The Complexity of Some Edge Deletion Problems

EHAB S. EL-MALLAH AND CHARLES J. COLBOURN

Abstract —**The edge deletion problem (EDP) corresponding to a given class H of graphs is to find the minimum number of edges whose deletion from a given graph G results in a subgraph G', $G' \in$ H. In this paper we extend previous complexity results by showing that the EDP corresponding to any class H of graphs in each of the following cases is NP-hard.**

- **(i) H is defined by a set of forbidden homeomorphs or minors in which every member is a 2-connected graph with minimum degree 3.**
- **(ii) H is defined by $K_4 - e$ as a forbidden homeomorph or minor.**
- **(iii) H is defined by $P_l, l \geq 3$, the simple path on l nodes, as a forbidden induced subgraph.**

I. Introduction

THE *edge* (*node*) *deletion problem EDP* (*NDP*) corresponding to a class H of graphs is to find the minimum number of edges (nodes) of a given graph G whose deletion results in a subgraph G', $G' \in$ H. Equivalently, an EDP (NDP) calls for finding a subgraph (an induced subgraph) G' of G satisfying a prescribed property and having the largest possible number of edges (nodes). We are interested in determining the complexity of various EDP's corresponding to classes of graphs for which some NP-complete problem can be solved efficiently.

This interest arises since numerous efficient algorithms for solving a variety of NP-complete partitioning, embedding and colouring problems have been developed recently (e.g. see [7] for a recent survey). Examples of such classes include trees, series-parallel graphs, Δ-Y graphs, planar graphs, and so on. Such algorithms can be combined with other efficient algorithms for computing subgraphs with certain prescribed properties to design various approximation algorithms to handle more complicated instances.

The following example illustrates some of the above ideas. Suppose a digital circuit N is to be implemented using circuit boards which can accommodate at most K integrated circuits (IC's) each. It is required to determine the minimum number of circuit boards that are necessary to implement N such that the total number of cross connections between different boards is at most D. The problem can be modelled by a graph $G=(V,E)$ whose nodes correspond to the (IC's) of N and in which two nodes are adjacent if the corresponding IC's are connected to each other.

The above partitioning problem is NP-hard (see for example [5, problem [ND14]]). To cope with this difficulty, we may try to compute a largest subgraph G' of G for which the problem can be solved efficiently (for example, a tree). Then, use the exact solution on G' as a bound on the required optimum solution. As expected, such a relaxation method can be used in various ways to develop more elaborate search algorithms.

Having motivated some interest in the EDP's we now present in Section II some graph theoretic definitions and notations required throughout this paper. In Section III we discuss some previously known results and introduce the main results of this paper. Sections IV–VI present proofs of the cases mentioned in the abstract. Finally, we draw some conclusions in Section VII.

II. Graph Theoretic Definitions and Notations

Most of the graph theoretic definitions used here appear in [3] and [6]. A graph $G=(V(G),E(G))$ is considered to be loopless and undirected unless otherwise specified. A graph is simple if it has no parallel edges. Some basic notations now follow. A graph G_1 is *smaller* than another graph G_2 if $|V(G_1)|<|V(G_2)|$ or if $|V(G_1)|=|V(G_2)|$ and $|E(G_1)|<|E(G_2)|$. If X is a subset of nodes (edges) then $G \setminus X$ (or $G-X$) denotes the graph obtained from G by removing X.

The set of nodes adjacent to a node v is denoted $N_G(v)$; its degree equals the number of edges incident to it. The minimum degree in G is denoted δ_G. Subscripts of a variable and qualifiers of a certain graph are at times omitted when no confusion can arise. K_n denotes the complete graph on n nodes. The notations '$\subseteq$' and '$\cong$' refer to the subgraph relationship and the isomorphism relationship, respectively.

