

Paths, Trees, and Flowers

“Jack Edmonds has been one of the creators of the field of combinatorial optimization and polyhedral combinatorics. His 1965 paper ‘Paths, Trees, and Flowers’ [1] was one of the first papers to suggest the possibility of establishing a mathematical theory of efficient combinatorial algorithms . . .” [from the award citation of the 1985 John von Neumann Theory Prize, awarded annually since 1975 by the Institute for Operations Research and the Management Sciences].

In 1962, the Chinese mathematician Mei-Ko Kwan proposed a method that could be used to minimize the lengths of routes walked by mail carriers. Much earlier, in 1736, the eminent Swiss mathematician Leonhard Euler, who had elevated calculus to unprecedented heights, had investigated whether there existed a walk across all seven bridges that connected two islands in the river Pregel with the rest of the city of Königsberg on the adjacent shores, a walk that would cross each bridge exactly once.

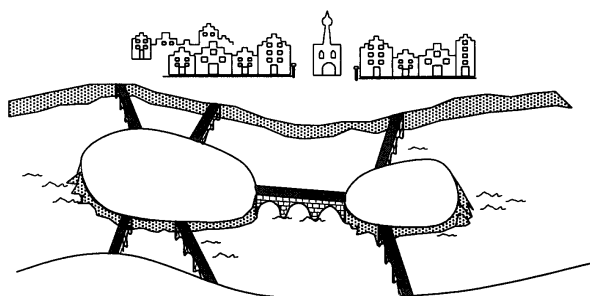


Fig. 1. The bridges of Königsberg.

Different as they may sound, these two inquiries have much in common. They both admit schematic representations based on the concept of *graphs*. Thus they are problems in *graph theory*, a twentieth century discipline which combines aspects of combinatorics and topology.

Given a set of items called *nodes* and a second set of items called *arcs*, a graph is defined as a relationship between such sets: For each arc, two of the nodes are specified to be *joined* by this arc. The actual, say physical, meaning of such nodes and arcs is not what distinguishes graphs. Only formal relationships count. In particular, the bridges of Königsberg may be consid-

ered as arcs, each joining a pair of the four nodes which correspond to the two islands and the two shores.

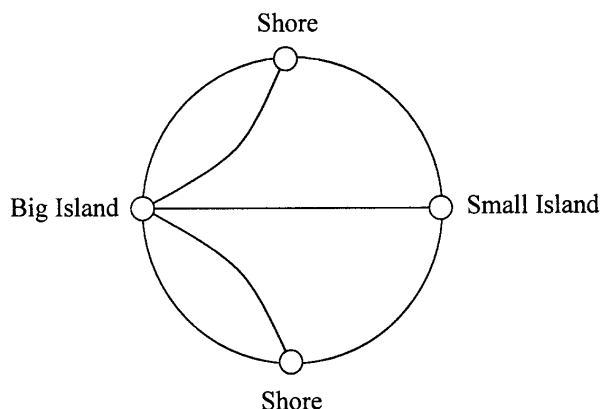


Fig. 2. The graph of the bridges of Königsberg.

The commonality between the two graph-theoretical problem formulations reaches deeper. Given any graph, for one of its nodes, say i_0 , find an arc a_1 that joins node i_0 and another node, say i_1 . One may try to repeat this process and find a second arc a_2 which joins node i_1 to a node i_2 and so on. This process would yield a sequence of nodes—some may be revisited—successive pairs of which are joined by arcs (Fig. 3). The first and the last node in that sequence are then connected by this

“path.”

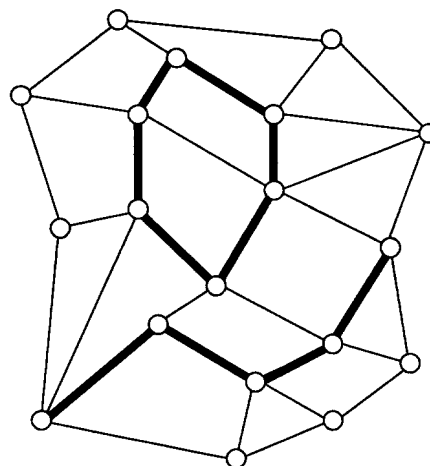


Fig. 3. Open and closed paths in the graph.

(Think of a route along street segments joining intersections.) If each pair of nodes in a graph can be connected by a path, then the whole graph is considered *connected*. A path in a graph is called *closed* if it returns to its starting point (Fig. 3).

