surface plots.

## B.7 Eggholder Function

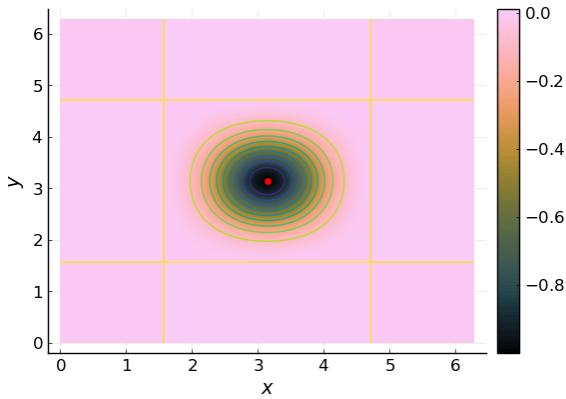
The **Eggholder function** is a widely-used non-convex function in the field of optimization, particularly for benchmarking optimization algorithms. The function is notorious for its multitude of local minima, presenting a complex search space that challenges the robustness and capability of an optimization algorithm to locate the global minimum.



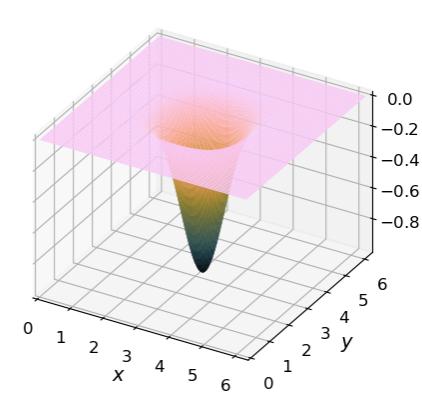
(a) Contour plot of the Easom function with the global minimum represented by the red dot



(b) Surface plot of the Easom function



(c) Contour plot of the Easom function in the vicinity of the global minimum,  $[0, 2\pi]$ .



(d) Surface plot of the Easom function in the vicinity of the global minimum,  $[0, 2\pi]$ .

Figure B.6: Contour and surface visualizations of the Easom function

**Definition B.7 (Eggholder Function).** Formally, the **Eggholder Function** is represented as a mapping  $\mathbb{R}^2 \rightarrow \mathbb{R}$ , and is mathematically defined as:

$$f(x, y) = -(y + 47) \sin \left( \sqrt{\left| \frac{x}{2} + (y + 47) \right|} \right) - x \sin \left( \sqrt{|x - (y + 47)|} \right) \quad (\text{B.7})$$

where the typical domain for evaluation spans  $-512 \leq x, y \leq 512$ .

The Eggholder function's global minimum resides at  $(512, 404.2319)$ , delivering a function value of  $f(x, y) = -959.6407$ .

Visual representations of the function can enhance understanding of its complexity. Figure fig. B.7 on the next page offers both contour and surface plots of the Eggholder function.

## B.8 Goldstein-Price Function

The **Goldstein-Price function**, believed to be proposed by individuals named Goldstein and Price, is a challenging multimodal function recognized for its landscape densely populated with local minima. This function serves as a standard benchmark in the field of optimization, testing the efficacy of various algorithms. The precise origins of this function, however, remain elusive in academic literature.

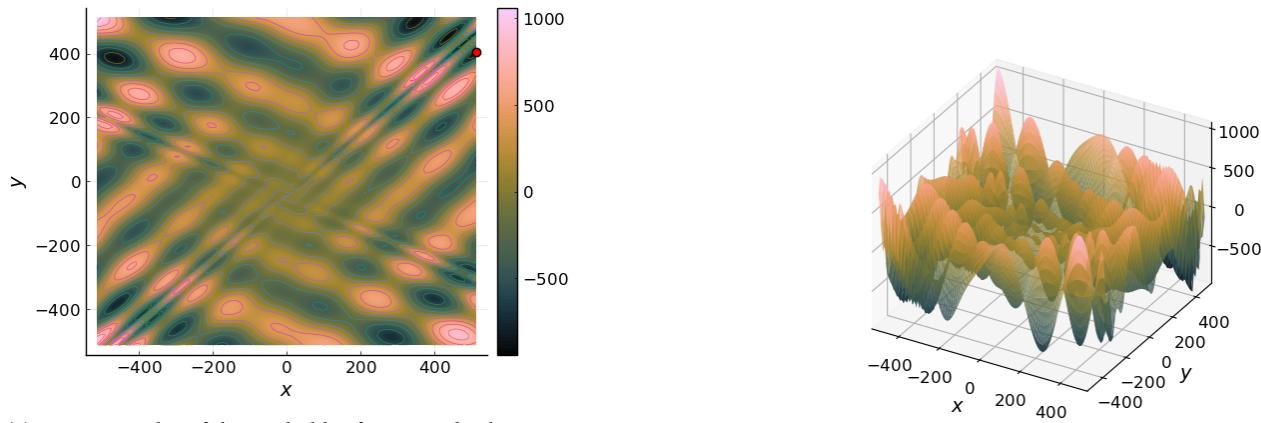


Figure B.7: Contour and surface visualizations of the Eggholder function, showcasing its intricate topology

**Definition B.8** (Goldstein-Price Function). *The Goldstein-Price function, denoted as  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , is formally articulated as follows:*

$$f(x, y) = [1 + (x + y + 1)^2 \cdot (19 - 14x + 3x^2 - 14y + 6xy + 3y^2)] \cdot [30 + (2x - 3y)^2 \cdot (18 - 32x + 12x^2 + 48y - 36xy + 27y^2)] \quad (\text{B.8})$$

The function's global minimum is located at  $f(x^*, y^*) = 3$  with  $(x^*, y^*) = (0, -1)$ . Evaluations of the Goldstein-Price function are typically performed within the range  $x, y \in [-2, 2]$ .

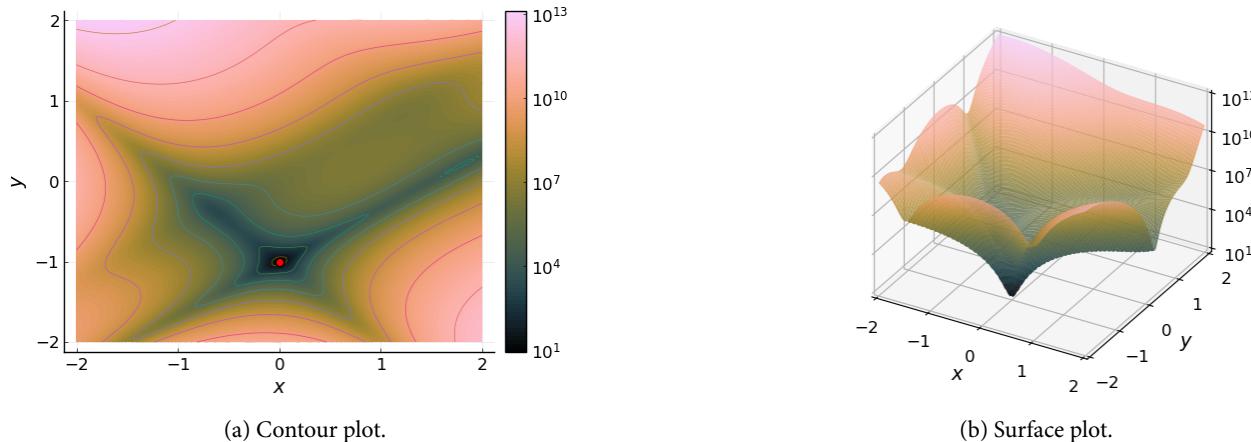


Figure B.8: Visual representations of the Goldstein-Price function

## B.9 Himmelblau's Function

The **Himmelblau's function**, attributed to David Mautner Himmelblau,<sup>3</sup> is a significant benchmark function in the realm of optimization, renowned for its complex multi-modal nature. Himmelblau, an American engineer, made substantial contributions to systems engineering and optimization theory.

<sup>3</sup>Himmelblau, D. (1972). "Applied Nonlinear Programming". McGraw-Hill. ISBN 0-07-028921-2.

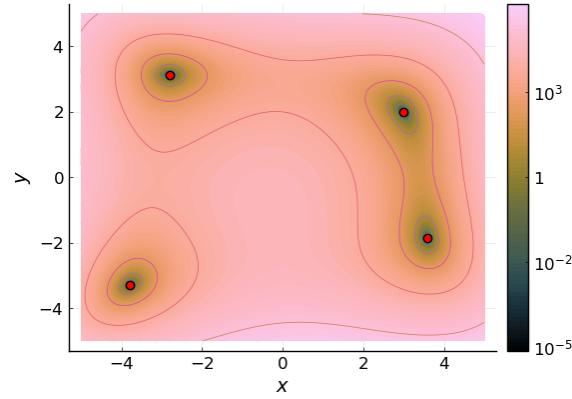
**Definition B.9** (Himmelblau's Function). *The Himmelblau's function, symbolized as  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , is formally described as:*

$$f(x, y) = (x^2 + y - 11)^2 + (x + y^2 - 7)^2 \quad (\text{B.9})$$

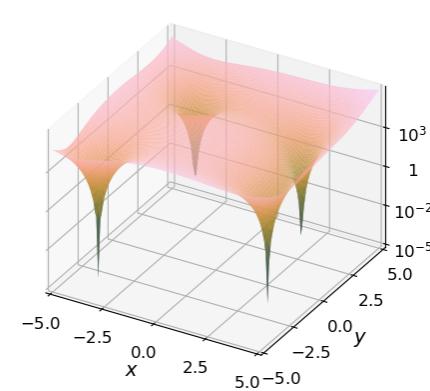
Here,  $x, y \in \mathbb{R}$  are the decision variables, with the domains  $\{x \mid -10 \leq x \leq 10\}$  and  $\{y \mid -10 \leq y \leq 10\}$ .