An edge $e=(u,v)$ is said to be *subdivided* if it is deleted and replaced by two new edges (u,w) and (w,v) incident to a new node w. A graph G' is said to be *homeomorphic from* a graph G if G' can be obtained from G by a possibly empty sequence of edge subdivisions. An edge $e=(u,v)$ is said to be *contracted* in G if e and all of its parallel edges are deleted and its end nodes are identified; the resulting

Manuscript received May 27, 1987; revised October 19, 1987. C. J. Colbourn's research was supported by NSERC Canada under Grant A0579. This paper was recommended by K. Thulasiraman and N. Deo, Guest Editors for this issue.

E. S. El-Mallah was with the University of Waterloo, Waterloo, Ont., Canada. He is now with the Department of Computing Science, University of Alberta, Edmonton, Alta., Canada T6G 2E1.

C. J. Colbourn is with the Departments of Computer Science and Combinatorics and Optimization, University of Waterloo, Waterloo, Ont., Canada.

IEEE Log Number 8718927.

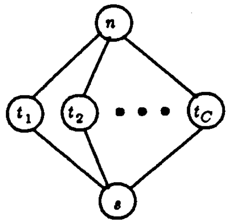
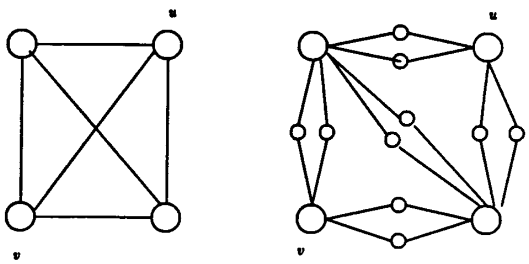


Fig. 1. A family of graphs $D_{n,s}(C)$, $C \geqslant 1$.



(a) (b)

Fig. 2. (a) Gf. (b) $MGf(u,v;2)$.

frequent references are made to particular members of the family $D_{n,s}(C)$, $C \geqslant 1$, illustrated in Fig. 1.

Let Gf be a simple 2-connected graph, $\delta_{Gf} \geqslant 3$, and (u,v) be any edge in Gf. Denote by $MGf(u,v;C)$ the graph obtained from Gf-(u,v) by attaching a copy of $D_{n,s}(C)$ at n and s to each edge (x,y) in Gf-(u,v), at x and y, then removing that edge. Fig. 2 illustrates an example of such a construction.

The following lemma follows easily, but for completeness we provide a proof.

Lemma 1. Let Gf be a simple 2-connected graph, $\delta_{Gf} \geqslant 3$, and (u,v) be any of its edges. Then, for all $C \geqslant 1$, $MGf(u,v;C)$ does not contain a subgraph contractable to Gf.

Proof: Let H be a subgraph of a graph $MGf(u,v;C)$, $C \geqslant 1$, that is contractible to a 2-connected minor H' with minimum degree 3. Denote by R the subset of $E(H)$ in which each edge is incident to some vertex v of degree 2 in H, $v \notin V(Gf)$. Furthermore, let $H \cdot R$ denote the graph obtained from H by contracting all edges in R. Since the degree of every vertex of H' is at least 3, it follows that $H' \leqslant_m H \cdot R$. However, by the construction of $MGf(u,v;C)$, we know that $H \cdot R$ is a proper subgraph of Gf. Hence, any such minor H' is smaller than Gf. □

Now, we are ready to prove our main theorem.

Proof of Theorem 1. Let $(\sigma, \mathbf{S})$ be an arbitrary instance of the exact 3-cover problem ($X3C$), where $\sigma = \{\sigma_1, \sigma_2, \cdots, \sigma_{3t}\}$ and $\mathbf{S} = \{S_i | 1 \leqslant i \leqslant r\}$ is a collection of 3-element subsets of σ. We may assume that $S_i \neq S_j$ for $i \neq j$ and that $r > t > 1$. If the first condition fails we can detect and remove any redundancy in $\mathbf{S}$ in polynomial time. If the latter fails (i.e. if $r \leqslant t$ or $t = 1$) the solution can be computed trivially. We show how to reduce any instance of the $X3C$ to a $H-$EDP($\mathbf{F}$) or a M-EDP($\mathbf{F}$) in polynomial time.

