Information Processing Letters 19 (1984) 125–129
North-Holland 19 October 1984

NON-PARTITIONABLE POINT SETS

David AVIS
School of Computer Science, McGill University, Montreal, Quebec H3A 2K6, Canada

Communicated by K. Ikeda
Received 12 June 1983
Revised 14 March 1984

Let S be a finite set of n points in d-dimensional space. S is α(n)-partitionable if there exists a set of d mutually intersecting hyperplanes that divide d-space into 2^d open regions, no $2^d - 1$ of which together contain more than α(n) points of S. Willard (1982) has shown that every set in 2-space is $\frac{3}{4}$n-partitionable. Yao (1983) has shown that every set in 3-space is $\frac{23}{24}$n-partitionable. It is shown here that there exist sets S of arbitrary cardinality in d-space, $d \geqslant 5$, for which $d^2 + 1$ open regions together contain at least $n - d^2$ points of S, in any partition by d intersecting hyperplanes. Further, at least $2^d - d^2 - 1$ open regions contain no points of S. This implies that the powerful balanced quad and oct-trees introduced by the above authors may not be generalized to balanced 2^d-trees in dimension at least 5.

Keywords: Discrete mathematics, data structures

1. Introduction

It is often useful to consider records in a database as points in space, the dimension of which corresponds to the number of data fields. Let S denote such a set of n points in d-dimensional space. A powerful data structure on this set of points can be obtained by dividing d-space into 2^d regions by a set of d intersecting hyperplanes, in such a way that each region contains some fixed fraction of the points in S. If such a division by hyperplanes exists, we say that S is partitionable. This data structure has been used in two and three dimensions to find efficient algorithms for various *range queries*, involving linear restraints on the keys. For example, Willard [3] has given a sublinear time algorithm for reporting all points lying in a given planar polygon. Yao [4] has given a sublinear time algorithm for reporting all points in a three-dimensional polytope. In this paper we are concerned with the question of whether these methods can be extended to higher dimensions.

We begin by making the notion of partitionability precise. Let α(n) be a function on the integers. A set of n points in d-space is α(n)-partitionable if there exists a set of d mutually intersecting hyperplanes that divide d-space into 2^d open regions, such that no $2^d - 1$ regions together contain more than α(n) points of S. Willard [3] has shown that every planar set S is $\frac{3}{4}$n-partitionable. Yao [4] extended this result to three dimensions by showing that every set in 3-space is $\frac{23}{24}$n-partitionable. We say that a set S is *partitionable* if it is α(n)-partitionable for some function α(n) such that $n - \alpha(n)$ is increasing in n, and *non-partitionable* otherwise. We will show that there exist non-partitionable sets in all dimensions greater than or equal to 5. Specifically, we will exhibit a set of points S of arbitrary cardinality n in d-space with the following properties:

Any partition of d-space by d intersecting hyperplanes has

(i) at most $d^2 + 1$ open regions containing at least $n - d^2$ of the points of S,

(ii) at least $2^d - d^2 - 1$ open regions containing no points of S.

Thus in 5-space there will be at least 6 regions containing no points of S. The situation in 4-space is not settled in general. For the point sets considered in this paper, the existence of a Gray code of length 16 with some special properties implies the existence of a partition such that each open region contains at least $\lfloor \frac{1}{16}n \rfloor$ points.

2. Non-partitionable sets

The point sets described in this section are the vertices of the so-called *cyclic polytopes* which play an important role in the theory of convex polyhedra. An accessible introduction to these fascinating polytopes is the paper by Gale [1]. Let $d \geqslant 5$ and $n \geqslant d$ be fixed. For any positive integer t let $s_t = (t, t^2, \ldots, t^d)$ and set

$$S = \{s_t : 1 \leqslant t \leqslant n\}. \tag{1}$$

Then, as shown in [1], S is the set of extreme points of a d-dimensional convex polytope, called a cyclic polytope. Let

$$f(x_1, x_2, \ldots, x_d) = \sum_{i=1}^{d} a_i x_i + a_{d+1}.$$

We denote by f the hyperplane $f(x_1, \ldots, x_d) = 0$ in d-space. The set S is partitioned by f according to the sign of the function

$$g(t) = \sum_{i=1}^{d} a_i t^i + a_{d+1}.$$

Integers t for which g(t) is positive correspond to points s_t in S that lie in one of the open half-spaces bounded by f, and integers t for which g(t) is negative correspond to points in the other open half-space. Integers for which $g(t) = 0$ correspond to points of S on f. Since g is a polynomial of degree d, it has some number k, which is at most d, of real roots. Therefore, f can contain at most d points of S. If k is zero, then all points of S lie on the same side of f. If k is positive, let $t_1 \leqslant t_2 \leqslant \cdots \leqslant t_k$ denote the real roots of g. Then S is partitioned by f into sets H, M and L as follows:

Case 1. $k = 0$. $H = S$, $M = L = \emptyset$.

