A decade of combinatorial optimization

Citation for published version (APA):

Document status and date:
Published: 01/01/1997

Publisher Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
Memorandum COSOR 97-09

A decade of combinatorial optimization

K.I. Aardal
S. van Hoesel
J.K. Lenstra
L. Stougie

Eindhoven, May 1997
The Netherlands
A Decade of Combinatorial Optimization

Karen Aardal
Utrecht University

Stan van Hoesel
University of Maastricht

Jan Karel Lenstra
Eindhoven University of Technology; CWI, Amsterdam

Leen Stougie
University of Amsterdam

Abstract
This paper offers a brief overview of the developments in combinatorial optimization during the past decade. We discuss improvements in polynomial-time algorithms for problems on graphs and networks, and review the methodological and computational progress in linear and integer optimization. Some of the more prominent software packages in these areas are mentioned. With respect to obtaining approximate solutions to NP-hard problems, we survey recent positive and negative results on polynomial-time approximability and summarize the advances in local search.

1 Introduction

Combinatorial optimization is involved with models and methods for optimization over discrete choices. It is rooted in the theory of linear programming, and has strong links with discrete mathematics, probability theory, algorithmic computer science, and complexity theory. Some problems in the area are relatively well understood and admit solution to optimality in polynomial time. Many others are NP-hard, and one is forced to go one of three ways. Either one chooses an enumerative method that is guaranteed to produce an optimal solution. Or one applies an approximation algorithm that runs in polynomial time. Or one resorts to some type of heuristic search technique, without any a priori guarantee in terms of solution quality or running time.

In the past decade we have seen significant progress on all these fronts. Network flow algorithms became more efficient, and so did algorithms for linear and convex optimization. For the hard problems, advances in polyhedral techniques extended the realm of true optimization methods. Performance bounds that can – or probably cannot – be met in polynomial time were tightened. And there has been a surge in the development of local search approaches. Remarkable aspects were the use of randomization in the design and analysis of algorithms, and the attention paid to on-line planning models.

Two developments outside the area stimulated research in combinatorial optimization. First, the continued increase in computing power strengthened the need for efficient algorithms. The ability to handle bigger problems made the distinction
between low and high order running times more pronounced and diminished the power of brute force. Second, at the application side, there has been an increasing confidence in the practical potential of optimization techniques. Large and difficult real-world problems that were out of reach ten years ago are now being solved. Notable examples occurred in airline crew scheduling, train timetabling, time-constrained vehicle routing, telecommunication network design, frequency allocation, VLSI layout synthesis, and statistical disclosure control.

Combinatorial optimization has established itself as a mature discipline of scientific interest and practical relevance. The selection of topics and references presented below has been governed by space constraints and personal bias. We apologize for all omissions.

2 Graphs and networks

Major improvements in running times of algorithms for specific graph related problems have been obtained during the last ten years. We highlight a few here. Quite a number of these results have been obtained via randomized algorithms, for which we refer to the book by Motwani and Raghavan [39].

Much research on designing faster algorithms for the maximum flow problem and the minimum cost flow problem was initiated by the work of Tardos [49], who found the first strongly polynomial algorithm for the minimum cost flow problem. Scaling of the input parameters and prefixing flows are the main ingredients of most of these new algorithms, but the design of efficient data structures has also had an important impact. For a network with \(n\) nodes and \(m\) arcs, the best known strongly polynomial algorithms for finding a maximum flow and a minimum cost flow have running times \(O(mn \log n)\) and \(O(m \log n(m + n \log n))\), respectively. A thorough treatment of these results is given by Ahuja et al. [6]. This book also describes improvements in solution times for problems such as shortest path and matching problems.

The minimum cut problem of finding a minimum weight set of arcs in a network whose removal would disconnect the network is dual to the maximum flow problem. Recently, new algorithms have been developed for this problem that do not exploit this duality. Nagamochi and Ibaraki [40], for instance, use edge contraction in their algorithm. Randomized edge contraction, introduced by Karger and Stein [32], leads to the fastest algorithm so far. An overview of these algorithms with a computational study is given by Chekuri et al. [17].

Interesting results have been obtained in determining polynomially solvable subclasses of generally NP-hard problems. Robertson and Seymour [45] proved an old conjecture of Wagner: for each set of graphs that is closed under taking minors, there exists a finite set of graphs that are forbidden to be minors of any graph in the set. This obstruction set can be enormously large, but its finiteness allows Robertson and Seymour to prove the existence of a polynomial-time algorithm for determining the tree-width of any graph in a class that is known to contain graphs of bounded tree-width only. The proof is non-constructive, and the algorithm may involve a large constant (of the order of the number of forbidden minors). Bodlaender [15] gives a good overview of the techniques involved together with some applications.

