UNIVERSITI PUTRA MALAYSIA

TOLERABLE CONSTRUCTIVE GRAPH-BASED HYPER-HEURISTIC ALGORITHM FOR EXAMINATION TIMETABLING

SHAHRZAD MOHAMMAD POUR

FSKTM 2009 11

TOLERABLE CONSTRUCTIVE GRAPH-BASED HYPER-HEURISTIC ALGORITHM FOR EXAMINATION TIMETABLING

By

SHAHRZAD MOHAMMAD POUR

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Master of Science

October 2009

To

Whom made me think

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirements for the degree of Master of Science

TOLERABLE CONSTRUCTIVE GRAPH-BASED HYPER-HEURISTIC ALGORITHM FOR EXAMINATION TIMETABLING

By

SHAHRZAD MOHAMMAD POUR

October 2009

Chairman: Abu bakar Md Sultan, PhD

Faculty: Computer Science and Information Technology

Examination Timetabling Problem (ETTP) is an NP-hard typical optimization problem faced by institutions and universities across the world. This nature leads to heuristic methods cover a large branch of researches in this area. On the other hand, the problem varies from one institution to another, depending on the size, structure and constraints of that institution. Therefore generality of the proposed methods is one of the major goals in solving timetabling problem nowadays. These methods are trying to keep generality while adding to factors of these methods. Hyperheuristic is one of these approaches which make the basis of this thesis.

In heuristic approaches getting stuck in local optimum is one of the propounded problems from early days. The main cause for the local optimal problem is that heuristic algorithms either focus on exploration (global improvement) rather than exploitation (local improvement) or vice versa.

The proposal to address the mentioned problem in this thesis is an extension to constructive Graph-based Hyperheuristic (GHH) algorithm presented in (Burke et al., 2007), where the researchers have not considered a dynamic hybridization of graph-based heuristics in their framework such that each low level heuristic is applied for scheduling fixed number of examinations for construction a solution (timetable) at each step. On the other hand, no supervision exists on manner of current heuristic on the solution such that it isn't clear scheduling of each exam based on order of current heuristic leads to improving or destroying the solution. By this way there is no legal control between exploration and exploitation of the search space in order to avoid getting stuck in local optimum. This study aims to use a dynamic mechanism so that algorithm makes a balance between exploration and exploitation of graph-based heuristic search space while keep the generality of the hyperheuristic approach.

In this study a new dynamic algorithm called Tolerable Graph-based Hyperheuristic (TGHH) is proposed with a new partial evaluation function and two embedded parameters; so that new partial evaluation function is designed to evaluate partial solution at each step in order to guide algorithm scheduling per exam with current heuristic is improving or destroying the solution. Good Tolerance parameter is introduced to control exploitation of heuristic search space and Bad Tolerance to balance exploration based on partial evaluation function value at each step.

The proposed algorithm has been tested on eight of benchmark datasets introduced by (Carter, Laporte and Lee, 1996). Different pair permutations of Tolerance parameters have been tuned in the algorithm and best pair is determined. The obtained results on five of the datasets are better than reported results by GHH presented in (Burke et al., 2007) and are in the range of published results by GHH on remained datasets. Obtaining solutions with less cost function implies previous results of other approaches were getting stuck in local optimum because now a solution has been achieved in another search space area with less violation of soft constraint that is closer to global optimum rather than previous results.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Master Sains

ALGORITMA KETAHANAN HIPERHEURISTIK BERASAKAN GRAF KONSTRUKTIF UNTUK PENJADUALAN PEPERIKSAAN

Oleh

SHAHRZAD MOHAMMAD POUR

Oktober 2009

Pengerusi: Abu Bakar Md Sultan, PhD

Fakulti: Sains Komputer Dan Teknologi Maklumat

Masalah penjadualan peperiksaan (ETTP) adalah masalah pengoptimuman NP-Hard tipikal yang dihadapi oleh institusi dan universiti seluruh dunia. Situasi ini mencetuskan kaedah heuristik yang merangkumi cabang yang luas kepada masalah ini. Selain itu masaalah penjadualan ini berbeza setiap institusi bergantung kepada saiz, struktur dan kekangan. Oleh itu ciri-ciri generaliti kaedah yang ditawarkan adalah merupakan matlamat utama kepada penyelesaian masalah penjadualan hari ini. Metod-metod ini cuba mengekalkan generaliti disebalik penambahan faktor-faktor ke atasnya. Hiperheuristik adalah salah satu kaedah sebegini yang menjadi asas utama tesis ini.

