

ON k -CLOSURE OPERATORS IN GRAPHS

BY

J. PŁONKA (WROCLAW)

0. Suppose that we want to go from some town a to another town b but the direct way from a to b is damaged. Then we look for some town c such that there exists a way from a to c and from c to b . If the way from a to c or from c to b is also not to use, then we go to c through some d or from c to b through some e , and so on. In this way we get some closure C_2 of the set $\{a, b\}$. We obtain a more general concept considering the following process: we build roads in some area starting from k different points a_1, a_2, \dots, a_k , where $k \geq 1$. We choose some point $c_1 \notin \{a_1, \dots, a_k\}$ and we build a road from any a_i to c_1 . Put $A_1 = \{c_1, a_1, \dots, a_k\}$. We choose a new point c_2 and we build k new ways from some different points $b_1, \dots, b_k \in A_1$ to c_2 . Put $A_2 = A_1 \cup \{c_2\}$. We choose a new point c_3 , build k new roads from some k points of A_2 to c_3 , and so on. In this way we obtain also some closure $C_k(\{a_1, \dots, a_k\})$ of the set $\{a_1, \dots, a_k\}$, where all points a_i and all points c_n , which we reach in a finite number of steps, belong to the set $C_k(\{a_1, \dots, a_k\})$. Obviously, the proper language for such a consideration is that of the theory of graphs. By a *graph* we mean a couple $\mathfrak{G} = (U; X)$, where U is a non-empty set, called the *set of vertices or points*, and X is a family of 2-element subsets of U , called the *set of edges*⁽¹⁾. The edges will be denoted by $[ab]$ instead of by $\{a, b\}$ to emphasize that we think about edges in a graph. We accept all definitions and the terminology from the cited book of Harary. In particular, if $[ab] \in X$, then we say that the vertices a, b are *adjacent* and we write $a \leftrightarrow b$, and if we draw the graph \mathfrak{G} , then we connect the points a and b . A sequence $a_1 a_2 \dots a_n$ of vertices is called a *simple chain* if all a_i are different and $a_i \leftrightarrow a_{i+1}$ for $i = 1, 2, \dots, n-1$.

In this paper we consider the k -closure operators C_k in graphs. In Section 1 we describe namely some properties of k -closure operators and, in particular, we prove that if $c \in C_2(\{a, b\})$ and $c \notin \{a, b\}$, then there exists a simple chain $a_1 a_2 \dots a_n c b_q \dots b_1$ of elements of $C_2(\{a, b\})$ such

⁽¹⁾ F. Harary, *Graph theory*, Reading, Mass., 1969.

that $a = a_1$ and $b = b_1$. In Section 2 we give some characterization of k -closed sets by means of some $(k+1)$ -ary relation in a set U . In Section 3 we study the notion of k - C -generation defined as follows: a graph $\mathfrak{G} = (U; X)$ is k - C -generated ($1 \leq k < |U|$) if for any $A \subseteq U$ such that $|A| = k$ we have $C_k(A) = U$. In Section 3 we answer the following question (see Theorem 5):

For $1 \leq k < n < \aleph_0$, what is the minimal integer m for which there exists a graph $\mathfrak{G} = (U; X)$ such that \mathfrak{G} is k - C -generated, $|U| = n$ and $|X| = m$?

1. Properties of k -closure operators. Let $\mathfrak{G} = (U; X)$ be a graph and let $A \subseteq U$. We say that a vertex $c \in U$ is k -reachable (k is a positive integer) from the set A if there exist k different vertices $a_1, \dots, a_k \in A$ such that $a_i \leftrightarrow c$ for $i = 1, 2, \dots, k$. If c is k -reachable from A , we write

$$c \xleftrightarrow[k]{} A.$$

We say that the set $A \subseteq U$ is k -closed in \mathfrak{G} or, briefly, k -closed, if A contains all vertices k -reachable from A .

We have

(i) *The set U is k -closed, the intersection of arbitrarily many k -closed sets is k -closed.*

(ii) *If $|A| < k$, then the set A is k -closed, and hence \emptyset is k -closed ($k = 1, 2, \dots$).*

For $A \subseteq U$ let us denote by $C_k(A)$ the smallest k -closed set containing A .