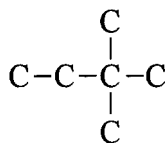
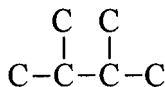
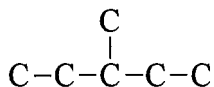
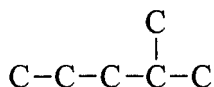
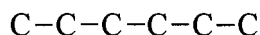


Fig. 4. The carbon graphs of the isomers of C_6H_{14} .

Mei-Ko Kwan's "Chinese Postman Problem," as it is now generally called (since first suggested by Alan Goldman, then of NBS), is to determine, in a given connected graph, a shortest closed path which traverses every arc at least once—delivers mail on every assigned street block. Euler also considers closed paths meeting all arcs, but aims at characterizing all graphs which accommodate an *Euler tour*, a closed path that crosses each arc exactly once. As Euler found, they are precisely those connected graphs in which each node attaches to an even number of arcs, or in other words, every node is of even *degree*. By this result, there can be no Euler tour

over the bridges of Königsberg. Euler's examination typifies the classical combinatorial query: Do certain constructs exist and if so, how many?

The following question also illustrates the flavor of classical combinatorics: how many isomers are there of the hydrocarbon $\text{C}_n\text{H}_{2n+2}$? Each such isomer (Fig. 4) is characterized by the joining patterns of the carbon atoms as nodes and bonds as arcs. The corresponding arrangements, therefore, represent graphs. These graphs have a very special property: For any two of their nodes there exists a unique connecting path without repeat arcs. Such a graph is called a

"tree."

Every isomer defines a tree of carbon atoms whose nodes are of degree four or less. Conversely, every such tree uniquely characterizes an isomer. In graph-theoretical terms, the question is: How many trees with n nodes are there with node degrees not exceeding four? (For $n = 40$, the number is 62,481,801,147,341, as determined by ingenious use of "generating functions.")

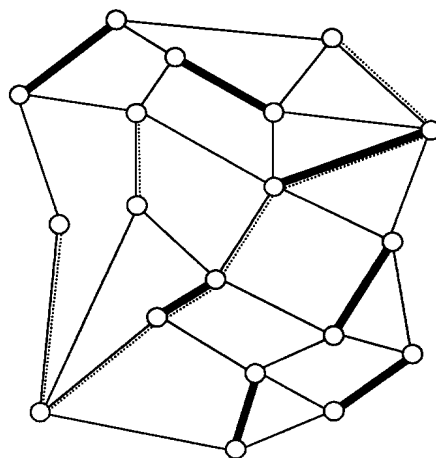


Fig. 5. A matching in a graph; two of the augmenting paths are highlighted by

Here is one more topic in the same vein. A matchmaker has a list of k single men and l single women. Her experience of many years enables her to discern which pairs are compatible. She considers a number of compatible matches that might be arranged simultaneously. How can she be sure that she arrived at the largest possible number of such matches?

A theorem by one of the founders of graph theory, the Hungarian mathematician Dénes König in 1931, suggests an answer to that question. He considered the

graph whose nodes are the matchmaker's clients and whose arcs join the compatible pairs, leading to the graph-theoretical concept of *matching*. A matching in a graph is any subset of its arcs such that no two arcs in the matching meet at a common node (Fig. 5).

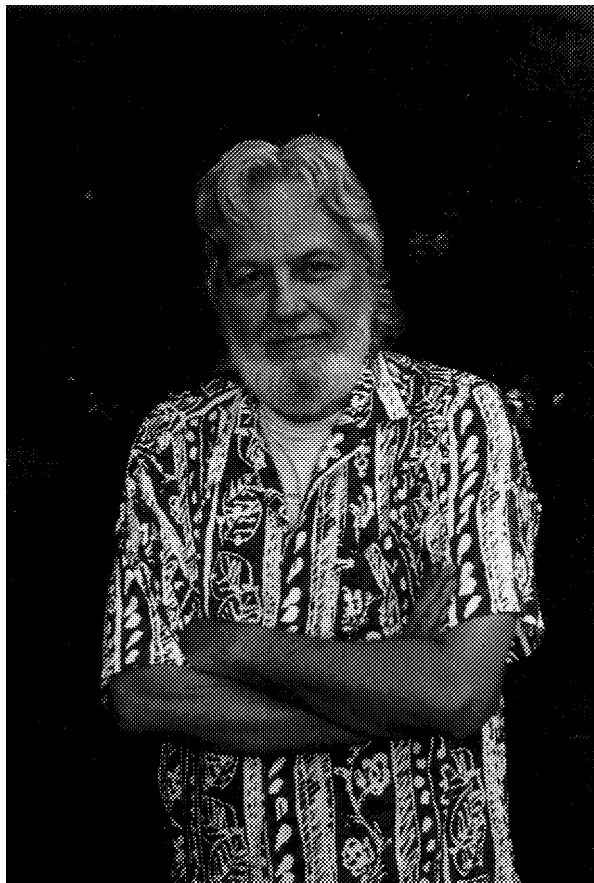


Fig. 6. Jack Edmonds.

