Discrete Mathematics 39 (1982) 123–127
North-Holland Publishing Company

BOOLEAN DISTANCE FOR GRAPHS

Frank HARARY
Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA

Robert A. MELTER*
Department of Mathematics, Southampton College of Long Island University, Southampton, NY 11968, USA

Uri N. PELED
Computer Science Department, Columbia University, New York, NY 10027, USA

Ioan TOMESCU
Faculty of Mathematics, University of Bucharest, Bucharest, Romania

Received 26 March 1980
Revised 26 March 1981

The boolean distance between two points x and y of a connected graph G is defined as the set of all points on all paths joining x and y in G ($\emptyset$ if $x=y$). It is determined in terms of the block-cutpoint graph of G, and shown to satisfy the triangle inequality $b(x,y)\subseteq b(x,z)\cup b(z,y)$. We denote by $B(G)$ the collection of distinct boolean distances of G and by $M(G)$ the multiset of the distances together with the number of occurrences of each of them. Then $|B(G)|=1+\binom{b+1}{2}$ where b is the number of blocks of G. A combinatorial characterization is given for $B(T)$ where T is a tree. Finally, G is reconstructible from $M(G)$ if and only if every block of G is a line or a triangle.

1. Boolean distance

All notation and terminology in this paper not defined below can be found in [1]. In particular a path does not have repeated points. If G is a connected graph, we define the *boolean distance* $b(x, y)$ between points x and y of G as follows: if $x=y$, then $b(x,y)=\emptyset$, and if $x\neq y$, then $b(x,y)$ is the set of all points on all paths joining x and y. The boolean distances of G can be determined by its block structure, as will be shown below. To this end recall that the block-cutpoint graph of G, bc(G), is the bipartite graph having as points the blocks and the cutpoints of G, in which block b is adjacent to cutpoint c if and only if $c\in b$ in G. For any point x of G, let $b(x)$ be x itself if x is a cutpoint of G and the unique block of G containing x if not. Since bc(G) is a tree [1, p. 37], for any points x, y of G there is a unique path joining $b(x)$ and $b(y)$ in bc(G), which will be denoted by $P(x, y)$. The study of the cutpoints of G on $P(x,y)$ suggested the concept of a "cutting

* The second author was a participant in the Exchange Program between the National Academy of Sciences of the U.S.A. and the Academy of the Socialist Republic of Romania in 1979.

center" of a tree in [2]. The following result relating the boolean distances of G to its block structure can now be stated.

Theorem 1. *For any distinct points x, y of G, $b(x, y)$ is the union of all blocks of G (considered as point-sets) lying on $P(x, y)$ in* $\mathrm{bc}(G)$.

Proof. The path $P(x, y)$ has the form $c_0, b_1, c_1, b_2, \ldots, c_{n-1}, b_n, c_{n+1}$ where the c_i are cutpoints and the b_i are blocks of G such that $c_i \in b_{i-1} \cap b_i$. The first cutpoint c_0 appears only if x is a cutpoint and then $c_0 = x$, otherwise $x \in b_1$, and similarly at the other end. First we prove the inclusion $b(x, y) \subseteq b_1 \cup \cdots \cup b_n$. If a path of G leaves a block, it cannot return to this block, because that would necessitate repeating a cutpoint. Therefore if P is any path joining x and y in G, then the sequence of blocks and cutpoints encountered by P is a path joining $b(x)$ and $b(y)$ in $\mathrm{bc}(G)$. But the latter path must be $P(x, y)$, and so all the points of P are contained in $b_1 \cup \cdots \cup b_n$. Now we prove the opposite inclusion $b_1 \cup \cdots \cup b_n \subseteq b(x, y)$. Let z be any point of b_i. Then by [1, p. 28] G has a path P joining c_i and c_{i+1} and containing z (if $i = 1$ and c_0 does not appear, then G has a path P joining x and c_1 and containing z, and similarly at the other end). Let Q be any path joining x and c_i and R any path joining c_{i+1} and y in G. Then by the previous argument, Q followed by P followed by R is a path in G, and this path joins x and y and contains z. □

As a corollary we can see that $b(x, y)$ is a *boolean metric* in the sense of [4].