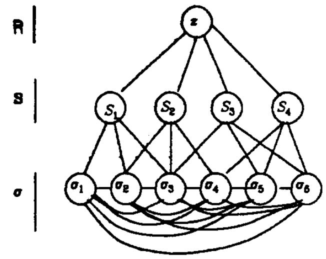


Fig. 3. G_0 corresponding to $\mathbf{S} = \{\{\sigma_1, \sigma_2, \sigma_3\}, \{\sigma_2, \sigma_3, \sigma_4\}, \{\sigma_3, \sigma_5, \sigma_6\}, \{\sigma_4, \sigma_5, \sigma_6\}\}$.

Let Gf, $Gf \in \mathbf{F}$, be a smallest member in $\mathbf{F}$. First, construct a graph G_0 in which every set S_i (element σ_j) is represented by a node labelled $S_i(\sigma_j)$. Let z be a new node. Thus $V(G_0) = \mathbf{R} \cup \mathbf{S} \cup \sigma$, where $\mathbf{R} = \{z\}$, $\mathbf{S} = \{S_i | 1 \leqslant i \leqslant r\}$, and $\sigma = \{\sigma_i | 1 \leqslant i \leqslant 3t\}$. In addition, let $E(G_0) = E_{\mathbf{RS}}(G_0) \cup E_{\mathbf{S}\sigma}(G_0) \cup E_{\sigma\sigma}(G_0)$, where

$$E_{\mathbf{RS}}(G_0) = \{(z, S_i) | 1 \leqslant i \leqslant r\}, \quad |E_{\mathbf{RS}}(G_0)| = r$$

$$E_{\mathbf{S}\sigma}(G_0) = \{(S_i, \sigma_j) | \sigma_j \in S_i\}, |E_{\mathbf{S}\sigma}(G_0)| = 3r$$

$$E_{\sigma\sigma}(G_0) = \{(\sigma_i, \sigma_j) | 1 \leqslant i < j \leqslant 3t\}, |E_{\sigma\sigma}(G_0)| = \binom{3t}{2}.$$

Fig. 3 illustrates the construction.

Second, construct a graph G from G_0 as follows. Let the constant

$$C_{S\sigma} = \binom{3t}{2}$$

be the number of edges in the layer $E_{\sigma\sigma}(G_0)$, denoted by $E_{\sigma\sigma}$. Also, let $C_{\mathbf{RS}}$ equal the sum of all edges in the layer $E_{\sigma\sigma}$ plus the number of edges that would exist if each edge in $E_{\mathbf{S}\sigma}(G_0)$ was replaced by a copy of $D_{n,s}(C_{\mathbf{S}\sigma})$. Thus $C_{\mathbf{RS}} = |E_{\sigma\sigma}| + 3r(2C_{\mathbf{S}\sigma}) = |E_{\sigma\sigma}|(1+6r)$. Further, attach a copy of $D_{n,s}(C_{\mathbf{S}\sigma})$, at n and s, to each edge (S_i, σ_j), at S_i and σ_j, respectively. Then, erase all (S_i, σ_j) edges. Denote the new set of edges by $E_{\mathbf{S}\sigma}$. Thus $|E_{\mathbf{S}\sigma}| = (3r)(2C_{\mathbf{S}\sigma}) = 6r|E_{\sigma\sigma}|$.

Finally, let (u,v) be any edge in Gf. Attach a copy of $MGf(u,v;C_{\mathbf{RS}})$, at u and v, to each edge (z, S_i) of G_0, at z and S_i respectively. Denote such a copy by $MGf(z, S_i; C_{\mathbf{RS}})$. Then erase all (z, S_i) edges. Denote the new set of edges by $E_{\mathbf{RS}}(G)$. Thus $|E_{\mathbf{RS}}| = r(|E(Gf)|-1)(2C_{\mathbf{RS}}) = 2r(|E(Gf)|-1)(|E_{\sigma\sigma}| + |E_{\mathbf{S}\sigma}|)$. Fig. 4 illustrates an example of this step.

