Applied Mathematical Sciences 175

The Linear Ordering Problem

Exact and Heuristic Methods in Combinatorial Optimization

Springer

Applied Mathematical Sciences
Volume 175

For further volumes:
http://www.springer.com/series/34

Rafael Martí • Gerhard Reinelt

The Linear Ordering Problem

Exact and Heuristic Methods in Combinatorial Optimization

Springer

Prof. Dr. Rafael Martí
Dpto Estadística e Inv. Operativa
Universitat de València
Dr. Moliner 50
46100 Valencia
Spain
rafael.marti@uv.es

Prof. Dr. Gerhard Reinelt
Universität Heidelberg
Inst. Informatik
Im Neuenheimer Feld 326
69120 Heidelberg
Germany
gerhard.reinelt@informatik.uni-heidelberg.de

Editors:

S.S. Antman
Department of Mathematics
and
Institute for Physical
Science and Technology
University of Maryland
College Park, MD 20742-4015
USA
ssa@math.umd.edu

J.E. Marsden
Control and Dynamical Systems, 107-81
California Institute of Technology
Pasadena, CA 91125
USA
marsden@cds.caltech.edu

L. Sirovich
Laboratory of Applied Mathematics
Department of Biomathematical Sciences
Mount Sinai School of Medicine
New York, NY 10029-6574
isirovich@rockefeller.edu

ISSN 0066-5452
ISBN 978-3-642-16728-7 e-ISBN 978-3-642-16729-4
DOI 10.1007/978-3-642-16729-4
Springer Heidelberg Dordrecht London New York

Mathematics Subject Classification (2010): 90Cxx, 90C57, 90C59, 0C90

Cover design: deblik, Berlin

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To Amparo Rico and Ximo Seró, el padrino,
for showing me the way.

Rafa Martí

Preface

The idea for writing this book came up when the authors met at the University of Valencia in 2005. While comparing our experiences with regard to various aspects of the linear ordering problem (LOP), we realized that most of the optimization technologies had been successfully applied to solve this problem. We also found that there were only a small number of books covering all state-of-the-art optimization methods for hard optimization problems (especially considering both exact methods and heuristics together). We thought that the LOP would make an ideal example to survey these methods applied to one problem and felt the time was ripe to embark on the project of writing this monograph.

Faced with the challenge of solving hard optimization problems that abound in the real world, classical methods often encounter serious difficulties. Important applications in business, engineering or economics cannot be tackled by the solution methods that have been the predominant focus of academic research throughout the past three decades. Exact and heuristic approaches are dramatically changing our ability to solve problems of practical significance and are extending the frontier of problems that can be handled effectively. In this text we describe state-of-the-art optimization methods, both exact and heuristic, for the LOP. We actually employ the LOP to illustrate current optimization technologies and the design of successful implementations of exact and heuristic procedures. Therefore, we do not limit the scope of this book to the LOP but, on the contrary, we provide the reader with the background and strategies in optimization to tackle different combinatorial problems.

This monograph is devoted to the LOP, its origins, applications, instances and especially to methods for its effective approximate or exact solution. Our intention is to provide basic principles and fundamental ideas and reflect the state-of-the-art of heuristic and exact methods, thus allowing the reader to create his or her personal successful applications of the solution methods. The book is meant to be of interest for researchers and practitioners in computer science, mathematics, operations research, management science, industrial engineering, and economics. It can be used as a textbook on issues of practical optimization in a master's course or as a reference resource for engineering optimization algorithms.

To make the book accessible to a wider audience, it is to a large extent self-contained, providing the reader with the basic definitions and concepts in optimization. However, in order to limit the size of this monograph we have not included extensive introductions. Readers interested in further details are referred to appropriate textbooks such as [4, 84, 102, 117, 118, 124].

The structure of this book is as follows. Chapter 1 provides an introduction to the problem and its applications and describes the set of benchmark instances which we are using for our computational experiments and which have been made publically available. Chapter 2 describes such basic heuristic methods such as construction and local searches. Chapter 3 expands on Chapter 2 and covers meta-heuristics in which the simple methods are now embedded in complex solution algorithms based on different paradigms, such as evolution or learning strategies. Chapter 4 discusses branch-and-bound, the principal approach for solving difficult problems to optimality. A special version based on polyhedral combinatorics, branch-and-cut, is presented in Chapter 5. Chapter 6 deals in more detail with the linear ordering polytope which is at the core of branch-and-cut algorithms. The book concludes with Chapter 7, where a number of further aspects of the LOP and potential issues for further research are described.

Rafael Martí's research was partially supported by the Ministerio de Ciencia e Innovación of Spain (Grant Refs. TIN2006-02696 and TIN2009-07516).