Corollary 1a. (1) *$b(x, y) = \emptyset$ if and only if $x = y$.*
(2) $b(x, y) = b(y, x)$.
(3) $b(x, y) \subseteq b(x, z) \cup b(z, y)$.

Proof. The first two statements are obvious, and third follows from Theorem 1. In fact for $x \neq y$ there is equality in (3) if and only if $b(z)$ appears in $P(x, y)$. □

2. Distance sets

The set of all boolean distances between points of G is called the *distance set* of G and is denoted by $B(G)$; it is understood that $\emptyset$ is always included as a boolean distance. Obviously $|B(G)| = 2$ if and only if G is a block. If G contains a cycle, boolean distances between distinct point-pairs may be equal. We write p for the number of points of G and b for the number of blocks, trusting that there will be no confusion between the symbols b and $b(x, y)$.

Theorem 2. *If G is a connected graph with b blocks, then $|B(G)| = 1 + \binom{b+1}{2}$. In particular $|B(G)| = 1 + \binom{p}{2}$ if and only if G is a tree.*

Proof. By Theorem 1, $B(G) - \{\emptyset\}$ is the set of unions of blocks of G (considered as point-sets) lying on paths of $\mathrm{bc}(G)$ beginning and ending in blocks of G.

Therefore $|B(G)|-1$ is equal to b (single blocks) plus $\binom{b}{2}$ (paths joining distinct blocks). The result on trees follows from this and from the fact that a connected graph has $p-1$ blocks if and only if it is a tree (certainly a tree has $p-1$ blocks, and if new lines are added to a tree, the number of blocks first decreases and then never increases). □

We remark that for almost all graphs G on p points $|B(G)|=2$ as $p\to\infty$, as it is observed in [3, p. 207] that almost all graphs are blocks. We also note that when $p\geq 3$, G is a star if and only if $B(G)-\{\emptyset\}$ contains only sets with two or three points. The next theorem characterizes the distance sets of trees.

Theorem 3. *Let X be an n-element set and let $\mathcal{F}\subseteq\mathcal{P}(X)$ be a collection of $\binom{n}{2}$ subsets of X. Then there exists a tree T with point-set X and $B(T)-\{\emptyset\}=\mathcal{F}$ if and only if the following three conditions are fulfilled:*

(i) *For any $F\in\mathcal{F}$, $|F|\geq 2$.*

(ii) *Any set F in $\mathcal{F}$ contains exactly $|F|-1$ 2-element subsets of $\mathcal{F}$. These 2-element subsets have the form $\{x_1,x_2\}$, $\{x_2,x_3\}$, $\dots$, $\{x_{k-1},x_k\}$, where $\{x_1,x_2,\dots,x_k\}=F$. We call x_1 and x_k end-elements of F.*

(iii) *If F_1, $F_2\in\mathcal{F}$ and $F_1\cap F_2=\{x\}$ where x is an end-element of both F_1 and F_2, then $F_1\cup F_2\in\mathcal{F}$.*

Proof. The necessity is obvious. In order to prove the sufficiency of the conditions, construct a graph G having point-set equal to X and line-set equal to the family of 2-element subsets of $\mathcal{F}$. Then G has no cycles, for if $x_0,x_1,\dots,x_{r-1}$ were the points of a cycle of G in that order, then $\{x_i,x_{i+1}\}\in\mathcal{F}$ for each i (indices mod r), hence by repeated use of (iii), $\{x_0,x_1,\dots,x_{r-1}\}\in\mathcal{F}$. Then by (ii) $\{x_0,x_1,\dots,x_{r-1}\}$ would have to contain exactly $r-1$ lines of G, but it contains at least r of them, a contradiction showing that G has no cycles. Now if any two points of G appeared more than once as end-elements, then by a standard argument G would contain a cycle, which is impossible. Hence there appear at most $\binom{n}{2}$ pairs of end-elements, so $|\mathcal{F}|\leq\binom{n}{2}$. But by assumption $|\mathcal{F}|=\binom{n}{2}$, and it follows that every two points of G appear as end-elements, and G is connected. Thus G is a tree and the point-sets of its paths are precisely the singletons and the members of $\mathcal{F}$. Hence $B(G)-\{\emptyset\}=\mathcal{F}$. □