By (i) and (ii) we have

(iii) *The couple $(U; C_k)$ is a closure system for $k = 1, 2, \dots$, i.e. for any $A, B \subseteq U$ we have*

$$A \subseteq C_k(A), \quad A \subseteq B \Rightarrow C_k(A) \subseteq C_k(B), \quad C_k(C_k(A)) = C_k^{\mathfrak{v}}(A).$$

(iv) *A set $A \subseteq U$ is k -closed in \mathfrak{G} if and only if $C_k(A) = A$.*

For $A \subseteq U$ let us denote by $R_k(A)$ the set of all vertices c for which there exists a sequence

$$(1) \quad c_1, \dots, c_s \quad (c_s = c)$$

such that

$$c_1 \xleftrightarrow[k]{} A, \quad c_2 \xleftrightarrow[k]{} A \cup \{c_1\}, \quad \dots, \quad c_s \xleftrightarrow[k]{} A \cup \{c_1, \dots, c_{s-1}\}.$$

We have

$$(v) \quad C_k^{\mathfrak{v}}(A) = A \cup R_k(A).$$

Proof. Obviously, $A \cup R_k(A) \subseteq C_k(A)$. It is enough to show that the set $A \cup R_k(A)$ is k -closed. Let

$$c \xleftrightarrow[k]{} A \cup R_k(A),$$

which means that there exist $c_1, \dots, c_k \in A \cup R_k(A)$ such that

$$c \leftrightarrow_k \{c_1, \dots, c_k\}.$$

If $\{c_1, \dots, c_k\} \subseteq A$, then $c \in R_k(A)$ and we are ready. Otherwise, we assume that $c_1, \dots, c_m \in R_k(A)$ ($1 \leq m \leq k$) and $c_{m+1}, \dots, c_k \in A \setminus R_k(A)$ if $m < k$. Let t_i denote the sequence $c_1^i, \dots, c_{s(t)}^i$ from (1) for the element c_i ($1 \leq i \leq m$). Then t_1, t_2, \dots, t_m, c is the sequence (1) for the element c , which means that $c \in A \cup R_k(A)$, i.e. $A \cup R_k(A)$ is k -closed.

(vi) $C_{k+1}(A) \subseteq C_k(A)$ ($1 \leq k \leq \aleph_0$).

This follows from the fact that if $c \leftrightarrow_{k+1} \{a_1, \dots, a_{k+1}\}$, then

$$c \leftrightarrow_k \{a_1, \dots, a_k\},$$

and from (v).

(vii) $C_1(A) = \bigcup_{a \in A} D(a)$, where $D(a)$ is the component containing a .

Let $\mathfrak{G} = (U; X)$ be a graph and let $A \subseteq U$. We define a sequence $D_0(A), D_1(A), \dots$ of subsets of U as follows:

We put $D_0(A) = A$. If we have already defined $D_n(A)$, we put

$$F_{n+1}(A) = \{u \in U : u \leftrightarrow_2 D_n(A)\} \quad \text{and} \quad D_{n+1}(A) = D_n(A) \cup F_{n+1}(A).$$

Obviously,

(viii) $a \in C_2(A)$ if and only if $a \in D_n(A)$ for some non-negative integer n .

LEMMA 1. If $A = \{a, b\} \subseteq U$ and $\{c, d\} \subseteq D_n(A)$, then there exist two vertex-disjoint simple chains $a_1 \dots a_p$ and $b_1 \dots b_q$ such that $\{a_1, \dots, a_p\} \subseteq D_n(A)$, $\{b_1, \dots, b_q\} \subseteq D_n(A)$, $a = a_1$, $b = b_1$ and $c = a_p$, $d = b_q$ or $c = b_q$, $d = a_p$.

Proof. We prove our lemma using induction on n . If $n = 0$, our chains are a and b . Let us assume that the lemma holds for any $\{e, f\} \subseteq D_n(A)$ for some $n \geq 0$. Let $\{c, d\} \subseteq D_{n+1}(A)$. We have 3 possibilities:

1° $\{c, d\} \subseteq D_n(A)$;

2° for one of the elements of $\{c, d\}$, e.g. c , we have $c \in F_{n+1}(A) \setminus D_n(A)$ and $d \in D_n(A)$;

3° $\{c, d\} \subseteq F_{n+1}(A) \setminus D_n(A)$.

In case 1° our lemma is true by induction.

In case 2° there exists $\{e, f\} \subseteq D_n(A)$ such that

$$c \leftrightarrow_2 \{e, f\}.$$

One of the elements e, f , say e , must be different from d . By induction we can assume that there exist disjoint simple chains $a_1 \dots a_p$ and $b_1 \dots b_q$ of elements of $D_n(A)$ such that $a = a_1$, $b = b_1$ and $e = a_p$, $d = b_q$ or

$e = b_q$ and $d = a_p$. Thus the chains $a_1 \dots a_p c$ and $b_1 \dots b_q$ or $a_1 \dots a_p$ and $b_1 \dots b_q c$ satisfy the requirements of the lemma for c, d .