Define

$$\gamma(Gf) = \begin{cases} 2t, & \text{if } Gf \cong K_4 \\ 3t, & \text{otherwise.} \end{cases}$$

We now show that $(\sigma, \mathbf{S})$ has an X3C if and only if G has a subgraph $G' = (V', E')$ such that:

$$|E| - |E'| \leqslant 3(r-t)|E_{\sigma\sigma}| + (|E_{\sigma\sigma}| - \gamma(Gf)) \qquad (1)$$

and secondly, if G' contains a subgraph G'' homeomorphic

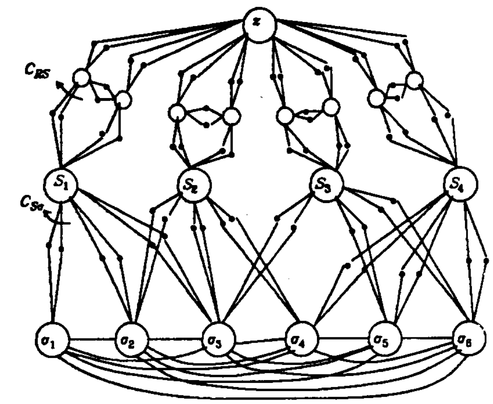


Fig. 4. G.

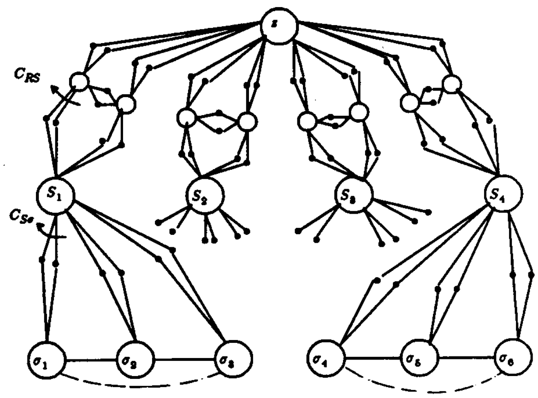


Fig. 5. G'.

from (or contractible to) some 2-connected graph H, $\delta_H \geq 3$, then H is smaller than Gf. This last condition ensures that G' is *F-h-free* (or *F-m-free*, respectively).

($\Rightarrow$) Suppose $(\sigma, \mathbf{S})$ has an $X3C$, say $S' = \{S_1, S_2, \cdots, S_t\}$. If $Gf \cong K_4$, then G has a subgraph G' analogous to the one illustrated in Fig. 5.

If $Gf \neq K_4$ then G' has t additional edges (the dotted lines in Fig. 5), and each such edge is associated with a 3-subset of σ-nodes which appears in the corresponding X3C. It is routine to verify that G' satisfies (1). To satisfy the second condition, it suffices to show that every 2-connected simple minor H, $\delta_H \geq 3$, of G' is smaller than Gf. Note that every 2-connected subgraph G'' of G' is either (1) a subgraph of some $MGf(z, S_i, C_{\mathbf{RS}})$, for some i, or (2) a subgraph that contains at most one S-node and at most three σ-nodes. In either case, one may verify using Lemma 1 that any 2-connected contraction H of G'', $\delta_H \geq 3$, is smaller than Gf.

($\Leftarrow$) Suppose G has a subgraph G' satisfying (1). The proof is to show how to construct a solution to the X3C. Denote $E_{\mathbf{S}\sigma} \cap E'$ and $E_{\sigma\sigma} \cap E'$ by $E'_{\mathbf{S}\sigma}$ and $E'_{\sigma\sigma}$, respectively. First, notice that in G' every σ-node can be reached by at most one S-node using a path of length 2. To see this, let $\mathbf{M}$ be the set of vertices incident to edges in $E_{\mathbf{RS}}(G)$. In addition, denote by $\mathbf{T}$ the subset of $\mathbf{M}$ having degree 2 in G. Then observe that if (u, w), $(w, v) \in E(G)$ where $u, v \in \mathbf{M} \backslash \mathbf{T}$ and $w \in \mathbf{T}$, then some (u, v)-path of length 2 appears in E'. Otherwise,

