In Pursuit of the Traveling Salesman: Mathematics at the Limits of Computation

A Review by Jan Karel Lenstra and David Shmoys

In Pursuit of the Traveling Salesman: Mathematics at the Limits of Computation
William J. Cook
Princeton University Press, 2012, xvi+228 pages
US$27.95 (hardcover)
ISBN 978-0-691-15270-7

A traveling salesman is faced with the problem of finding the shortest route that takes him through each of a number of cities and back home again. Being easy to state and hard to solve, and having a popular name and many applications, the traveling salesman problem has become representative of the field of combinatorial optimization. It embodies the quest for finding the best alternative out of a finite but huge number of possibilities. For all of these attributes, it is the perfect vehicle to attract a broader audience to one of the more dynamic areas of mathematics research over the past fifty years; this is the goal of the book under review.

Jan Karel Lenstra is emeritus professor of optimization at Eindhoven University of Technology and former director of Centrum Wiskunde & Informatica in Amsterdam. His email address is Jan.Karel.Lenstra@cwi.nl

David Shmoys is Laibe/Acheson Professor of Business Management and Leadership Studies and director of the School of Operations Research and Information Engineering, Cornell University. His email address is shmoys@cs.cornell.edu.

With Eugene Lawler and Alexander Rinnooy Kan, they edited the book The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization *(Wiley, 1985).*
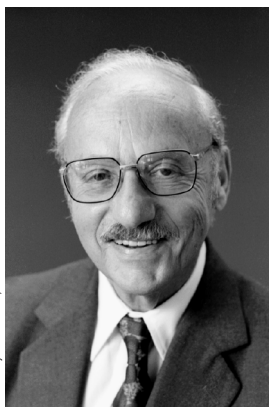
The traveling salesman problem, or TSP for short, poses both a complexity question and a computational challenge. The complexity question is a fundamental one: Can an optimal solution be found in polynomial time or, equivalently, is P equal to NP? The answer to this question is worth eternal fame and, as a side benefit, one million dollars. The computational challenge is one of algorithm development and engineering: As long as we have to assume that P is not equal to NP, how well can we do in finding good or even provably optimal solutions to large instances of the TSP?

The TSP as a mathematical problem was discussed by Karl Menger in Vienna and by Hassler Whitney in Princeton in the 1930s. The 1940s brought a practical interest in solving optimization problems that occurred in a logistical or industrial setting. George Dantzig proposed the linear optimization model for such problems and developed the simplex method for their solution. With this major advance came the realization of its limitations. Many decision problems of a combinatorial nature, for example, can be cast in linear terms with the additional constraint that the variables can take on only integral values. The TSP is a case in point. In fact, many of the concepts and techniques of combinatorial optimization were originally conceived for the TSP.

> *The TSP poses both a complexity question and a computational challenge.*

The algorithmic history of the TSP started with the seminal 1954 paper of (top to bottom) George B. Dantzig, Delbert Ray Fulkerson, and Selmer M. Johnson.

The algorithmic history of the TSP starts in 1954 with a seminal paper by Dantzig, Fulkerson, and Johnson. They solved a problem defined on forty-two cities in the US to optimality essentially by solving a series of more and more constrained linear optimization problems.

For each link between two cities, a 0-1 variable indicates whether it is chosen to be included in the tour or not. The constraints say that each city must be incident to two chosen links and that those links together must form a connected tour. The objective measures the total length of the links chosen. The authors start by solving a relaxed problem without the integrality and connectivity constraints, which yields a solution that is not a tour. They then identify by hand a violated connectivity constraint, add it to the problem, and solve again. By adding seven such constraints and two ad hoc constraints to ensure integrality, they obtained an optimal solution, a truly remarkable feat in view of the exponential number of constraints they could choose from. The added constraints are called *cutting planes*, since they are linear inequalities that separate an infeasible optimum from the feasible region.