Next to these classical graph problems, there are several problems in planning and
scheduling that can be viewed as problems on graphs. One well-studied problem of this kind is the uncapacitated lot sizing problem. In 1958 Wagner and Whitin proposed an $O(T^2)$ dynamic programming algorithm, where $T$ is the number of time periods in the planning horizon. It lasted more than thirty years before a better algorithm was found. In the early 1990's three groups simultaneously developed algorithms with running time $O(T\log T)$; see [5], [24], [52].

Important for the implementation of graph related algorithms is the availability of software packages. The most prominent software library is LEDA, A Library of Efficient Datatypes and Algorithms, developed by Melhorn and Näher [38]. It is implemented by a C++ class library, and incorporates many efficient data structures and algorithms. LEDA is available at ftp://ftp.mpi-sb.mpg.de in directory pub/LEDA.

3 Linear optimization

The main developments in linear optimization have sprouted from the work of Karmarkar [33], who started a wave of research on so-called interior point methods. Both theoretical and practical advances were accomplished over the past ten years, and by now some interior point methods are competitive with the celebrated simplex method. An interesting overview and discussion of the use of simplex and interior point methods can be found in the ORSA Journal on Computing 6.1 (1994). The book by Roos et al. [46] gives a comprehensive treatment of interior point methods for linear optimization. Interior point methods have also been developed for convex optimization problems. The application of interior point type methods to semidefinite optimization has led to results that have proved particularly useful in the design of approximation algorithms for certain combinatorial optimization problems; see Section 5.

A new line of research is the development of randomized algorithms for the search of an optimal basic feasible solution. The main open question here is if there exist randomized algorithms that solve linear optimization problems in strongly polynomial expected running time. Though this question has not been resolved yet, major steps have been taken. The fastest randomized algorithm is due to Kalai [31], and has expected running time $O(n^2m + b\sqrt{n\log n}\log m)$, where $n$ is the number of variables, $m$ the number of constraints, and $b$ a constant independent of the input. It is in essence a randomized simplex algorithm. For a review of research in this direction we refer to Chapter 9 of the book by Motwani and Raghavan [39].

With respect to deterministic simplex algorithms, many improvements in practical performance have been achieved. Many of these improvements have been implemented in the state-of-the-art software package CPLEX [21]. CPLEX also contains an interior point method.

To enhance user-friendliness of software for linear and integer optimization, modeling languages that allow for representation of variables and constraints in a set-based format are very useful. Leading computer packages for modeling are AMPL [25] and AIMMS [14].
4 Integer optimization

The most commonly used technique for solving (mixed) integer programs is still branch-and-bound. The quality of the available upper and lower bounds on the optimal value of the considered instance is the decisive factor for success of this tree search technique. A lower bound on the optimal value (assuming a minimization problem) is obtained from a relaxation of the integer program. In the past ten to fifteen years attention has shifted from Lagrangian relaxation to linear programming relaxation, since the latter type of relaxation can be strengthened more easily by using cutting planes. Combining cutting planes and Lagrangian relaxation usually causes convergence problems. Moreover, good LP solvers, such as CPLEX, that allow for addition of rows are nowadays available.

The theory of cutting planes in the form of valid inequalities that define facets of the convex hull of feasible solutions to an instance, was mainly developed prior to this past decade. During the past ten years, however, an enormous amount of more problem specific results have been obtained. Moreover, surprisingly large instances have been solved using a mixture of cutting plane algorithms and branch-and-bound. For recent surveys we refer to Aardal and Van Hoesel [1], [2], and to Chapter 3 of [22]. Similar developments have been attained for column generation methods, which can be viewed as dual to cutting plane techniques. For a survey we refer to Barnhart et al. [11].

A new development of the last decade is the theoretical quality analysis of cutting planes. Negative results for some classes of cutting planes have been reported by Goemans [26]. He evaluated the worst-case improvement resulting from adding several of the known classes of facets for the traveling salesman polytope to the subtour polyhedron, i.e., the set of vectors satisfying the so-called subtour elimination constraints.

