

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA INFORMÁTICA

**BINARY GLOBAL-BEST HARMONY SEARCH ALGORITHM FOR
SOLVING SET-COVERING PROBLEM**

JUAN AGUSTÍN SALAS FERNÁNDEZ

THESIS TO APPLY FOR THE MASTER'S DEGREE
IN INFORMATIC ENGINEERING

JULY, 2016

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA INFORMÁTICA

BINARY GLOBAL-BEST HARMONY SEARCH ALGORITHM FOR SOLVING SET-COVERING PROBLEM

JUAN AGUSTÍN SALAS FERNÁNDEZ

Dr. Broderick Crawford Labrín
Master Thesis Advisor

PhD. Ricardo Soto de Giorgis
Master Thesis Co-Advisor

THESIS TO APPLY FOR THE MASTER'S DEGREE
IN INFORMATIC ENGINEERING

JULY, 2016

Abstract

In this thesis, we propose a Binary Global-Best Harmony Search (BGBHS) Algorithm to solve different instances of the Set Covering Problem (SCP). The SCP belongs to the \mathcal{NP} -hard class problems, and has become a classic combinatorial optimization problem due to its many practical applications. In this document we consider applying BGBHS and Modified BGBHS supported in adaptive adjustment of probability parameter p in the Bernoulli trials when the harmonies are created. The different results presented in this thesis show that our algorithm is a good alternative at a low cost to solve the SCP.

Keywords: Binary Global-Best Harmony Search, Metaheuristic, Set Covering Problem.

Table of Contents

1	Introduction	1
2	Main goal	1
2.1	Specific objectives	1
3	Theoretical Framework	2
3.1	Metaheuristics	2
3.1.1	Related terms	3
3.1.2	Framework for development	4
3.2	Set Covering Problem	5
3.2.1	SCP formulation	5
3.2.2	SCP sample solution	6
3.3	Traditional Harmony Search and some variants	8
3.3.1	Harmony Search Metaheuristic	8
3.3.2	Global-Best Harmony Search Metaheuristic	9
3.3.3	Binary Global-Best Harmony Search Metaheuristic	9
4	Algorithm Implementation	11
5	Proposed Improvements	12
6	Results	13
7	Conclusion	14
8	Appendix	15
8.1	Instance 4.1	37
8.2	Instance 4.2	38
8.3	Instance 4.3	39
8.4	Instance 4.4	40
8.5	Instance 4.5	41
8.6	Instance 4.6	42
8.7	Instance 4.7	43
8.8	Instance 4.8	44
8.9	Instance 4.9	45
8.10	Instance 4.10	46
8.11	Instance 5.1	47
8.12	Instance 5.2	48
8.13	Instance 5.3	49
8.14	Instance 5.4	50
8.15	Instance 5.5	51
8.16	Instance 5.6	52
8.17	Instance 5.7	53
8.18	Instance 5.8	54
8.19	Instance 5.9	55
8.20	Instance 5.10	56
8.21	Instance 6.1	57
8.22	Instance 6.2	58
8.23	Instance 6.3	59
8.24	Instance 6.4	60
8.25	Instance 6.5	61
8.26	Instance A.1	62
8.27	Instance A.2	63
8.28	Instance A.3	64

8.29 Instance A.4	65
8.30 Instance A.5	66
8.31 Instance B.1	67
8.32 Instance B.2	68
8.33 Instance B.3	69
8.34 Instance B.4	70
8.35 Instance B.5	71
8.36 Instance C.1	72
8.37 Instance C.2	73
8.38 Instance C.3	74
8.39 Instance C.4	75
8.40 Instance C.5	76
8.41 Instance D.1	77
8.42 Instance D.2	78
8.43 Instance D.3	79
8.44 Instance D.4	80
8.45 Instance D.5	81

List of Figures

1	Classification of metaheuristic	3
2	Trajectory-based method	3
3	Set Covering Problem example.	6
4	Set Covering Problem solution.	7
5	Guidelines for solving SCP.	7
6	Harmony Memory Matrix	9
7	Instance 4.1.	15
8	Instance 4.2.	15
9	Instance 4.3.	16
10	Instance 4.4.	16
11	Instance 4.5.	17
12	Instance 4.6.	17
13	Instance 4.7.	18
14	Instance 4.8.	18
15	Instance 4.9.	19
16	Instance 4.10.	19
17	Instance 5.1.	20
18	Instance 5.2.	20
19	Instance 5.3.	21
20	Instance 5.4.	21
21	Instance 5.5.	22
22	Instance 5.6.	22
23	Instance 5.7.	23
24	Instance 5.8.	23
25	Instance 5.9.	24
26	Instance 5.10.	24
27	Instance 6.1.	25
28	Instance 6.2.	25
29	Instance 6.3.	26
30	Instance 6.4.	26
31	Instance 6.5.	27
32	Instance A.1.	27
33	Instance A.2.	28
34	Instance A.3.	28
35	Instance A.4.	29
36	Instance A.5.	29
37	Instance B.1.	30
38	Instance B.2.	30
39	Instance B.3.	31
40	Instance B.4.	31
41	Instance B.5.	32
42	Instance C.1.	32
43	Instance C.2.	33
44	Instance C.3.	33
45	Instance C.4.	34
46	Instance C.5.	34
47	Instance D.1.	35
48	Instance D.2.	35
49	Instance D.3.	36

50 Instance D.4. 36

51 Instance D.5. 37

List of Tables

1	HS - Musical Notes	8
2	HS components - Musician	8
3	HS components - Pitch range	8
4	HS components - Solution	9
5	HS components - Aesthetics audience	9
6	Instance 4.1	37
7	Instance 4.2	38
8	Instance 4.3	39
9	Instance 4.4	40
10	Instance 4.5	41
11	Instance 4.6	42
12	Instance 4.7	43
13	Instance 4.8	44
14	Instance 4.9	45
15	Instance 4.10	46
16	Instance 5.1	47
17	Instance 5.2	48
18	Instance 5.3	49
19	Instance 5.4	50
20	Instance 5.5	51
21	Instance 5.6	52
22	Instance 5.7	53
23	Instance 5.8	54
24	Instance 5.9	55
25	Instance 5.10	56
26	Instance 6.1	57
27	Instance 6.2	58
28	Instance 6.3	59
29	Instance 6.4	60
30	Instance 6.5	61
31	Instance A.1	62
32	Instance A.2	63
33	Instance A.3	64
34	Instance A.3	65
35	Instance A.5	66
36	Instance B.1	67
37	Instance B.2	68
38	Instance B.3	69
39	Instance B.4	70
40	Instance B.5	71
41	Instance C.1	72
42	Instance C.2	73
43	Instance C.3	74
44	Instance C.4	75
45	Instance C.5	76
46	Instance D.1	77
47	Instance D.2	78
48	Instance D.3	79
49	Instance D.4	80

50 Instance D.5 81

1 Introduction

Every process is potentially optimizable. The vast majority of companies is involved in solving optimization problems every day [1]. Indeed, many challenging applications in science and industry can be modelled as optimization problems.

Metaheuristics are a branch of optimization in computer science that apply mathematics related to algorithms and computational complexity theory. Metaheuristics are raising a large interest in diverse technologies, industries, and services since they proved to be efficient algorithms in solving a wide range of complex real-life optimization problems in different domains.

The instantiation of a metaheuristic requires to choose among a set of different possible components and to assign specific values to all free parameters. Instantiations are defined, by the author, as a beginning configuration, and its relevance turns crucial on solving the SCP. [2].

The SCP is a well-known mathematical problem, which tries to cover a set of needs at the lowest possible cost. SCP is very important in practice, as it has been used to model a large range of problems like scheduling, manufacturing, service planning, information retrieval, among others.

In an attempt to solve the SCP, a variation of Harmony Search metaheuristic in binary domain was designed. This method has been tested using SCP instances from the OR-library [3] and this results have shown that the developed algorithm generates good solutions for each instance.

2 Main goal

To solve the SCP using the Binary Global-Best HS metaheuristic algorithm.

2.1 Specific objectives

- To solve the SCP with the BGBHS metaheuristic and to compare the results with benchmark of Beasley
- To introduce new operators that make the results closer to the global optimum.
- To search nearest optimal solutions to problems in the SCP.
- To compare and to analyze the results using different resolution strategies.

3 Theoretical Framework

Optimization problems are encountered in many domains: BigData, engineering, logistics, and business. An optimization problem may be defined by the couple (S, f) , where S represents the set of feasible solutions, and $f : S \Rightarrow \mathbb{R}$ the objective function to optimize. The objective function assigns to every solution $s \in S$ of the search space a real number indicating its worth. The objective function f allows to define a total order relation between any pair of solutions in the search space.

The global optimum is defined as a solution $s^* \in S$ and it has a better objective function than all solutions of the search space, i.e., $\forall s \in S, f(s^*) \leq f(s)$.

3.1 Metaheuristics

Definition

The *meta* and *heuristic* are Greek words, *meta* it means *higherlevel* and *heuristics* means *to find, to know or to discover*. Metaheuristics are a set of intelligent strategies to enhance the efficiency of heuristic procedures.

A metaheuristic will be successful on a given optimization problem if it can provide a balance between exploration and exploitation. Exploration means to generate diverse solutions, associating them to different points of the search space on the global scale, while exploitation means to search in a local region by exploiting the information that a current good solution has been found in this region. Exploration is often directed by use of randomization, which enables an algorithm to have the ability to jump between different bounded spaces and gives to the algorithm the capability to search optimal solutions globally.

Classification

The different metaheuristic approaches can be characterized by different aspects concerning the search path they follow or how memory is exploited [4]. The majority of these algorithms has a stochastic behavior and mimics biological or physical processes. This presented classification was proposed by Beheshti et al. [5], and can be understood better as shown in the (Figure 1).

Nature-inspired vs. non-nature inspiration This class is based on the origin of algorithm. The majority of metaheuristics are nature-inspired algorithms such as Black Hole (BH) [6], Particle Swarm Optimization (PSO) [7] and Genetic Algorithms (GA) [8]. Also, some of them are non-nature-inspired algorithms like Iterated Local Search (ILS) [9] and TabuSearch (TS) [10].

Population-based vs. single-point search There is a certain group of metaheuristics, that can be classified by the number of solutions in their lifecycle or execution, such as Trajectory methods, which are algorithms working based on a single solution at any time (Figure 2). On the other hand, Population-based algorithms perform searches with multiple initial points in a parallel style. Examples of these metaheuristics can be: Harmony Search (HS) [11], GA [12] and PSO.

Dynamic vs. static objective function Another way of classifying metaheuristics, is by the way of utilizing the objective function. Some algorithms maintain the objective function intact throughout the execution cycle, while others modify the objective function according to information collected at runtime.

An example of the second case presented, is Guided Local Search (GLS) [13]. The idea behind this approach is to escape from local optima by changing the search landscape.

One vs. various neighborhood structures The majority of metaheuristic algorithms apply one single neighborhood structure. The fitness landscape topology remain unaltered thru the course of the algorithm while others, like Variable Neighborhood Search (VNS) [14], employ a set of neighborhood structures. This latter structure gives the possibility to diversify the search by swapping between different fitness landscapes.

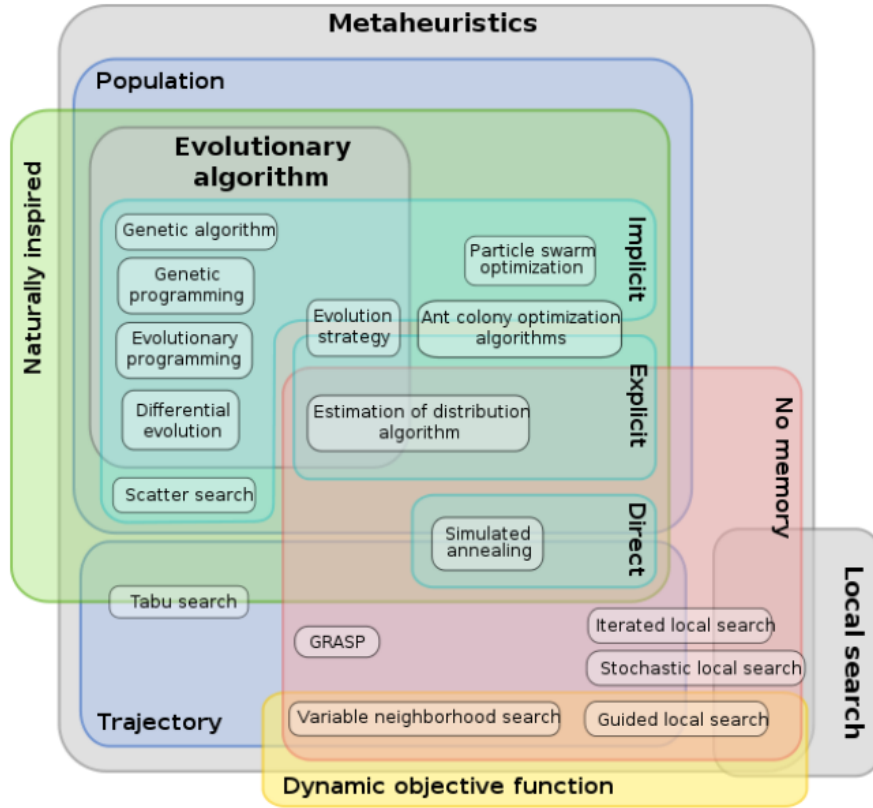


Fig. 1: Classification of metaheuristic

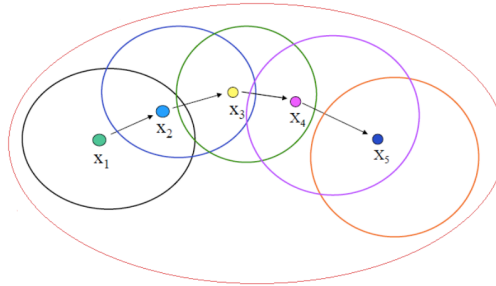


Fig. 2: Trajectory-based method

Memory usage vs. memoryless methods One of the most interesting variables to classify a metaheuristic is undoubtedly use of memory. Short term usually is different from long term memory. The first kind usually keeps track of recently performed moves, or taken decisions. The second is usually an accumulation of synthetic parameters about the search.