Similar concepts are that of a *packing* or *independent set*, namely, a subset of the nodes no two of which are connected by an arc, and of a *cover*, namely, a subset of the nodes that meets every arc.

The matchmaker's graph has the special property that each of its nodes is of one of two kinds, male or female, and every arc connects nodes of different gender. For such a *bipartite* graph, König's theorem states that the number of arcs in a maximum matching equals the number of nodes in a minimum cover. However, this deep result ignores the problem of actually finding a minimum cover in order to prove the optimality of a matching.

Given a particular matching in any graph, an *exposed node* is one that is not met by any arc of the matching.

An *alternating path*, that is, a path whose every other arc is part of the matching, is called *augmenting* if it connects two exposed nodes. Indeed, switching arcs in and out of the matching along an augmenting path results in a new matching with one more arc in it. It is a 1957 theorem by the French mathematician Claude Berge that a larger matching exists not merely if, but also only if, the current matching admits an augmenting path. The classical graph theorist would look at this elegant characterization of maximum matchings and ask: what more needs to be said?

That outlook had been changing during and after World War II. The extensive planning needs, military and civilian, encountered during the war and post-war years now required finding actual solutions to many graph-theoretical and combinatorial problems, but with a new slant: Instead of asking questions about existence or numbers of solutions, there was now a call for "optimal" solutions, crucial in such areas as logistics, traffic and transportation planning, scheduling of jobs, machines, or airline crews, facility location, microchip design, just to name a few. George B. Dantzig's celebrated "Simplex Algorithm" for linear programming was a key achievement of this era.

The Chinese Postman is a case in point. He does not care about how many tours of his street blocks there are to choose from nor whether there are indeed Euler tours. He cares about delivering his mail while walking the shortest possible distance. (Admittedly, mail carriers, Chinese and otherwise, don't really think in terms of the Chinese Postman Problem. But how about garbage collection in a big city?)

And then there were computers, and the expectation that those electronic wizards would handle applied combinatorial optimizations even if faced with the large numbers of variables which typically rendered manual computations impractical.

Enter Jack Edmonds. Edmonds did his undergraduate work at the George Washington University. He recounts—as one of the pioneers featured in the volume *History of Mathematical Programming: A Collection of Personal Reminiscences* [20]—his varied interests and activities of that period. Thus he designed toys and games with expectations of monetary rewards, which unfortunately did not materialize. A stint as a copy boy at the Washington Post found him at the night desk during President Eisenhower's heart attack in 1955. Mathematics, however, survived as his overriding interest. Fascinated by the study of polytopes by the Canadian Mathematician H. S. M. Coxeter, Edmonds' master thesis at the University of Maryland (1959) addressed the problem of embedding graphs into surfaces [6].

In 1959 he joined NBS, became a founding member of Alan Goldman's newly created Operations Research Section in 1961, and was steered towards precisely the endeavor of developing optimization algorithms for problems in graph theory and combinatorics. In particular, he was drawn to two fundamental problems: the "Maximum Packing Problem" and the "Minimum Cover Problem" of determining largest packings and smallest covers, respectively, in a given graph [2].

Edmonds quickly recognized the major challenge of that task, a challenge that he called "the curse of exponentiality." While plainly "finite," many of the known graph-theoretical algorithms required *exponential effort*, or their authors had not detailed their procedures in a way that avoided such exponentiality.

Consider the "Maximum Matching Problem" of finding the largest matching in a given graph. Recall that a matching can be enlarged whenever there is an augmenting path, that is, an alternating path connecting two exposed nodes. As a consequence, there exists an "algorithm" which in a finite number of steps—everything is finite here—determines whether a matching is maximum or shows how to improve it. "All" that's needed is to examine, at each step, the alternating paths originating at exposed nodes. It's just that there are so darn many of them!