Another surprising theoretical result in polyhedral combinatorics is due to Lovász and Schrijver [37], who developed an algorithm for obtaining a sequence of tighter and tighter relaxations of integer 0-1 programs. The algorithm iterates the following steps. First, each constraint of the considered problem is multiplied by each variable \( x_j \) and its complement \( 1 - x_j \) \((j = 1, \ldots, n)\). The resulting quadratic program is then linearized by replacing the nonlinear terms \( x_j x_j \) by new variables \( y_{ij} \). This linear formulation is finally projected onto the space of the original variables. Lovász and Schrijver showed that this procedure needs to be repeated at most \( n \) times before the convex hull of feasible solutions is obtained. Balas et al. [10] showed that it is sufficient to multiply each constraint by a single variable \( x_j \) and its complement at a time.

For branch-and-bound algorithms powerful and quite flexible software packages have been developed. We mention MINTO [47] and ABACUS [50]. MINTO contains more tools such as preprocessing and generic valid inequalities, whereas ABACUS has the advantage that it is written in C++.

Apart from the further development of existing solution techniques, also two new techniques for integer optimization received much attention in the last decade. The first algorithm we mention, developed by H.W. Lenstra [35], is older than ten years, but served as an inspiration for further developments. Lenstra's algorithm was developed to show that the problem of determining whether the polyhedron \( K = \{ x \in \mathbb{R}^n : Ax \leq b \} \) contains an integer vector \( z \), can be solved in polynomial time if \( n \) is fixed. First, the
algorithm finds a transformation \( \tau \) such that the polyhedron \( \tau K \) has a “spherical” appearance. If the basis of the lattice \( \tau \mathbb{Z}^n \) has short and near-orthogonal vectors, then the membership problem can be solved recursively by branching on a number of parallel hyperplanes. The number of such hyperplanes can be proved to be bounded by a constant depending only on \( n \). For any lattice such a basis exists and can be found in polynomial time starting from an arbitrary basis by using basis reduction; see Lenstra et al. [34]. Lovász and Scarf [36] designed a “generalized” basis reduction algorithm, which works directly on the polyhedron instead of using approximations such as Lenstra does. The advantage of their method is that less information is lost, the disadvantage is that it uses considerably more computational steps. Cook et al. [20] implemented the Lovász-Scarf algorithm and solved some previously unsolved integer programming problems. Barvinok [12] generalized Lenstra’s result and proved that the number of integer points in a polyhedron can be counted in polynomial time if the dimension is fixed.

Another new technique, based on the theory of Gröbner bases, was already known in computational algebraic geometry, and was introduced for solving integer optimization problems by Conti and Traverso [19]. It amounts to translating the integer programming problem into an algebraic membership problem. The Gröbner bases are used to guide the generalized division that decides the membership. Advances in applicability of these methods are due mainly to Thomas [51]. Their current practical power is restricted by the size of the Gröbner bases, which is large for most problems. Due to their structure such methods have advantages over other more conventional IP methods in solving stochastic integer programming problems; see Schultz et al. [48]. Computer packages for computing Gröbner bases are available, e.g., CoCoa [16] and MACAULAY [13].

5 Polynomial-time approximation

As an alternative to solving NP-hard combinatorial optimization problems to optimality, which may be very time consuming, a stream of research has concentrated on designing polynomial-time algorithms that aim at good approximations for such problems. A widely accepted quality measure of such approximations is the performance guarantee, i.e., an upper bound on the ratio between the approximate solution value and the optimal one. A comprehensive and up-to-date survey of the theory of approximation algorithms is provided in the book edited by Hochbaum [28].

Some of the major achievements in this field are based on a combination of relaxation and randomization. Goemans and Williamson (see [27] and Chapter 11 of [28]) designed approximation algorithms that solve appropriately chosen relaxations of mathematical programming formulations of the considered combinatorial problems, and then round the obtained solution in a randomized way. The rounding can be derandomized yielding deterministic approximation algorithms. In particular, Goemans and Williamson use semidefinite optimization relaxations to design algorithms with very good performance guarantees for the problem of finding a maximum cardinality cut in a graph and the problem of finding the maximum number of simultaneously satisfiable clauses in a Boolean expression with at most two literals per clause.
A remarkable result was obtained by Arora [8]. He developed a polynomial-time approximation scheme for the traveling salesman problem (TSP) in the Euclidean space. Here we notice that Christofides' algorithm of 1976 [18], with its performance guarantee of 3/2, is still the best polynomial approximation algorithm for the TSP whose distances are symmetric and satisfy the triangle inequality.

Apart from the above positive sounds on approximation, there has also been a breakthrough on the negative side, in the sense of non-approximability of optimal solutions of some problems. Papadimitriou and Yannakakis [42] defined a class of maximization problems for the purpose of distinguishing problems whose optimal solutions are hard to approximate within arbitrarily small ratio. This class called MAXSNP has a two-sided polynomial reduction defined on it under which it is closed. Given a Boolean expression in conjunctive normal form, the problem MAXSAT of finding a truth assignment to the variables that satisfies the maximum number of clauses is complete for this class.