3.1.1 Related terms

Operator: Unitary procedure for transforming information or implement the behavior of the algorithm.

Solution: Array of n columns containing a solution for a given problem. In binary case, possible values are 0 and 1.

Constrain: Conditions to be met to find a viable solution.

Benchmark: Optimal set of known problem instances to validate the propose algorithm.

Objective Function: Implements the mathematical expression representing the problem to solve. The guiding objective function is related to the goal to achieve.

Fitness: Value resulting by applying the objective function to a found solution.

Matrix of Costs: Consist of n columns vector containing the cost associated to each problem variable.

Optimal Value: Solution with the best fitness.

Domain: Set of feasible values for the variable.

Matrix A: Matrix containing the restrictions for the given problem.

RPD: Relative Percentage Deviation.

Harmony Memory: Memory space which includes the population of the solution vectors.

Harmony Memory Size: Defines the amount of harmonies that can be stored in HM.

Harmony Memory Consideration Rate (HMCR): In memory consideration, the value of decision variable x'_1 is randomly selected from the historical values, other decision variables, $(x'_2, x'_3, \dots, x'_N)$ are sequentially selected in the same manner with probability where $\text{HMCR} \in (0,1)$.

Pitch Adjusting Rate (PAR): Each decision variable x'_i assigned a value by memory considerations is pitch adjusted with the probability of PAR where $\text{PAR} \in (0,1)$.

3.1.2 Framework for development In an effort to maintain consistency in the structure of the proposed meta-heuristic development, a well-defined model is followed, in order to structure the development steps of the proposed technique (Figure 5).

At first step, the problem is modeled and, based on its properties, application requirements are defined. In the particular case of this investigation the SCP is defined and validated against instances of benchmark. The number of iterations has been considered, but runtime has not.

The second step basically corresponds to design the metaheuristic. In the case of this research, problems will be treated with variants of Harmony Search. That is to say not designed from the ground up, but certain elements are incorporated to improve the algorithm standard behavior. These elements are selected once known the nature of the problem to be solved and the metaheuristic to be used. Elements like the objective function to use and classification (Population based) are set by default.

The third step is to adopt a strategy of development of the technique or metaheuristic. In this research, it has been chosen to develop metaheuristics from scratch, thereby achieving maximum control of it. It was used for this purpose Python 2.7 as programming language and PyCharm as IDE.

The fourth step is handling optimization parameters. There are two types of parameter tuning: Online and Off line. In this research, management of both types is performed to achieve good results. Specifically improving the metaheuristic proposal is based on the dynamic variation of the parameter that generates solutions. The following sections will go into detail on this subject.

The fifth and final step is to design experiments, get results and compare improvements. Then perform changes in parameters or operators so as to keep improving convergence, achieving an optimum balance between exploration and exploitation. This framework adopts an iterative approach because when step five is completed, systems can return to step one, two or three to seek a better solution.

3.2 Set Covering Problem

3.2.1 SCP formulation The SCP is a well-known mathematical problem, which tries to cover a set of needs at the lowest possible cost. The SCP was included in the list of 21 \mathcal{NP} -complete problems of Karp [15]. There are many practical uses for this problem, such as: crew scheduling [16, 17], location of emergency facilities [18, 19], production planning in industry [20, 21, 22], vehicle routing [23, 24], ship scheduling [25, 26], network attack or defense [27], assembly line balancing [28, 29], traffic assignment in satellite communication systems [30, 31], simplifying boolean expressions [32], the calculation of bounds in integer programs [33], information retrieval [34], political districting [35], crew scheduling problems in airlines [36], among others. The SCP can be formulated as follows:

$$\text{Minimize } Z = \sum_{j=1}^n c_j x_j \quad (1)$$

Subject to:

$$\sum_{j=1}^n a_{ij} x_j \geq 1 \quad \forall i \in I \quad (2)$$

$$x_j \in \{0, 1\} \quad \forall j \in J \quad (3)$$

Let $A = (a_{ij})$ be a $m \times n$ 0-1 matrix with $I = \{1, \dots, m\}$ and $J = \{1, \dots, n\}$ be the row and column sets respectively. We say that column j can be cover a row i if $a_{ij} = 1$. Where c_j is a nonnegative value that represents the cost of selecting the column j and x_j is a decision variable, it can be 1 if column j is selected or 0 otherwise. The objective is to find a minimum cost subset $S \subseteq J$, such that each row $i \in I$ is covered by at least one column $j \in S$. In the following section, we present a simple way to understand the SCP, through an example.

3.2.2 SCP sample solution Imagine that an ambulance station can meet the needs of an geographic zone. Similarly the ambulance station can cover all the needs of the nearby areas. For example, if a station is built in Zone 1 (Figure 3) ambulance station can meet the needs of neighboring areas, that is, it could also cover: Zone 1, Zone 2, Zone 3 and Zone 4. This can be appreciated in equation (5).

In this example, we must fulfill the need to cover the geographical areas defined in accordance with the restrictions. The restriction of this case is that all areas must be covered by at least one ambulance station and the goal is to minimize the number of stations built, the cost of building a station is the same for all areas. The x_j variable represents the area j which is 1 if the ambulance station is built, and will be 0 if not. As above, it can be formulated as follows:

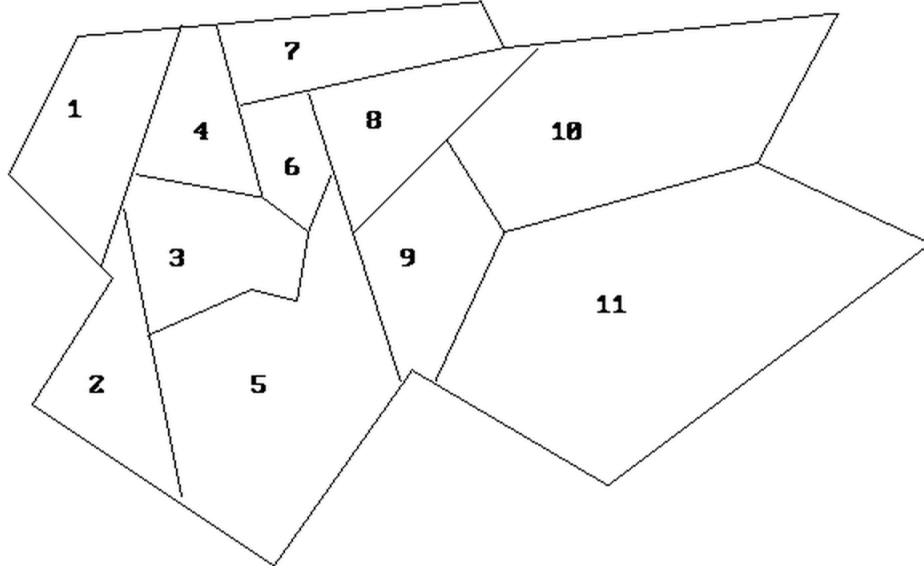


Fig. 3: Set Covering Problem example.

$$\text{Min } c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 + c_5x_5 + c_6x_6 + c_7x_7 + c_8x_8 + c_9x_9 + c_{10}x_{10} + c_{11}x_{11} \quad (4)$$

Subject to:

$$x_1 + x_2 + x_3 + x_4 \geq 1 \quad (5)$$

$$x_1 + x_2 + x_3 + x_5 \geq 1 \quad (6)$$

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 \geq 1 \quad (7)$$

$$x_1 + x_3 + x_4 + x_6 + x_7 \geq 1 \quad (8)$$

$$x_2 + x_3 + x_5 + x_6 + x_8 + x_9 \geq 1 \quad (9)$$

$$x_3 + x_4 + x_5 + x_6 + x_7 + x_8 \geq 1 \quad (10)$$

$$x_4 + x_5 + x_6 + x_7 + x_8 \geq 1 \quad (11)$$

$$x_5 + x_6 + x_7 + x_8 + x_9 + x_{10} \geq 1 \quad (12)$$

$$x_5 + x_8 + x_9 + x_{10} + x_{11} \geq 1 \quad (13)$$

$$x_8 + x_9 + x_{10} + x_{11} \geq 1 \quad (14)$$

$$x_9 + x_{10} + x_{11} \geq 1 \quad (15)$$

The first constraint (5) indicates that to cover zone 1, it is possible to locate a station in the same area or in the border. The following restriction is for zone 2 and so on. One possible optimal solution for this problem is to locate ambulance stations in zones 3, 8 and 9. That is, $x_3 = x_8 = x_9 = 1$ y $x_1 = x_2 = x_4 = x_5 = x_6 = x_7 = x_{10} = x_{11} = 0$. As shown in (Figure 4).

The author propose solve the SCP, with a variation metaheuristic Harmony Search (HS) called Binary Global-Best Harmony Search to obtain satisfactory solutions within a reasonable time. HS mimics the process of musical improvisation, where musicians make adjustments in tone to achieve aesthetic harmony.

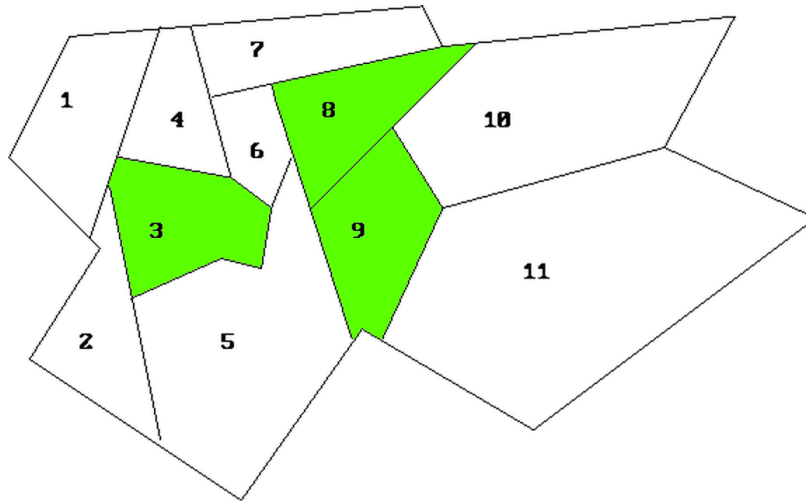


Fig. 4: Set Covering Problem solution.

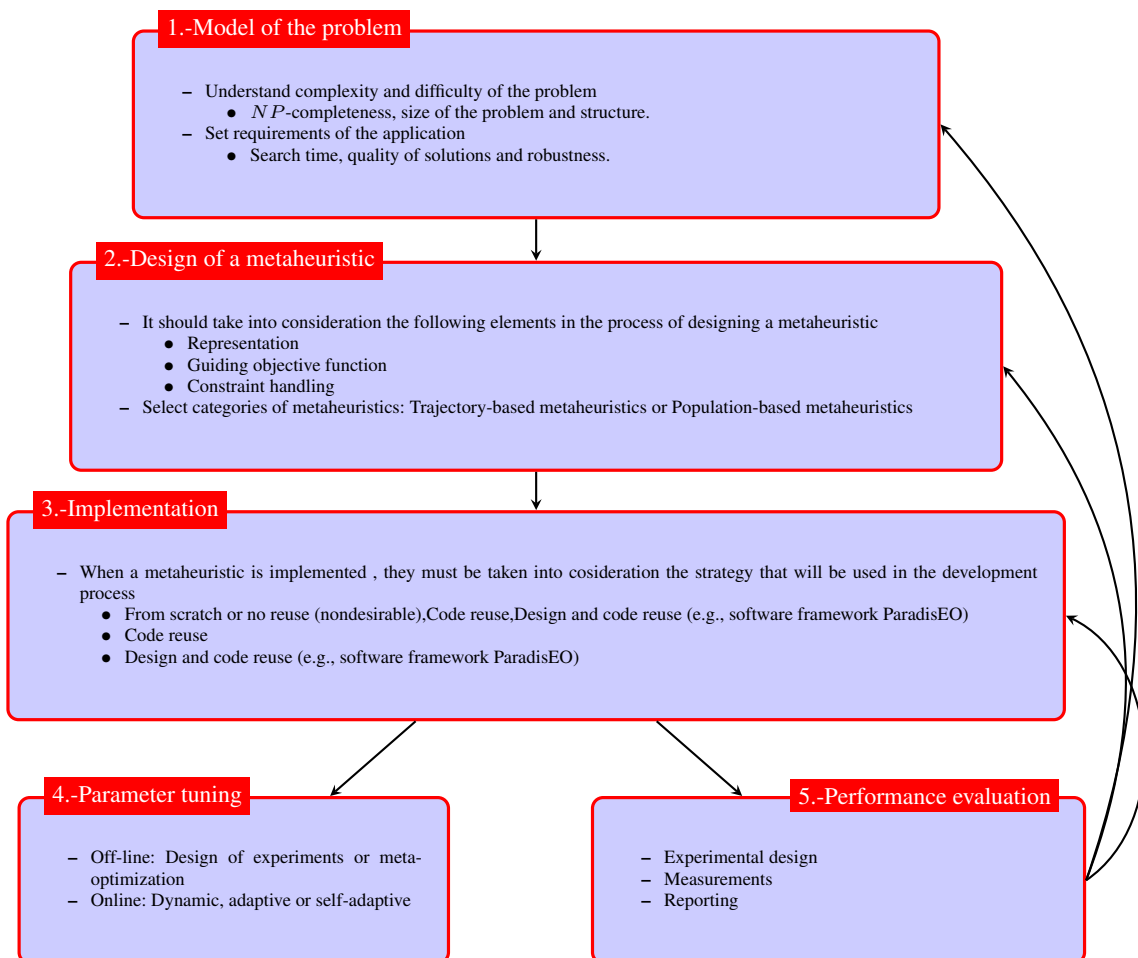


Fig. 5: Guidelines for solving SCP.