Terperangkap dalam optima awalan merupakan masalah utama pendekatan heuristik semenjak ianya diperkenalkan. Punca utamanya kerana algoritma berasaskan tempatan samada memokus kepada penerokaan (peningkatan global) atau pemecahan (peningkatan local) atau sebaliknya.

Proposal terhadap masalah di atas dalam tesis ini adalah lanjutan daripada Constructive Graph-based Hypherheuristc Framework oleh (Burke et al., 2007), dimana penyelidik tidak melihat kepada kacukan dinamik antara heuristik berasaskan graf dalam rangka kerja mereka seperti setiap heuristik aras bawah digunakan untuk menjadual bilangan peperiksaan tertentu bagi membina penyelesaian di setiap peringkat. Di sebahagian lain pula, tiada kewujudan kawalan kepada heuristik semasa atas penyelesaian yang mencetuskan peningkatan atau penghapusan penyelesaian. Melalui cara ini tiada kawalan sah antara penerokaan dan pemecahan ruang carian bagi mengelakkan terperangkap dalam optima awalan. Kajian ini bermatlamat menggunakan mekanisma dinamik yang membolehkan algoritma membuat imbangan antara penerokaan dan pemecahan ruang carian berasaskan graf semasa mengekalkan genelaliti kaedah hiperheuristik.

Dalam kajian ini algoritma dinamik baru dipanggil Algoritma Ketahanan Hiperheuristik berasaskan Graf (TGHH) diperkenalkan dengan satu fungsi penilaian separa dan dua umpukan parameter; jadi fungsi penilaian separa ini direkabentuk untuk menilai penyelesaian separa pada setiap peringkat bagi memandu algoritma penjadualan dengan heuristik semasa agar penyelesaian dipertingkatkan atau dihapuskan. Parameter Ketahanan baik diperkenalkan untuk mengawal pemecahan ruang carian heuristik dan ketahanan buruk pula mengimbang penerokaan berasaskan nilai fungsi penilaian separa di setiap peringkat.

Algoritma yang diperkenalkan telah diuji kepada lapan set data yang diperkenalkan (Carter, Laporte dan Lee, 1996). Pasangan permutasi kepada parameter-parameter kebolehtahanan ditala dalam algoritma dan keputusannya dibincangkan. Keputusan yang dilaporkan adalah dalam julat yang hampir sama diperolehi oleh algoritma terkini dan sebahagian keputusan dari set data adalah lebih baik dari yang dihasilkan oleh GHH seperti dilaporkan dalam (Burke et al., 2007) dan dalam julat keputusan oleh GHH bagi set data yang selebihnya. Dapatan dari penyelesaian ini dengan fungsi kos yang rendah menunjukan keputusan dari kaedah sebelum telah tersekat dalam masalah optima setempat kerana tiada penyelesaian dicapai dalam ruang carian yang lain dengan pelanggaran kekangan rendah yang kecil menghampiri optima global.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr Abu Bakar Md Sultan for his valuable comments and advice through the course of this research. His encouragement and professional review helped this thesis and other technical papers to be further improved.

My further gratitude goes to Associate Prof Md. Nasir Bin Sulaiman and Associate Prof. Dr Ramlan b Mahmud for their great help and technical advices.

Also, my eternal gratitude is owed to my family who have been supportive in everything I have done. In particular, I would like to thank my mother, Azam for her never ending love and support. I am highly indebted to my sister, Rabeah for her understanding, encouragement and support during my study.

I also want to thank of all my second family members in Malaysia, including all my friends for providing me with great friendship and experience in my academic and social life. Specially, I owe gratitude to my friend, Bahram for his impressive help in my thesis and his expressions when things seemed not to be in track.

Finally, thanks God for giving me another opportunity to know myself by living in Malaysia.