$$|E| - |E'| \geq C_{\mathbf{RS}} = |E_{\sigma\sigma}|(1 + 6r) > |E_{\sigma\sigma}|(1 + 3(r - t)) - \gamma(Gf)$$

a contradiction. Now, in E' there may be no (S_i, σ_j, S_k)-path, contained in $E'_{\mathbf{S}\sigma} \cup E'_{\sigma\sigma}$, joining any two distinct nodes S_i and S_k. Otherwise, G' has a homeomorph from Gf whose edges are contained in $\{MGf(z, S_i, C_{\mathbf{RS}}) \cup MGf(z, S_k, C_{\mathbf{RS}})\}$ plus that particular (S_i, σ_j, S_k)-path.

Second, notice that every σ-node, say σ_i, is reached from some node in $\mathbf{S}$, say S_j, by some path of length 2. To derive a contradiction, assume that the number of σ-nodes that are reachable in that way is strictly less than $3t$. Then,

$$|E| - |E'| \geq C_{S\sigma}(1 + 3(r - t)) > |E_{\sigma\sigma}|(1 + 3(r - t)) - \gamma(Gf)$$

a contradiction. Thus G' satisfies (1) only if $|E'_{\sigma\sigma}| \geq \gamma(Gf)$. It remains to show that this latter condition, together with the two conditions mentioned above, imply the existence of a solution to the X3C. For this purpose, denote by a_i, $0 \leq i \leq 3$, the number of nodes of $\mathbf{S}$, each of which reaches exactly i σ-nodes in G' by a path of length two. Thus showing that $a_0 = r - t$, $a_1 = a_2 = 0$, and $a_3 = t$ in G' imply a solution for the X3C. For $Gf \neq K_4$, $|E'_{\sigma\sigma}| \geq \gamma(Gf)$ only if

$$a_2 + 3a_3 \geq 3t \tag{2.a}$$

otherwise, $|E'_{\sigma\sigma}| \geq \gamma(Gf)$ only if

$$a_2 + 2a_3 \geq 2t. \tag{2.b}$$

In addition, since each σ-node is reachable from some S-node, we have

$$a_1 + 2a_2 + 3a_3 = 3t. \tag{3}$$

Furthermore, at least t nodes of $\mathbf{S}$ are required to cover all of the σ-nodes. Thus

$$a_1 + a_2 + a_3 \geq t. \tag{4}$$

It can be seen that (2.a) and (3) are satisfied only if $a_3 = t$. For the second case, when $Gf \cong K_4$, subtracting (4) from (3) yields $a_2 + 2a_3 \leq 2t$, which implies that (2.b) can only be satisfied as an equation. Thus

$$a_2 + 2a_3 = 2t. \tag{2.b)'$$

Multiplying (2.b)' by 3/2 and subtracting the result from (3) yields $a_1 + (1/2)a_2 = 0$. Thus $a_1 = a_2 = 0$ and $a_3 = t$. This completes the proof. □

The above theorem strengthens a theorem by [13] and extends another theorem by [2]. It reproves, in a unified setting the NP-completeness for finding largest subgraphs that are planar graphs, outerplanar graphs, series-parallel graphs and Δ-Y reducible graphs (e.g., [2], [12], [13], [16]).

V. The Complexity of the EDP on Cactuses

A *cactus* is a graph having no homeomorph from $K_4 - e$, the complete graph on four nodes missing one edge. Cactuses have a tree-like structure as indicated in Lemma 2 that can be easily verified. Moreover, all cactuses are series-parallel graphs. Throughout this section, $C(G)$ denotes the set of all cycles of a graph G.

Lemma 2. The following are equivalent.

1) $G = (V, E)$ is a cactus.

2) Every 2-connected component of G is either an edge or a simple cycle.

3) G is a graph having $c(G)$ components, $|C(G)|$ cycles and $|E| = |V| - c(G) + |C(G)|$ edges.