3.3 Traditional Harmony Search and some variants

3.3.1 Harmony Search Metaheuristic Harmony Search (HS) is a population-based metaheuristic algorithm inspired from the musical process of searching for a perfect state of Harmony or Aesthetic Quality of a Harmony (AQH). The HS was proposed by Z. W. Geem et al. [37]. In other words, the main idea of the HS metaheuristic is to mimic the process performed by musicians when they try to play a beautiful harmony.

For the purpose of properly understand what is looking like a good solution in this metaheuristic, we must know the meaning of AQH. The AQH in an instrument is essentially determined by its pitch (or frequency), sound quality, and amplitude (or loudness). The sound quality is mainly determined by the harmonic content that is in turn, determined by the waveforms or modulations of the sound signal. However, the harmonics that it can generate will largely depend on the pitch or frequency range of the particular instrument [38]. Different notes have different frequencies. For example, the note A has a fundamental frequency $f_0 = 440Hz$. The fundamental frequency of each note can be seen in (Table 1).

Musical Note	Frequency
C	261,625565 Hz
D	293,664768 Hz
E	329,627557 Hz
F	349,228231 Hz
G	391,995436 Hz
A	440,000000 Hz
B	493,883301 Hz

Table 1: HS - Musical Notes

Given the above, it is established that a good harmony has good AQH. The pitch of each musical instrument determines the AQH, just as the fitness function values determines the quality of solution.

In the music improvisation process, all musicians (Table 2) sound pitches (Table 3) within possible range together to make one harmony (Table 4), If all pitches make a good harmony, each musician stores in his memory that experience and possibility of making a good harmony (Table 5) in increased next time. The same thing in optimization: the initial solution is generated randomly from decision variables within the possible range, If the objective function values of these decision variables is good to make promising solution, then the possibility to make a good solution is increased next time.

In this document, the researcher set its focus on studying the classical SCP problem, given by equation (1), where we want minimize the cost of solution.

The traditional HS metaheuristic has 5 steps [39]. They will be reviewed in detail and then will be proposed the change to the original metaheuristic.


	Each musician represents a decision variable, according to the example shown, there would be 11 musicians, since there are 11 decision variables $x_1 \dots x_{11}$.
---	---

Table 2: HS components - Musician


	The pitch range of the instrument represents the range of values that can take a decision variable. Given the nature of the problem, the possible values are $\{0, 1\}$
---	---

Table 3: HS components - Pitch range

Step 1: Initialize the problem and algorithm parameters.

The optimization problem is defined as shown in equation (1) where the goal is to minimize Z subject to equations (2) and (3). x_{iL} and x_{iU} are the lower and upper bounds for decision variables. The required parameters to solve the optimization problem, that is, Harmony Memory Size (HMS) or number of solution vectors in the harmony memory,


	Musical harmony at a certain time, corresponds to a solution at a certain iteration.
---	--

Table 4: HS components - Solution


	Aesthetics audience, judges whether harmony is good or not. In the problem it refers to the objective function.
---	---

Table 5: HS components - Aesthetics audience

Harmony Memory Consideration Rate (HMCR) which determines the rate of selecting the value from the memory and Pitch Adjusting Rate (PAR) which determines the probability of local improvement and number of improvisations (NI) are given when the metaheuristic begins.

Step 2: Initialize the harmony memory.

An initial HM is filled with a population of Harmony Memory Size (HMS). We can formally say the following: $x_i^j = x_{iL} + rand()(x_{iU} - x_{iL})$, $j = 1, 2, \dots, HMS$. Where $rand()$ is a random from a uniform distribution of $[0, 1]$.

Step 3: Improve a new harmony.

Harmonies are generated randomly and the details of the procedure to improvise harmony x_i' can be given in Algorithm 1. According to the process, vectors are generated as follows (x^1, \dots, x^{HMS}) . Vectors results, make up a matrix such as that shown in (Figure 6):

$$HM = \begin{bmatrix} x_1^1 & \dots & x_n^1 \\ \vdots & \ddots & \vdots \\ x_1^{HMS} & \dots & x_n^{HMS} \end{bmatrix}$$

Fig. 6: Harmony Memory Matrix

Step 4: Update harmony memory.

If the new generated harmony $x' = (x'_1, x'_2, \dots, x'_n)$ has a better fitness than the worst one x_{worst} in HM, then replace the worst harmony with the new one $x_{worst} \leftarrow x'$; otherwise, go to the next step.

Step 5: Check the stop criteria.

If the stopping criterion (maximum number of iterations NI) is satisfied, computation is terminated. Otherwise, step 3 is repeated.

3.3.2 Global-Best Harmony Search Metaheuristic To further improve the convergence performance of HS and overcome some shortcomings of HS, a new variant of HS, called Global-Best Harmony Search (GHS), was proposed by Omran and Mahdavi [40]. First, the GHS dynamically updates parameter PAR according to equation (16):

$$PAR(t) = PAR_{min} + \frac{PAR_{max} - PAR_{min}}{NI} t \quad (16)$$

where $PAR(t)$ represents the pitch adjusting rate at generation t , PAR_{min} and PAR_{max} are the minimum and maximum adjusting rate, respectively. The parameter t is the iterative variable, and parameter NI is the number of improvisations.

3.3.3 Binary Global-Best Harmony Search Metaheuristic The HS is good at identifying the high performance regions of the solution space in a reasonable time, but poor at performing local search [41]. Namely, there is an unbalance between the exploration and the exploitation of HS. Furthermore, HS designed for continuous space cannot be directly used to solve discrete combinatorial optimization problems.

Algorithm 1: Generating new harmony

```
1 for  $i \leftarrow 1$  to  $n$  do
2   if  $\text{rand}() \leq \text{HMCR}$  then
3      $x'_i = x_i^j$  ( $j = 1, 2, \dots, \text{HMS}$ ) //memory consideration
4     if  $\text{rand}() \leq \text{PAR}$  then  $x'_i = x_i \pm r(bw)$ ;
5   else
6      $x'_i = x_{iL} + \text{rand}() (x_{iU} - x_{iL})$  //random selection
```

In order to overcome the drawbacks of HS, a novel binary global-best harmony search (BGHS) is designed for binary optimization problems.

Owing to better performance of GHS, some modifications to GHS are introduced to further enhance the convergence performance of GHS. Then a novel binary coded GHS, a two-phase repair operator, and a greedy selection mechanism are integrated into the BGHS, and they are described in detail as follows.

Algorithm 2: Generating a new greedy harmony

```
1 Function new_aggressive_harmony ()
2   CostVector  $\leftarrow$  getCostVector()
3   ProfitVector  $\leftarrow$  getMu_j(CostVector)
4   AMatrix  $\leftarrow$  getAMatrix()
5   foreach row in AMatrix do
6     RowCost  $\leftarrow$  row * ProfitVector LowCost  $\leftarrow$  getMaxValue(ProfitVector)
```

Algorithm 3: Repair operator ADD and DROP

```
1: //ADD Phase
2:  $M \leftarrow 1, 2, \dots, m$ 
3:  $A_i \sum_{j=1}^n a_{ij}x_j, i \in M$ 
4: for  $j \leftarrow 1$  to  $n$  do
5:   if  $x_j = 0$  and  $\exists i \in M, A_i < 1$  then
6:      $x_j \leftarrow 1$ 
7:      $A_i \leftarrow A_i + a_{ij}$ 
8:   end if
9: end for
10: //DROP Phase
11: for  $j \leftarrow n$  to  $1$  do
12:   if  $x_j = 1$  and  $\exists i \in M, A_i - a_{ij} \geq 1$  then
13:      $x_j \leftarrow 0$ 
14:      $A_i \leftarrow A_i - a_{ij}$ 
15:   end if
16: end for
```

4 Algorithm Implementation

Implementacion

5 Proposed Improvements

improvements.tex

6 Results

Resultados

7 Conclusion

Conclusión

8 Appendix

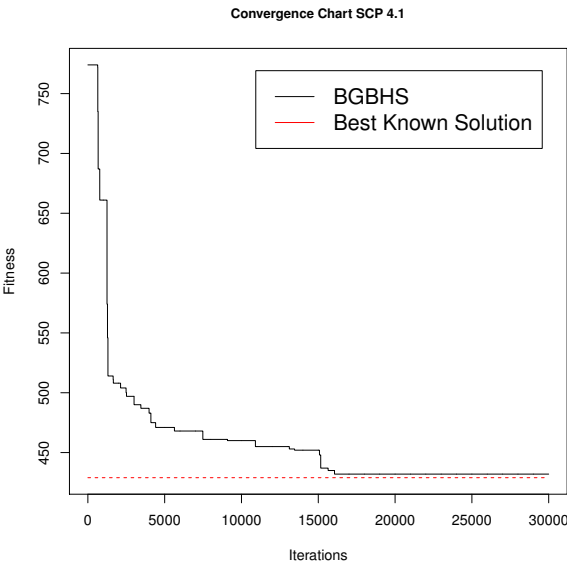


Fig. 7: Instance 4.1.

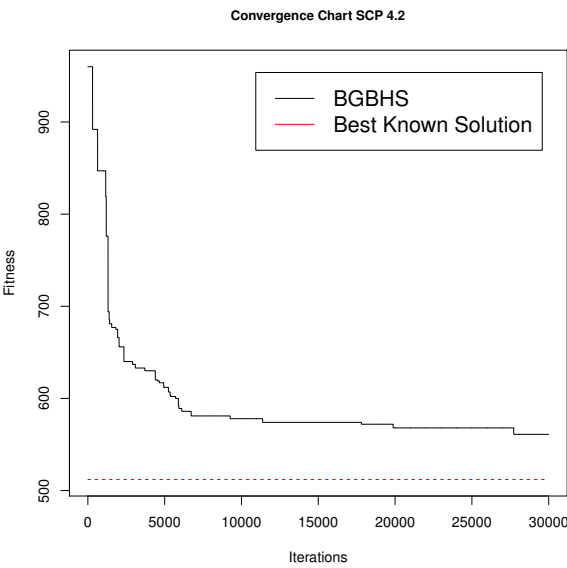


Fig. 8: Instance 4.2.

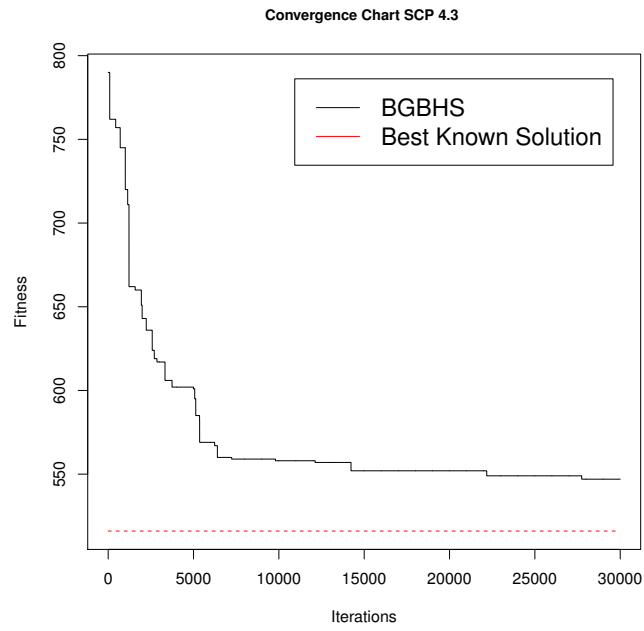


Fig. 9: Instance 4.3.