In a general framework, computational problems have a natural *size*, for instance, the number of nodes or arcs in a graph. For algorithms, computational effort can then be defined as the number of basic computational steps, such as individual additions or multiplications, and depends on problem size. If computational effort increases as fast or faster than exponentially with problem size, then it is said to require exponential effort. *Polynomial time*, by contrast,—Edmonds used the term *good*—prevails if computational effort is bounded by a power d of problem size n . The notation $O(n^d)$ is used to indicate that level of *complexity*.

Regardless of technological progress, computers are helpless when faced with exponential effort. To wit, the "Traveling Salesman Problem." Here a connected graph is given along with a particular length for each arc. What is wanted is a closed path of least total length that visits every node. The sequence of visits essentially defines such a round trip. Consider, for instance, the 48 state capitals of the contiguous United States and Washington, DC, as nodes of a graph with an arc between any two of them. From a base in Washington, all those capitals are to be visited while minimizing the total distance traveled. Thus there are $48!$ (factorial) round trips in the running. A future supercomputer, spending 1 nanosecond per trip, would require more than 3×10^{44} years to examine all possible trips.

Stressing the integral role of complexity, Edmonds became the leading proponent of a new direction: to develop good algorithms for problems in graph theory and combinatorics (or to identify problems for which such algorithms can be proven not to exist). This has spawned a new area of research that has grown and flourished for four decades and is still going strong. It is not by accident that a graph-theoretical optimization problem, namely, the Traveling Salesman Problem, is now frequently used as a complexity standard, calling a problem *NP-complete* if it requires computational effort equivalent to that of the Traveling Salesman Problem.

In 1961, while attending a summer workshop at the RAND corporation in Santa Monica, California, Jack Edmonds discovered a good algorithm for the Maximum Matching Problem, whose complexity he conservatively pegged at $O(n^4)$.

That algorithm is described in the paper *Paths, Trees, and Flowers* [1]. In its title, the words "paths" and "trees" refer to the standard graph-theoretical concepts. The algorithms augmenting paths are found by a "tree search" combined with a sophisticated process of shrinking certain subgraphs called blossoms into single nodes of a reduced graph. Hence the term

"flowers"

in the title.

Why was it a breakthrough? The answer is that all good graph-theoretical algorithms known at the time addressed "unimodular" problems such as the "Shortest Path" and "Network Flow" problems, the rigorous proof for the latter having been given by Edmonds with collaboration by Richard M. Karp [13]. These are problems that could be formulated as integrality-preserving linear programs, which by themselves did not create good algorithms but indicated the potential for such. Edmonds' matching algorithm was the very first instance of a good algorithm for a problem outside that mold.

In addition, *Paths, Trees, and Flowers* contributed a major theoretical result: a generalization of König's theorem that holds for matchings in all kinds of graphs, not just bipartite ones.

Edmonds also conjectured in this paper that both the Maximum Packing Problem and the Minimum Cover Problem were intrinsically harder than the Maximum Matching Problem. Indeed, both of the former problems were subsequently shown to be NP-complete.

In one of his seminal papers [3-10] published in the *Journal of Research of the National Bureau of Standards*, Edmonds [7] extended his algorithm to find matchings that optimize the sum of given arc weights, and, perhaps more importantly, he laid the foundation

for a polyhedral interpretation of matching theory which was pursued, for instance, in the doctoral thesis by William R. Pulleyblank [15] advised by Edmonds.

Subsequently, Edmonds found other good algorithms, for instance, in his path-breaking research on combinatorial abstractions of matrices, called *matroids* [4, 12, 14-17, 19]. And, last but not least, he and Ellis L. Johnson used the matching paradigm to arrive at a first good algorithm for the Chinese Postman Problem [18].

In 1969, Edmonds accepted a professorship of mathematics at the University of Waterloo, where a list of distinguished doctoral students is testament to his special gift of guiding and motivating young mathematicians. He remains to this day an active and highly influential researcher in the field of graph theory and combinatorics.

Why is it important to identify even a few graph-theoretical and combinatorial problems with good solution algorithms, when there is such a great variety of real-life optimization tasks, most of them defined in a less clear-cut fashion? The utility of good algorithms for idealized problems and their theory is that they suggest generalizations, variations, promising avenues of attack, treatable approximations, iterative applications, and also flag problem formulations best to avoid. In all these roles, Edmonds' matching algorithm has been an indispensable and inspirational part of the toolkit for combinatorial optimization and its multiple applications to modern technology.