Case 2. $k = 2m + 1$ is odd.

$$H = \{s_t : 1 \leqslant t < t_1\} \cup \bigcup_{i=1}^{m} \{s_t : t_{2i} < t < t_{2i+1}\},$$

$$L = \{s_t : t_k < t \leqslant n\} \cup \bigcup_{i=1}^{m} \{s_t : t_{2i-1} < t < t_{2i}\},$$

$$M = S - (H \cup L).$$

Case 3. $k = 2m$ is even.

$$H = \{s_t : 1 \leqslant t < t_1\} \cup \{s_t : t_k < t \leqslant n\} \cup \bigcup_{i=1}^{m-1} \{s_t : t_{2i} < t < t_{2i+1}\},$$

$$L = \bigcup_{i=1}^{m} \{s_t : t_{2i-1} < t < t_{2i}\},$$

$$M = S - (H \cup L).$$

The crucial observation is that the sets H and L together contain at most $d + 1$ consecutive strings of points from S. This fact is required in proving the following theorem.

Theorem. *Let* $F = \{f_1, \ldots, f_d\}$ *be a set of hyperplanes in* d*-space that intersect at a common point, and let* S *be defined as in* (1). *Then at least* $2^d - d^2 - 1$ *open regions generated by* F *contain no points of* S.

Proof. For each i, $i = 1, 2, \ldots, d$, define g_i, H_i, M_i, L_i with respect to f_i in the same way that g, H, M, L were defined with respect to f. Let

$$h_i(t) = \begin{cases} 0, & s_t \in H_i, \\ 1, & s_t \in L_i, \\ 2, & s_t \in M_i, \end{cases} \qquad t = 1, 2, \ldots, n.$$

Then s_t lies in an open region generated by F if and only if $h_i(t) = 0$ or 1, $i = 1, \ldots, d$. Furthermore, two points s_t and s_u lie in the same open region if and only if $h_i(t) = h_i(u) = 0$ or 1, $i = 1, \ldots, d$. We now construct a $d \times n$ matrix $A = [a_{ij}]$ as follows:

$$a_{ij} = h_i(j), \quad i = 1, \ldots, d, \; j = 1, \ldots, n.$$

This construction of a typical row of A is illustrated in Fig. 1 for $d = 3$ and $n = 10$. From the preceding remarks, each column j of A consisting entirely of zeros and ones corresponds to a point s_j of S lying in an open region of F. Such a column is called a *binary* column. The number of different binary columns of A is the number of open regions generated by F containing points of S. The remarks preceding the Theorem show that each row of A can contain at most $d+1$ consecutive strings of zeroes and ones. Starting from column 1 of A, find the first binary column. Let this be column i. Thus point s_i lies in an open region of F. Continue scanning the columns of A until a different binary column is found. In at least one row, one binary digit has been interchanged. We now find a binary column that is different from the first two. Again a binary interchange must have occurred in at least one row of A. We precede in this way until all columns of A have been scanned. Suppose we have found some number, $r+1$, of different binary columns in A. This implies that r open regions of F contain points of S. By the above discussion, at least r binary interchanges have occurred. Each row of A can contribute at most d such interchanges, and there are d rows. Therefore, we can have at most d^2 interchanges and hence d^2+1 different binary columns of A. The Theorem now follows. In case d is even we can conclude that A has at most d^2 different binary columns. This is because if each row contributes exactly d interchanges, the first binary column in A will be the same as the last one. □

As previously remarked, each hyperplane in F can contain at most d points of S. Thus at most d^2 points of S are contained in hyperplanes in F, and hence we have the following corollary.

Corollary. *At most* d^2+1 *open regions generated by* F *contain at least* $n-d^2$ *points of* S.

Therefore, we have shown that S is not partitionable if $d \geq 5$.

3. Partitions in 4-space

In this section we show that the sets S described in the last section are $\frac{15}{16}$n-partitionable in 4-space. By examining the proof of the Theorem for $d=4$, we see that a set S in 4-space defined by (1) is partitionable only if there exists a 4×16 binary matrix A with the following properties:

- P_1: A is a binary matrix and every column of A is distinct.
- P_2: Each row of A consists of at most 5 strings of zeroes and ones.
- P_3: Two consecutive columns of A differ in exactly one row.