Arora et al. [9] gave a strong justification for investigating these concepts. They showed that there cannot exist a polynomial-time approximation scheme for MAXSAT unless $P = NP$. The proof is based on an alternative definition of NP in terms of randomized certificate verification based on fingerprinting methodology. This important result implies that for any MAXSNP-complete problem there must be a threshold value strictly greater than 1 on the achievable polynomial-time performance guarantee. For an overview of specific results in this direction, we refer to Chapter 10 of [28].

In sequencing and scheduling, techniques based on linear programming and rounding led to surprising performance guarantees for the off-line and on-line minimization of total (weighted) completion time on a single machine and on parallel machines, and for the minimization of makespan on parallel machines subject to communication delays. An investigation of the complexity of finding very short schedules yielded lower bounds on the polynomial-time approximability of several scheduling problems, including the job shop scheduling problem. For specific results and references, we refer to Chapter 12 of [22].

The previous paragraphs concerned the worst-case approach to approximation. A complementary approach is average-case or probabilistic analysis, a research field that started more than twenty years ago. The main developments in this field during the last decade were based on discovering the possibility to exploit existing results from probability theory. Empirical process theory provided tools for the analysis of the optimal solution value of a series of number problems; see Piersma [43]. Martingale theory allowed for relatively elegant asymptotic characterizations of optimal solution values of several problems; see Rhee and Talagrand [44]. Finally we mention the rather complete probabilistic analysis of bin-packing algorithms, presented in Chapter 2 of [28].

Next to these developments for optimization problems, a breakthrough in approximation was accomplished for counting problems, again based on randomization. Counting combinatorial structures such as the number of Hamiltonian cycles in a graph is obviously harder than just deciding on the presence of the structure. Jerrum et al. [30] showed the equivalence between approximate counting and approximate sampling for a wide class of combinatorial structures. Building on work by Aldous [7], they use Markov chains to simulate random (uniform) sampling of the structures, and proved
that these "mix rapidly". As a first result Jerrum and Sinclair [29] devised a fully polynomial randomized approximation scheme (FPRAS) for counting perfect matchings in dense graphs, whose vertices have degree at least half of the total number of vertices. The non-dense graph case is still open. Another prominent result in this direction is an FPRAS for computing volumes of convex bodies by Dyer et al. [23]. A series of subsequent papers have given schemes with increasingly better running times. For an overview we refer to Chapter 12 of [28].

6 Local search

For many years heuristic search approaches have been used throughout science and engineering. Their performance was generally considered to be satisfactory, partly based on experience, partly based on a belief in some physical or biological analogy, which was not always supported by familiarity with what has been achieved in mathematics.

Still, in the past decade local search has reinforced its position as a standard approach in combinatorial optimization. Problem size or lack of analytical insight may prohibit the application of true optimization algorithms. Polynomial-time approximation algorithms may give inferior solutions, and their performance bounds, if they can be obtained at all, may be meaningless in practice. Local search is a robust way to obtain good solutions to real problems in reasonable time.

Simulated annealing has established itself as a relatively straightforward technique that performs very well when given enough time. Tabu search requires more tuning but often less running time. Genetic algorithms are not known to perform well in a pure form, even when problem solutions allow a natural string representation, but hybrid forms in which offspring are subjected to iterative improvement are promising. Neural networks have many applications, which, however, seem to fall outside the realm of optimization.

Many aspects of local search are discussed in the book edited by Aarts and Lenstra [3]. We see three main lines of advance. First, a theory of the computational complexity and performance analysis of local search is now emerging; see Chapters 2 and 3 of [3]. Second, neighborhoods embodying problem-specific knowledge and data structures supporting incremental computations are being used in rather sophisticated implementations. Third, some of the more successful search strategies are hybrids, which combine local search with a constructive method, with tree search or, again, with local search. The shifting bottleneck procedure for job shop scheduling of Adams et al. [4] is a constructive rule that reoptimizes partial schedules along the way. The shop scheduling algorithms of Nowicki and Smutnicki (see, e.g., [41] and Chapter 11 of [3]) apply tabu search and jump back to previously considered promising but rejected moves; many other combinations of local search and tree search have been proposed. Johnson's iterated Lin-Kernighan algorithm for the TSP (see Chapter 8 of [3]) is a nested form of local search, which applies 4-exchanges to local optima resulting from variable-depth search.
References