In case 3° there exist $\{e, f\} \subseteq D_n(A)$ and $\{g, h\} \subseteq D_n(A)$ such that

$$c \leftrightarrow_2 \{e, f\} \quad \text{and} \quad d \leftrightarrow_2 \{g, h\}.$$

We can assume that $e \neq h$. By induction we can assume that there exist two disjoint simple chains $a_1 \dots a_p$ and $b_1 \dots b_q$ of elements of $D_n(A)$, where $a = a_1$, $b = b_1$ and $e = a_p$, $h = b_q$ or $e = b_q$, $h = a_p$. Then the sequences $a_1 \dots a_p c$, $b_1 \dots b_q d$ or $a_1 \dots a_p d$, $b_1 \dots b_q c$ satisfy the requirements of the lemma.

COROLLARY 1. *If $A = \{a, b\} \subseteq U$ and $\{c, d\} \subseteq C_2(A)$, then there exist two vertex-disjoint simple chains of elements of $C_2(A)$ connecting a with c and b with d , respectively, or connecting a with d and b with c .*

THEOREM 1. *If $A = \{a, b\} \subseteq U$ and $c \in C_2(A)$, $c \notin A$, then there exists a simple chain $a_1 \dots a_p c b_q \dots b_1$, where*

$$a = a_1, \quad b = b_1 \quad \text{and} \quad \{a_1, \dots, a_p, b_1, \dots, b_q\} \subseteq C_2(A).$$

Proof. Let n be the smallest integer such that $c \in D_n(A)$. Thus $c \in F_n(A) \setminus D_{n-1}(A)$ and there exist $u, v \in D_{n-1}(A)$ such that

$$c \leftrightarrow_2 \{u, v\}.$$

By Lemma 1 there exist disjoint simple chains $a_1 \dots a_p$ and $b_1 \dots b_q$ of elements of $D_{n-1}(A)$, where $a = a_1$, $b = b_1$, $u = a_p$, $v = b_q$ or $u = b_q$, $v = a_p$. Assume that $u = a_p$ and $v = b_q$. Thus the chain $a_1 \dots a_p c b_q \dots b_1$ satisfies the requirements of the theorem.

2. A characterization of k -closed sets in graphs. Let $\mathfrak{G} = (U; X)$ be a graph. In the graph \mathfrak{G} we define a relation r_k ($k = 1, 2, \dots$) as follows:

$$\langle a_1, \dots, a_k, b \rangle \in r_k \text{ iff } a_1, \dots, a_k \text{ are all different and } b \leftrightarrow_k \{a_1, \dots, a_k\}.$$

We write $r_k(a_1, \dots, a_k, b)$ instead of $\langle a_1, \dots, a_k, b \rangle \in r_k$. Let U be a set, $|U| \geq 1$. A $(k+1)$ -ary relation r in U is called the *relation of k -reachability* in U if r satisfies the following conditions:

- (2) If $\langle a_1, \dots, a_k, b \rangle \in r$, then the elements a_1, \dots, a_k, b are all different.
- (3) $r(x_1, \dots, x_k, y) \Rightarrow r(x_{i_1}, \dots, x_{i_k}, y)$, where (i_1, \dots, i_k) is an arbitrary permutation of $1, \dots, k$.
- (4) If $a_1, \dots, a_k, b \in U$ and for any i ($1 \leq i \leq k$) there exist $a_2^i, \dots, a_k^i \in U$ such that $\langle a_i, a_2^i, \dots, a_k^i, b \rangle \in r$ or $\langle b, a_2^i, \dots, a_k^i, a_i \rangle \in r$, then $\langle a_1, \dots, a_k, b \rangle \in r$.

