ON THE PROBLEM OF STEINER

Z. A. Melzak

(received January 3, 1961)

1. There is a well-known elementary problem:

(S_3) Given a triangle T with the vertices a_1, a_2, a_3, to find in the plane of T the point p which minimizes the sum of the distances $|pa_1| + |pa_2| + |pa_3|$.

p, called the Steiner point of T, is unique: if an angle of T is $\geq 2\pi/3$ then p is its vertex, otherwise p lies inside T and the sides of T subtend at p the angle $2\pi/3$. In the latter case p is called the S-point of T, and it can be found by the following simple construction: let a_{12} be the third vertex of the equilateral triangle whose other two vertices are a_1 and a_2, and whose interior does not overlap that of T, let C be the circle through a_1, a_2, a_{12}; then p is the intersection of C and the straight segment $a_{12}a_3$. It is easily proved that any one of the three ellipses through p with two of the vertices of T as foci is tangent at p to the circle through p about the third vertex of T.

A generalization of (S_3) is the problem:

(P_n) Given a convex polygon P with the vertices $a_1, \ldots, a_n$, to find in the plane of P the point q which minimizes the sum $\Sigma_{j=1}^{n} |qa_j|$.

q shares with the point p of (S_3) the tangency property. Given $k \geq 1$ points $f_1, \ldots, f_k$ in the plane, a k-ellipse with the foci $f_1, \ldots, f_k$ is the locus of points x in the plane, for which $\Sigma_{j=1}^{k} |xf_j| = \text{const}$. If a k-ellipse is a locus with at least two points then it is a closed smooth convex curve. One can show now that a k-ellipse, $1 \leq k < n$, with any k consecutive

https://doi.org/10.4153/CMB-1961-016-2 Published online by Cambridge University Press

vertices of P as foci, and passing through q, is tangent at q to the (n - k)-ellipse through q whose foci are the remaining vertices of P. This proves, among other things, the uniqueness of q.

A much greater interest attaches to a different generalization of (S_3), known as the Steiner problem:

(S_n) Given n points $a_1, \ldots, a_n$ in the plane, $n \geq 3$, to construct the shortest tree(s) whose vertices contain these n points.

For our purpose a tree may be defined as follows. Given N points $b_1, \ldots, b_N$ in the plane, a tree U on the vertices $b_1, \ldots, b_N$ is any set consisting of some of the $\binom{N}{2}$ closed straight segments $b_i b_j$, with the property that any two vertices can be joined by a sequence of segments belonging to U in one and only one way. A segment $b_i b_j$ in U is called a branch, the length L(U) of U is the sum of the lengths of its branches, $\{b_i\}$ is the set of all vertices sending branches to the vertex b_i, and $w(b_i)$ is their number. It will be observed that a tree may be self-intersecting.

There are two other problems similar to (S_n) and (P_n):

(C_n) To connect n given points in the plane by the shortest tree whose vertices are these n points, and

(T_n) Given n points $a_1, \ldots, a_n$ in the plane, to find the shortest unbranched tree with these n points as vertices.

A tree is unbranched if $w(a) \leq 2$ for everyone of its vertices a. The problem (C_n) is known as the problem of the shortest spanning subtree or the shortest connecting network, while (T_n) is related to the travelling salesman problem.

It is possible also to formulate a problem including (S_n), (P_n), (C_n), (T_n) as special cases:

$(S_{n\alpha\beta\gamma})$ Given three real numbers α, β, γ, and n points $a_1, \ldots, a_n$ in the plane, to find an integer k and k points $p_1, \ldots, p_k$, and to construct the tree U on the vertices $a_1, \ldots, a_n, p_1, \ldots, p_k$ so as to minimize the sum

https://doi.org/10.4153/CMB-1961-016-2 Published online by Cambridge University Press

$$L(U) + \alpha \Sigma_{j=1}^{n} w(a_j) + \beta \Sigma_{j=1}^{k} w(p_j) + \gamma k.$$

We obtain (S_n) when $\alpha = \beta = \gamma = 0$, (C_n) when $\alpha = 0$ and $\max(\beta, \gamma) >> 1$, (P_n) when $\beta = 0$ and $\alpha > \gamma >> 1$ (and the points $a_1, \ldots, a_n$ are the vertices of a convex polygon), and (T_n) when $\max(\beta, \gamma) >> \alpha >> 1$.