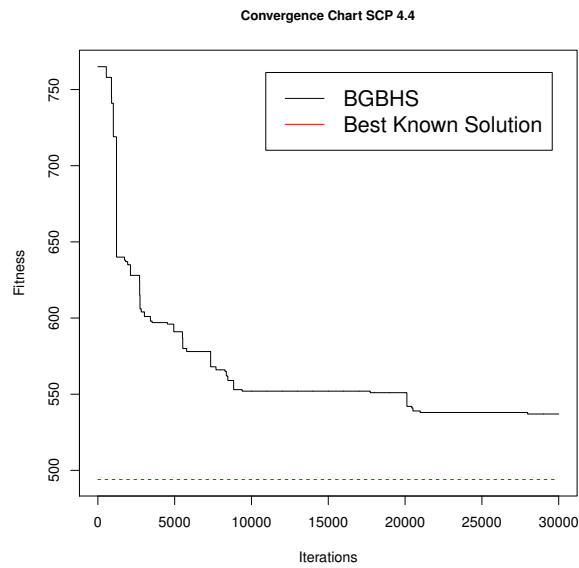


Fig. 10: Instance 4.4.

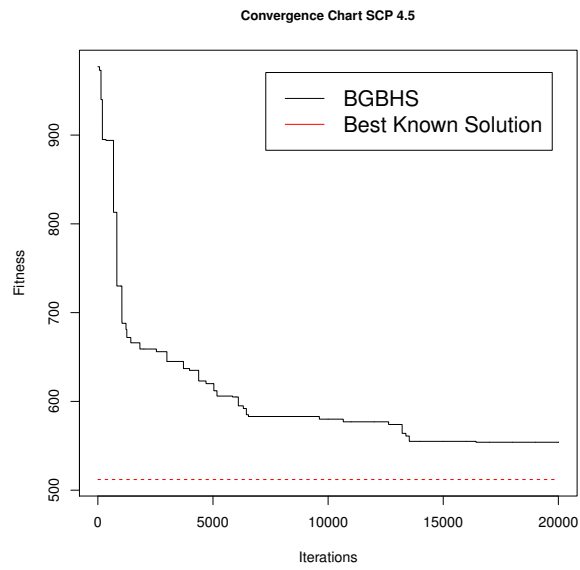


Fig. 11: Instance 4.5.

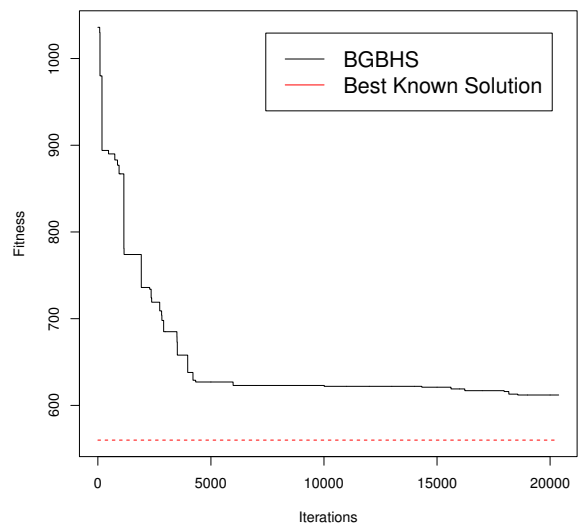


Fig. 12: Instance 4.6.

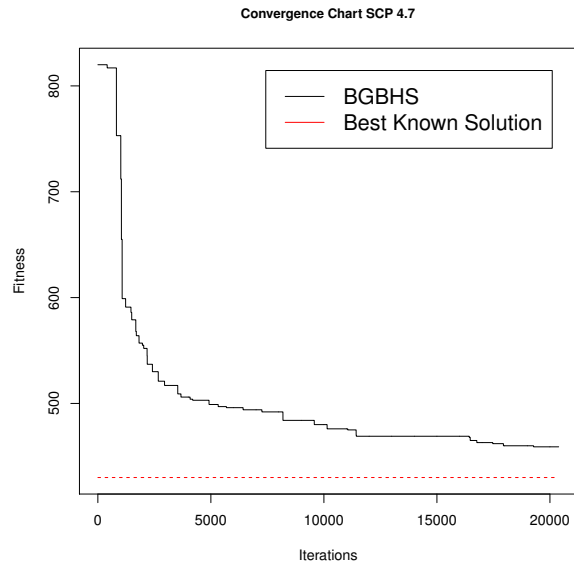


Fig. 13: Instance 4.7.

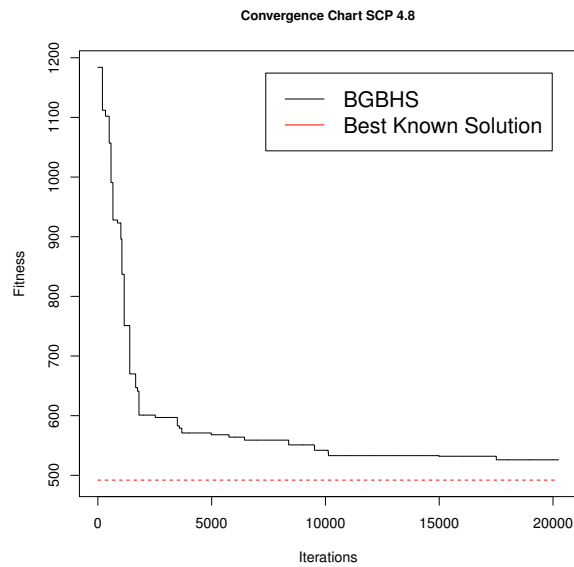


Fig. 14: Instance 4.8.

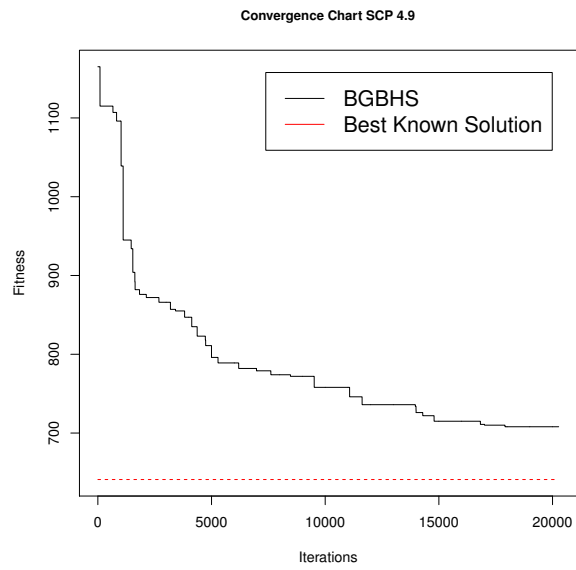


Fig. 15: Instance 4.9.

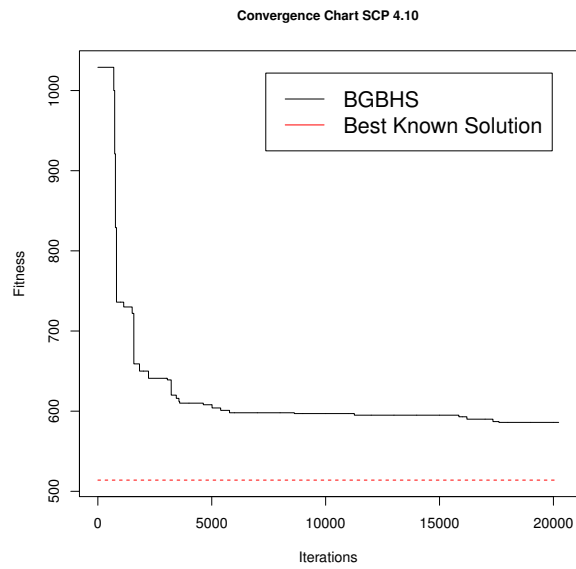


Fig. 16: Instance 4.10.

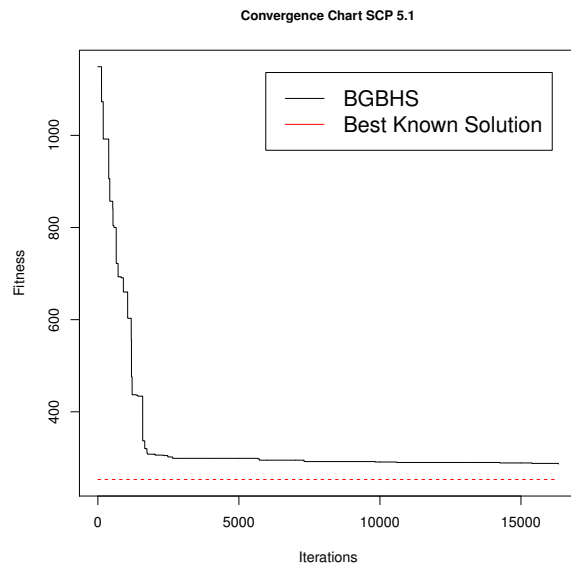


Fig. 17: Instance 5.1.

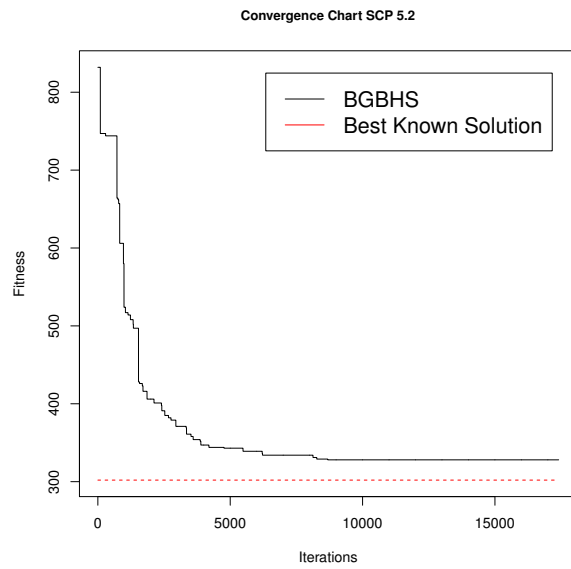


Fig. 18: Instance 5.2.

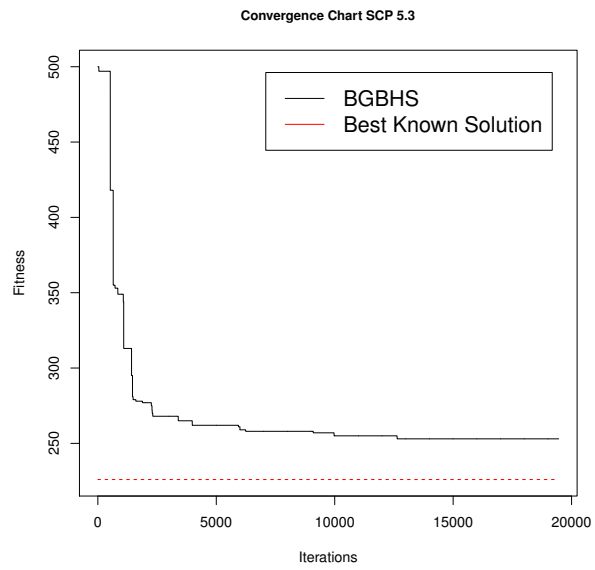


Fig. 19: Instance 5.3.

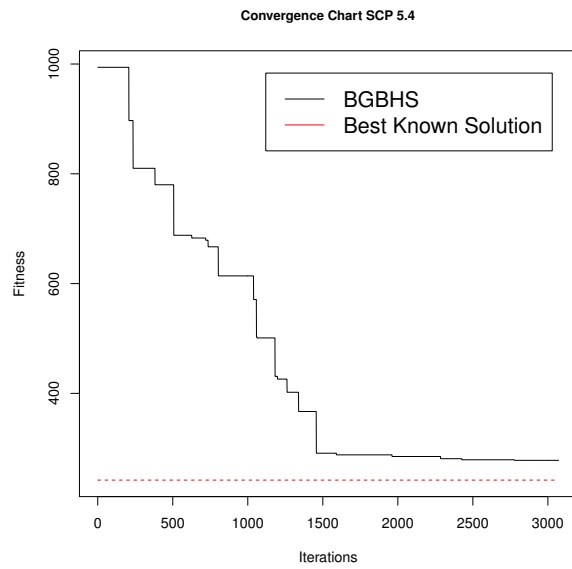


Fig. 20: Instance 5.4.

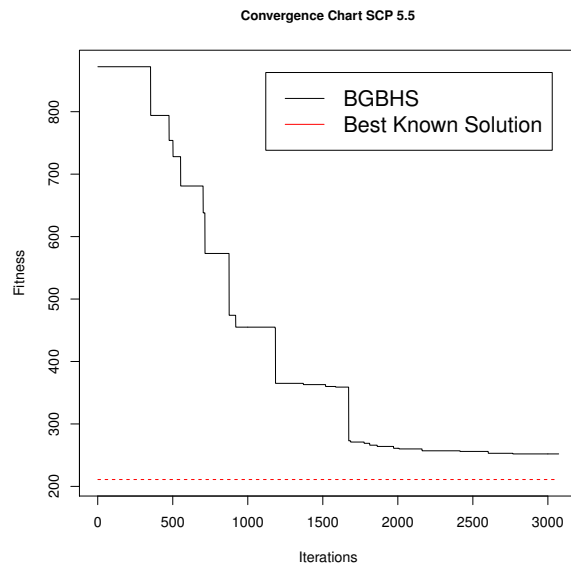


Fig. 21: Instance 5.5.