Distinguished by its four minima, located at  $(3, 2)$ ,  $(-2.805118, 3.131312)$ ,  $(-3.779310, -3.283186)$ , and  $(3.584428, -1.848126)$ , all roughly equating to zero, this function showcases its complexity.

Figure B.9 portrays the multi-modal landscape of Himmelblau's function through contour and surface plots, underlining its inherent intricacy, and as such, its utility in the evaluation of optimization techniques.



(a) Contour plot of Himmelblau's function, the minima are signified by the red dots



(b) Surface plot of Himmelblau's function

Figure B.9: The detailed multi-modal structure of Himmelblau's function illustrated through contour and surface plots

## B.10 Hölder Table Function

The **Hölder Table function** is a two-dimensional real-valued function employed in various branches of mathematical analysis. Due to its intriguing properties and complex topography, it serves as a valuable tool for function approximation and numerical analysis studies.

**Definition B.10** (Hölder Table function). *The Hölder Table function, designated as  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , is described by the equation*

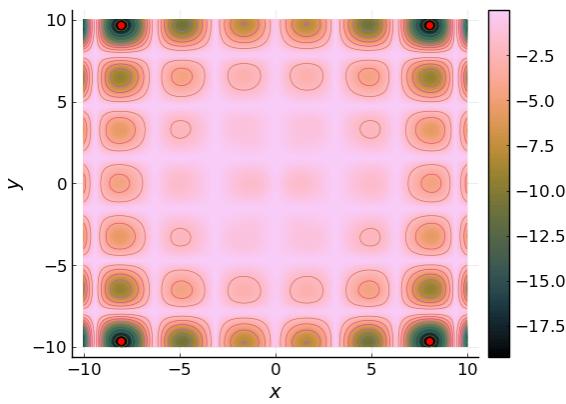
$$f(x, y) = - \left| \sin(x) \cos(y) \exp \left( \left| 1 - \frac{\sqrt{x^2 + y^2}}{\pi} \right| \right) \right| \quad (\text{B.10})$$

This function is defined for  $-10 \leq x, y \leq 10$ .

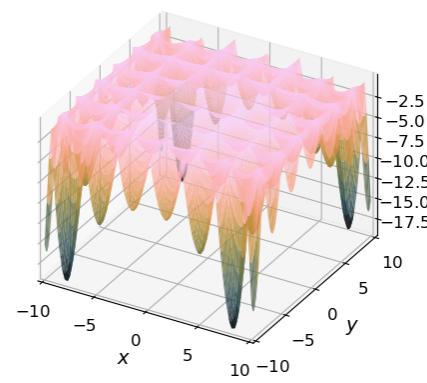
The Hölder Table function peaks globally at  $f(x^*, y^*) = 19.2085$  for  $x^* = \pm 8.05502$  and  $y^* = \pm 9.66459$ . Its intricately structured landscape can be vividly illustrated through contour and surface plots, as depicted in fig. B.10 on the following page.

## B.11 Lévi Function N.13

The **Lévi function N.13** is a noteworthy two-dimensional function frequently employed in the field of optimization algorithms for performance testing. Its complex, sinuous landscape, teeming with numerous local minima, presents a significant challenge to optimization procedures.



(a) Contour visualization of the Hölder Table function. The global minimum points are denoted by red dots.



(b) Three-dimensional surface representation of the Hölder Table function.

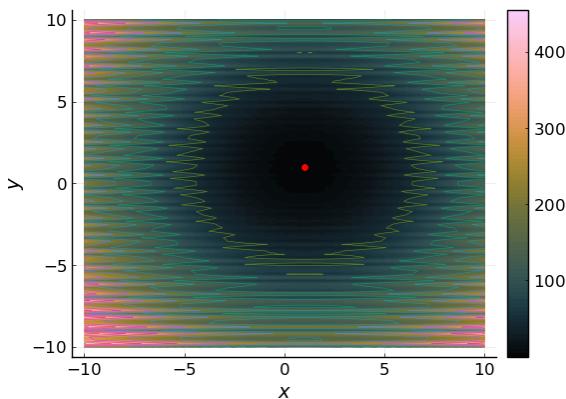
Figure B.10: The Complex Topography of the Hölder Table Function: Contour and Surface Representations

**Definition B.11** (Lévi Function N.13). *The Lévi function N.13, denoted as  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , is formally defined as follows:*

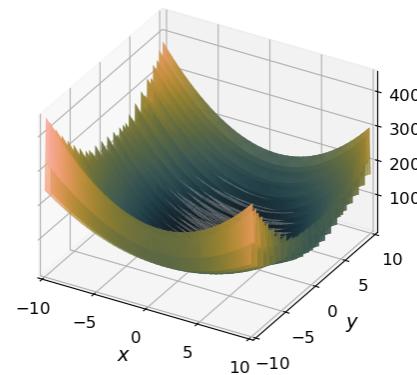
$$f(x, y) = \sin^2(3\pi x) + (x - 1)^2 \cdot (1 + \sin^2(3\pi y)) + (y - 1)^2 \cdot (1 + \sin^2(2\pi y)) \quad (\text{B.11})$$

where  $x, y \in \mathbb{R}$  are the decision variables.

The Lévi function N.13 finds its global minimum at  $f(1, 1) = 0$ . The complex topology of this function is visually captured in the contour and surface plots shown in Figure B.11.



(a) Contour plot of the Lévi function N.13



(b) Surface plot of the Lévi function N.13

Figure B.11: Contour and Surface Representations of the Lévi Function N.13

## B.12 Matyas Function

The **Matyas function**, known for its simplicity and convex nature, is a standard test problem in the field of optimization algorithms. Despite its apparent simplicity, it provides invaluable insights into an algorithm's performance and behavior.

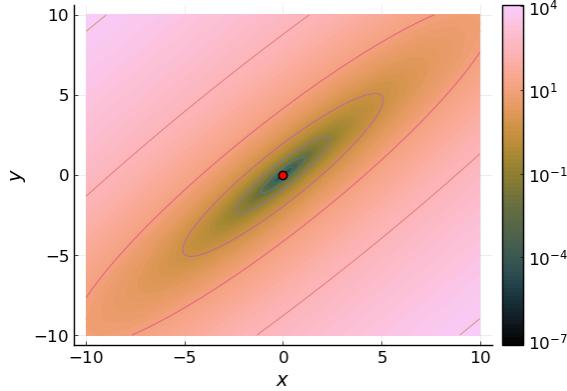
**Definition B.12** (Matyas Function). *The **Matyas function**, denoted as  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , is formulated as:*

$$f(x, y) = 0.26(x^2 + y^2) - 0.48xy \quad (\text{B.12})$$

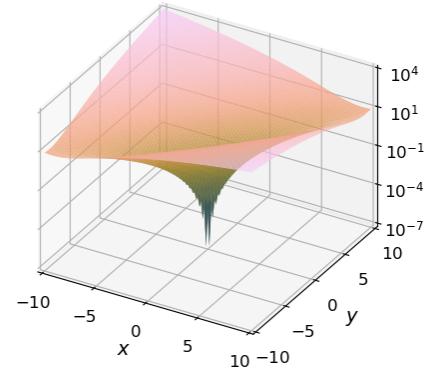
where:

- $x, y \in \mathbb{R}$  denote the decision variables.

The Matyas function reaches its global minimum at the origin, with  $f(0, 0) = 0$ . The contour and surface visualizations of the Matyas function, offering perspectives on its topographical attributes, are presented in fig. B.12.



(a) Contour plot of the Matyas function with a red dot indicating the global minimum



(b) Surface plot of the Matyas function

Figure B.12: Contour and Surface Visualizations of the Matyas Function

## B.13 McCormick Function

The field of optimization research frequently employs the **McCormick function** as an insightful benchmark function. This function owes its name to the researcher, Garth P. McCormick, who first employed it in his seminal study on factorable nonconvex programs.<sup>4</sup> The McCormick function is renowned for its oscillatory properties and its complex landscape featuring a single global minimum amidst numerous local minima. This complex landscape can pose a significant challenge to optimization algorithms, as they risk being ensnared in the local minima.