We offer now an economic interpretation of the problem $(S_{n\alpha\beta\gamma})$; this will explain how the above four problems arise as special cases of $(S_{n\alpha\beta\gamma})$ and also point out some possible applications. Let the points $a_1, \ldots, a_n$ represent n cities and let the tree U represent a system of roads connecting the cities. Let a point at which s roads meet, $s \geq 3$, be called an s-junction. Suppose that the cost of building one unit of length of the road is 1 (in some monetary units), that a city s-junction costs $s\alpha$, any other s-junction costs $s\beta$, and in addition there is a fixed charge γ for each junction outside of a city. Now $(S_{n\alpha\beta\gamma})$ is formally identical with asking: what is the cheapest system of roads that connects the n cities?

Suppose next that $\alpha = 0$ and that $\max(\beta, \gamma)$ is sufficiently large. This puts a great premium on avoiding any junctions outside of the cities, and one obtains the problem (C_n). Similarly, if $\beta = 0$ and $\alpha > \gamma$, where γ is sufficiently high, then there is a premium on avoiding junctions in the cities and also on keeping the number k of new junctions possibly low. For suitable α it will follow that the most economical system will have exactly one s-junction, and since $\beta = 0$, that is, since the number s does not influence the total cost, we obtain the problem (P_n). The case of (T_n) can be handled similarly.

Of the several problems mentioned (P_n) is elementary, (C_n) is completely solved [1], [2], [3], (T_n) can be solved, in principle at least, by trial and error, being discrete, and $(S_{n\alpha\beta\gamma})$ is apparently too hard to be attacked in its full generality. Several necessary conditions are known for (S_n) [4], but they do not suffice to construct the solution(s). Call a geometrical construction Euclidean if it uses only the ruler and compass in the traditional sense. We shall prove

THEOREM 1. For every n there exists a finite sequence of Euclidean constructions yielding all the minimizing trees of the problem (S_n).

https://doi.org/10.4153/CMB-1961-016-2 Published online by Cambridge University Press

2. Let U be a minimizing tree of (S_n), then

(1) U has the vertices $a_1, \ldots, a_n, s_1, \ldots, s_k$,

(2) U is non-selfintersecting,

(3) $w(s_i) = 3$, $1 \leq i \leq k$,

(4) each s_i, $1 \leq i \leq k$, is the S-point of the triangle $\{s_i\}$,

(5) $w(a_j) \leq 3$, $1 \leq j \leq n$,

(6) $0 \leq k \leq n - 2$.

It is understood that when $k = 0$ the only vertices of U are $a_1, \ldots, a_n$. The conditions (2) - (5) are easy consequences of the solution for (S_3) given in the previous section, and (6) is given, although not proved, in [4], p. 361. It appears that the conditions (1) - (6) sum up the total present knowledge about the minimizing trees of (S_n). We let $A = \{a_1, \ldots, a_n\}$.

We shall call every tree U which satisfies (1) - (6) an S-tree. To bring out the dependence on k we shall also call an S-tree with $n + k$ vertices an S_k-tree. It follows from (6) that every S-tree is an S_k-tree for some k, $0 \leq k \leq n - 2$. The set $\{s_1, \ldots, s_k\}$ of the vertices of an S_k-tree, other than those in A, will be called its S_k^n-set; it is empty when $k = 0$.

3. LEMMA 1. The number $N(n, k)$ of S_k^n-sets is finite for every n and k, and every such set can be obtained by a Euclidean construction.