APPROVAL

I certify that an Examination Committee has met on **date of viva** to conduct the final examination of **Shahrzad Mohammad Pour** on her **Master of Science** thesis entitled " **TOLERABLE CONSTRUCTIVE GRAPH-BASED HYPER-HEURISTIC ALGORITHM FOR EXAMINATION TIMETABLING**" in accordance with Universiti Pertanian Malaysia (Higher Degree) Act 1980 and Universiti Pertanian Malaysia (Higher Degree) Regulations 1981. The Committee recommends that the candidate be awarded the relevant degree. Members of the Examination Committee are as follows:

Chairman, PhD
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Examiner 1, PhD
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Internal Examiner)

Examiner 2, PhD
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(External Examiner)

External Examiner, PhD
Faculty of Science and Technology
(External Examiner)

HASANAH MOHD GHAZALI, PhD
Professor/Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia
Date:

This thesis submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirements for the degree of Master of Science. Members of the Supervisory Committee were as follows:

Abu Bakar Md. Sultan, PhD
Senior Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Md. Nasir Bin Sulaiman, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

HASANAH MOHD GHAZALI, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 11 February 2010

DECLARATION

I hereby declare that the thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at UPM or other institutions.

____

Shahrzad Mohammad Pour
Date:

TABLE OF CONTENTS

				Page
ABSTRACT				iii
ABSTRAK				vi
ACKNOWLEDGMENT				ix
APPROVAL				x
DECLARATION				xii
LIST OF TABLES				xvi
LIST OF FIGURES				xvii
LIST OF ABRIVATIONS				xx
CHAPTER				
1	**INTRODUCTION**			1
2	**LITERATURE REVIEW**			11
	2.1	Introduction		11
	2.2	Timetabling Definitions		12
	2.3	Timetabling Problems		13
		2.3.1	Educational Timetabling	14
		2.3.2	Examination timetabling	16
	2.4	Graph-based Techniques		18
		2.4.1	Hybridisation with Graph heuristics	21
	2.5	Metaheuristic Approaches		23
	2.6	Hyperheuristics		25
		2.6.1	Hyperheuristic Framework	26
		2.6.2	Improvement and Constructive Techniques	30
		2.6.3	Single Hyperheuristics	33
		2.6.4	Hybrid hyperheuristic	36
		2.6.5	Graph-based Hyperheuristic (GHH)	38
	2.7	Strategies of Previous Works to Solve Local Optimum		43
		2.7.1	Local Search Based Techniques	43
		2.7.2	Population-Based Techniques:	45
		2.7.3	Metaheuristics	46
		2.7.4	Hyperheuristics	47
	2.8	Summary		50
3	**RESEARCH METHODOLOGY**			52
	3.1	Introduction		52
	3.2	Steps of Methodology		52
	3.3	STEP 1: Literature Review		54
	3.4	STEP 2: Design of Proposed Hyperheuristic Algorithm		55
		3.4.1	System Framework	56
		3.4.2	Key Issues in Partial Hyperheuristic Framework	57
		3.4.3	Low-Level Heuristics	59

		3.4.4 Selection and Apply Heuristics	60
	3.5	STEP 3: Implementation	63
		3.5.1 Datasets	63
		3.5.2 Evaluation Function	65
		3.5.3 Complete Evaluation Function	67
		3.5.4 Experimental Design	68
	3.6	STEP 4: Comparison Results	70
	3.7	Summary	70
4	**PROPOSED DYNAMIC HYPERHEURISTIC**		71
	4.1	Introduction	71
	4.2	Customized Partial Hyperheuristic Framework	72
	4.3	Definition of Tolerance Parameters	73
		4.3.1 Good Tolerance	73
		4.3.2 Bad Tolerance	74
	4.4	Implementation of Tolerance Parameters	75
		4.4.1 Good Tolerance Parameter	75
		4.4.2 Bad Tolerance Parameter	76
	4.5	Tolerance Parameters in Practice	77
	4.6	Tolerable Graph-based Hyperheuristic (TGHH) Algorithm	80
	4.7	Pseudo-code of TGHH	83
	4.8	Tackling Local Optimum during Balancing Process	84
	4.9	Balancing between Exploration and Exploitation	87
	4.10	Minimize Getting Stuck in Local Optimum	87
	4.11	Proposed Partial Evaluation Function	87
		4.11.1 Evaluating Partial Solution Based on Scheduled Exams	88
		4.11.2 Evaluating Partial Solution Based on Unscheduled Exams	89
		4.11.3 Design Final Evaluation Function Formula	95
	4.12	Summary	99
5	**RESULTS AND DISCUSSION**		101
	5.1	Introduction	101
	5.2	Investigation Effect of Good and Bad Tolerance Parameters	102
	5.3	Results of TGHH by Testing Different Permutations of the Parameters on Each Dataset	107
	5.4	Best Penalty and Average Penalty by TGHH	117
	5.5	Comparison with Other Approaches	118
	5.6	Overall Discussion	124
		5.6.1 Dynamicity of the TGHH	126
		5.6.2 Generality from Point of View of Hyperheuristic Framework	127
		5.6.3 Generality from Point of View of Proposed TGHH	127
	5.7	Summary	129
6	**CONCLUSION**		131