**Definition B.13** (McCormick Function). *The **McCormick Function**, symbolized as  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , is expressed mathematically as follows:*

$$f(x, y) = \sin(x + y) + (x - y)^2 - 1.5x + 2.5y + 1 \quad (\text{B.13})$$

*The McCormick function is typically evaluated within the range  $x, y \in [-1.5, 4]$  and  $[-3, 4]$  respectively.*

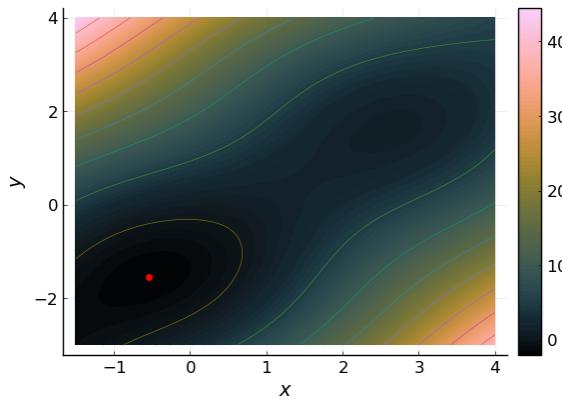
The global minimum of the McCormick function is found at  $f(-0.54719, -1.54719) \approx -1.9133$ . Visualizations of the McCormick function in the form of a contour plot and a surface plot are provided in fig. B.13 on the following page.

## B.14 Rastrigin Function

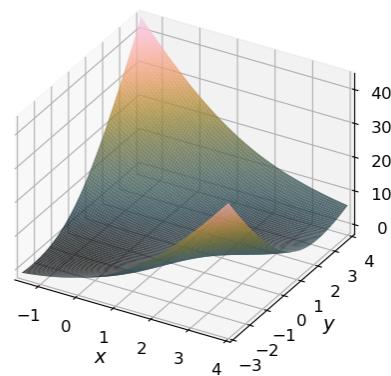
The Rastrigin function, first proposed by Rastrigin in 1974<sup>5</sup>, is a prominent non-convex function utilized as a benchmark for optimization algorithms. This function is a classic example of non-linear multimodal optimization problems, known for their

<sup>4</sup>McCormick, Garth P. "Computability of Global Solutions to Factorable Nonconvex Programs: Part I — Convex Underestimating Problems." Mathematical Programming 10, no. 1 (December 1976): 147–75. <https://doi.org/10.1007/BF01580665>.

<sup>5</sup>Rastrigin, L. A. (1974). "Systems of extremal control." Mir, Moscow.



(a) Contour plot of the McCormick Function. The location of the global minimum is designated by a red dot.



(b) Surface plot of the McCormick Function

Figure B.13: Visual representations of the McCormick Function.

complexity due to the abundance of local minima.

**Definition B.14** (Rastrigin Function). *The Rastrigin function, denoted as  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , is formulated as:*

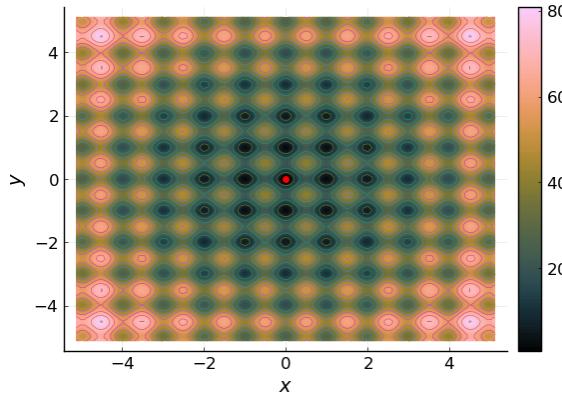
$$f(\mathbf{x}) = An + \sum_{i=1}^n [\mathbf{x}_i^2 - A \cos(2\pi\mathbf{x}_i)] \quad (\text{B.14})$$

Here,

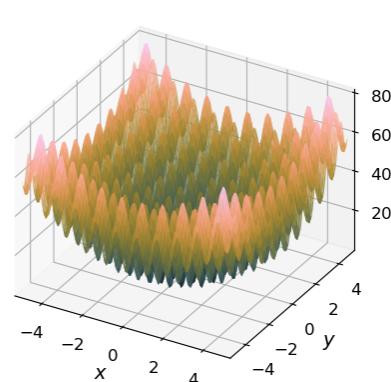
- $n$  is the number of dimensions.
- $\mathbf{x}$  is a vector composed of  $n$  real numbers.
- $A$  is a constant, typically set to 10.

The global minimum of the Rastrigin function is situated at  $\mathbf{x}^* = (0, \dots, 0)$ , with  $f(\mathbf{x}^*) = 0$ . It's particularly challenging due to its numerous local minima, distributed regularly throughout the search space. The function is predominantly evaluated within the hypercube  $\mathbf{x} \in [-5.12, 5.12]^n$ .

Visualizations of the Rastrigin function for  $n = 2$  are presented in fig. B.14, showcasing both the contour plot and the 3D surface plot.



(a) Contour plot of the Rastrigin Function.



(b) 3D surface plot of the Rastrigin Function.

Figure B.14: Visualizations of the Rastrigin Function for  $n = 2$ , with the global minimum marked by a red dot.

## B.15 The Rosenbrock Function

Introduced by Howard H. Rosenbrock in 1960<sup>6</sup>, the Rosenbrock function is a non-convex function that serves as a benchmark for optimization algorithms. Despite its simplicity, it poses a challenge due to its characteristic landscape, a long, narrow, parabolic-shaped flat valley. While finding the valley is relatively simple, converging to the global minimum within it is notably difficult.

**Definition B.15** (Rosenbrock function). *The Rosenbrock function, denoted as  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , is defined by the following equation:*

$$f(\mathbf{x}) = \sum_{i=1}^{n-1} [100(\mathbf{x}_{i+1} - \mathbf{x}_i^2)^2 + (1 - \mathbf{x}_i)^2] \quad (\text{B.15})$$

In this equation:

- $n$  is the number of dimensions.
- $\mathbf{x}$  is a vector composed of  $n$  real numbers.

The Rosenbrock function's global minimum is at  $f(\mathbf{x}^*) = 0$ , corresponding to the point  $\mathbf{x}^* = (1, \dots, 1)$ . To illustrate the function's behavior for  $n = 2$ , fig. B.15 presents a contour plot and a surface plot.

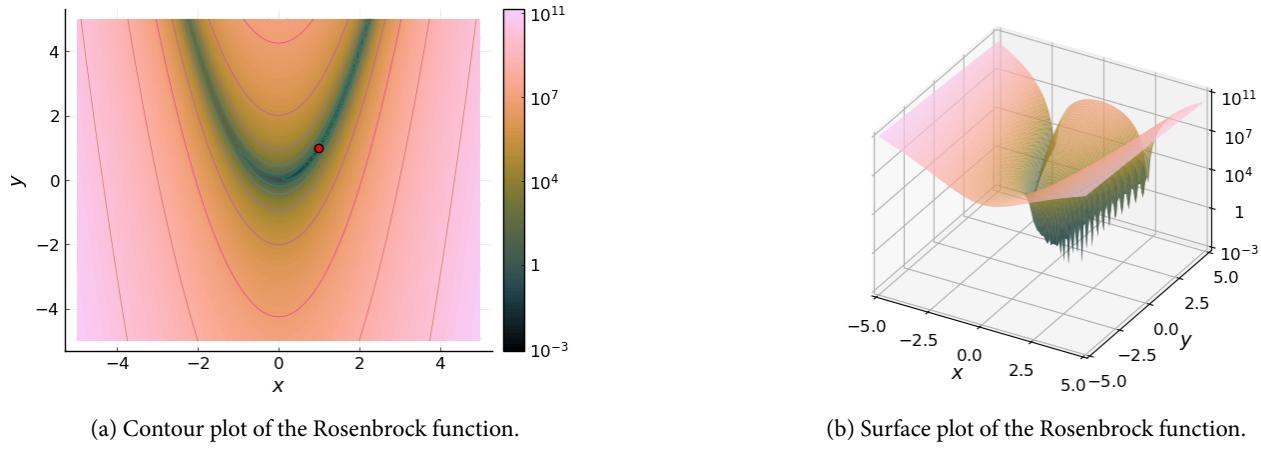


Figure B.15: Rosenbrock Function for  $n = 2$ , the red dot indicates the global minimum.

## B.16 Schaffer Function N.2

The **Schaffer Function N.2** is a well-established mathematical function frequently used in the realm of algorithm testing. It serves as a performance benchmark for a wide array of optimization algorithms, particularly those grounded in evolutionary computation and swarm intelligence principles.