3. Reconstructibility from boolean distances

The collection of boolean distances of G can be regarded as a multiset by taking the multiplicity of the sets of points into account. For example, $\emptyset$ has multiplicity p and the set of endpoints of a bridge has multiplicity 1. We thus define the *boolean distance multiset* $M(G)$ as the pair $(B(G),m)$, where m is the function

$$m : B(G) \to \left\{1, 2, \ldots, \binom{p}{2}\right\}$$

that associates with each set $S \in B(G)$ the number of unordered pairs $\{x, y\}$ of points of G such that $b(x, y) = S$. A graph G with given point-set is said to be *reconstructible from its boolean distance multiset* if G is uniquely determined by $M(G)$, i.e., there is a procedure to identify the lines of G using only $M(G)$.

We now determine the multiplicities of the blocks of G considered as point-sets.

Theorem 4. *A set $S \in B(G)$ has multiplicity $m(S) = \binom{|S|}{2}$ if and only if S induces a block of G.*

Proof. Clearly we may assume $S \neq \emptyset$. Then by Theorem 1, bc(G) has a unique path of the form $b_1, c_1, \ldots, c_{n-1}, b_n$, where the b_i are blocks and the c_i cutpoints of G, such that $S = b_1 \cup \cdots \cup b_n$. Thus S induces a block of G if and only if $n = 1$. If $n = 1$, then

$$m(S) = \binom{|b_1|}{2} = \binom{|S|}{2}.$$

If $n = 2$, then

$$m(S) = \tfrac{1}{2}(|b_1| - 1)(|b_2| - 1) < \binom{|S|}{2}.$$

If $n \geq 3$, then

$$m(S) = \tfrac{1}{2}|b_1| \cdot |b_n| < \binom{|S|}{2}. \quad \square$$

We define the *block completion* $K(G)$ as the graph obtained by replacing each block of G by a complete subgraph on the same set of points. Thus $K(G)$ is a 'block graph'; see [1, p. 29]. Obviously G and $K(G)$ have the same cutpoints. We then have the following corollary of Theorem 4.

Corollary 4a. *For any connected graph G, the block completion $K(G)$ is reconstructible from the multiset $M(G)$.*

Proof. The blocks of G are uniquely determined from the condition $m(S) = \binom{|S|}{2}$, and then two points are adjacent in $K(G)$ if and only if they belong to the same block of G. $\square$

We conclude with the following corollary showing which graphs G are reconstructible from $M(G)$.

Corollary 4b. *A connected graph G is reconstructible from $M(G)$ if and only if G has no cycle of length greater than* 3.

Proof. Assume that G contains a cycle C_n of length $n \geq 4$. Then C_n is contained in some block H having at least four points. If H is complete we denote by G_1 the graph obtained from G by deleting an arbitrary line of H. If H is not complete we denote by G_1 the graph obtained from G by adding a line between two nonadjacent points of H. In both cases G and G_1 have the same cutpoints and blocks (considered as point-sets). Hence $\mathrm{bc}(G) = \mathrm{bc}(G_1)$ and $M(G) = M(G_1)$, so G is not reconstructible from $M(G)$. Conversely, assume that G has no cycle of length greater than 3. We show that all blocks of G are lines or triangles. For otherwise there is a block H with at least four points and the longest cycle of H contains exactly three points, say x, y and z. Then z, say, is adjacent to a fourth point t of H, and there is a path $P = (t, \ldots, y)$ not containing z. If x is not a point of P, then $(t, \ldots, y, x, z, t)$ is a cycle of length greater than 3, and if x is a point of P, then $(t, \ldots, x, y, z, t)$ is such a cycle. This contradiction proves that the blocks of G are lines or triangles. Therefore $K(G) = G$ and by Corollary 4a, G is reconstructible from $M(G)$. □

References

[1] F. Harary Graph Theory (Addison-Wesley, Reading, MA, 1969).
[2] F. Harary and P. Ostrand, The cutting center theorem for trees, Discrete Math. 1 (1971) 7–18.
[3] F. Harary and E.M. Palmer, Graphical Enumeration (Academic Press, New York, 1973).
[4] R.A. Melter, Boolean valued rings and boolean metric spaces, Arch. Math. 15 (1964) 354–363.