FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE D
COIMBRA

Pedro Miguel Duque Rodrigues

Principled Modelling Of The Google Hash Code Problems For Meta-Heuristics

Dissertation in the context of the Master in Informatics Engineering, Specialization in Intelligent Systems, advised by Professor Alexandre B. Jesus and Professor Carlos M. Fonseca, and presented to the Faculty of Sciences and Technology / Department of Informatics Engineering.

September 2023

The work presented in this thesis was carried out in the Algorithms and Optimization Laboratory of the Adaptive Computation group of the Centre for Informatics and Systems of the University of Coimbra and financially supported by the Foundation for Science and Technology under a scholarship with reference UIDP/00326/2020.

© 2023 Pedro Rodrigues

Abstract

Acknowledgments

Contents

1 Introduction **1**

1.1 Motivation 1
1.2 Goals & Scope 3
1.3 Contributions 4
1.4 Software 5
1.5 Outline 5

Acronyms **6**

Bibliography **7**

List of Figures

1.1 Modelling and Problem Solving 2

List of Tables

Chapter 1

Introduction

"I've got a bad feeling about this..."

— Han Solo, *Star Wars*

1.1 Motivation

Optimization (O) problems are ubiquitous in real-world scenarios. Consider, for example, the task planning a road trip, which involves various optimization challenges. These range from selecting the best route considering factors like fuel efficiency, travel time, and budget limitations, to efficiently packing luggage. Solving these problems entails finding good feasible solutions, ideally in most time-efficient manner possible.

Tackling an optimization problem typically begins by understanding the problem and its details. Then, a strategy is employed to solve the problem. This highlights the two main steps: first, describing the problem's details to construct a representation (model), and second, using the gathered information to employ a strategy (solver) aimed at discovering feasible solutions.

In fact, this approach serves as the foundation for one of the most widely employed methods used to address optimization problems, wherein a standard mathematical formulation (model) describing the problem is developed. Subsequently, one of the numerous available generic solvers, *e.g.* "simplex", is utilized to find solutions. Notably, "simplex", being an exact solver, ensures the optimality of the obtained solutions.

Although this method for solving optimization problems is clearly established in the community there is recently growing interest in how to employ heuristic and meta-heuristic strategies to solve these problems, specifically, in the

Combinatorial Optimization (CO) field.

Heuristic (H) methods are a set of, often problem-specific, procedures that attempt to quickly solve a problem and provide a helpful "rule of thumb" for achieving reasonably good results, usually in a greedy fashion. Meta-Heuristic (MH) methods combine several subordinate heuristics thus being more generic and can be applied to broad a range of problems. Natural processes and phenomena, such as collective behavior, natural selection, and some physical properties of materials, inspire several meta-heuristic search processes making them flexible and adaptable, although more computationally intensive than more conventional heuristics. Remarkably, both strategies are generally more time-efficient than exact approaches, though without an associated guarantee of optimality.

Given the nature of Heuristic and Meta-Heuristic methods, and the inherent diversity of problems, crafting universal meta-heuristic solvers is challenging. Unlike mathematical optimization, there's no established modelling framework for meta-heuristics that follows the same principled approach, possibly due to skepticism about its feasibility and effectiveness.

The establishment of this framework would standardize problem-solving approaches, facilitate the reuse of meta-heuristic methods, and distinctly separate the tasks of problem modeling and solver development. Moreover, it would provide researchers and practitioners with a useful tool to experimentally assess the performance of Meta-Heuristic methods across a range of diverse problems.

Figure 1.1: Modelling and Problem Solving

In summary, the concept of a modeling framework for problem tackling can be visually depicted as shown in Figure 1.1. For a given problem, a computational model that describes the problem is generated and provided to a machine, which then employs a meta-heuristic solver to obtain solutions. This approach is built upon prior work [1, 3] that has initiated the formalization of this objective.

Simultaneously, alongside the development and application of meta-heuristic strategies to address combinatorial optimization problems, there exists a community interest in constructing a collection of benchmark optimization problems that hold both theoretical and practical significance. The Google Hash

Code Competition problems, arguably, present themselves as suitable candidates.