P_1 follows from the fact that every open region generated by a set of four intersecting hyperplanes must contain some point of S. P_2 follows from the fact that each 4th degree polynomial has at most four roots. P_3 follows from the fact that 15 binary interchanges are required, there are three rows of A that can contribute four interchanges and one row that can contribute three interchanges. Thus

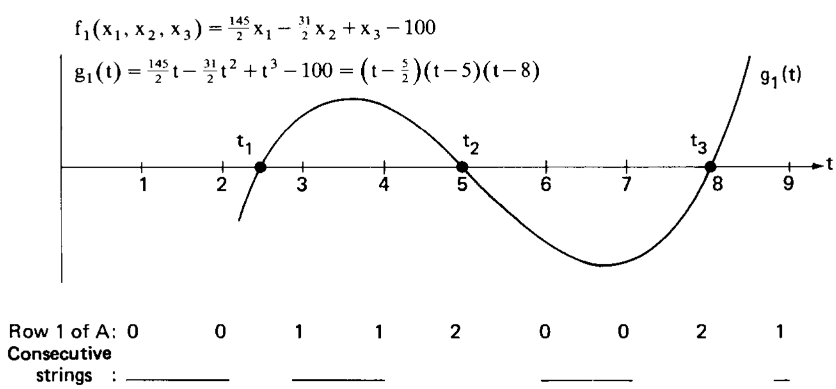

Fig. 1. Illustration of the proof of the Theorem.

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

$$\begin{array}{cccccccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 \end{array}$$

Fig. 2. Matrix satisfying P_1, P_2, P_3.

there must be exactly one interchange between each pair of columns.

Properties P_1 and P_3 define a Gray code. Gilbert [2] has published a list of all Gray codes of length 16. Of the nine such codes, one satisfies P_2, and is shown in Fig. 2. From this matrix, we can construct a $\frac{15}{16}$n-partition of the set S for any n. For example, suppose n = 16. Then we can define the functions g_1, g_2, g_3, g_4 as follows:

$$g_1(t) = 16(t - \tfrac{3}{2})(t - \tfrac{7}{2})(t - \tfrac{13}{2})(t - \tfrac{25}{2})/6825,$$

$$g_2(t) = 16(t - \tfrac{5}{2})(t - \tfrac{17}{2})(t - \tfrac{23}{2})(t - \tfrac{27}{2})/52785,$$

$$g_3(t) = -8(t - \tfrac{9}{2})(t - \tfrac{21}{2})(t - \tfrac{29}{2})/5481,$$

$$g_4(t) = 16(t - \tfrac{11}{2})(t - \tfrac{15}{2})(t - \tfrac{19}{2})(t - \tfrac{31}{2})/97185.$$

From the above, f_1, f_2, f_3, f_4 may be readily computed. Each function g_i is normalized so that $g_i(0) = 1$. Thus all hyperplanes intersect at the origin, and each open region contains exactly one point of S. Similar partitions can be generated for any value of n, and for more general cyclic polytopes defined, for any n positive numbers $0 < t_1 < t_2 < \cdots < t_n$, by $S = \{s_{t_i} \mid i = 1, \ldots, n\}$.

4. Additional results

We now state informally some related results. Detailed statements and proofs will be included in a later paper. In the plane, Willard showed that the first line in a $\frac{3}{4}$n-partition can be chosen to be any bisector of the given point set. Similarly, in 3-space, Yao showed that the first plane may be chosen to be any bisecting plane of the given point set. In 4-space, consider the set S defined in (1) of Section 2, and let

$$f_1(x_1, x_2, x_3, x_4) = x_1 - \tfrac{1}{2}(n + 1).$$

Then $f_1 = 0$ is a bisector of S. We can show that using f_1 and any three additional planes, S is non-partitionable. Therefore, in 4-space, the initial bisecting plane cannot be chosen arbitrarily.

Our final result concerns simultaneous partitioning of two equal cardinality point sets. Firstly, we recall the following corollary of Helly's theorem (see [4]).

Centre Theorem. *Let* S *be a set of* n *points in* d*-space. A point* x *exists in* d*-space, called a centre of* S, *such that on each side of any hyperplane* H *passing through* x *there are at least* $n/(d+1)$ *of the given points (including points lying on the hyperplane* H *itself).*

Let S_1 and S_2 be two planar point sets, each of cardinality n, with a common centre. Then it is shown in [4] that there exist two lines that simultaneously form a $\frac{11}{12}$n-partition for both S_1 and S_2. We can show that this result does not extend to three dimensions. For arbitrary n, there exist n point sets S_1 and S_2 with a common centre that cannot be simultaneously partitioned by any set of three planes.

Note added in proof

F. Yao has informed the author (personal communication) that she has proved that every point set in 3-space is $\frac{7}{8}$n-partitionable.

Acknowledgment

This work was carried out whilst the author was on leave at the University of Tokyo, and was supported by a grant from the Japan Society for the Promotion of Science. The author is grateful to Professor M. Iri for providing a conducive research environment.

References

[1] D. Gale, Neighborly and cyclic polytopes, in: V. Klee, ed., Convexity, Proc. Symp. on Pure Mathematics 7 (1963) 225–232.

[2] E.N. Gilbert, Gray codes and paths on the *n*-cube, Bell Syst. Tech. J. (1958) 815–826.

[3] D.E. Willard, Polygon retrieval, SIAM J. Comput. 11 (1982) 149–165.

[4] F. Yao, 3-space partition and its applications, Proc. 15th STOC (1983) 258–263.