Prepared by Christoph Witzgall with help by Ronald Boisvert, Geraldine Cheok, Saul Gass, Alan Goldman, and James Lawrence.

Bibliography

- [1] Jack Edmonds, Paths, Trees, and Flowers, *Canad. J. Math.* **17**, 449-467 (1965).
- [2] Jack Edmonds, Covers and Packings in a Family of Sets, *Bull. Amer. Math. Soc.* **68**, 494-499 (1962).
- [3] Jack Edmonds, Existence of k-Edge Connected Ordinary Graphs with Prescribed Degrees, *J. Res. Natl. Bur. Stand.* **68B**, 73-74 (1964).
- [4] Jack Edmonds, Minimum Partition of a Matroid into Independent Subsets, *J. Res. Natl. Bur. Stand.* **69B**, 67-72 (1965).
- [5] Jack Edmonds, Lehman's Switching Game and a Theorem of Tutte and Nash-Williams, *J. Res. Natl. Bur. Stand.* **69B**, 73-77 (1965).
- [6] Jack Edmonds, On the Surface Duality of Linear Graphs, *J. Res. Natl. Bur. Stand.* **69B**, 121-123 (1965).
- [7] Jack Edmonds, Maximum Matching and a Polyhedron with 0,1-Vertices, *J. Res. Natl. Bur. Stand.* **69B**, 125-130 (1965).
- [8] Jack Edmonds and D. R. Fulkerson, Transversals and Matroid Partition, *J. Res. Natl. Bur. Stand.* **69B**, 147-153 (1965).
- [9] Jack Edmonds, Optimum Branchings, *J. Res. Natl. Bur. Stand.* **71B**, 233-240 (1967).
- [10] Jack Edmonds, Systems of Distinct Representatives and Linear Algebra, *J. Res. Natl. Bur. Stand.* **71B**, 241-245 (1967).
- [11] Jack Edmonds, Matroid Partition, in *Mathematics of the Decision Sciences, Part I*, (Proceedings of the Fifth Summer Seminar on the Mathematics of the Decision Sciences, Stanford University, Stanford, CA, 1967), American Mathematical Society, Providence, RI (1968) pp. 335-345.
- [12] Jack Edmonds, Submodular Functions, Matroids, and Certain Polyhedra, in *Combinatorial Structures and their Applications* (Proceedings of the Calgary International Conference on Combinatorial Structures and their Applications, University of Calgary, Calgary, Alberta, Canada, 1969), Gordon and Breach, New York (1970) pp. 69-81.
- [13] Jack Edmonds and Richard M. Karp, Theoretical Improvements in Algorithmic Efficiency for Network Flow Problems, in *Combinatorial Structures and their Applications* (Proceedings of the Calgary International Conference on Combinatorial Structures and their Applications, University of Calgary, Calgary, Alberta, Canada, 1969), Gordon and Breach, New York (1970) pp. 93-96.
- [14] Jack Edmonds, Matroids and the Greedy Algorithm, *Math. Programming* **1**, 127-136 (1971).
- [15] W. Pulleyblank and Jack Edmonds, Facets of 1-Matching Polyhedra, in *Hypergraph Seminar* (Proceedings of the First Working Seminar on Hypergraphs, Ohio State Univ., Columbus, Ohio, 1972) Lecture Notes in Mathematics Vol. 411, Springer-Verlag, Berlin (1974) pp. 214-242.
- [16] Jack Edmonds, Edge-Disjoint Branchings, in *Combinatorial Algorithms* (Courant Computer Science Symposium 9, Naval Postgraduate School, Monterey, CA, 1972), Algorithmics Press, New York (1973) pp. 91-96.
- [17] Peyton Young and Jack Edmonds, Matroid Designs, *J. Res. Natl. Bur. Stand.* **77B**, 15-44 (1973).
- [18] Jack Edmonds and Ellis L. Johnson, Matching, Euler Tours, and the Chinese Postman, *Math. Programming* **5**, 88-124 (1973).
- [19] Jack Edmonds, Matroid Intersection, in *Discrete Optimization I* (Proceedings of the Advanced Research Institute on Discrete Optimization and Systems Applications, Alberta, Canada, 1977); *Ann. Discrete Math.* **4**, 39-49 (1979).
- [20] Jan Karel Lenstra, Alexander H. G. Rinnooy Kan, Alexander Schrijver, *History of Mathematical Programming, A Collection of Personal Reminiscences*, North-Holland, Amsterdam (1991).