In spite of this success, solution approaches to the TSP took a different course, that of implicit enumeration in the style of the backtrack methods that were proposed at the time for many combinatorial search problems. Eastman gave the first enumerative method for the TSP in his PhD thes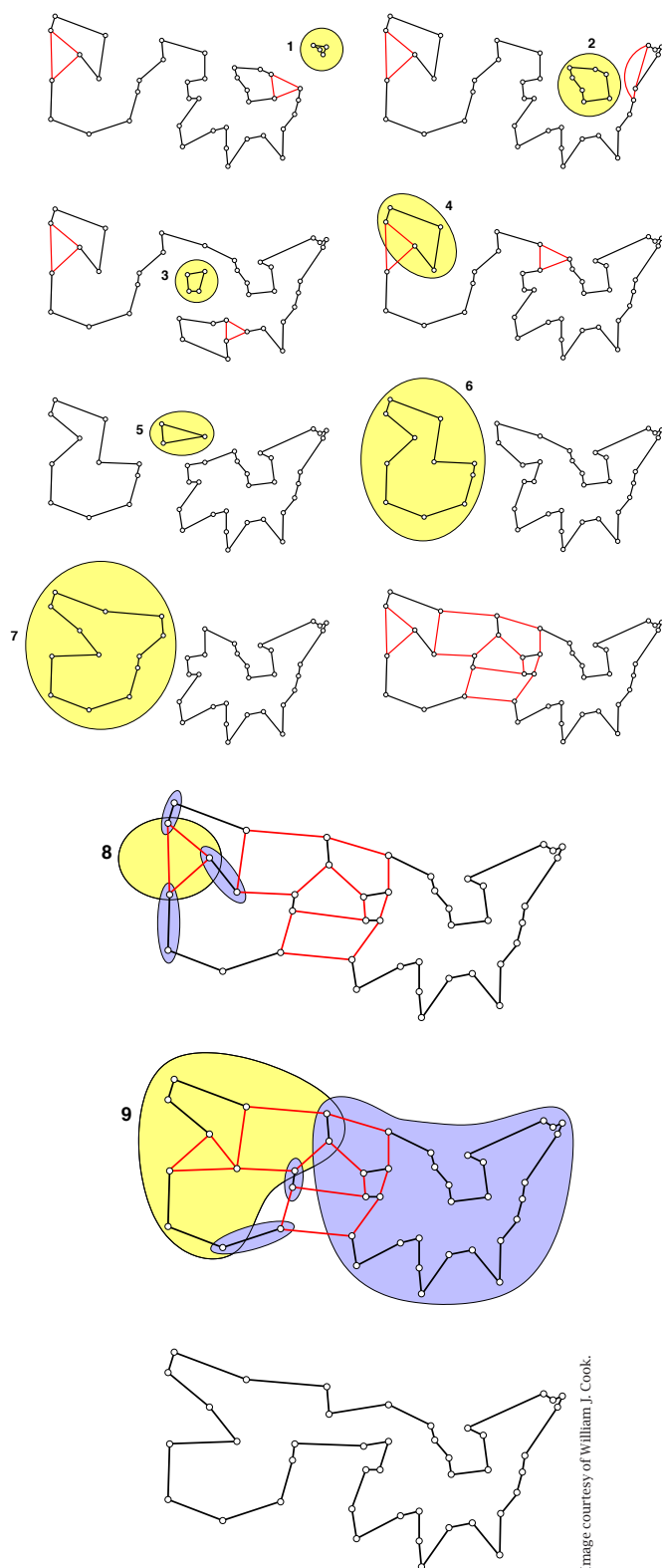is in 1958. The approach was made popular in 1963 by Little, Murty, Sweeney, and Karel, who coined the term *branch-and-bound*. The early methods of this type employ weak bounds and rely on branching to locate the optimum tour. Computational success was limited to problem instances with twenty or perhaps thirty cities. A drastic change of scene occurred with the branch-and-bound algorithm of Held and Karp in 1971. They developed a combinatorial bound which can be viewed as a way to approximate the Lagrangian relaxation of the degree constraints and which, at best, reaches the value of the linear optimization relaxation that satisfies all connectiv-

These diagrams illustrate how Dantzig, Fulkerson, and Johnson ingeniously solved the 42-city problem by successively adding seven connectivity constraints and two ad hoc constraints to a linear relaxation of the problem. The edges drawn in black carry the value 1, the edges drawn in red carry the value ½.

Botticelli's *The Birth of Venus* **as a 140,000-city TSP. The cities are placed in proportion to the various shades of gray in the original image.**

ity constraints. It was the first systematic algorithm that computationally dominated the heuristic approach of Dantzig et al.

Around the same time, three strands of research were initiated that have had a huge impact on the study of the TSP: the development of complexity theory, the analysis of approximation algorithms, and the investigation of local search methods. We will briefly discuss these before turning to the rebirth of the cutting plane approach.

The work of Stephen Cook, Richard Karp, and Leonid Levin in the early 1970s on the computational complexity of combinatorial problems must be viewed as the single most important achievement in algorithmic mathematics. Using the notion of solvability in polynomial time as a concept for easiness, it explained why many problems are probably not easy. The theory brought order into chaos and had a reassuring effect, since we no longer had to blame ourselves for our inability to tackle hard problems. Very special cases of the TSP such as the problem of finding a Hamiltonian circuit in grid graphs have been shown to be *NP*-complete.

This work reinforced the performance analysis of fast approximation algorithms, a line of research that had been initiated by Ronald Graham in 1966 for multiprocessor scheduling. For the TSP where distances are symmetric and satisfy the triangle inequality, there is a host of simple methods for producing a tour that is no longer than twice the optimum. The first improvement and still essentially the last word on the issue is Nicos Christofides's algorithm of 1975, with a performance guarantee of 3/2. There is hope to reduce this to 4/3, but only tiny steps have been made in this direction. The 4/3-conjecture states that the lower bound of Held and Karp is always at least 3/4 of the optimum or, equivalently, that the solution to the linear optimization relaxation satisfying all connectivity constraints can be rounded to an integral solution of cost at most 4/3 as much.