**Definition B.16** (Schaffer Function N.2). *The Schaffer Function N.2 is defined over a two-dimensional domain, mapping  $\mathbb{R}^2$  to  $\mathbb{R}$ . The function, denoted as  $f(\mathbf{x})$ , is expressed as follows:*

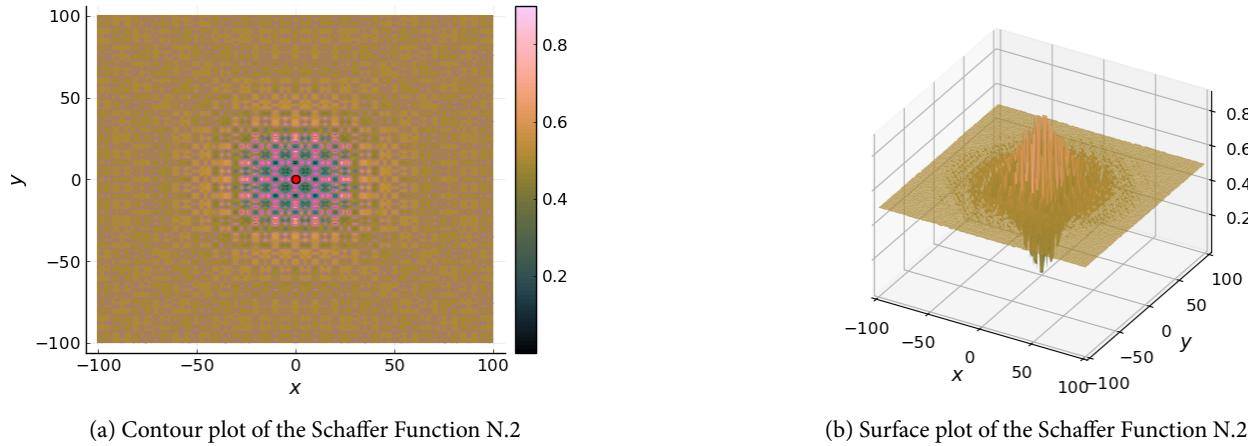
$$f(\mathbf{x}) = 0.5 + \frac{\sin^2(\sqrt{x_1^2 + x_2^2}) - 0.5}{(1.0 + 0.001 \cdot (x_1^2 + x_2^2))^2}$$

Here,  $\mathbf{x} = (x_1, x_2)$  represents a point in the two-dimensional domain, with  $x_1, x_2$  each ranging within the interval [-100, 100].

<sup>6</sup>Rosenbrock, H.H. (1960). "An automatic method for finding the greatest or least value of a function". The Computer Journal. 3 (3): 175-184. doi:10.1093/comjnl/3.3.175. ISSN 0010-4620.

The Schaffer Function N.2 achieves its global minimum at  $f(0, 0) = 0$ . A visual exploration of this function can be enhanced by both contour and surface plots, providing a clearer understanding of its global minimum and topological characteristics.

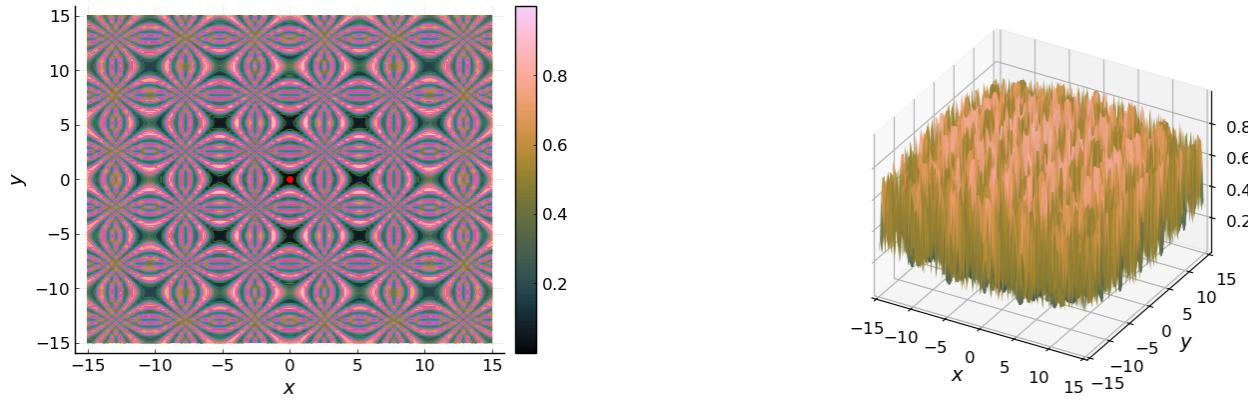
fig. B.17 illustrate the *Schaffer Function N.2*.



(a) Contour plot of the Schaffer Function N.2

(b) Surface plot of the Schaffer Function N.2

Figure B.16: a



(a) Close-up contour plot of the Schaffer Function N.2 in the vicinity of the global minimum.

(b) Close-up surface plot of the Schaffer Function N.2 in the vicinity of the global minimum.

Figure B.17: Visualization of the Schaffer Function N.2. The contour and surface plots illustrate the function's topology, with the global minimum denoted by a red dot. The close-up contour and surface plots provide a more precise view of the global minimum and the immediate surroundings.

## B.17 Schaffer Function N.4

The **Schaffer Function N.4** is a mathematical function used as a performance test problem for optimization algorithms. It is particularly suited for testing algorithms that need to optimize complex, non-linear problems with many local minima.

**Definition B.17** (Schaffer Function N.4). *Schaffer Function N.4 is defined as follows:  
For  $x \in \mathbb{R}^2$ , the function is given by:*

$$f(x) = 0.5 + \frac{\cos^2(\sin(|x^2 - y^2|)) - 0.5}{[1 + 0.001(x^2 + y^2)]^2}$$

Where:

- $x$  and  $y$  are decision variables.
- The function has four global minima of 0.292579, located at  $0, \pm 1.25313$  and  $\pm 1.25313, 0$ .

The characteristic feature of Schaffer N.4 is its landscape with several local minima and four global minima. This structure poses a significant challenge to optimization algorithms, especially those prone to being trapped in local minima.

Figure B.18 provides a visualization of the Schaffer Function N.4. The contour and surface plots show the function's topology, with the red dots indicating the locations of the global minima. The close-up contour and surface plots offer a detailed view of the global minima and their immediate surroundings, highlighting the intricacy of the function's landscape.

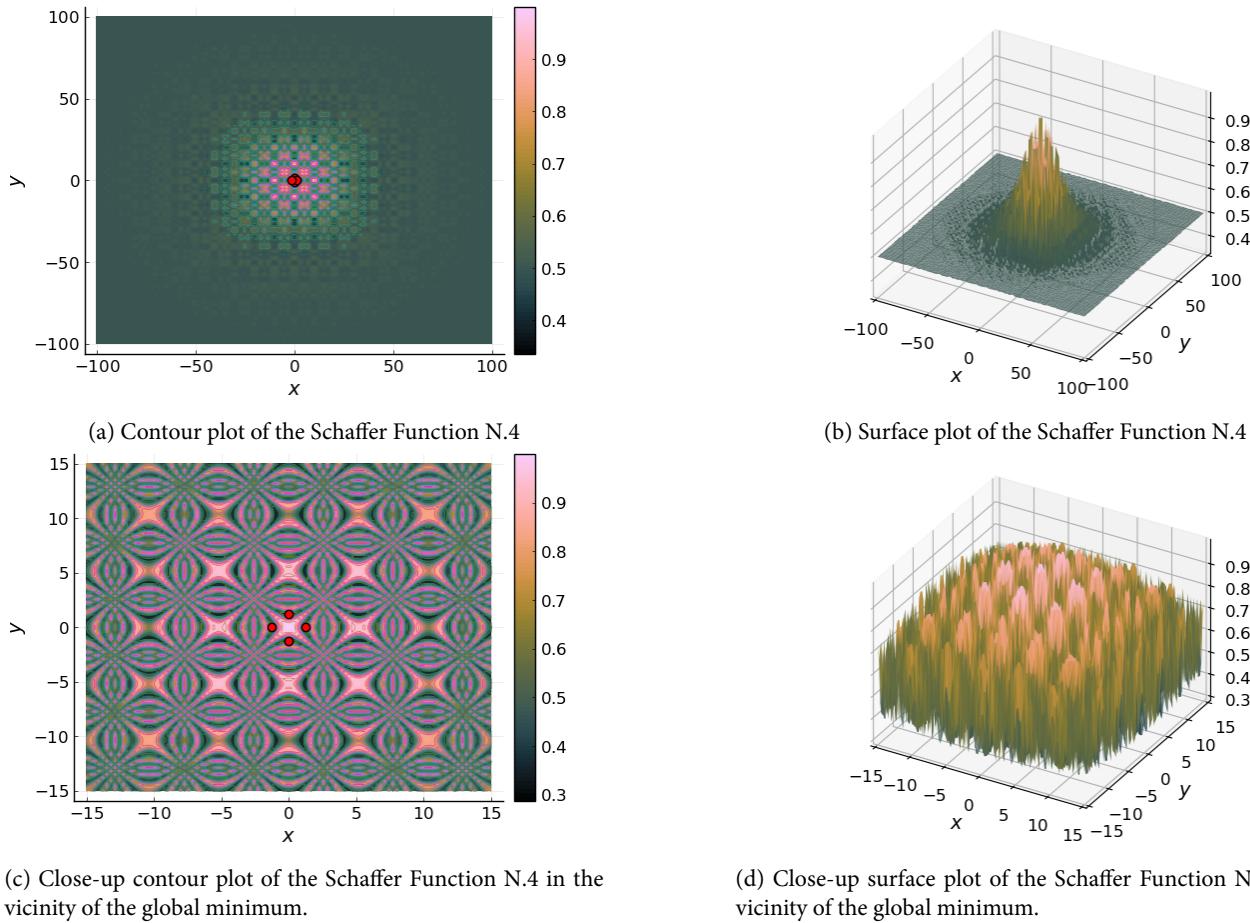


Figure B.18: Visualization of the Schaffer Function N.4. The contour and surface plots illustrate the function's topology, with the global minimum denoted by a red dot. The close-up contour and surface plots provide a more precise view of the global minimum and the immediate surroundings.