The Hash Code programming competition, formerly hosted annually by Google, challenged teams of up to four members to solve intricate problems within a four-hour time frame. Participants were allowed to use any tools, resources, and programming languages of their choice. These problems often drew inspiration from real-world issues and engineering challenges, such as vehicle routing, task scheduling, and Wi-Fi router placement. Essentially, can be classified as "open" research problems. Importantly, if not entirely, the majority of these problems can be classified as Combinatorial Optimization problems.

Given the pertinence of these problems, they serve as apt benchmarks for the evaluation of meta-heuristics. Moreover, they offer a suitable approach to assess the feasibility of the aforementioned principled modeling approach. Not only do these problems present challenges from theoretical and practical standpoints, but they are also have well written and detailed statements. The competition context in which they were presented also provides a wealth of empirical data regarding the best solutions attained thus far.

1.2 Goals & Scope

This thesis focuses on creating, implementing, and evaluating meta-heuristic approaches using a principled modelling framework. The Google HashCode Problems serve as benchmarks for our study.

Our approach encompasses two primary aspects. First, from a modeling perspective, we aim to expand upon the modelling concepts that have been partially explored in previous research [1, 2, 3]. Our primary objective is to solidify existing concepts while introducing additional functionality, both in conceptual understanding and practical application.

Additionally, we aim to construct models for the Hash Code problems. These models will not only be analyzed in this thesis but will also serve as illustrative examples for explaining and teaching the modeling concepts. Furthermore, they will enable a critical evaluation of the merits and shortcomings of this structured approach in comparison to more ad-hoc and traditional methods of problem-solving.

Secondly, the development of state-of-the-art meta-heuristic solvers. This development enables a thorough examination of the models and the overarching strategy. This assessment will encompass an analysis of both performance metrics and the quality of solutions achieved through this approach.

As such, the main research questions we outline for this thesis are:

R1. Can we formalize the existing ideas explored by previous work on the modelling framework and produce a practical implementation, potentially contributing with new features?

R2. Can we develop general-purpose meta-heuristic solvers and utilities with respect to the principled modelling framework implementation?

R3. Can Google Hash Code problems be solved using a modelling approach?

1.3 Contributions

The main contributions of this thesis related to the aforementioned research questions, are as follows:

C1. With the existing research on principled modeling for meta-heuristics [1, 2, 3], this document aims to consolidate and formalize a comprehensive specification. Our objective is to encapsulate all the concepts and developments made thus far. Additionally, we have created a practical Python implementation of this framework. In essence, both in the formalization and implementation, we endeavor to synthesize the existing ideas concerning modeling for constructive and local search techniques. These techniques are integral components of meta-heuristic solvers and from our perspective, they should be integrated into a model, as we will elaborate upon in this thesis.

C2. We developed several meta-heuristic solvers and utilities both for gathering the solutions and for testing the developed models. Given that these are general-purpose they can work with any model that is developed under the practical implementation of the framework we devised.

C3. After a careful selection process we selected two Google Hash Code problems for which some models for each of the problems were developed that explore the different properties of the problems in an attempt to both obtain the best solutions possible. Those models not only provide a practical example on how to model a problem but also explore the framework capabilities in great detail.

1.4 Software

The following software resulted from the development of this thesis and has been distributed under an open source license.

S1. Python Framework (to be continued...)

S2. Models and Experiments Code (to be continued...)

1.5 Outline

The remainder of this thesis is organized as follows...

Acronyms

AC Adaptive Computation. ii

ALGO Algorithms and Optimization Laboratory. ii

CISUC Centre for Informatics and Systems of the University of Coimbra. ii

CO Combinatorial Optimization. 2, 3

FCT Foundation for Science and Technology. ii

H Heuristic. 2

MH Meta-Heuristic. 2

O Optimization. 1

Bibliography

[1] Ana Vieira. "Uma plataforma para a avaliação experimental de meta-heurísticas". Doctoral Thesis. University of Algarve, Portugal, 2009. URL: https://sapientia.ualg.pt/handle/10400.1/518 (visited on 01/10/2023).

[2] Carlos M. Fonseca. *Nasf4nio*. Oct. 27, 2021. URL: https://github.com/cmfonseca/nasf4nio (visited on 01/15/2023).

[3] Samuel Barroca do Outeiro. "An Application Programming Interface for Constructive Search". Msc Thesis. University of Coimbra, Portugal, Nov. 9, 2021. URL: https://estudogeral.sib.uc.pt/handle/10316/98261 (visited on 01/10/2023).