A natural special case is the TSP where each city corresponds to a point in some d-dimensional space and the distance between two cities is given by the Euclidean distance between them. For the 2-dimensional case, Arora and Mitchell independently developed rather different polynomial-time approximation schemes which compute, for any constant $\varepsilon > 0$, a tour no longer than $1 + \varepsilon$ times the optimum. Supplementary to the worst-case approach is the probabilistic analysis of TSP heuristics pursued by Karp in 1977, which extends the classic work of Beardwood, Halton, and Hammersley published in 1959 on the

shortest length of Hamiltonian circuits through n random points in the d-dimensional cube.

In 1965 Shen Lin's paper on link-exchange methods for the TSP appeared. The method makes simple local improvements in a given tour and repeats the process until no further improvements are possible and a local optimum has been found. Such methods tend to yield good tours, without ironclad guarantees regarding solution quality or running time. In 1973, with Brian Kernighan, Shen published a more involved and problem-specific local search method for the TSP, which has excellent experimental performance. The approach has developed further in two directions. One is the emergence of generic heuristic search methods, such as simulated annealing, tabu search, genetic algorithms, and artificial neural networks, which nowadays come together under the heading of *metaheuristics* and which can be relied upon for obtaining reasonably good solutions for hard problems in the absence of deep insights into specific problem structure. Another line of work is the development of highly engineered approaches for specific problems. For the TSP the present world champions are Keld Helsgaun's extension of Lin-Kernighan and Yiuchi Nagata's genetic TSP algorithm. These methods come extremely close to the optimum for problems with tens of thousands of cities.

> *The ultimate question remains: How do we solve large instances to proven optimality?*

In spite of the hardness of the TSP and the existence of good methods that get close to the optimum, the ultimate question remains: How do we solve large instances to proven optimality? After the work of Held and Karp, researchers returned to the cutting plane approach with startling success. Optimization records jumped from 80 to 120 to 318 to 666 cities. A landmark was set by the *branch-and-cut* algorithm of Padberg and Rinaldi in 1987. They solved a 2,392-city problem to optimality by branch-and-bound with bounds obtained from linear relaxations strengthened by connectivity constraints and many more TSP-specific cutting planes. The bottom line of the approach is to use strong bounds derived from an in-depth study of the TSP polytope and to branch only if necessary. It was taken further by the team of David Applegate, Bob Bixby, Vašek Chvátal, and Bill Cook. They combined strong bounds with strong branching, where several options for branching are tried out and evaluated up to a certain depth before a final branching decision is taken. Their work resulted in the Concorde code and the solution to optimality of TSP instances up to 85,900 cities.

It should be noted that TSP history runs parallel to the development of algorithms for general integer linear

The writing is relaxed and entertaining; the presentation is excellent.

optimization. Ralph Gomory proposed a cutting plane algorithm for integer optimization in 1958. Despite initial success, it was overshadowed by the branch-and-bound algorithm of Land and Doig of 1960, which uses simple linear relaxations as lower bounds. But present state-of-the-art software for integer optimization combines both approaches and heavily relies on Gomory's ideas.

Applegate, Bixby, Chvátal, and Cook presented a detailed account of their work on the TSP in their book *The Traveling Salesman Problem: A Computational Study* (Princeton University Press, 2006). It was awarded the INFORMS Lanchester Prize in 2007.

In the book under review, Cook gives a broader, more popular, and less technical exposition of the history of TSP research. The core of the book is the story of branch-and-cut, starting from the Dantzig-Fulkerson-Johnson paper of 1954 and culminating in the Concorde code of Applegate, Bixby, Chvátal, and Cook. It is a lucid account which enables the reader with a mathematical mind to appreciate the development of ideas and concepts that led to a truly remarkable feat of algorithm engineering. The emphasis is, of course, on intuition rather than proofs. The reader who wishes to see the i's dotted and the t's crossed is directed elsewhere.

This main part of the book is preceded by a chapter on approximation and local search and is followed by a chapter on computational complexity. Again, the exposition is not for the faint-hearted but provides informative and concise conceptual histories. Cook discusses many more facets of the TSP. In the earlier chapters he provides a broad context in which he unearths origins of the problem that were unknown to the present reviewers, and he describes its myriad of applications. In the later chapters he describes how children and animals handle instances of the TSP, and he discovers the TSP in works of art, transcending well beyond the realm of mathematics.

Cook's book covers all aspects of the TSP. It starts from its earliest history and reaches up to the limits of current research. The writing is relaxed and entertaining; the presentation is excellent. We greatly enjoyed reading it.