## B.18 The Sphere Function

The sphere function serves as a prominent benchmark problem in the realm of optimization algorithms. This convex function's simplicity and well-defined nature make it ideal for gauging the performance of such algorithms.

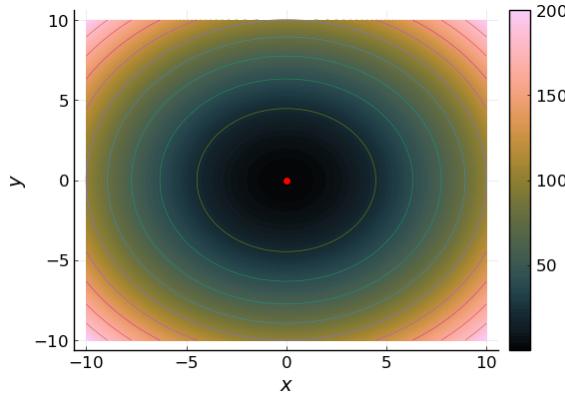
**Definition B.18** (Sphere Function). *The **sphere function**, denoted as  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , is formally described as:*

$$f(\mathbf{x}) = \sum_{i=1}^n \mathbf{x}_i^2 \quad (\text{B.16})$$

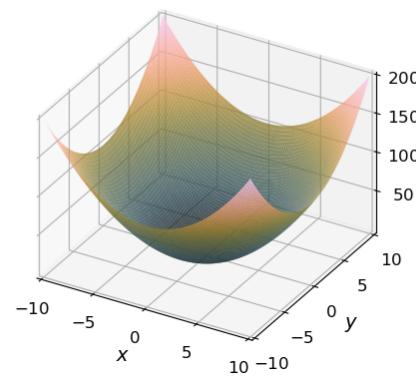
Herein:

- $n \in \mathbb{N}$  signifies the number of dimensions.
- $\mathbf{x}_i \in \mathbb{R}$  represents the  $i$ -th element of the vector  $\mathbf{x}$ .

The global minimum of the sphere function occurs at  $f(\mathbf{x}^*) = 0$  with the input vector  $\mathbf{x}^* = (0, \dots, 0)$ . The contour and surface plots showcasing the behavior of the sphere function for a two-dimensional input ( $n = 2$ ) are illustrated in fig. B.19.



(a) Contour plot of the Sphere Function.



(b) Surface plot of the Sphere Function.

Figure B.19: Visual Representation of the Sphere Function with  $n = 2$

## B.19 Styblinski-Tang function

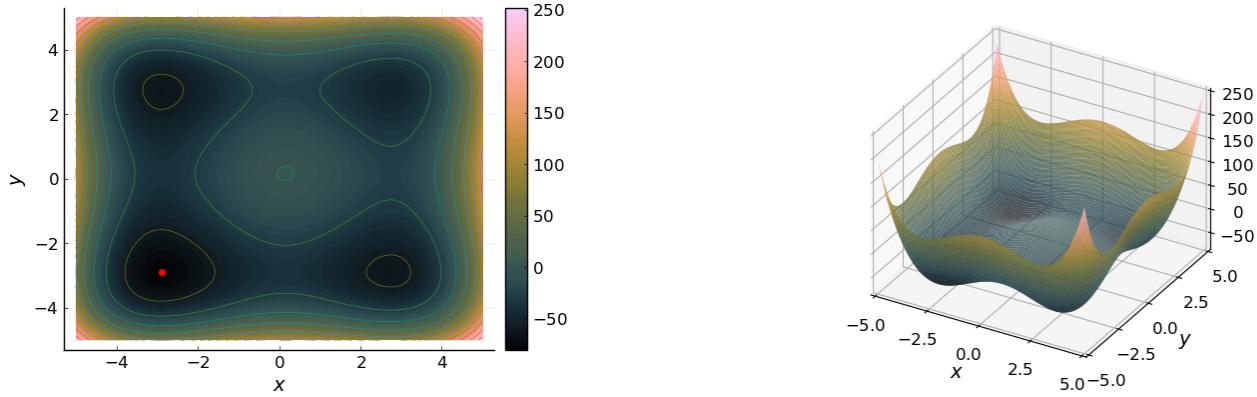
The **Styblinski-Tang function** is a non-convex function used as a performance test problem for optimization algorithms. This function is known for its complex and intricate landscape, presenting numerous local minima that pose challenges to optimization algorithms.

**Definition B.19** (Styblinski-Tang function). *The Styblinski-Tang function is defined in two dimensions, with the variables  $x$  and  $y$ .*

$$f(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^n (\mathbf{x}_i^4 - 16\mathbf{x}_i^2 + 5\mathbf{x}_i) \quad (\text{B.17})$$

where  $x$  and  $y$  are any real numbers.

The intriguing properties of this function have made it a common benchmark in the study of algorithm performance, particularly in the field of evolutionary computation and swarm intelligence.



(a) Contour plot of the Styblinski-Tang function. The intricate patterns show the complex landscape of the function. The red dot signifies the global minimum.

(b) Three-dimensional surface plot of the Styblinski-Tang function. The plot illustrates the numerous local minima and the complex topography of the function.

Figure B.20: Visualizations of the Styblinski-Tang function.

## B.20 Three-Hump Camel Function

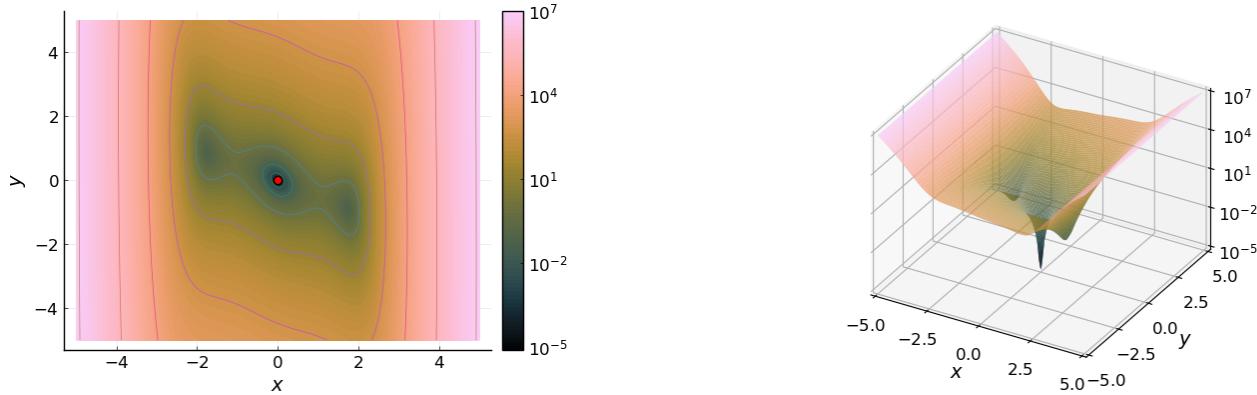
The **Three-Hump Camel function**, colloquially termed as the camel-back function, serves as a standard benchmark in the optimization algorithms testing landscape. This two-dimensional function earns its name from the characteristic tri-modal hump visual pattern it presents in a three-dimensional space, bearing resemblance to a camel's back.

**Definition B.20** (Three-Hump Camel Function). *The Three-Hump Camel function, depicted as  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , is formally expressed as:*

$$f(x, y) = 2x^2 - 1.05x^4 + \frac{x^6}{6} + xy + y^2 \quad (\text{B.18})$$

where  $-5 \leq x \leq 5$  and  $-5 \leq y \leq 5$ .

The Three-Hump Camel function exhibits its global minimum at the origin,  $f(0, 0) = 0$ . Although its form appears straightforward, the function's multiple local optima pose considerable challenges for optimization algorithms, making it an excellent test case.



(a) Contour plot of the Three-Hump Camel function. The red dot signifies the global minimum.

(b) Surface plot showcasing the characteristic tri-modal humps of the Three-Hump Camel function.

Figure B.21: Contour and surface visualizations of the Three-Hump Camel function



# Appendix C

## Additional Listings

This appendix contains additional listings of the source code used in this thesis that are not essential to the understanding of the thesis. The listings are included here for completeness.