We are in debt to many people, but in particular to some very good friends and colleagues who helped us to gain a deeper understanding of the linear ordering problem: Vicente Campos, Thomas, Christof, Angel Corberán, Carlos García, Fred Glover, Martin Grötschel, Michael Jünger, Manuel Laguna and Dionisio Pérez. The proofreading by Cara Cocking, Elena Fernández, Héctor Fraire, Marcus Oswald, Rodolfo Pazos, Hanna Seitz, Markus Speth and Pei Wang is particularly acknowledged. Finally, our special thanks go to Abraham Duarte, who implemented most of the methods described in Chapters 2 and 3 and performed some of the associated experiments.

Valencia, Heidelberg
October 2010

Rafael Martí
Gerhard Reinelt

Contents

Chapter 1
Introduction

Abstract The linear ordering problem (LOP) is one of the classical combinatorial optimization problems which was already classified as NP-hard in 1979 by Garey and Johnson [50]. It has received considerable attention in various application areas ranging from archeology and scheduling to economics and even mathematical psychology. Solution methods for the LOP have been proposed since 1958, when Chenery and Watanabe outlined some ideas on how to obtain solutions for this problem. The interest in this problem has continued over the years, resulting in the book [111] and many recent papers in scientific journals. This chapter surveys the main LOP applications and instances. We have compiled a comprehensive set of benchmark problems including all problem instances which have so far been used for conducting computational experiments. Furthermore we have included new instances. All of them form the new benchmark library `LOLIB`. We will use them in the next chapters to report our experiments with heuristics, meta-heuristics and exact approaches for the LOP.

1.1 Basic definitions

In its graph version the LOP is defined as follows. Let $D_n = (V_n, A_n)$ denote the complete digraph on n nodes, i.e., the directed graph with node set $V_n = \{1, 2, \ldots, n\}$ and the property that for every pair of nodes i and j there is an arc (i, j) from i to j and an arc (j, i) from j to i. A *tournament* (or *spanning tournament*) T in A_n consists of a subset of arcs containing for every pair of nodes i and j either arc (i, j) or arc (j, i), but not both. A (*spanning*) *acyclic tournament* is a tournament without directed cycles, i.e., not containing an arc set of the form $\{(v_1, v_2), (v_2, v_3), \ldots, (v_k, v_1)\}$ for some $k > 1$ and distinct nodes $v_1, v_2, \ldots, v_k$.

A *linear ordering* of the nodes $\{1, 2, \ldots, n\}$ is a ranking of the nodes given as linear sequence, or equivalently, as a permutation of the nodes. We denote the linear ordering that ranks node v_1 first, v_2 second, etc., and v_n last by $\langle v_1, v_2 \ldots, v_n \rangle$ and write $v_i \prec v_j$ if node v_i is ranked before node v_j. If σ denotes a linear

R. Martí and G. Reinelt, *The Linear Ordering Problem*, Exact and Heuristic Methods in Combinatorial Optimization 175, DOI: 10.1007/978-3-642-16729-4_1,

ordering, then $\sigma(i)$ gives the position of node i in this ordering. We will also consider *partial orderings* where only a subset of the nodes is ranked or only some pairs are compared.

It is easy to see that an acyclic tournament T in A_n corresponds to a linear ordering of the nodes of V_n and vice versa: the node ranked first is the one without entering arcs in T, the node ranked second is the one with one entering arc (namely from the node ranked first), etc., and the node ranked last is the one without leaving arcs in T.

Usually, ordering relations are weighted and we have weights c_{ij} giving the benefit or cost resulting when node i is ranked before node j or, equivalently, when the arc (i,j) is contained in the acyclic tournament. The (weighted) *linear ordering problem* is defined as follows.

Linear ordering problem

Given the complete directed graph $D_n = (V_n, A_n)$ with arc weights c_{ij} for every pair $i,j \in V_n$, compute a spanning acyclic tournament T in A_n such that $\sum_{(i,j)\in T} c_{ij}$ is as large as possible.

Alternatively, the LOP can be defined as a matrix problem, the so-called *triangulation problem.*

Triangulation problem

Let an (n,n)-matrix $H = (H_{ij})$ be given. Determine a simultaneous permutation of the rows and columns of H such that the sum of superdiagonal entries becomes as large as possible.

Obviously, by setting arc weights $c_{ij} = H_{ij}$ for the complete digraph D_n, the triangulation problem for H can be solved as a linear ordering problem in D_n. Conversely, a linear odering problem for D_n can be transformed to a triangulation problem for an (n,n)-matrix H by setting $H_{ij} = c_{ij}$ and the diagonal entries $H_{ii} = 0$.