REFERENCES 137
BIODATA OF STUDENT 145
LIST OF PUBLICATIONS 146

LIST OF TABLES

Table		Page
2.1	Summary of Approaches Strategy to Avoid Local Optimum	48
2.2	Overall View of Dynamic Approaches for Solving ETTP	49
3.1	Key Issues in Design of Hyperheuristic Framework	57
3.2	Characteristics of Benchmark Exam Timetabling Problem Datasets	64
3.3	Experimental Design	68
5.1	Different Permutations of Good and Bad Tolerance Tested in TGHH algorithm for each Dataset	103
5.2	Summary of Effect of different Good and Bad Tolerance by TGHH on obtained Results	115
5.3	Effect of Different Values of Good and Bad Tolerances by TGHH	125

LIST OF FIGURES

Figure	Page
1.1 General Hyperheuristic Framework	4
1.2 Study Module	7
2.1 Classification of Metaheuristic Approaches	24
2.2 Hyperheuristic Framework	27
2.3 Issues for Designing and Developing of a Hyperheuristic Framework	29
2.4 General View of Improvement and Constructive Techniques in Hyperheuristic Framework	31
2.5 Hyperheuristic Framework When Dealing with Partial Solution or Complete Solution	32
2.6 Process of Selecting and Applying Heuristics in GHH	40
2.7 Mechanism of Generating Heuristic List in GHH	42
3.1 Steps of Methodology	53
3.2 General Hyperheuristic Framework with Partial Solutions	56
3.3 Examples of Graph low level heuristics	59
3.4 Difference between Complete and Partial Evaluation Function	66
4.1 Customized Partial hyper-heuristic framework	72
4.2 A Scenario of Proposed Algorithm when Good Tolerance is *5* and Bad Tolerance is *3*	78
4.3 Proposed Tolerable Hyperheuristic Black Box	82
4.4 Pseudo-code of TGHH	83
4.5 Tackling Local Optimum during Balancing Process	85
4.6 Pseudo-code for Calculating Number of Required Swaps in Partial Evaluation Function Based on Unscheduled Exams	93

4.7 Example of SD list and unscheduled exams list based on time Feasibility 94

4.8 Example of calculating *f(ec)* by bubble sort 95

4.9 Example of *K1* and *K2* coefficients in partial evaluation formula 97

5.1 Generating Different Permutations of Good and Bad Tolerances 103

5.2 Effect of increasing Good Tolerance on Average and Minimum cost on *hec92* 10

5.3 Effect of increasing Bad Tolerance on Average and Minimum cost on *hec92* 105

5.4 Average and Minimum Cost obtained by TGHH with 36 Pair Permutations of Good and Bad Tolerance Parameters on *hec92* Dataset 108

5.5 Average and Minimum Cost obtained by TGHH with 36 Pair Permutations of Good and Bad Tolerance Parameters on *ear83* Dataset 109

5.6 Average and Minimum Cost obtained by TGHH with 36 Pair Permutations of Good and Bad Tolerance Parameters on *kfu93* Dataset 110

5.7 Average and Minimum Cost obtained by TGHH with 36 Pair Permutations of Good and Bad Tolerance Parameters on *lse91* Dataset 111

5.8 Average and Minimum Cost obtained by TGHH with 36 Pair Permutations of Good and Bad Tolerance Parameters on *sta83* Dataset 112

5.9 Average and Minimum Cost obtained by TGHH with 36 Pair Permutations of Good and Bad Tolerance Parameters on *tre92* Dataset 113