The EDP on cactuses is motivated since the EDP on trees (i.e., H-EDP($\{K_3\}$)) is trivial but the EDP on series-parallel graphs (i.e., H-EDP($\{K_4\}$)) is NP-complete. Thus, knowing the complexity status of this problem on cactuses becomes of interest. We make the following notes. First, a maximum cactus subgraph of any connected graph is a connected spanning subgraph. Lemma 2(3) then implies that the number of edges of a maximum cactus subgraph of a given graph is defined by the number of cycles contained in it. Moreover, in a search for a maximum cactus subgraph G' of a graph G one may restrict the attention to cactuses in which each cycle does not have a chord (an edge between two nonconsecutive nodes of the cycle) in $E(G) \setminus E(G')$. This latter observation can be verified by constructing a new cactus G'' in which every cycle is chordless from a given cactus G', $|E(G'')| = |E(G')|$, which violates this property. The main theorem of this section now follows.

Theorem 2. EDP($\{K_4 - e\}$) is NP-complete.

Proof: Membership in NP is straightforward. It remains to show that the problem is NP-hard. Let $F(X) = \bigwedge_{q=1}^{k} C[q]$ be an arbitrary instance of the 3-satisfiability problem (3-SAT) on r variables $X = \{x_i | 1 \leqslant i \leqslant r\}$ (or $2r$ *literals*). We may assume that $C[i] \neq C[j]$ for $i \neq j$, $1 \leqslant i, j \leqslant k$, and that literals of each clause correspond to distinct variables. A graph $G = (V, E)$ is constructed as follows. Each variable in X is represented by a graph $GA_i[4k]$ isomorphic to the one illustrated in Fig. 6, called a gadget hereafter. Any $GA_i[4k]$ contains $2k$ 4-sided faces called the *truth* faces. A truth face containing a b-node is called a *true* face otherwise it is a *false* face. In addition, there are $2k$ 3-sided faces called *isolation* faces. Edges incident to the *center* node r_i are called *rays*.

Each clause $C[q]$, $1 \leqslant q \leqslant k$, is represented by a new edge $(A[q], B[q])$ incident to the two new nodes $A[q]$ and $B[q]$. Moreover, node $A[q]$ is made adjacent to node $a_i[q]$ whenever variable x_i appears in clause $C[q]$. Finally, node $B[q]$ is made adjacent to node $b_i[q]$ ($\bar{b}_i[q]$) whenever literal $\bar{x}_i(x_i)$ appears in $C[q]$. Thus associated with each clause there are three 4-sided chordless cycles: $\{(A[q], a_i[q], b_i[q] (\text{or } \bar{b}_i[q]), B[q], A[q] : \bar{x}_i \text{ (or } x_i) \in C[q])\}$ called *clause*-cycles. Thus $|V| = r(6k+1) + 2k$ and $|E| = k(10r + 7)$. For convenience, a boldface letter is used to denote the set of nodes that are basically labelled by that letter (e.g., $\mathbf{A} = \bigcup_{q=1}^{k} A[q]$). Denote by E_{GG} the subset of edges incident to some clause-node and some gadget-node. Fig. 7 illustrates a construction of G for a given $F(X)$. We next show that F is satisfiable if and only if G has a cactus subgraph G' containing at least $k(2r+1)$ cycles.

($\Rightarrow$) Let T be a satisfying truth assignment for F. G has a connected cactus G' containing k true (false) faces and k isolation faces of GA_i whenever variable x_i is true (false) in T. In addition, since T satisfies F at least one literal, say $x_i(\bar{x}_i)$, in clause $C[q]$ satisfies $C[q]$. G' includes the following clause-cycle: $(A[q], \mathrm{Ia}_i[q], b_i[q]$ (or $\bar{b}_i[q]$), $B[q], A[q])$; literal $\bar{x}_i$ (or x_i) satisfies $C[q]$. Fig. 8 il-

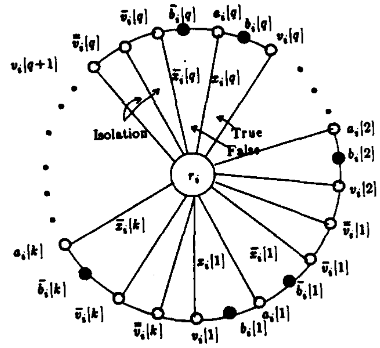


Fig. 6. A family of graphs $GA_i[4k]$.