Consider as an example the (5,5)-matrix

$$H = \begin{pmatrix} 0 & 16 & 11 & 15 & 7 \\ 21 & 0 & 14 & 15 & 9 \\ 26 & 23 & 0 & 26 & 12 \\ 22 & 22 & 11 & 0 & 13 \\ 30 & 28 & 25 & 24 & 0 \end{pmatrix}.$$

The sum of its superdiagonal elements is 138. An optimum triangulation is obtained if the original numbering $(1,2,3,4,5)$ of the rows and columns is changed to $(5,3,4,2,1)$, i.e., the original element H_{12} becomes element $H_{\sigma(1)\sigma(2)} = \tilde{H}_{54}$ in the permuted matrix. Thus the optimal triangulation of H is

$$\tilde{H} = \begin{pmatrix} 0 & 25 & 24 & 28 & 30 \\ 12 & 0 & 26 & 23 & 26 \\ 13 & 11 & 0 & 22 & 22 \\ 9 & 14 & 15 & 0 & 21 \\ 7 & 11 & 15 & 16 & 0 \end{pmatrix}.$$

Now the sum of superdiagonal elements is 247.

1.2 Applications of the Linear Ordering Problem

We review some of the many applications of the linear ordering problem.

1.2.1 Equivalent Graph Problems

The *acyclic subdigraph problem* (ASP) is defined as follows. Given a directed graph $D = (V, A)$ with arc weights d_{ij}, for all $(i, j) \in A$, determine a subset $B \subseteq A$ which contains no directed cycles and has maximum weight $d(B) = \sum_{(i,j)\in B} d_{ij}$.

It can easily be seen that this problem is equivalent to the LOP. For a given ASP define a LOP on D_n, where $n = |V|$, by setting for $1 \le i, j \le n$, $i \ne j$:

$$c_{ij} = \begin{cases} \max\{0, d_{ij}\}, & \text{if } (i, j) \in A, \\ 0, & \text{otherwise.} \end{cases}$$

If T is a tournament of maximum weight, then $B = \{(i, j) \in T \cap A \mid c_{ij} > 0\}$ is an acyclic subdigraph of D of maximum weight. In the opposite direction, by adding a suitably large constant, we can transform a given LOP into an equivalent one where all weights are strictly positive. Then an acyclic subdigraph of maximum weight is a tournament.

The *feedback arc set problem* (FBAP) in a weighted digraph $D = (V, A)$ consists of finding an arc set B of minimum weight such that $A \setminus B$ is acyclic, i.e., such that B is a so-called *feedback arc set* intersecting every dicycle of D. Obviously, FBAP and ASP are equivalent because they are complementary questions.

Fig. 1.1 shows a digraph on 9 nodes where the arcs of a minimum feedback arc set are drawn as dotted lines. If the six arcs of the feedback arc set are removed, we obtain an acyclic arc set.

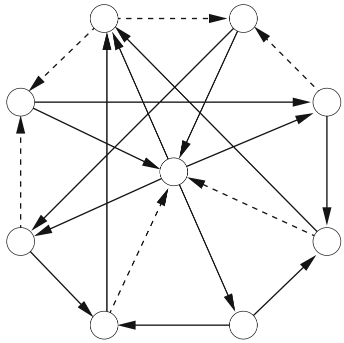

Fig. 1.1 A digraph with minimum feedback arc set

1.2.2 Related Graph Problems

There are some further problems dealing with acyclic subdigraphs. The *node induced acyclic subdigraph problem* asks for a node set $W \subseteq V$ such that the subdigraph $(W, A(W))$ is acyclic. (Here $A(W)$ denotes the set of arcs with both end nodes in W.) The problem can be defined either with node weights d, and $d(W)$ is to be maximized, or with arc weights c where $c(A(W))$ has to be maximum. Analogously, the *feedback node set problem* is to find a set $W \subseteq V$ such that $(V \setminus W, A(V \setminus W))$ is acyclic. Here, sums of node weights or arc weights have to be minimized.

The request that solution digraphs have to be node induced adds a further complexity. These problems cannot be transformed to a pure linear ordering problem and are even more difficult.

1.2.3 Aggregation of Individual Preferences

Linear ordering problems may occur whenever rankings of some objects are to be determined. Consider for example the following situation. A set of n objects $O_1, O_2, \ldots, O_n$ is given which have to be rated by m persons according to their individual preferences. Then a ranking of these objects is to be found which reflects these single rankings as closely as possible. The first question to be answered is how the individual rankings can be obtained. One solution is a pairwise comparison experiment. For any pair O_i and O_j, $1 \leq i < j \leq n$, of objects each person decides whether O_i should be preferred to O_j or vice versa. The results of these $m\binom{n}{2}$ comparisons are stored in an (n,n)-matrix $H = (H_{ij})$ where H_{ij} = number of persons preferring object O_i to object O_j. A ranking of these objects which infers as few contradictions to the individual rankings as possible can be obtained by triangulating H. It should be remarked that there are various statistical methods to aggregate single preference relations to one relation.