THEOREM 2. *The relation $r(x_1, \dots, x_k, y)$ in U is the relation of k -reachability in U if and only if there exists a graph $\mathfrak{G} = (U; X)$ and r is equal to r_k in \mathfrak{G} .*

Proof. The sufficiency is obvious, since $\langle a_1, \dots, a_k, b \rangle \in r_k$ means

$$b \xleftrightarrow[k]{k} \{a_1, \dots, a_k\}.$$

Suppose that r satisfies (2)-(4). We define the graph $\mathfrak{G} = (U; X)$ forming the set X from all $[uv]$, $u, v \in U$, for which there exist elements u_2, \dots, u_k such that $\langle u, u_2, \dots, u_k, v \rangle \in r$. So, if $\langle a_1, \dots, a_k, b \rangle \in r$, then b is k -reachable from $\{a_1, \dots, a_k\}$ by (2). Suppose that b is k -reachable from $\{a_1, \dots, a_k\}$. By the definition of X the assumptions of (4) are satisfied. Hence, by (4), $\langle a_1, \dots, a_k, b \rangle \in r$. Thus $r = r_k$ in \mathfrak{G} , which completes the proof.

Let U be a non-empty set and let $A \subseteq 2^U$. If $r(x_1, \dots, x_{k+1})$ ($k \geq 1$) is a relation of k -reachability in U , we denote by $[2^U]^r$ the family obtained from 2^U by removing from it all sets A for which there exists $\langle a_1, \dots, a_k, b \rangle \in r$, where $\{a_1, \dots, a_k\} \subseteq A$, $b \notin A$.

THEOREM 3. *A is a family of all k -closed sets of some graph $\mathfrak{G} = (U; X)$ if and only if $A = [2^U]^r$ for some $(k+1)$ -ary relation of k -reachability in U .*

Proof. To prove the necessity it is enough to take as r the relation r_k in \mathfrak{G} . To prove the sufficiency it is enough to define a graph \mathfrak{G} as in Theorem 2. Then $r = r_k$ in \mathfrak{G} and a set B belongs to $[2^U]^r$ iff B is k -closed in \mathfrak{G} .

3. k -connectivity in graphs. Let $\mathfrak{G} = (U; X)$ be a graph, where $|U| = a > 1$ (a need not be finite). We say that the graph \mathfrak{G} is k - C -generated ($1 \leq k < \min(\aleph_0, a)$) if for any $A \subseteq U$ such that $|A| = k$ we have $C_k(A) = U$. Obviously, \mathfrak{G} is 1- C -generated if and only if it is connected. If \mathfrak{G} is k - C -generated, then it need not be $(k+1)$ - C -generated: the graph in Fig. 1 is 1- C -generated but it is not 2- C -generated.

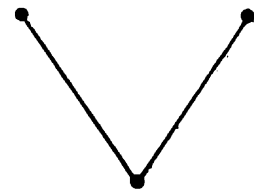


Fig. 1

However, we have

(ix) *If $\mathfrak{G} = (U; X)$ is a graph, where $|U| = a > 2$ and \mathfrak{G} is k - C -generated for some k , $1 < k < \min(\aleph_0, a)$, then \mathfrak{G} is m - C -generated for any m , $1 \leq m < k$.*

Proof. Obviously, it is enough to show that \mathfrak{G} is $(k-1)$ - C -generated. Let us take an arbitrary set $A_0 = \{a_1, \dots, a_{k-1}\} \subseteq U$. We have to prove that $C_{k-1}(A_0) = U$. Since $k < a$, there exists in U an element, say $a_k \notin A_0$.

Put $A_1 = \{a_1, \dots, a_k\}$. Since $k < a$, there exists in U an element $c_0 \notin A_1$ such that

$$c_0 \leftrightarrow_k A_1.$$

Thus $c_0 \leftrightarrow_{k-1} A_0$, which means that $c_0 \in C_{k-1}(A_0)$. So

$$(5) \quad C_{k-1}(A_0) = C_{k-1}(A), \quad \text{where } A = \{a_1, \dots, a_{k-1}, c_0\}.$$

Since \mathfrak{G} is k - C -generated, we have $C_k(A) = U$, hence, by (vi), $C_{k-1}(A) = U$. Thus by (5) we get $C_{k-1}(A_0) = U$.

We say that a graph $\mathfrak{G} = (U; X)$ is *edge-minimal k - C -generated* if \mathfrak{G} is k - C -generated and no graph $\mathfrak{G}' = (U; X')$, where $X' \subset X$ and $X' \neq X$, is k - C -generated. For any two positive integers k, n , where $1 \leq k < n$, we write

$$\varphi(n, k) = kn - \binom{k+1}{2}.$$

The function $\varphi(n, k)$ will play an important role in our further considerations.