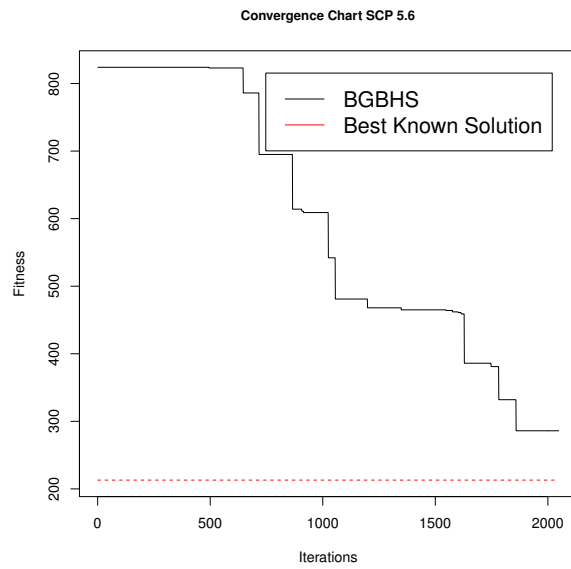


Fig. 22: Instance 5.6.

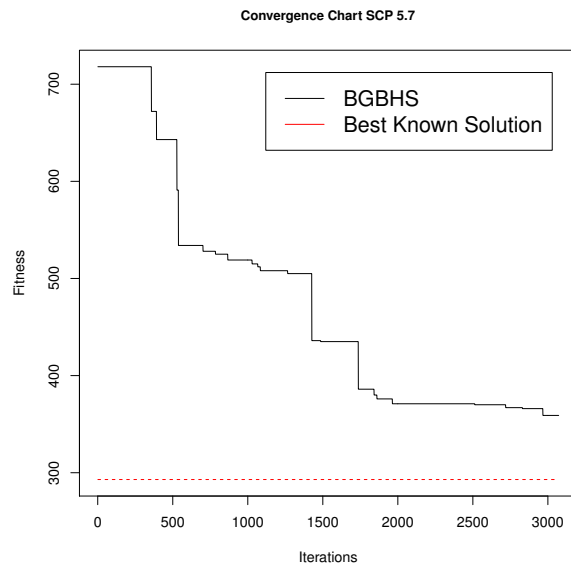


Fig. 23: Instance 5.7.

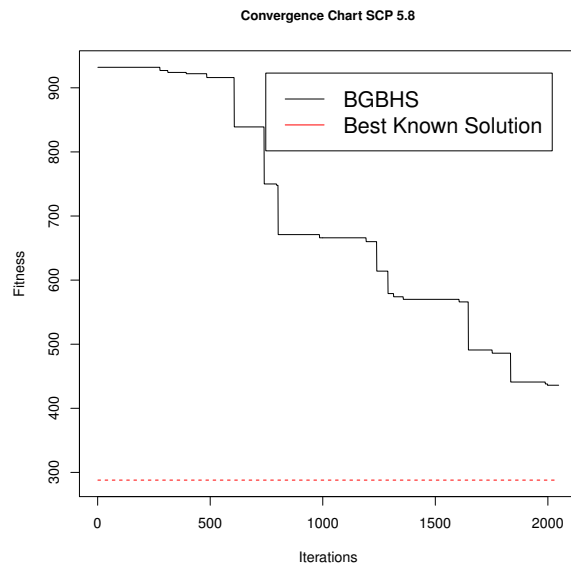


Fig. 24: Instance 5.8.

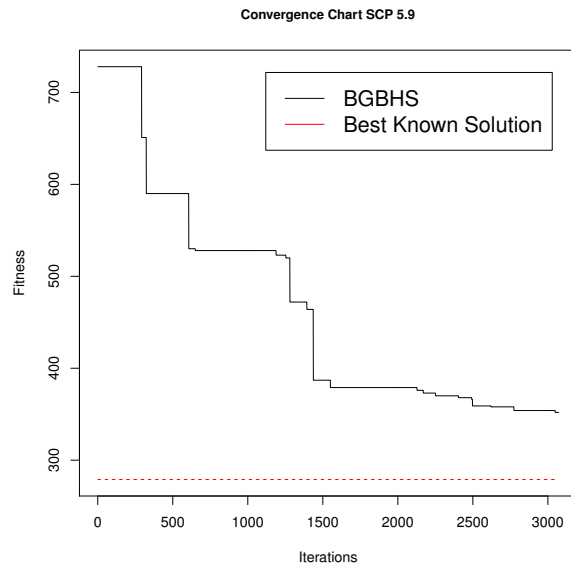


Fig. 25: Instance 5.9.

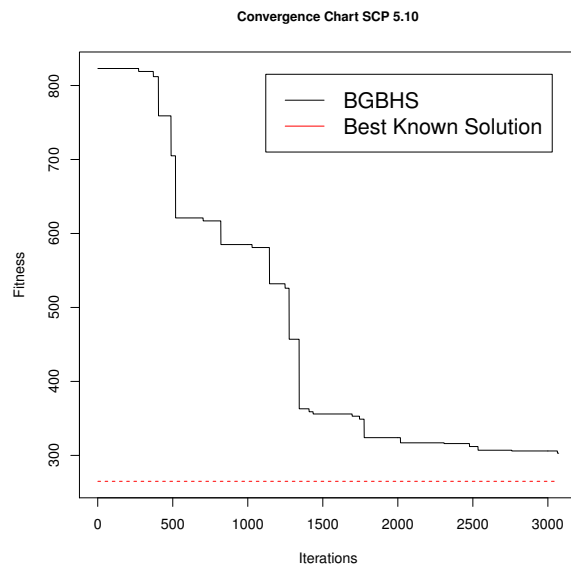


Fig. 26: Instance 5.10.

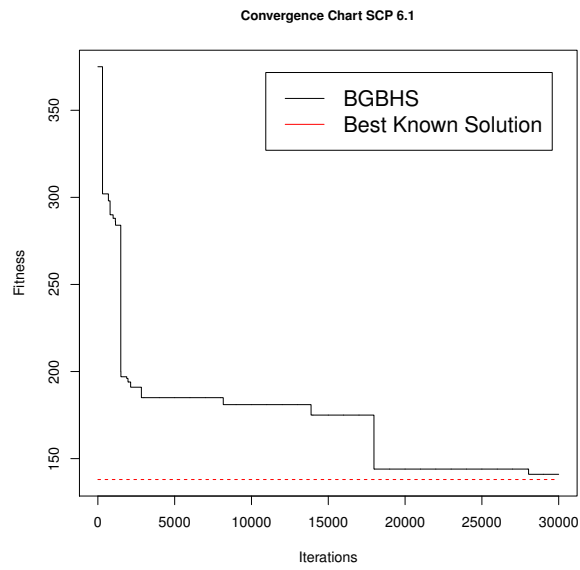


Fig. 27: Instance 6.1.

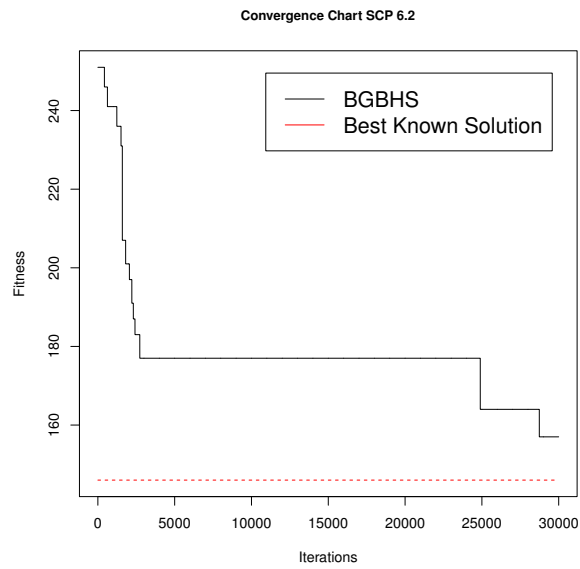


Fig. 28: Instance 6.2.

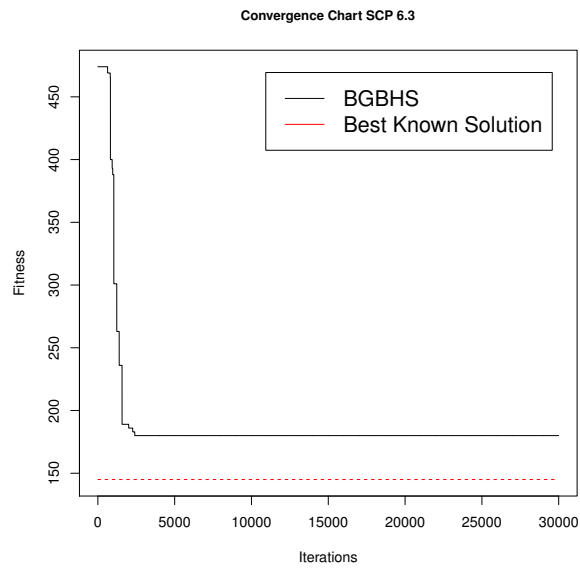


Fig. 29: Instance 6.3.

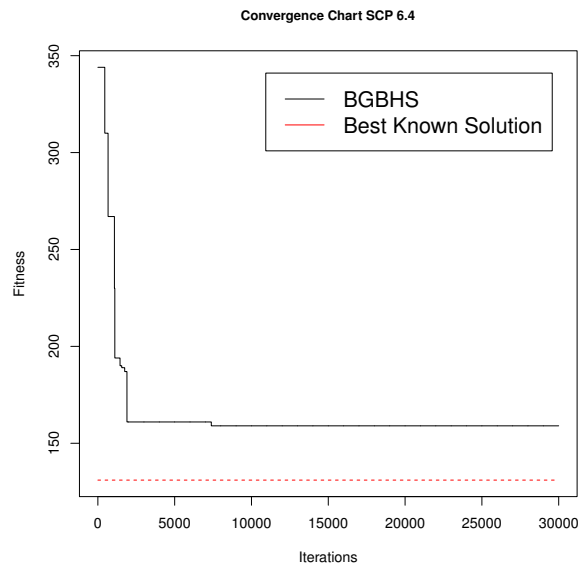


Fig. 30: Instance 6.4.

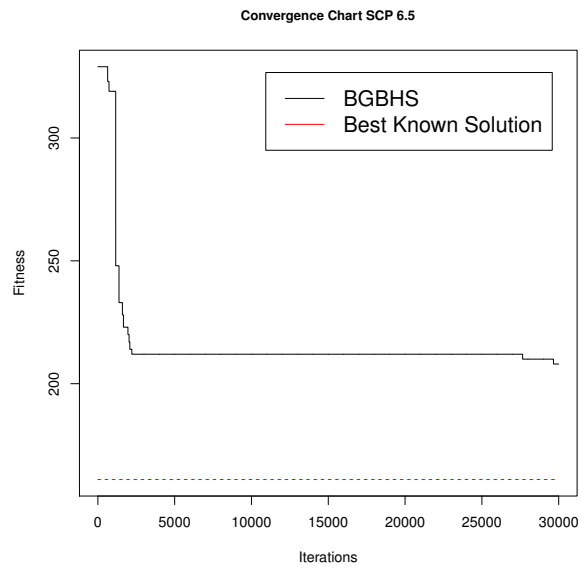


Fig. 31: Instance 6.5.

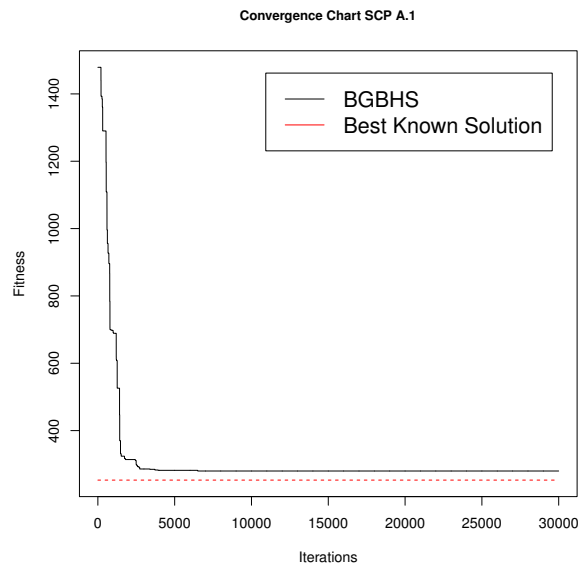


Fig. 32: Instance A.1.

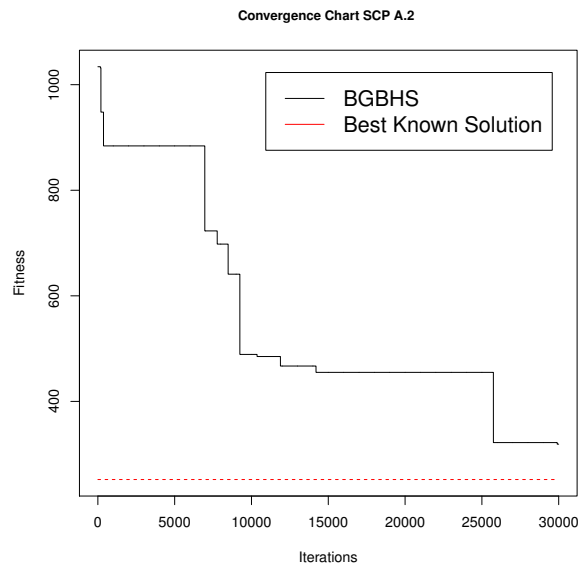


Fig. 33: Instance A.2.