5.10 Average and Minimum Cost obtained by TGHH with 36 Pair Permutations of Good and Bad Tolerance Parameters on *ute92* Dataset 114

5.11 Average and Minimum Cost obtained by TGHH with *36* Pair Permutations of Good and Bad Tolerance Parameters on *york83* Dataset 114

5.12 Best & Average Cost obtained by TGHH on benchmark Datasets 117

5.13 Comparing Best Obtained Result by TGHH with Other Approaches on *ear83* dataset 118

5.14 Comparing Best Obtained Result by TGHH with Other Approaches on *hec92* dataset 119

5.15 Comparing Best Obtained Result by TGHH with Other Approaches on *kfu93* dataset 119

5.16 Comparing Best Obtained Result by TGHH with Other Approaches on *lse91* dataset 120

5.17 Comparing Best Obtained Result by TGHH with Other Approaches on *sta83* dataset 120

5.18 Comparing Best Obtained Result by TGHH with Other Approaches on *tre92* dataset 121

5.19 Comparing Best Obtained Result by TGHH with Other Approaches on *ute92* dataset 121

5.20 Comparing Best Obtained Result by TGHH with Other Approaches on *york83* dataset 122

LIST OF ABBREVIATIONS

ETTP	Examination Timetabling Problem
NP	Nondeterministic Polynomial-time
GHH	Graph-based Hyper-Heuristic
CLP	Constraints Logic Programming
TS	Tabu Search
SA	Simulated Annealing
HC	Hill Climbing
EA	Evolutionary Algorithms
GA	Genetic Algorithm
GHH	Graph-based Hyperheuristic
TGHH	Tolerable Graph-based Hyperheuristic
SD	Saturation Degree
LD	Largest Degree
LWD	Largest Weighted Degree
LE	Largest Enrolment
CD	Color Degree
CBR	Case-Based Reasoning

CHAPTER 1

INTRODUCTION

1.1 Background

Examination timetabling problem (ETTP) as a subclass of educational timetabling is one the most famous problem which has taken a lot of efforts by researchers to solve it until now. At least once a year, schools and universities have to solve an instance of the timetabling problem whose manual solution requires a lot of manpower. It would be desirable to have a program that schedules courses and/or exams instead of a human.

On the other hand the problem varies from one institution to another depending on the size, constraints, type of the problem and their objectives. It can be inferred that the solution which is appropriate for one institution, may not work at others. Therefore the generality of the proposed methods is one of the major goals in solving timetabling problem.

On the other hand, many real-life problems lead naturally to combinatorial search which is a very computationally intensive task. Unfortunately, no general method exists for solving this kind of problems efficiently. The Automated construction of Examination timetables is a typical combinatorial optimization known as NP (Nondeterministic Polynomial-time) hard problem due to large-scale computationally, multi-constrained and belonging to combinatorial optimization. There is no linear exact method to solve the problems which fall under this category of combinatorial optimization. Of course constructing of an initial solution (timetable) is not problem, the issue is improvement of solutions and obtaining an optimum solution.

Due to NP-hard nature of ETTP and more generally educational timetabling problem, the heuristic methods cover a large branch of researches in this area. In computer science, a heuristic algorithm or simply a heuristic is an algorithm that ignores whether the solution to the problem can be proven to be correct, but which usually produces a good solution or solves a simpler problem that contains or intersects with the solution of the more complex problem. Heuristics are typically used when there is no known way to find an optimal solution, or when it is desirable to give up finding the optimal solution for an improvement in run time (Pearl and Judea, 1984).

In search algorithms two conflicting aspects are termed `exploration' and `exploitation'. Exploration is an algorithm's ability to search broadly through the problem's search space and exploitation is an algorithm's ability to search locally around good solutions that have been found previously. Proper control of global exploration and local exploitation is crucial in heuristic approaches in order to avoid local optimum.

The basic heuristic method is Hill Climbing (HC) or iterative improvement which repeatedly moves to a solution better than the current one until it finds a local optimum (i.e. a solution which is better than all others in its neighborhood). Since only improving moves are accepted, hill climbing tends to get stuck fairly in local optimum, which may be much worse than the global optimum. To overcome this, modern heuristics (or metaheuristics) are equipped with some way of getting away local optima. The idea is to accept a solution even if it is worse than the current one in order to find better solutions later in the search process. Of course the local optimum is not solved completely yet. The main cause for the local optimal problem in metaheuristics is that algorithms don't make a harmony on exploitation (local improvement) and exploration (global improvement) in search space of solutions.