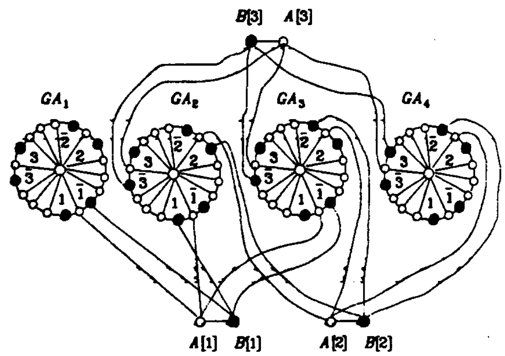


Fig. 7. A graph constructed from $F = (x_1 \vee \bar{x}_2 \vee x_3)(\bar{x}_2 \vee x_3 \vee x_4) \wedge (x_2 \vee x_3 \vee \bar{x}_4)$, $k = 3$ and $r = 4$.

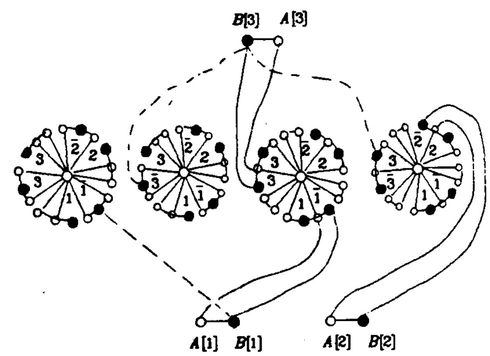


Fig. 8. A cactus corresponding to the assignment $x_1 = x_2 = false$, $x_3 = x_4 =$ true containing $k(2r+1) = 3(8+1) = 27$ cycles.

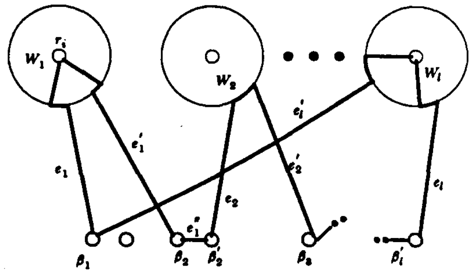


Fig. 9. The structure of a gadget-to-clause cycle.

lustrates one possible solution for the example illustrated in Fig. 7. The dotted edges are chosen arbitrary to connect isolated components of solid edges.

For purpose of showing the other direction in the proof it is convenient to introduce the following definitions. Recall that a cycle is called *chordless* in G' if it has no chord in $E(G)\setminus E(G')$. Call a cycle C of G which contains at least one edge from the set E_{CG} a *gadget-to-clause* cycle. Note that any such cycle contains an even number of edges of E_{CG}. Denote by $C(l)$ the set of all such cycles in G in which each cycle contains exactly $2l$ edges from E_{CG}. One may verify that any cycle $C, C \in C(l)$, can be written as (see Fig. 9):

$$C=(\beta_1, e_1, W_1, e_1', \beta_2, e_1'', \beta_2', e_2, W_2, \cdots, e_l, W_l, e_l', \beta_1), \quad 1 \leqslant l \leqslant k.$$

Here,

$$\beta_i \in A \cup B, \qquad 1 \leqslant i \leqslant l$$

$$e_i, e_i' \in E_{CG}, \qquad 1 \leqslant i \leqslant l$$

W_i, $\quad i \leqslant 1 \leqslant l$, *is a walk whose elements belong to one particular gadget*

$$e_i'' \in \{\phi\} \cup \{(A[q], B[q]) : 1 \leqslant q \leqslant k\}$$

and

$$\beta_i' \in A \cup B, \qquad 1 \leqslant i < l, \text{ or } \phi \text{ if } e_i'' = \phi.$$

Moreover, we say that a cactus subgraph G' of G *defects* an edge e if either $e \in E(G')$ or e appears as a cut-edge in every super cactus G'' of G'.