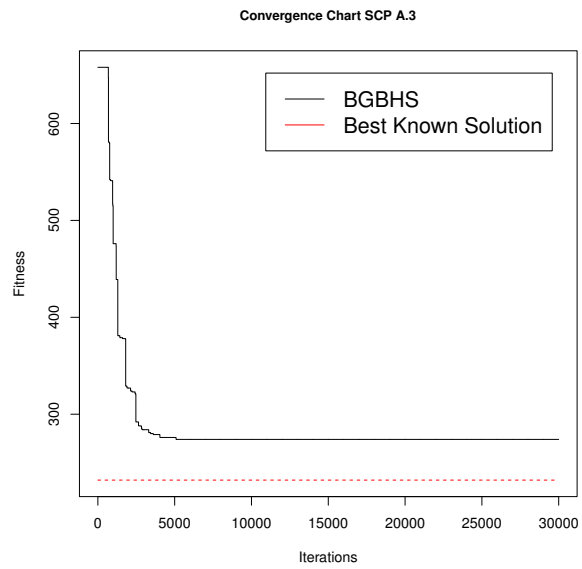


Fig. 34: Instance A.3.

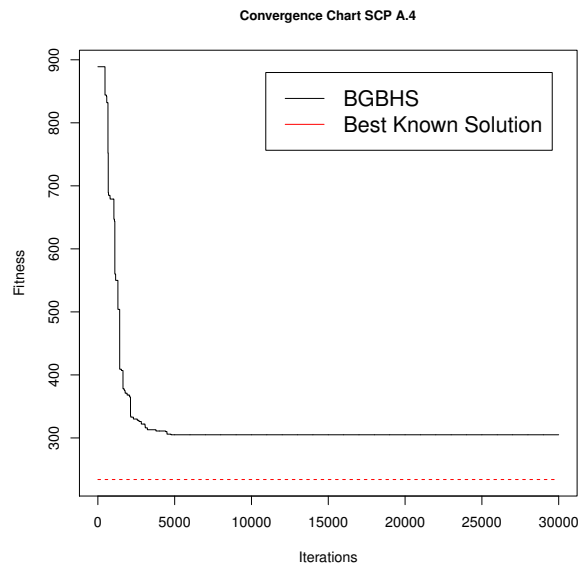


Fig. 35: Instance A.4.

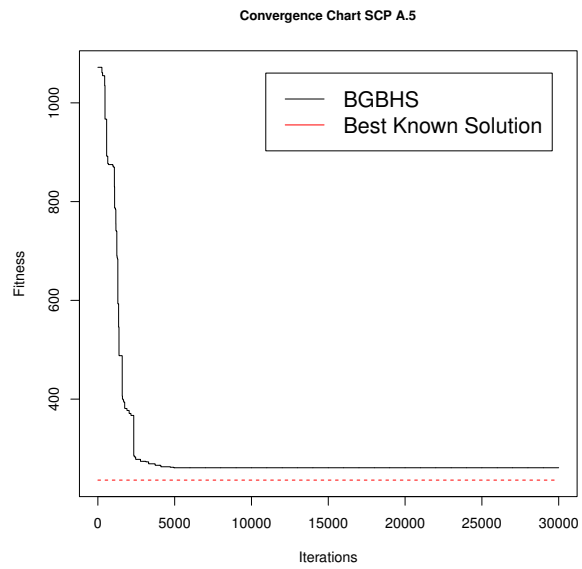


Fig. 36: Instance A.5.

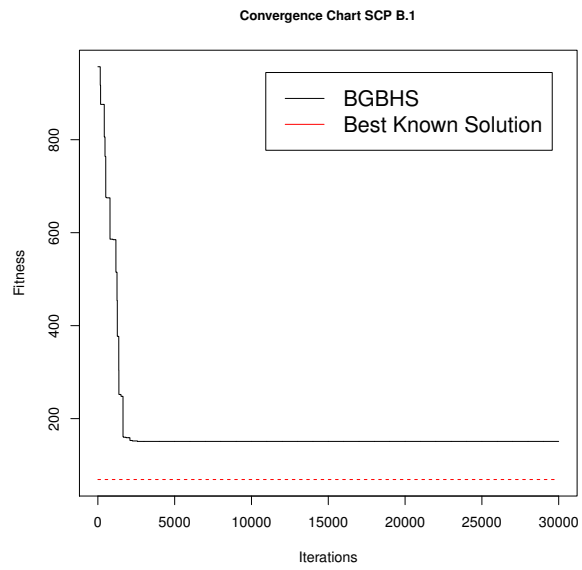


Fig. 37: Instance B.1.

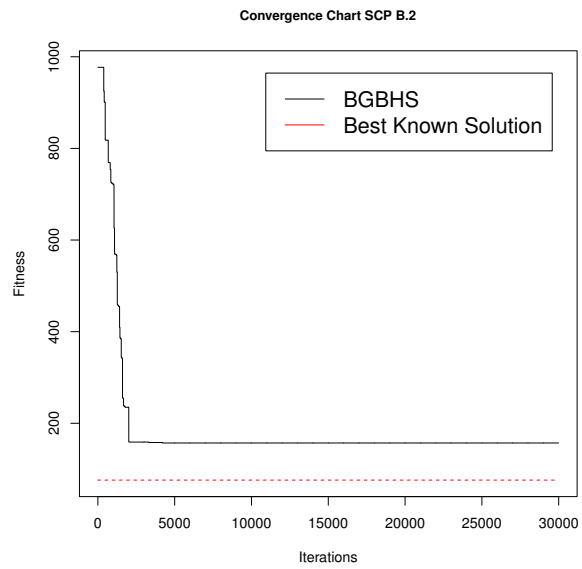


Fig. 38: Instance B.2.

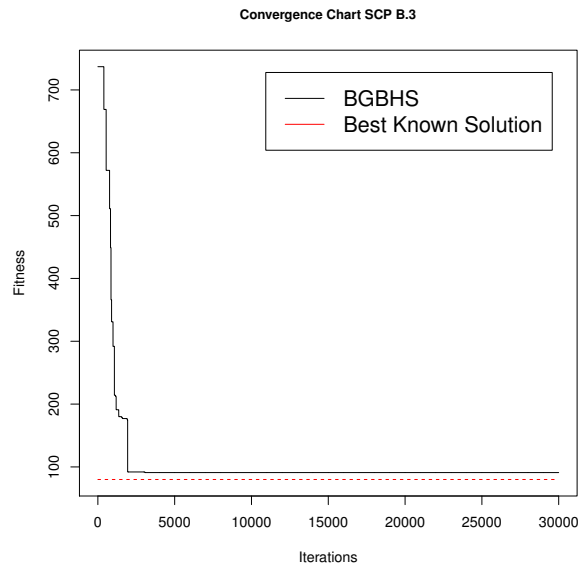


Fig. 39: Instance B.3.

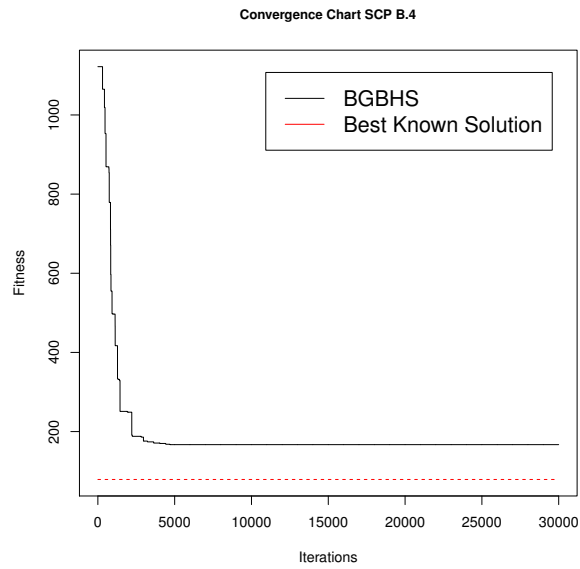


Fig. 40: Instance B.4.

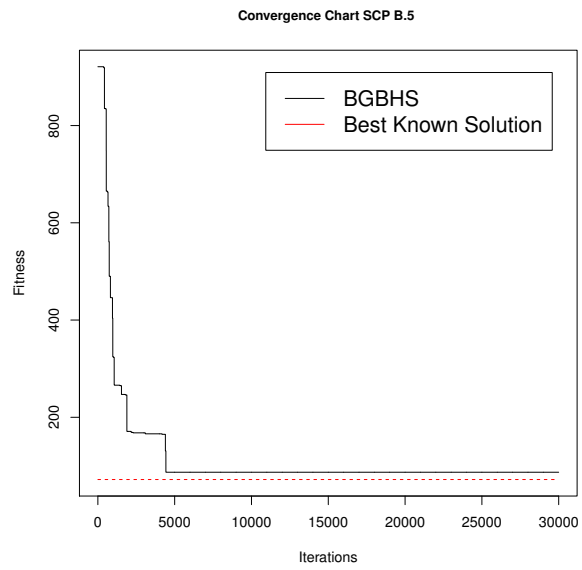


Fig. 41: Instance B.5.

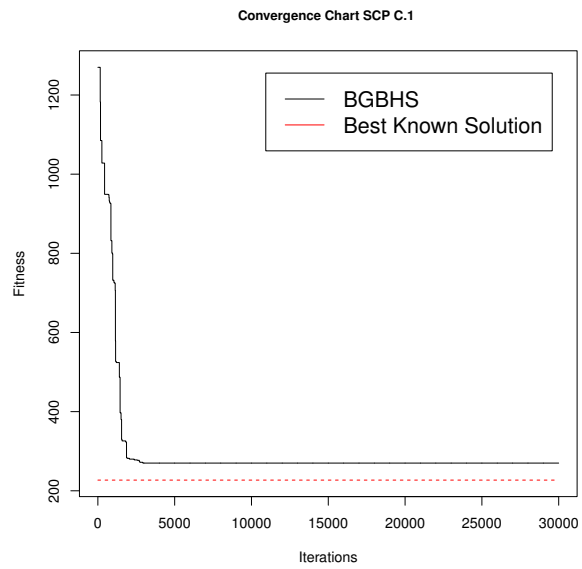


Fig. 42: Instance C.1.

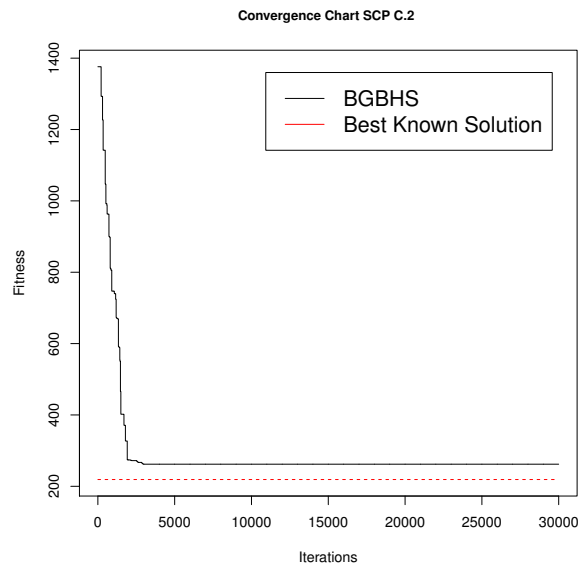


Fig. 43: Instance C.2.

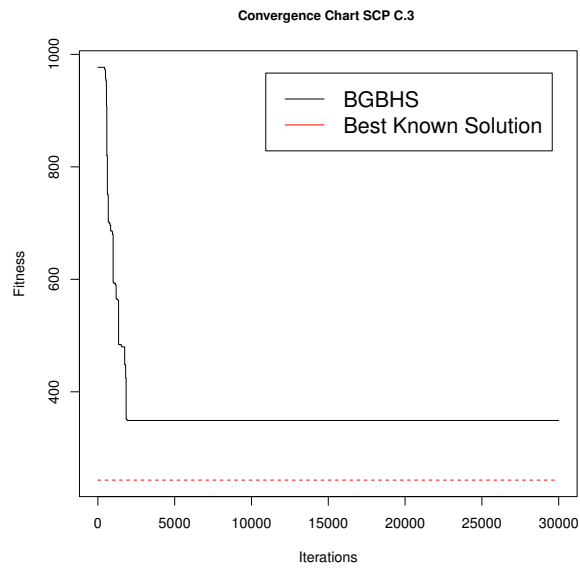


Fig. 44: Instance C.3.