This will be proved by induction. We have first $N(n, 0) = 1$ by definition. Let $k = 1$, then any S_1^n-set consists of a single point s which is by (4) an S-point of some triangle with the vertices in A. Therefore $N(n, 1) \leq \binom{n}{3}$ and each S_1^n-set can be found by the Euclidean construction of the problem (S_3). Suppose now that the lemma has been proved for $k = 1, \ldots, K$, $K \leq n - 3$, for every n. Consider a particular S_{K+1}^n-set Y. There must be in Y a point s such that $\{s\}$ includes at least two points of A, say a_{i_1} and a_{i_2}. Let b be the third point in $\{s\}$; b may be either in A or in Y. Let a be the intersection of the circle C through a_{i_1}, a_{i_2}, and s, and the extension of the straight

https://doi.org/10.4153/CMB-1961-016-2 Published online by Cambridge University Press

segment sb beyond s. Then it follows, as in the construction for (S_3), that a is the third vertex of the equilateral triangle with two vertices a_{i_1} and a_{i_2}, or rather, that it is the third vertex of one of the two such triangles. Since there are two such triangles and since there are $\binom{n}{2}$ ways of selecting a_{i_1} and a_{i_2} in A, it follows that there are n(n - 1) possibilities for a. The crucial point is that a, or rather the possible a's, can be found by Euclidean constructions based on the set A alone. It is clear now that the K members of Y, other than s, form an S^{n+1}_{K}-set for the points $a_1, \ldots, a_n$, a. Hence

(7) $$N(n, K+1) \le n(n-1)\, N(n+1, K).$$

Since it has been shown that $N(n, 1) \le \binom{n}{3}$ it follows from (7) that

(8) $$N(n,k) \le [(n+k-2)!\,(n+k-3)!]\binom{n+k-1}{3}/[(n-1)!\,(n-2)!],$$

and each of these N(n, k) sets can be found by Euclidean constructions. This proves the lemma.

LEMMA 2. There is a finite number of S-trees and each one can be obtained by a Euclidean construction.

Consider a particular S^n_k-set $\{s_1, \ldots, s_k\}$. This, together with A, gives the set of n + k vertices. It is known, [5], that there are N^{N-2} trees on N distinct vertices. Hence there are $(n+k)^{n+k-2}$ trees on the given n + k vertices. Since there are by lemma 1 N(n, k) possibilities for the S^n_k-set, there are together

$$\Sigma^{n-2}_{k=0} N(n,k)\,(n+k)^{n+k-2}$$

trees to examine. However, given any one of these trees, it is clearly possible by Euclidean constructions to decide whether it is an S-tree or not.

4. The proof of theorem 1 follows immediately. We first

construct all the S-trees; by lemma 2 this can be done. The process of finding the length of an S-tree, as well as the process of finding the minimum length of a finite number of trees, can be carried out by Euclidean constructions. Since the S-trees comprise the minimizing tree(s) of the problem (S_n), the proof is complete.

It is obvious that our algorithm, although effective, is extremely redundant and inefficient. On the other hand, the full power of the conditions (1) - (6), and other similar ones that can be derived, has not been used. Work is currently in progress on a practicable algorithm, by means of which the problem (S_n) can be solved with the aid of an automatic computer for, say, $n = 30$. The preliminary results seem to indicate that the number of operations necessary is of the order $p(n)$, where p is the partition function for unrestricted partitions. It is hoped to be able to report the results of this work in the near future.

5. The author acknowledges gratefully the encouragement and suggestions given to him by Dr. R. C. Prim of the Bell Telephone Laboratories.

REFERENCES

1. O. Boruvka, On a minimal problem, Prace Moravske Pridovedecke Spolecnosti 3 (1926).

2. J.B. Kruskal, On the shortest spanning subtree of a graph, Proc. Amer. Math. Soc. 7 (1956), 48-50.

3. R.C. Prim, Shortest Connecting Networks, Bell System Tech. J. 31, 1398-1401.

4. R. Courant and H. Robbins, What is Mathematics?, (New York, 1941).

5. J. Riordan, Introduction to Combinatorial Analysis, (New York, 1959).

University of British Columbia