This area of application is the oldest one of the LOP. In 1959 Kemeny [77] posed the following problem (*Kemeny's problem*). Suppose that there are m persons and each person i, $i = 1,\ldots,m$, has ranked n objects by giving a linear ordering T_i of the objects. Which common linear ordering aggregates the individual orderings in the best possible way? We can solve this problem as a linear ordering problem by setting c_{ij} = number of persons preferring object O_i to object O_j. Note that this is basically the problem stated above, but this time the relative ranking of the objects by each single person is consistent (which is not assumed above).

Slater [119], in 1961, asked for the minimum number of arcs that have to be reversed to convert a given tournament T into an acyclic tournament. In the context of preferences, the input now is a collection of rankings for all pairs i and j of objects stating whether i should be preferred to j or vice versa and the problem is to find a the maximum number of pairwise rankings without contradiction. Also *Slater's problem* can also be solved as a LOP, namely by setting

$$c_{ij} = \begin{cases} 1, & \text{if } (i,j) \in T, \\ 0, & \text{otherwise.} \end{cases}$$

Questions of this type naturally occur in the context of voting (How should a fair distribution of seats to parties be computed from the votes of the electors?) and have already been studied in the 18th century by Condorcet [37].

1.2.4 Binary Choice Probabilities

Let S_n denote the set of all permutations of $\{1,2,\ldots,n\}$ and let P be a probability distribution on S_n.

Define the *induced (binary choice) probability system* p for $\{1,2,\ldots,n\}$ as the mapping $p : \{1,2\ldots,n\} \times \{1,2\ldots,n\} \setminus \{(i,i) \mid i = 1,2,\ldots,n\} \to [0,1]$ where

$$p(i,j) = \sum_{S \in S_n,\ i \prec j \text{ in } S} P(S).$$

The question of whether a given vector p is a vector of binary choice probabilities according to this definition is of great importance in mathematical psychology and the theory of social choice (see [48] for a survey).

In fact, the set of binary choice vectors is exactly the linear ordering polytope which will play a prominent role later in this book.

1.2.5 Triangulation of Input-Output Tables

One field of practical importance in economics is *input-output analysis*. It was pioneered by Leontief [88, 89] who was awarded the Nobel Prize in 1973 for his fundamental achievements. The central component of input-output analysis is the so-called *input-output table* which represents the dependencies between the different branches of an economy. To make up an input-output table the economy of a country is divided into sectors, each representing a special branch of the economy. An input-output table shows the transactions between the single sectors in a certain year. To be comparable with each other all amounts are given in monetary values. Input-output analysis is used for forecasting the development of industries and for structural planning (see [69] for an introductory survey).

Triangulation is a means for a descriptive analysis of the transactions between the sectors. In a simple model of production structure the flow of goods begins in sectors producing raw material, then sectors of manufacturing follow, and in the last stage goods for consumption and investments are produced. A real economy, of course, does not show such a strict linearity in the interindustrial connections, here there are flows between almost any sectors. Nevertheless it can be observed that the main stream of flows indeed goes from primary stage sectors via the manufacturing sectors to the sectors of final demand. Triangulation is a method for determining a hierarchy of all sectors such that the amount of flow incompatible with this hierarchy (i.e., from sectors ranked lower to sectors ranked higher) is as small as possible. Such rankings allow interpretations of the industrial structure of a country and comparisons between different countries.

1.2.6 Optimal Weighted Ancestry Relationships

This application from anthropology has been published in [56]. Consider a cemetery consisting of many individual gravesites. Every gravesite contains artifacts made of different pottery types. As gravesites sink over the years and are reused, it is a reasonable assumption that the depth of a pottery type is related to its age. So every gravesite gives a partial ordering of the pottery types contained in it. These partial orderings may not be consistent in the sense that pairs of pottery types may be ranked differently depending on the gravesite. The task of computing a global ordering with as few contradictions as possible amounts to solving a linear ordering problem in the complete directed graph where the nodes correspond to the pottery types and the arc weights are aggregations of the individual partial orderings. In [56] several possibilities for assigning arc weights are discussed and a simple heuristic for deriving an ordering is presented.

1.2.7 Ranking in Sports Tournaments

In many soccer leagues each team plays each other team twice. The winner of a match gets three points, in case of a tie both teams get one point. In the standard procedure, the final ranking of the teams in the championship is made up by adding these points and breaking ties by considering the goals scored. Another