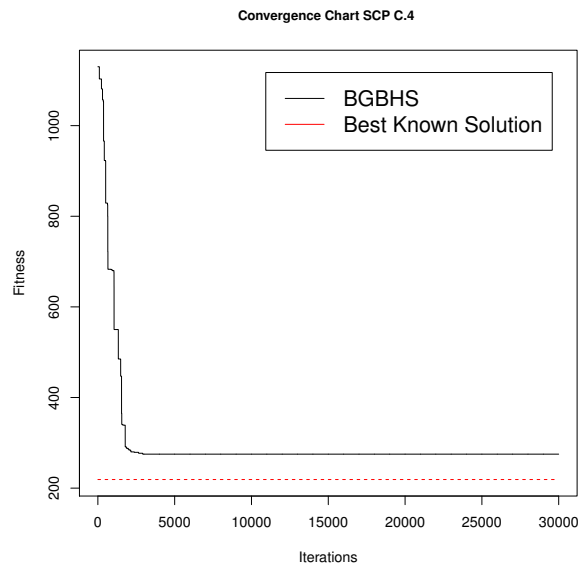


Fig. 45: Instance C.4.

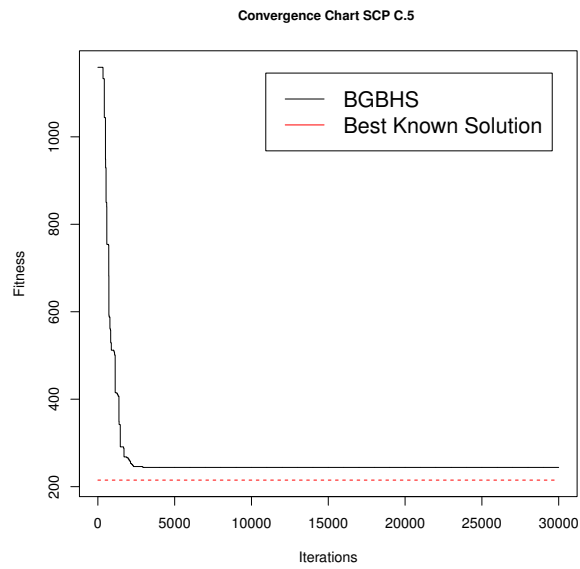


Fig. 46: Instance C.5.

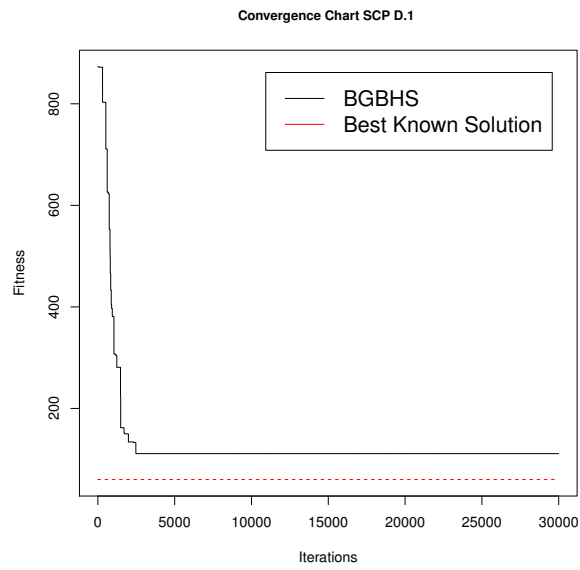


Fig. 47: Instance D.1.

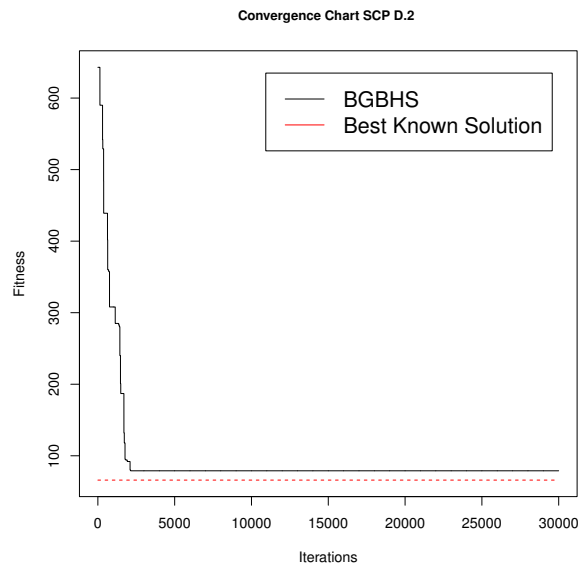


Fig. 48: Instance D.2.

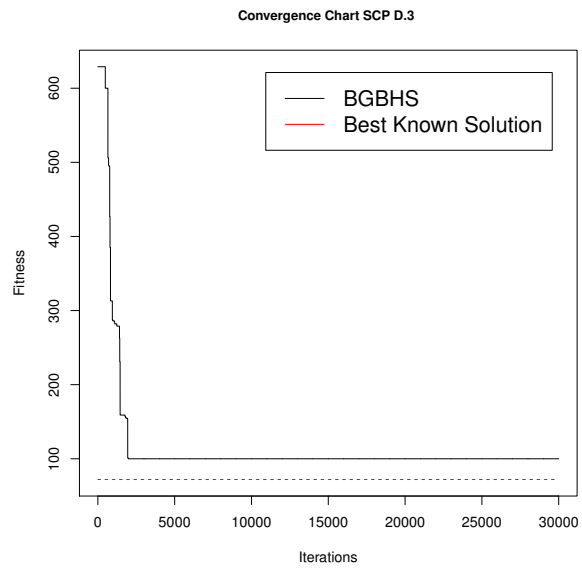


Fig. 49: Instance D.3.

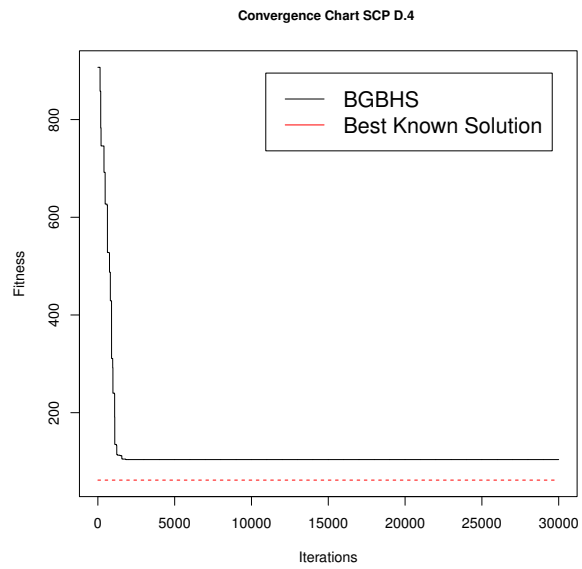


Fig. 50: Instance D.4.

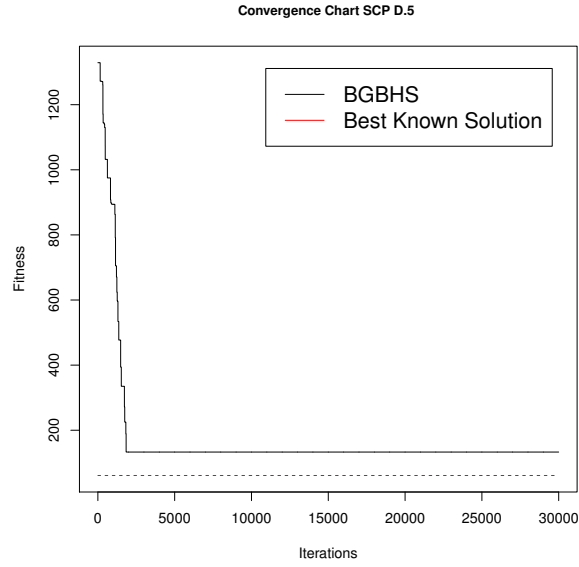


Fig. 51: Instance D.5.

8.1 Instance 4.1

Method	Optimum	Min	Max	Avg	RPD
BGBHS	429	534			24.475524475524477
BGBHS	429	553			28.904428904428904
BGBHS	429	533			24.242424242424242
BGBHS	429	600			39.86013986013986
BGBHS	429	438			2.0979020979020979
BGBHS	429	460			7.2261072261072261
BGBHS	429	455			6.0606060606060606
BGBHS	429	470			9.5571095571095572
BGBHS	429	533			24.242424242424242
BGBHS	429	532			24.009324009324008
BGBHS IMPROVED	429	451			5.1282051282051286
BGBHS IMPROVED	429	440			2.5641025641025643
BGBHS IMPROVED	429	460			7.2261072261072261
BGBHS IMPROVED	429	445			3.7296037296037294
BGBHS IMPROVED	429	470			9.5571095571095572
BGBHS IMPROVED	429	430			0.23310023310023309
BGBHS IMPROVED	429	430			0.23310023310023309
BGBHS IMPROVED	429	432			0.69930069930069927
BGBHS IMPROVED	429	435			1.3986013986013985
BGBHS IMPROVED	429	430			0.23310023310023309

Table 6: Instance 4.1

8.2 Instance 4.2

Method	Optimum	Min	Max	Avg	RPD
BGBHS	512	632			23.4375
BGBHS	512	662			29.296875
BGBHS	512	660			28.90625
BGBHS	512	602			17.578125
BGBHS	512	641			25.1953125
BGBHS	512	644			25.78125
BGBHS	512	642			25.390625
BGBHS	512	647			26.3671875
BGBHS	512	641			25.1953125
BGBHS	512	623			21.6796875
BGBHS IMPROVED	512	550			7.421875
BGBHS IMPROVED	512	560			9.375
BGBHS IMPROVED	512	563			9.9609375
BGBHS IMPROVED	512	520			1.5625
BGBHS IMPROVED	512	523			2.1484375
BGBHS IMPROVED	512	555			8.3984375
BGBHS IMPROVED	512	524			2.34375
BGBHS IMPROVED	512	518			1.171875
BGBHS IMPROVED	512	519			1.3671875
BGBHS IMPROVED	512	520			1.5625

Table 7: Instance 4.2

8.3 Instance 4.3

Method	Optimum	Min	Max	Avg	RPD
BGBHS	516				-100
BGBHS	516				-100
BGBHS	516				-100
BGBHS	516				-100
BGBHS	516				-100
BGBHS	516				-100
BGBHS	516				-100
BGBHS	516				-100
BGBHS	516				-100
BGBHS	516				-100
BGBHS IMPROVED	516	613			18.7984496124031
BGBHS IMPROVED	516				-100
BGBHS IMPROVED	516				-100
BGBHS IMPROVED	516				-100
BGBHS IMPROVED	516				-100
BGBHS IMPROVED	516				-100
BGBHS IMPROVED	516				-100
BGBHS IMPROVED	516				-100
BGBHS IMPROVED	516				-100
BGBHS IMPROVED	516				-100

Table 8: Instance 4.3

8.4 Instance 4.4

Method	Optimum	Min	Max	Avg	RPD
BGBHS	494				-100
BGBHS	494				-100
BGBHS	494				-100
BGBHS	494				-100
BGBHS	494				-100
BGBHS	494				-100
BGBHS	494				-100
BGBHS	494				-100
BGBHS	494				-100
BGBHS	494				-100
BGBHS IMPROVED	494				-100
BGBHS IMPROVED	494				-100
BGBHS IMPROVED	494				-100
BGBHS IMPROVED	494				-100
BGBHS IMPROVED	494				-100
BGBHS IMPROVED	494				-100
BGBHS IMPROVED	494				-100
BGBHS IMPROVED	494				-100
BGBHS IMPROVED	494				-100
BGBHS IMPROVED	494				-100

Table 9: Instance 4.4

8.5 Instance 4.5

Method	Optimum	Min	Max	Avg	RPD
BGBHS	512				-100
BGBHS	512				-100
BGBHS	512				-100
BGBHS	512				-100
BGBHS	512				-100
BGBHS	512				-100
BGBHS	512				-100
BGBHS	512				-100
BGBHS	512				-100
BGBHS	512				-100
BGBHS IMPROVED	512				-100
BGBHS IMPROVED	512				-100
BGBHS IMPROVED	512				-100
BGBHS IMPROVED	512				-100
BGBHS IMPROVED	512				-100
BGBHS IMPROVED	512				-100
BGBHS IMPROVED	512				-100
BGBHS IMPROVED	512				-100
BGBHS IMPROVED	512				-100
BGBHS IMPROVED	512				-100

Table 10: Instance 4.5

8.6 Instance 4.6

Method	Optimum	Min	Max	Avg	RPD
BGBHS	560				-100
BGBHS	560				-100
BGBHS	560				-100
BGBHS	560				-100
BGBHS	560				-100
BGBHS	560				-100
BGBHS	560				-100
BGBHS	560				-100
BGBHS	560				-100
BGBHS	560				-100
BGBHS IMPROVED	560				-100
BGBHS IMPROVED	560				-100
BGBHS IMPROVED	560				-100
BGBHS IMPROVED	560				-100
BGBHS IMPROVED	560				-100
BGBHS IMPROVED	560				-100
BGBHS IMPROVED	560				-100
BGBHS IMPROVED	560				-100
BGBHS IMPROVED	560				-100
BGBHS IMPROVED	560				-100