Listing C.1– Calculation of  $|\mathbb{T}_{\leq 5}(\mathcal{T}, \mathcal{F})|$  for  $\mathcal{T} = \{x, c\}$  and the set  $\mathcal{F} = \{+, -, \times, /, \sin, \cos, \text{pow}\}$  using the *Julia* programming language.

```
arities = [2, 2, 2, 2, 1, 1, 2] # A({+, -, ×, ÷, sin, cos, pow}) = {2, 2, 2, 2, 1, 1, 2}
terminals_size = 8 # |T| = |{x, 1, 2, 3, 4, 5, 6, 7}| = 8
# |T_{≤h}|
t_leq(h::Int)::Int128 = if h == 0 # |T| if h = 0
    terminals_size
else # ( ∑_{h=0}^{h-1} ∑_{f ∈ F} |T_h|^{A(f)} ) + |T| if h > 0
    c_sum = terminals_size
    for i = 0:h - 1
        c_sum = c_sum + sum(t(i) .^ arities)
    end
    c_sum
end
# |T_h|
t(h::Int)::Int128 = if h == 0 # |T| if h = 0
    terminals_size
else # ∑_{f ∈ F} |T_{h-1}|^{A(f)} if h > 0
    sum(t(h - 1) .^ arities)
end
res = t_leq(5) # |T_{≤5}(T, F)|
```

Listing C.2– Minimal implementation of the *One Max* problem using the *Keen* framework.

```
fun main() {
    Core.random = Random(11)
    val result = engine(
        { gt: Genotype<Boolean, BoolGene> -> gt.flatten().count { it }.toDouble() },
        genotype { chromosome { booleans { size = 20; truesProbability = 0.5 } } })
    populationSize = 20
    selector = TournamentSelector(sampleSize = 3)
    alterers = listOf(
```

```

        BitFlipMutator(probability = 0.03),
        SinglePointCrossover(probability = 0.2)
    )
    limits = listOf(TargetFitness(20.0))
}.evolve()
println("Target fitness reached at generation ${result.generation}")
println("Best individual is ${result.best.genotype}")
println("with fitness ${result.best.fitness}")
}

```

Listing C.3– Room Scheduling problem using the *Keen* framework.

```

private data class Meeting(val start: Int, val end: Int)

private val meetings =
    listOf(
        Meeting(start = 1, end = 3),
        Meeting(start = 2, end = 3),
        Meeting(start = 5, end = 6),
        Meeting(start = 7, end = 9),
        Meeting(start = 4, end = 7),
        Meeting(start = 8, end = 10),
        Meeting(start = 2, end = 7),
        Meeting(start = 3, end = 4),
        Meeting(start = 1, end = 5),
        Meeting(start = 3, end = 6),
        Meeting(start = 4, end = 5)
    )

private fun fitnessFunction(genotype: Genotype<Int, IntGene>): Double {
    // We can access the genotype components by index as it is a matrix.
    val rooms = meetings.groupBy { genotype[meetings.indexOf(it)][0].value }
    val conflicts = rooms.values.sumOf { meetingList ->
        val table = IntArray(size = 10)
        meetingList.forEach { meeting ->
            // The ..< operator is equivalent to the range: [start, end)
            for (i in meeting.start..<meeting.end) {
                table[i]++
            }
        }
        table.count { it > 1 }
    }
    // Fitness is penalized by the number of conflicts.
    return rooms.size.toDouble() + conflicts
}

private const val POPULATION_SIZE = 100

fun main() {
    Domain.random = Random(420)
    val summary = EvolutionSummary<Int, IntGene>()
    val plotter = EvolutionPlotter<Int, IntGene>()
}

```

```

    val engine = evolutionEngine(
        ::fitnessFunction,
        genotype {
            repeat(meetings.size) {
                chromosomeOf {
                    ints {
                        size = 1
                        ranges += meetings.indices.first..meetings.indices.last
                    }
                }
            }
        })
    populationSize = POPULATION_SIZE
    ranker = FitnessMinRanker()
    alterers += listOf(RandomMutator(individualRate = 0.06), SinglePointCrossover(
        ↪ chromosomeRate = 0.2))
    limits += listOf(SteadyGenerations(generations = 20), GenerationLimit(generations =
        ↪ 100))
    listeners += plotter + summary // Add both listeners to the engine
}
engine.evolve()
summary.display()
val schedule = MutableList(meetings.size) { mutableListOf<Meeting>() }
meetings.forEachIndexed { index, meeting ->
    val room = summary.fittest.genotype[index][0].value
    schedule[room] += meeting
}
schedule.forEachIndexed { index, meetings ->
    println("Room $index: $meetings")
}
plotter.display()
}

```

The above code will produce the following output:

```

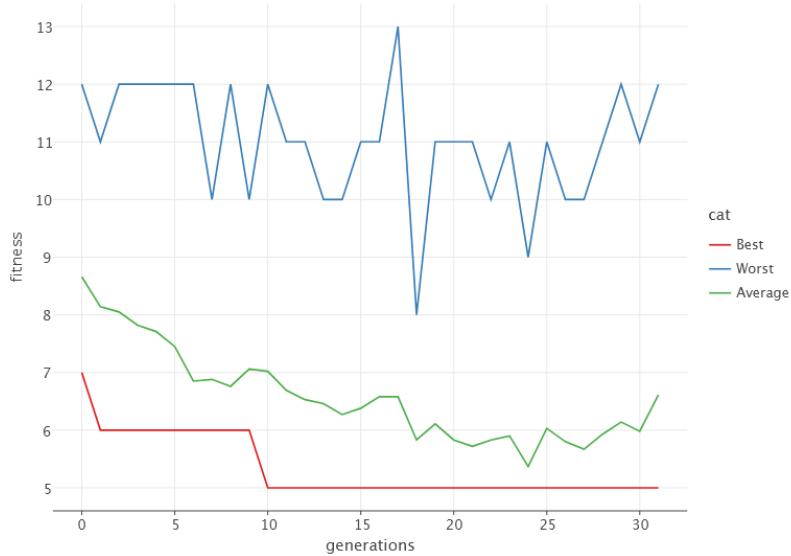
----- Evolution Summary -----
|--> Initialization time: 43 ms
----- Evaluation Times -----
|--> Average: 0.5 ms
|--> Max: 10 ms
|--> Min: 0 ms
----- Selection Times -----
|   |--> Offspring Selection
|   |   |--> Average: 0.15625 ms
|   |   |--> Max: 5 ms
|   |   |--> Min: 0 ms
|   |--> Survivor Selection
|   |   |--> Average: 0.0 ms
|   |   |--> Max: 0 ms
|   |   |--> Min: 0 ms
----- Alteration Times -----
|--> Average: 1.4375 ms
|--> Max: 17 ms
|--> Min: 0 ms

```

```

----- Evolution Results -----
|--> Total time: 222 ms
|--> Average generation time: 5.53125 ms
|--> Max generation time: 99 ms
|--> Min generation time: 1 ms
|--> Generation: 32
|--> Steady generations: 21
|--> Fittest: Genotype(chromosomes=[[7], [5], [10], [7], [7], [5], [4], [7], [10], [0], [5]])
|--> Best fitness: 5.0
Room 0: [Meeting(start=3, end=6)]
Room 1: []
Room 2: []
Room 3: []
Room 4: [Meeting(start=2, end=7)]
Room 5: [Meeting(start=2, end=3), Meeting(start=8, end=10), Meeting(start=4, end=5)]
Room 6: []
Room 7: [Meeting(start=1, end=3), Meeting(start=7, end=9), Meeting(start=4, end=7),
→ Meeting(start=3, end=4)]
Room 8: []
Room 9: []
Room 10: [Meeting(start=5, end=6), Meeting(start=1, end=5)]

```

Figure C.1: Evolution plot for the *Room Scheduling* problem.

Listing C.4– Example of using the Fun class to implement the addition operation

```

// Using the Fun class
var add = Fun<Double>("+", 2) { it[0] + it[1] }
// Extending the Fun class
class Add : Fun<Double>("+", 2, { it[0] + it[1] })
var add = Add()
// As an anonymous object
var add = object : Fun<Double>("+", 2, { it[0] + it[1] }) {}

```

Listing C.5– Demonstration of the use of *Strait-Jakt*'s constraint DSL to validate preconditions inside a *Probability Selector*.

```
private fun selectByProbabilities(
    population: Population<T, G>,
    probabilities: List<Double>,
    count: Int,
): List<Individual<T, G>> {
    constraints {
        "The cumulative probabilities list must have the same size as the population" {
            probabilities must HaveSize(population.size)
        }
        "The last cumulative probability must be 1.0" {
            probabilities.last() must BeEqualTo(1.0)
        }
    }
    return List(count) {
        population[probabilities.indexOfFirst { Domain.random.nextDouble() <= it }]
    }
}
```



## Appendix D

# Keen Open Source License: BSD 2-Clause License

**Copyright 2023 Ignacio Slater M.**

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.