Generally metaheuristic approaches generate good results. They are suitable when the goal is generating high quality solutions. On the other hand, they are problem-specific and tailor-made nature approaches so that if they are applied to another problem or even another instance of the same problem, lots of effort will be demanded for changing programming and implementation due to match them. Therefore applying of approaches which work at a higher level of generality in different kinds of problem will be justified.

The new generation of heuristic methods is hyperheuristic approaches introduced by (Burke et al., 2003). The development of hyper-heuristics is motivated by the goal of raising level of generality for automatically solving a range of problems.

Hyper-heuristics can be defined to be heuristics which choose between heuristics in order to solve a given optimization problem at a higher level. It means that they don't optimize solutions directly. They work by way of an operator (a low level heuristic). This places a hyper-heuristic at a higher level of abstraction and generality rather than most current studies of metaheuristics. A number of hyper-heuristics have been developed over the past few years (Cowling, Kendall and Soubeiga, 2000; Ayob and Kendall, 2003; Burke et al., 2007).

Figure 1.1 indicates the general hyper-heuristic framework introduced by (Soubeiga, 2003). Hyper-heuristics can be considered as black box systems, which take the problem instance and several low level heuristics as methods which can produce the result independent of the problem characteristics.

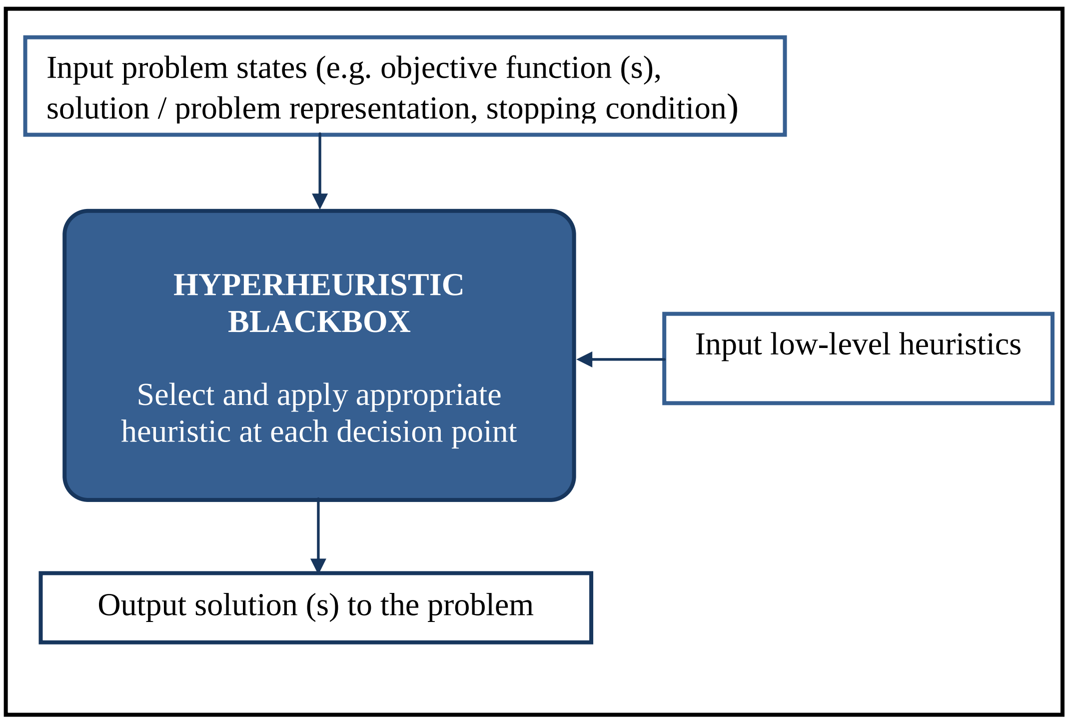

Figure 1.1: General Hyperheuristic Framework

In this concept, hyper-heuristics use only non problem-specific data provided by each low level heuristic in order to select and apply them to candidate solution (Burke et al., 2003).