Table 11: Instance 4.6

8.7 Instance 4.7

Method	Optimum	Min	Max	Avg	RPD
BGBHS	430				-100
BGBHS	430				-100
BGBHS	430				-100
BGBHS	430				-100
BGBHS	430				-100
BGBHS	430				-100
BGBHS	430				-100
BGBHS	430				-100
BGBHS	430				-100
BGBHS	430				-100
BGBHS IMPROVED	430				-100
BGBHS IMPROVED	430				-100
BGBHS IMPROVED	430				-100
BGBHS IMPROVED	430				-100
BGBHS IMPROVED	430				-100
BGBHS IMPROVED	430				-100
BGBHS IMPROVED	430				-100
BGBHS IMPROVED	430				-100
BGBHS IMPROVED	430				-100
BGBHS IMPROVED	430				-100

Table 12: Instance 4.7

8.8 Instance 4.8

Method	Optimum	Min	Max	Avg	RPD
BGBHS	492				-100
BGBHS	492				-100
BGBHS	492				-100
BGBHS	492				-100
BGBHS	492				-100
BGBHS	492				-100
BGBHS	492				-100
BGBHS	492				-100
BGBHS	492				-100
BGBHS	492				-100
BGBHS IMPROVED	492				-100
BGBHS IMPROVED	492				-100
BGBHS IMPROVED	492				-100
BGBHS IMPROVED	492				-100
BGBHS IMPROVED	492				-100
BGBHS IMPROVED	492				-100
BGBHS IMPROVED	492				-100
BGBHS IMPROVED	492				-100
BGBHS IMPROVED	492				-100
BGBHS IMPROVED	492				-100

Table 13: Instance 4.8

8.9 Instance 4.9

Method	Optimum	Min	Max	Avg	RPD
BGBHS	641				-100
BGBHS	641				-100
BGBHS	641				-100
BGBHS	641				-100
BGBHS	641				-100
BGBHS	641				-100
BGBHS	641				-100
BGBHS	641				-100
BGBHS	641				-100
BGBHS	641				-100
BGBHS IMPROVED	641				-100
BGBHS IMPROVED	641				-100
BGBHS IMPROVED	641				-100
BGBHS IMPROVED	641				-100
BGBHS IMPROVED	641				-100
BGBHS IMPROVED	641				-100
BGBHS IMPROVED	641				-100
BGBHS IMPROVED	641				-100
BGBHS IMPROVED	641				-100
BGBHS IMPROVED	641				-100

Table 14: Instance 4.9

8.10 Instance 4.10

Table 15: Instance 4.10

8.11 Instance 5.1

Table 16: Instance 5.1

8.12 Instance 5.2

Table 17: Instance 5.2

8.13 Instance 5.3

Table 18: Instance 5.3

8.14 Instance 5.4

Table 19: Instance 5.4

8.15 Instance 5.5

Table 20: Instance 5.5

8.16 Instance 5.6

Table 21: Instance 5.6

8.17 Instance 5.7

Table 22: Instance 5.7

8.18 Instance 5.8

Table 23: Instance 5.8

8.19 Instance 5.9

Table 24: Instance 5.9

8.20 Instance 5.10

Table 25: Instance 5.10

8.21 Instance 6.1

Table 26: Instance 6.1

8.22 Instance 6.2

Table 27: Instance 6.2

8.23 Instance 6.3

Table 28: Instance 6.3

8.24 Instance 6.4

Table 29: Instance 6.4

8.25 Instance 6.5

Table 30: Instance 6.5

8.26 Instance A.1

Table 31: Instance A.1

8.27 Instance A.2

Table 32: Instance A.2

8.28 Instance A.3

Table 33: Instance A.3

8.29 Instance A.4

Table 34: Instance A.3

8.30 Instance A.5

Table 35: Instance A.5

8.31 Instance B.1

Table 36: Instance B.1

8.32 Instance B.2

Table 37: Instance B.2

8.33 Instance B.3

Table 38: Instance B.3

8.34 Instance B.4

Table 39: Instance B.4

8.35 Instance B.5

Table 40: Instance B.5

8.36 Instance C.1

Table 41: Instance C.1

8.37 Instance C.2

Table 42: Instance C.2

8.38 Instance C.3

Table 43: Instance C.3

8.39 Instance C.4

Table 44: Instance C.4

8.40 Instance C.5

Table 45: Instance C.5

8.41 Instance D.1

Table 46: Instance D.1

8.42 Instance D.2

Table 47: Instance D.2

8.43 Instance D.3

Table 48: Instance D.3

8.44 Instance D.4

Table 49: Instance D.4

8.45 Instance D.5

Table 50: Instance D.5

References

1. E. Talbi, *Metaheuristics - From Design to Implementation*. Wiley, 2009.
2. M. Birattari, T. Stützle, L. Paquete, and K. Varrentrapp, "A racing algorithm for configuring metaheuristics," in *GECCO 2002: Proceedings of the Genetic and Evolutionary Computation Conference, New York, USA, 9-13 July 2002* (W. B. Langdon, E. Cantú-Paz, K. E. Mathias, R. Roy, D. Davis, R. Poli, K. Balakrishnan, V. Honavar, G. Rudolph, J. Wegener, L. Bull, M. A. Potter, A. C. Schultz, J. F. Miller, E. K. Burke, and N. Jonoska, eds.), pp. 11–18, Morgan Kaufmann, 2002.
3. J. E. Beasley, "OR-library: distributing test problems by electronic mail," *Journal of the Operational Research Society*, vol. 41, no. 11, pp. 1069–1072, 1990.
4. M. Birattari, L. Paquete, T. Stützle, and K. Varrentrapp, "Classification of Metaheuristics and Design of Experiments for the Analysis of Components."
5. Z. Beheshti and S. M. Hj. Shamsuddin, "Capso: Centripetal accelerated particle swarm optimization," *Inf. Sci.*, vol. 258, pp. 54–79, Feb. 2014.
6. Á. G. Rubio, B. Crawford, R. Soto, A. Jaramillo, S. M. Villablanca, J. Salas, and E. Olguín, *An Binary Black Hole Algorithm to Solve Set Covering Problem*, pp. 873–883. Cham: Springer International Publishing, 2016.
7. O. Durán, N. Rodriguez, and L. A. Consalter, "Collaborative particle swarm optimization with a data mining technique for manufacturing cell design," *Expert Syst. Appl.*, vol. 37, pp. 1563–1567, Mar. 2010.
8. B. Crawford, R. Soto, W. Palma, F. Johnson, F. Paredes, and E. Olguín, "A 2-level approach for the set covering problem: Parameter tuning of artificial bee colony algorithm by using genetic algorithm," in *Advances in Swarm Intelligence - 5th International Conference, ICSI 2014, Hefei, China, October 17-20, 2014, Proceedings, Part I* (Y. Tan, Y. Shi, and C. A. C. Coello, eds.), vol. 8794 of *Lecture Notes in Computer Science*, pp. 189–196, Springer, 2014.
9. R. Aringhieri, A. Grosso, P. Hosteins, and R. Scatamacchia, "Local search metaheuristics for the critical node problem," *Networks*, vol. 67, no. 3, pp. 209–221, 2016.
10. R. Soto, B. Crawford, C. Galleguillos, E. Monfroy, and F. Paredes, "A hybrid ac3-tabu search algorithm for solving sudoku puzzles," *Expert Syst. Appl.*, vol. 40, no. 15, pp. 5817–5821, 2013.
11. K. Al-Ajmi and M. El-Abd, "Implementing a population-based harmony search algorithm on graphic processing units," in *IEEE 27th Canadian Conference on Electrical and Computer Engineering, CCECE 2014, Toronto, ON, Canada, May 4-7, 2014*, pp. 1–5, IEEE, 2014.
12. S. Aupetit, N. Monmarché, and M. Slimane, *Hidden Markov Models Training Using Population-based Metaheuristics*, pp. 415–438. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008.
13. P. Vansteenwegen, W. Souffriau, G. V. Berghe, and D. V. Oudheusden, "A guided local search metaheuristic for the team orienteering problem," *European Journal of Operational Research*, vol. 196, no. 1, pp. 118–127, 2009.
14. B. Sarasola, K. F. Doerner, V. Schmid, and E. Alba, "Variable neighborhood search for the stochastic and dynamic vehicle routing problem," *Annals OR*, vol. 236, no. 2, pp. 425–461, 2016.
15. R. M. Karp, "Reducibility among combinatorial problems," in *50 Years of Integer Programming 1958-2008 - From the Early Years to the State-of-the-Art* (M. Jünger, T. M. Liebling, D. Naddef, G. L. Nemhauser, W. R. Pulleyblank, G. Reinelt, G. Rinaldi, and L. A. Wolsey, eds.), pp. 219–241, Springer, 2010.
16. A. I. Ali and H. Thiagarajan, "A network relaxation based enumeration algorithm for set partitioning," *European Journal of Operational Research*, vol. 38, no. 1, pp. 76–85, 1989.
17. J. J. Bartholdi, "A guaranteed-accuracy round-off algorithm for cyclic scheduling and set covering," *Operations Research*, vol. 29, no. 3, pp. 501–510, 1981.
18. W. Walker, "Using the set-covering problem to assign fire companies to fire houses," *Operations Research*, vol. 22, pp. 275–277, 1974.
19. F. J. Vasko and G. R. Wilson, "Using a facility location algorithm to solve large set covering problems," *Operations Research Letters*, vol. 3, no. 2, pp. 85–90, 1984.
20. F. J. Vasko, F. E. Wolf, and K. L. Stott, "Optimal selection of ingot sizes via set covering," *Operations Research*, vol. 35, pp. 346–353, June 1987.
21. F. J. Vasko, F. E. Wolf, and K. L. Stott, "A set covering approach to metallurgical grade assignment," *European Journal of Operational Research*, vol. 38, no. 1, pp. 27–34, 1989.
22. F. J. Vasko, F. E. Wolf, K. L. Stott, and J. W. Scheirer, "Selecting optimal ingot sizes for bethlehem steel," *Interfaces*, vol. 19, no. 1, pp. 68–84, 1989.
23. M. L. Balinski and R. E. Quandt, "On an integer program for a delivery problem," *Operations Research*, vol. 12, no. 2, pp. 300–304, 1964.
24. B. A. Foster and D. M. Ryan, "An integer programming approach to the vehicle scheduling problem," *Operations Research*, vol. 27, pp. 367–384, 1976.

25. M. L. Fisher and M. B. Rosenwein, "An interactive optimization system for bulk-cargo ship scheduling," *Naval Research Logistics*, vol. 36, no. 1, pp. 27–42, 1989.
26. M. Bellmore, H. J. Geenberg, and J. J. Jarvis, "Multi-commodity disconnecting sets," *Management Science*, vol. 16, no. 6, pp. B427–B433, 1970.
27. M. Bellmore and H. D. Ratliff, "Optimal defense of multi-commodity networks," *Management Science*, vol. 18, no. 4-part-i, pp. B174–B185, 1971.
28. B. A. Freeman and J. V. Jucker, "The line balancing problem," *Journal of Industrial Engineering*, vol. 18, pp. 361–364, 1967.
29. M. E. Salveson, "The assembly line balancing problem," *Journal of Industrial Engineering*, vol. 6, pp. 18–25, 1955.
30. C. C. Ribeiro, M. Minoux, and M. C. Penna, "An optimal column-generation-with-ranking algorithm for very large scale set partitioning problems in traffic assignment," *European Journal of Operational Research*, vol. 41, no. 2, pp. 232–239, 1989.
31. S. Ceria, P. Nobile, and A. Sassano, "A lagrangian-based heuristic for large-scale set covering problems," *Mathematical Programming*, vol. 81, no. 2, pp. 215–228, 1998.
32. M. A. Breuer, "Simplification of the covering problem with application to boolean expressions," *Journal of the Association for Computing Machinery*, vol. 17, pp. 166–181, Jan. 1970.
33. N. Christofides, "Zero-one programming using non-binary tree-search," *Computer Journal*, vol. 14, no. 4, pp. 418–421, 1971.
34. R. H. Day, "Letter to the editor—on optimal extracting from a multiple file data storage system: An application of integer programming," *Operations Research*, vol. 13, no. 3, pp. 482–494, 1965.
35. R. S. Garfinkel and G. L. Nemhauser, "Optimal political districting by implicit enumeration techniques," *Management Science*, vol. 16, no. 8, pp. B495–B508, 1970.
36. E. Housos and T. Elmroth, "Automatic optimization of subproblems in scheduling airline crews," *Interfaces*, vol. 27, no. 5, pp. 68–77, 1997.
37. Z. W. Geem, J. Kim, and G. V. Loganathan, "A new heuristic optimization algorithm: Harmony search," *Simulation*, vol. 76, no. 2, pp. 60–68, 2001.
38. Z. W. Geem, *Music-Inspired Harmony Search Algorithm: Theory and Applications*. Springer Publishing Company, Incorporated, 1st ed., 2009.
39. D. Zou, L. Gao, S. Li, and J. Wu, "Solving 0-1 knapsack problem by a novel global harmony search algorithm," *Appl. Soft Comput.*, vol. 11, no. 2, pp. 1556–1564, 2011.
40. M. G. H. Omran and M. Mahdavi, "Global-best harmony search," *Applied Mathematics and Computation*, vol. 198, no. 2, pp. 643–656, 2008.
41. W. Xiang, M. An, Y. Li, R. He, and J. Zhang, "An improved global-best harmony search algorithm for faster optimization," *Expert Syst. Appl.*, vol. 41, no. 13, pp. 5788–5803, 2014.