# Bibliography

- [1] A. M. Turing. "I.—COMPUTING MACHINERY AND INTELLIGENCE". In: *Mind* LIX.236 (Oct. 1, 1950), pp. 433–460. ISSN: 1460-2113, 0026-4423. doi: [10.1093/mind/LIX.236.433](https://doi.org/10.1093/mind/LIX.236.433). URL: <https://academic.oup.com/mind/article/LIX/236/433/986238> (visited on 06/10/2023).
- [2] H. H. Rosenbrock. "An Automatic Method for Finding the Greatest or Least Value of a Function". In: *The Computer Journal* 3.3 (Mar. 1, 1960), pp. 175–184. ISSN: 0010-4620, 1460-2067. doi: [10.1093/comjnl/3.3.175](https://doi.org/10.1093/comjnl/3.3.175). URL: <https://academic.oup.com/comjnl/article-lookup/doi/10.1093/comjnl/3.3.175> (visited on 07/04/2023).
- [3] Nils Aall Barricelli. "Numerical Testing of Evolution Theories: Part I Theoretical Introduction and Basic Tests". In: *Acta Biotheoretica* 16.1-2 (Mar. 1962), pp. 69–98. ISSN: 0001-5342, 1572-8358. doi: [10.1007/BF01556771](https://doi.org/10.1007/BF01556771). URL: <http://link.springer.com/10.1007/BF01556771> (visited on 06/10/2023).
- [4] Sartaj Sahni. "Approximate Algorithms for the 0/1 Knapsack Problem". In: *Journal of the ACM (JACM)* 22.1 (1975), pp. 115–124.
- [5] David E Goldberg and Robert Lingle. "Alleles, Loci, and the Traveling Salesman Problem". In: *Proceedings of an international conference on genetic algorithms and their applications* (1985), pp. 154–159.
- [6] John H. Holland. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. 1st MIT Press ed. Complex Adaptive Systems. Cambridge, Mass: MIT Press, 1992. 211 pp. ISBN: 978-0-262-08213-6 978-0-262-58111-0.
- [7] John R. Koza. *Genetic Programming. I: On the Programming of Computers by Means of Natural Selection*. Complex Adaptive Systems. Cambridge, Mass.: MIT Press, 1992. 819 pp. ISBN: 978-0-262-52791-0.
- [8] Peter J Angeline and Jordan B Pollack. "The Evolutionary Induction of Subroutines". In: *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*. Bloomington, Indiana, 1992, pp. 236–241.
- [9] John R. Koza. *Genetic Programming II: Automatic Discovery of Reusable Programs*. Complex Adaptive Systems. Cambridge, Mass: MIT Press, 1994. 746 pp. ISBN: 978-0-262-11189-8.
- [10] Peter J. Angeline. "Genetic Programming and Emergent Intelligence". In: *Advances in Genetic Programming*. Advances in Genetic Programming 1. Cambridge (Mass.) London: MIT press, 1994, pp. 75–97. ISBN: 978-0-262-11188-1.
- [11] J.P. Rosca and D.H. Ballard. "Learning by Adapting Representations in Genetic Programming". In: *Proceedings of the First IEEE Conference on Evolutionary Computation. IEEE World Congress on Computational Intelligence*. First IEEE Conference on Evolutionary Computation. IEEE World Congress on Computational Intelligence. Orlando, FL, USA: IEEE, 1994, pp. 407–412. ISBN: 978-0-7803-1899-1. doi: [10.1109/ICEC.1994.349916](https://doi.org/10.1109/ICEC.1994.349916). URL: <http://ieeexplore.ieee.org/document/349916/> (visited on 06/23/2023).
- [12] Tobias Bickle and Lothar Thiele. "A Mathematical Analysis of Tournament Selection." In: *ICGA* 95 (1995), pp. 9–15.
- [13] Shigeyoshi Tsutsui, Masayuki Yamamura, and T. Higuchi. "Multi-Parent Recombination with Simplex Crossover in Real-Coded Genetic Algorithms". In: *Proceedings of Genetic and Evolutionary Computation Conference* (Jan. 1, 1999).
- [14] R. Andonov, V. Poirriez, and S. Rajopadhye. "Unbounded Knapsack Problem: Dynamic Programming Revisited". In: *European Journal of Operational Research* 123.2 (June 2000), pp. 394–407. ISSN: 03772217. doi: [10.1016/S0377-2217\(99\)00265-9](https://doi.org/10.1016/S0377-2217(99)00265-9). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0377221799002659> (visited on 08/09/2023).
- [15] I. Kushchuk. "Genetic Programming and Evolutionary Generalization". In: *IEEE Transactions on Evolutionary Computation* 6.5 (Oct. 2002), pp. 431–442. ISSN: 1941-0026. doi: [10.1109/TEVC.2002.805038](https://doi.org/10.1109/TEVC.2002.805038).

- [16] W. B. Langdon and Riccardo Poli. *Foundations of Genetic Programming*. Berlin ; New York: Springer, 2002. 260 pp. ISBN: 978-3-540-42451-2.
- [17] Kalyanmoy Deb, Pawan Zope, and Abhishek Jain. "Distributed Computing of Pareto-Optimal Solutions with Evolutionary Algorithms". In: *Evolutionary Multi-Criterion Optimization*. Ed. by Carlos M. Fonseca et al. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2003, pp. 534–549. ISBN: 978-3-540-36970-7. doi: [10.1007/3-540-36970-8\\_38](https://doi.org/10.1007/3-540-36970-8_38).
- [18] Kenneth A. De Jong. *Evolutionary Computation: A Unified Approach*. 2006. ISBN: 978-0-262-25598-1. URL: <https://ieeexplore.ieee.org/servlet/opac?bknumber=6267245>.
- [19] *The MIT License*. Open Source Initiative. Oct. 31, 2006. URL: <https://opensource.org/license/mit/> (visited on 07/26/2023).
- [20] Dipankar Dasgupta et al. "A Comparison of Multiobjective Evolutionary Algorithms with Informed Initialization and Kuhn-Munkres Algorithm for the Sailor Assignment Problem". In: *Proceedings of the 10th Annual Conference Companion on Genetic and Evolutionary Computation*. GECCO '08. New York, NY, USA: Association for Computing Machinery, July 12, 2008, pp. 2129–2134. ISBN: 978-1-60558-131-6. doi: [10.1145/1388969.1389035](https://doi.org/10.1145/1388969.1389035). URL: <https://doi.org/10.1145/1388969.1389035> (visited on 08/13/2023).
- [21] Riccardo Poli et al. *A Field Guide to Genetic Programming*. [Morrisville, NC: Lulu Press], 2008. 233 pp. ISBN: 978-1-4092-0073-4.
- [22] Lucas de P. Veronese and Renato A. Krohling. "Differential Evolution Algorithm on the GPU with C-CUDA". In: *IEEE Congress on Evolutionary Computation*. IEEE Congress on Evolutionary Computation. July 2010, pp. 1–7. doi: [10.1109/CEC.2010.5586219](https://doi.org/10.1109/CEC.2010.5586219).
- [23] Travis Desell et al. "An Analysis of Massively Distributed Evolutionary Algorithms". In: *IEEE Congress on Evolutionary Computation*. IEEE Congress on Evolutionary Computation. July 2010, pp. 1–8. doi: [10.1109/CEC.2010.5586073](https://doi.org/10.1109/CEC.2010.5586073).
- [24] Thé Van Luong, Nouredine Melab, and El-Ghazali Talbi. "GPU-based Island Model for Evolutionary Algorithms". In: *Proceedings of the 12th Annual Conference on Genetic and Evolutionary Computation*. GECCO '10. New York, NY, USA: Association for Computing Machinery, July 7, 2010, pp. 1089–1096. ISBN: 978-1-4503-0072-8. doi: [10.1145/1830483.1830685](https://doi.org/10.1145/1830483.1830685). URL: <https://doi.org/10.1145/1830483.1830685> (visited on 08/08/2023).
- [25] Thé Van Luong, Nouredine Melab, and El-Ghazali Talbi. "Parallel Hybrid Evolutionary Algorithms on GPU". In: *IEEE Congress on Evolutionary Computation*. IEEE Congress on Evolutionary Computation. July 2010, pp. 1–8. doi: [10.1109/CEC.2010.5586403](https://doi.org/10.1109/CEC.2010.5586403).
- [26] Xinjie Yu and Mitsuo Gen. *Introduction to Evolutionary Algorithms*. Decision Engineering. London ; New York: Springer, 2010. 418 pp. ISBN: 978-1-84996-128-8 978-1-84996-129-5.
- [27] Saber M. Elsayed, Ruhul A. Sarker, and Daryl L. Essam. "GA with a New Multi-Parent Crossover for Solving IEEE-CEC2011 Competition Problems". In: *2011 IEEE Congress of Evolutionary Computation (CEC)*. 2011 IEEE Congress of Evolutionary Computation (CEC). June 2011, pp. 1034–1040. doi: [10.1109/CEC.2011.5949731](https://doi.org/10.1109/CEC.2011.5949731).
- [28] Otman Abdoun, Jaafar Abouchabaka, and Chakir Tajani. *Analyzing the Performance of Mutation Operators to Solve the Travelling Salesman Problem*. Mar. 14, 2012. doi: [10.48550/arXiv.1203.3099](https://doi.org/10.48550/arXiv.1203.3099). arXiv: [1203.3099 \[cs\]](https://arxiv.org/abs/1203.3099). URL: [http://arxiv.org/abs/1203.3099](https://arxiv.org/abs/1203.3099) (visited on 08/22/2023). preprint.
- [29] Daniel Shiffman. *The Nature of Code*. Version 1.0, generated December 6, 2012. s.l.: Selbstverl., 2012. 498 pp. ISBN: 978-0-9859308-0-6.
- [30] Mateusz Guzek et al. "Multi-Objective Evolutionary Algorithms for Energy-Aware Scheduling on Distributed Computing Systems". In: *Applied Soft Computing* 24 (Nov. 1, 2014), pp. 432–446. ISSN: 1568-4946. doi: [10.1016/j.asoc.2014.07.010](https://doi.org/10.1016/j.asoc.2014.07.010). URL: <https://www.sciencedirect.com/science/article/pii/S1568494614003408> (visited on 08/08/2023).
- [31] Paweł Liskowski et al. "Comparison of Semantic-aware Selection Methods in Genetic Programming". In: *Proceedings of the Companion Publication of the 2015 Annual Conference on Genetic and Evolutionary Computation*. GECCO '15: Genetic and Evolutionary Computation Conference. Madrid Spain: ACM, July 11, 2015, pp. 1301–1307. ISBN: 978-1-4503-3488-4. doi: [10.1145/2739482.2768505](https://doi.org/10.1145/2739482.2768505). URL: <https://dl.acm.org/doi/10.1145/2739482.2768505> (visited on 06/21/2023).