THEOREM 4. *For any positive integer k and any cardinal a , where $k < a$, there exists a graph $\mathfrak{G}_a^k = (U; X)$ which is edge-minimal k - C -generated, for which $|U| = a$, and such that $|X| = a$ if $a \geq \aleph_0$ and $|X| = \varphi(n, k)$ if $a = n < \aleph_0$.*

Proof. Let us take a set U , where $|U| = a$, and a subset $U_0 \subset U$, where $|U_0| = k$. Let us put $X = X_0 \cup X_1$, where X_0 consists of all two-element subsets of U_0 , and X_1 consists of all two-element subsets $[uv]$ with $u \in U_0$ and $v \in U \setminus U_0$. If $a \geq \aleph_0$, then, obviously, $|X| = a$. If $a = n < \aleph_0$, then

$$|X| = |X_0| + |X_1| = \binom{k}{2} + k(n-k) = \varphi(n, k).$$

The proof that \mathfrak{G}_a^k is edge-minimal k - C -generated is left to the reader.

A graph $\mathfrak{G} = (U; X)$ is called *complete* if X consists of all two-element subsets of U .

LEMMA 2. *If $\mathfrak{G} = (U; X)$ is a k - C -generated graph, $|U| = n$ ($1 \leq k < n < \aleph_0$), then \mathfrak{G} contains a complete subgraph $\mathfrak{G}_0 = (U_0, X_0)$ such that $|U_0| = k$.*

Proof. If $k = 1$, it is enough to put $U_0 = \{a\}$ for arbitrary $a \in U$. If $k > 1$, let us choose different elements $a_1, b_2, \dots, b_k \in U$ and put $A_1 = \{a_1, b_2, \dots, b_k\}$. Since $k < n$ and \mathfrak{G} is k - C -generated, there must exist $a_2 \in U \setminus A_1$ such that

$$a_2 \leftrightarrow_k A_1.$$

If $k = 2$, we put $U_0 = \{a_1, a_2\}$ and we are ready.

If $k > 2$, then we continue our construction, i.e., if we have already constructed

$$A_i = \{a_1, a_2, \dots, a_i, b_{i+1}, \dots, b_k\},$$

where the subgraph generated by $\{a_1, \dots, a_i\}$ is complete, then — since \mathfrak{G} is k - \mathcal{C} -generated and $k < n$ — we can find an element $a_{i+1} \in U \setminus A_i$ such that

$$a_{i+1} \xleftrightarrow[k]{} A_i.$$

So we put

$$A_{i+1} = \{a_1, \dots, a_{i+1}, b_{i+2}, \dots, b_k\}$$

and we note that the subgraph generated by $\{a_1, \dots, a_{i+1}\}$ is complete. Now it is visible that after k steps we obtain the set $U_0 = \{a_1, \dots, a_k\}$ which generates a complete subgraph of \mathfrak{G} .

LEMMA 3. *If $\mathfrak{G} = (U; X)$ is a k - \mathcal{C} -generated graph, $|U| = n$ ($1 \leq k < n < \aleph_0$), then $|X| \geq \varphi(n, k)$.*

Proof. Let us take the set U_0 from Lemma 2. Denote by X_0 the set of all edges connecting vertices in U_0 . By Lemma 2 we know that

$$X_0 = \bigcup_{\substack{a, b \in U_0 \\ a \neq b}} \{a, b\}.$$

Thus $|X_0| = \binom{k}{2}$. Since $U \setminus U_0 \neq \emptyset$ and \mathfrak{G} is k - \mathcal{C} -generated, there must exist an element $c_1 \in U \setminus U_0$ such that

$$c_1 \xleftrightarrow[k]{} U_0.$$

Thus in X we have k new edges $[a_i c_1]$ ($i = 1, 2, \dots, k$). Let us put $B_1 = \{a_1, \dots, a_k, c_1\}$. If $n - k > 1$, there must exist $c_2 \in U \setminus B_1$ and elements $d_1, \dots, d_k \in B_1$ such that

$$c_2 \xleftrightarrow[k]{} \{d_1, \dots, d_k\}.$$

Thus in X we have k new edges $[c_2 d_i]$, and so on. It is now easy to see that after $n - k$ steps we exhaust the set U and

$$|X| \geq |X_0| + k(n - k) = \binom{k}{2} + k(n - k) = \varphi(n, k).$$

By Theorem 4 and Lemma 3 we get

THEOREM 5. *The number $\varphi(n, k)$ is equal to the smallest integer m for which there exists a k - \mathcal{C} -generated graph $\mathfrak{G} = (U; X)$ with $|U| = n$, $|X| = m$ ($1 \leq k < n < \aleph_0$).*

COROLLARY 2. *If $n > 1$, then any edge-minimal $(n - 1)$ - C -generated graph having n vertices is complete.*

In fact, the graph must have at least $\binom{n}{2}$ edges, which means that it is complete.