- [32] Clara Schmitt et al. "Half a Billion Simulations: Evolutionary Algorithms and Distributed Computing for Calibrating the Simpoplocal Geographical Model". In: *Environment and Planning B: Planning and Design* 42.2 (Apr. 1, 2015), pp. 300–315. ISSN: 0265-8135. DOI: [10.1068/b130064p](https://doi.org/10.1068/b130064p). URL: <https://doi.org/10.1068/b130064p> (visited on 08/08/2023).
- [33] Qi Chen, Mengjie Zhang, and Bing Xue. "Feature Selection to Improve Generalization of Genetic Programming for High-Dimensional Symbolic Regression". In: *IEEE Transactions on Evolutionary Computation* 21.5 (Oct. 2017), pp. 792–806. ISSN: 1941-0026. DOI: [10.1109/TEVC.2017.2683489](https://doi.org/10.1109/TEVC.2017.2683489).
- [34] Anas Arram and Masri Ayob. "A Novel Multi-Parent Order Crossover in Genetic Algorithm for Combinatorial Optimization Problems". In: *Computers & Industrial Engineering* 133 (July 1, 2019), pp. 267–274. ISSN: 0360-8352. DOI: [10.1016/j.cie.2019.05.012](https://doi.org/10.1016/j.cie.2019.05.012). URL: <https://www.sciencedirect.com/science/article/pii/S0360835219302773> (visited on 08/24/2023).
- [35] Sergii Yaroshko and Svitlana Yaroshko. "Multithreaded Evolutionary Computing". In: *2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON)*. 2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON). July 2019, pp. 1041–1045. DOI: [10.1109/UKRCON.2019.8879863](https://doi.org/10.1109/UKRCON.2019.8879863).
- [36] Alexandre Bergel. *Agile Artificial Intelligence in Pharo: Implementing Neural Networks, Genetic Algorithms, and Neuroevolution*. For Professionals by Professionals. New York: Apress, 2020. 386 pp. ISBN: 978-1-4842-5384-7 978-1-4842-5383-0.
- [37] Alexandre Bergel and Ignacio Slater Muñoz. "Beacon: Automated Test Generation for Stack-Trace Reproduction Using Genetic Algorithms". In: *2021 IEEE/ACM 14th International Workshop on Search-Based Software Testing (SBST)*. 2021 IEEE/ACM 14th International Workshop on Search-Based Software Testing (SBST). Madrid, Spain: IEEE, May 2021, pp. 1–7. ISBN: 978-1-66544-571-9. DOI: [10.1109/SBST52555.2021.00007](https://doi.org/10.1109/SBST52555.2021.00007). URL: <https://ieeexplore.ieee.org/document/9476241/> (visited on 06/07/2023).
- [38] Qi Chen, Bing Xue, and Mengjie Zhang. "Rademacher Complexity for Enhancing the Generalization of Genetic Programming for Symbolic Regression". In: *IEEE Transactions on Cybernetics* 52.4 (Apr. 2022), pp. 2382–2395. ISSN: 2168-2275. DOI: [10.1109/TCYB.2020.3004361](https://doi.org/10.1109/TCYB.2020.3004361).
- [39] Zhengyuan Wang, Hui Zhang, and Yali Li. "Fast Polynomial Time Approximate Solution for 0-1 Knapsack Problem". In: *Computational Intelligence and Neuroscience* 2022 (Oct. 22, 2022), e1266529. ISSN: 1687-5265. DOI: [10.1155/2022/1266529](https://doi.org/10.1155/2022/1266529). URL: <https://www.hindawi.com/journals/cin/2022/1266529/> (visited on 08/09/2023).
- [40] djakobovic. *ECF - Evolutionary Computation Framework*. Mar. 13, 2023. URL: <https://github.com/djakobovic/ECF> (visited on 08/09/2023).
- [41] Diego Giacomelli. *GeneticSharp*. June 7, 2023. URL: <https://github.com/giacomelli/GeneticSharp> (visited on 06/08/2023).
- [42] *Kotlin Telegram Bot*. Kotlin Telegram Bot, Aug. 3, 2023. URL: <https://github.com/kotlin-telegram-bot/kotlin-telegram-bot> (visited on 08/03/2023).
- [43] *Kotlinx.Html*. Kotlin, Aug. 3, 2023. URL: <https://github.com/Kotlin/kotlinx.html> (visited on 08/03/2023).
- [44] *Pagmo*. European Space Agency, Aug. 8, 2023. URL: <https://github.com/esa/pagmo2> (visited on 08/09/2023).
- [45] Christian S. Perone. *Perone/Pyevolve*. May 23, 2023. URL: <https://github.com/perone/Pyevolve> (visited on 06/08/2023).
- [46] *Pygmo*. European Space Agency, June 4, 2023. URL: <https://github.com/esa/pygmo2> (visited on 06/08/2023).
- [47] *The Latest Programming Languages Statistics & Trends For 2023 • GITNUX*. Feb. 1, 2023. URL: <https://blog.gitnux.com/programming-languages-statistics/> (visited on 07/21/2023).
- [48] *Android's Kotlin-first Approach*. Android Developers. URL: <https://developer.android.com/kotlin/first> (visited on 07/21/2023).
- [49] *Calling Java from Kotlin | Kotlin*. Kotlin Help. URL: <https://kotlinlang.org/docs/java-interop.html> (visited on 07/21/2023).
- [50] *Calling Kotlin from Java | Kotlin*. Kotlin Help. URL: <https://kotlinlang.org/docs/java-to-kotlin-interop.html> (visited on 07/21/2023).
- [51] Fort Collins. "A Comparison of Genetic Sequencing Operators". In: (), p. 8.

- [52] **DEAP Documentation — DEAP 1.3.3 Documentation.** URL: <https://deap.readthedocs.io/en/master/> (visited on 06/08/2023).
- [53] **Easy\_ga/Src at Main · RubenRubioM/Easy\_ga.** GitHub. URL: [https://github.com/RubenRubioM/easy\\_ga](https://github.com/RubenRubioM/easy_ga) (visited on 06/13/2023).
- [54] **ECJ.** URL: <https://cs.gmu.edu/~eclab/projects/ecj/> (visited on 06/08/2023).
- [55] **EvolvingObjects: Welcome to Evolving Objects.** URL: <https://eodev.sourceforge.net/eo/doc/html/index.html> (visited on 06/08/2023).
- [56] **Genevo - Rust.** URL: <https://docs.rs/genevo/latest/genevo/index.html> (visited on 06/13/2023).
- [57] **GNU Lesser General Public License v3.0 - GNU Project - Free Software Foundation.** URL: <https://www.gnu.org/licenses/lgpl-3.0.en.html> (visited on 07/26/2023).
- [58] **Inspyred: Bio-inspired Algorithms in Python — Inspyred 1.0.1 Documentation.** URL: <https://aarongarrett.github.io/inspyred/> (visited on 06/08/2023).
- [59] **Kotest | Kotest.** URL: <https://kotest.io/> (visited on 08/03/2023).
- [60] **Kotlin Programming Language.** Kotlin. URL: <https://kotlinlang.org/> (visited on 07/21/2023).
- [61] **Kotlin-Java Interop Guide.** Android Developers. URL: <https://developer.android.com/kotlin/interop> (visited on 07/21/2023).
- [62] **One Max Problem — DEAP 1.3.3 Documentation.** URL: [https://deap.readthedocs.io/en/master/examples/ga\\_onemax.html](https://deap.readthedocs.io/en/master/examples/ga_onemax.html) (visited on 06/11/2023).
- [63] **Schlatterbeck/Pgapack: Parallel Genetic Algorithm Library Originally by David Levine from Argonne National Laboratory.** URL: <https://github.com/schlatterbeck/pgapack/tree/master> (visited on 08/09/2023).
- [64] **Schlatterbeck/Pgapy: Python Wrapper for Pgapack, the Parallel Genetic Algorithm Library.** URL: <https://github.com/schlatterbeck/pgapy> (visited on 08/09/2023).
- [65] Ignacio Slater Muñoz. **Keen | Kotlin Genetic Algorithm Framework.** URL: <https://github.com/r8vnhill/keen>.
- [66] **Stream (Java Platform SE 8 ).** URL: <https://docs.oracle.com/javase/8/docs/api/java/util/stream/Stream.html> (visited on 07/26/2023).
- [67] **The Top Programming Languages.** The State of the Octoverse. URL: <https://octoverse.github.com/2022/top-programming-languages> (visited on 07/21/2023).
- [68] **Type-Safe Builders | Kotlin.** Kotlin Help. URL: <https://kotlinlang.org/docs/type-safe-builders.html> (visited on 07/31/2023).
- [69] Franz Wilhelmstötter. **Jenetics: Java Genetic Algorithm Library.** jenetics.io. URL: <https://jenetics.io/> (visited on 06/08/2023).