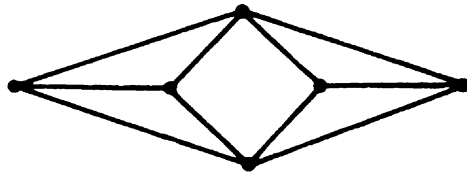


Fig. 2

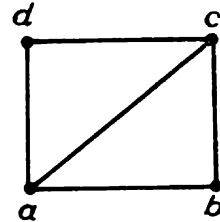


Fig. 3

Remark 1. One could suspect that any edge-minimal k - C -generated graph having exactly n vertices, where $1 \leq k < n < \aleph_0$, has $\varphi(n, k)$ edges. This, however, is not true, since the graph in Fig. 2 is edge-minimal 2- C -generated, has 6 vertices and 10 edges although $\varphi(6, 2) = 9$.

Remark 2. If $k = 2$, then the number 6 is the minimum of all such numbers n for which there exists a 2- C -generated edge-minimal graph $\mathfrak{G} = (U; X)$, where $|U| = n > 2$ and $|X| > \varphi(n, 2)$. This follows from the existence of the graph in Fig. 2 and from

THEOREM 6. *If $\mathfrak{G} = (U; X)$ is a 2- C -generated edge-minimal graph, where $|U| = n$, $3 \leq n \leq 5$, then $|X| = \varphi(n, 2)$.*

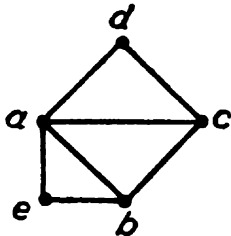


Fig. 4

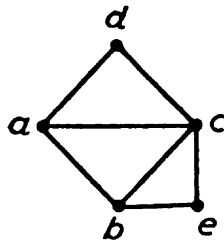


Fig. 5

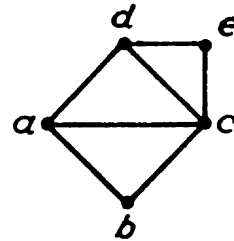


Fig. 6

Proof. By Lemma 2 there exists in \mathfrak{G} at least 1 edge, say $[ab]$, where $a, b \in U$. Since $n \geq 3$, there exists $c \in U$ such that

$$c \leftrightarrow_2 \{a, b\}.$$

If $n = 3$, then \mathfrak{G} is a triangle and $|X| = 3 = \varphi(3, 2)$. If $n > 3$, then there exists $d \in U$ such that

$$d \leftrightarrow_2 \{a, b, c\}.$$

Thus, if $n = 4$, we obtain (up to isomorphism) a graph from Fig. 3 having 5 edges and $\varphi(4, 2) = 5$.

Let $n = 5$ and let e be the 5-th vertex of \mathfrak{G} . Since \mathfrak{G} is k - \mathcal{C} -generated, we have (up to isomorphism) one of the cases from Figs. 4-9.

In cases 4-8 we obtain edge-minimal 2- \mathcal{C} -generated graphs having $\varphi(5, 2) = 7$ edges. Case 9 leads to a contradiction. In fact, the graph

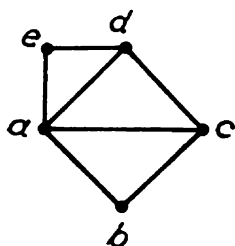


Fig. 7

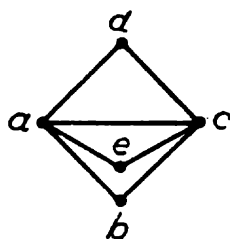


Fig. 8

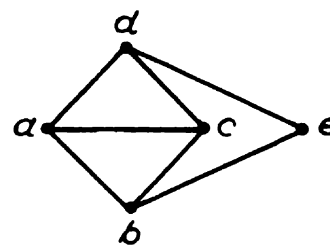


Fig. 9

from Fig. 9 is not 2- \mathcal{C} -generated. However, if we add the edge $[ce]$, then the edge $[be]$ can be removed. If we add the edge $[ae]$, then the edge $[be]$ can be removed. If we add the edge $[bd]$, then the edge $[ab]$ can be removed. Thus, case 9 does not lead to an edge-minimal 2- \mathcal{C} -generated graph.

Requ par la Rédaction le 20. 12. 1977