Claim I. Let C be any chordless gadget-to-clause cycle in $C(l)$ then

1) If $l=1$ then C is a clause cycle.
2) If $l \geqslant 2$ then C defects at least $2l$ rays in G.

Proof of Claim I:

(1) If $l=1$ then $C_l=(\beta_1, e_1, W_1, e_1', \beta_2, e_1'', \beta_1)$. Thus β_1 and β_2 are the two nodes representing some clause, say $C[q]$. Let GA_i be the host of W_1. Either $(a_i[q], b_i[q]) \in E(W_1)$ or $(a_i[q], \bar{b}_i[q]) \in E(W_1)$, and in each case there is nothing to prove. Otherwise, C_l would have a chord in $E(G)\setminus E(C_l)$, a contradiction.

(2) It is sufficient to show that each walk W_i of C defects at least two rays from its hosting gadget. Observe that, for $l \geqslant 2$, the first and the last elements of W_i are two $\{a, b, \bar{b}\}$-nodes belonging to two truth faces of two different clauses, otherwise, C has a chord. Figure 10 depicts two possible cases of W_i. The remaining cases are similar.

Claim II. If GA_i has at least $2k_i', k_i' \leqslant 2k$, defective rays by some gadget-to-clause cycles in some cactus G' then $|C(G\bar{A}_i \cap G')| \leqslant 2k - k_i'$.

Proof of Claim II: Any cycle in $C(GA_i \cap G')$ uses exactly two rays and no ray may appear in more than one cycle. Thus

$$|C(GA_i \cap G')| \leqslant \frac{4k-2k_i'}{2} = 2k - k_i'. \qquad \square$$

Now, we are ready to prove the other direction.

($\Leftarrow$) Suppose G has a cactus subgraph G' having at least $k(2r+1)$ chordless cycles. Let $C(l, G')$ be the subset of $C(l)$ in G'. Thus

$$C(G') = \bigcup_{i=1}^{r} C(GA_i \cap G') \bigcup_{l=1}^{k} C(l, G').$$

It is sufficient to show that

$$|C(GA_i \cap G')| = 2k, \qquad 1 \leqslant i \leqslant r \tag{1}$$

$$|C(1, G')| = k \tag{2}$$

and $|C(l, G')| = 0$, $2 \leqslant l \leqslant k$.

Equation (1) ensures that every cycle included in G' is either a true (false) face or an isolation face of GA_i. A truth assignment of variable x_i is then chosen to agree with the truth faces selected. In addition, (2) ensures that exactly one clause cycle is selected for each clause. This proves that $F(X)$ is satisfiable. Let $2k_i', k_i' \leqslant 2k$, be the largest number not exceeding the number of defective rays in GA_i caused by $\cup_{l=2}^{k} C(l, G')$. G' is a solution only if

$$|C(G')| = \sum_{i=1}^{r} |C(GA_i \cap G')| + \sum_{l=1}^{k} |C(l, G')| \geqslant k(2r+1).$$

However, by Claim II we have

$$\sum_{i=1}^{r} |C(GA_i \cap G')| \leqslant \sum_{i=1}^{r} (2k - k_i') = 2kr - \sum_{i=1}^{r} k_i'.$$

In addition, by Claim I(1) and the assumption that the cycles of G' are chordless we know that every cycle in $C(1, G')$ is a clause-cycle. Hence, $|C(1, G')| \leqslant k$. We then may conclude that

$$\sum_{l=2}^{k} (|C(l, G')|) - \sum_{i=1}^{r} k_i' \geqslant 0. \tag{3}$$

However, by Claim I(2) we have

$$\sum_{i=1}^{r} 2k_i' \geqslant \sum_{l=2}^{k} 2l|C(l, G')|.$$

This implies

$$\sum_{i=1}^{r} k_i' \geqslant 2 \sum_{l=2}^{k} |C(l, G')|.$$

Hence, (3) holds if and only if $\Sigma_{l=2}^{k} |C(l, G')| = 0$. Consequently, $k_i' = 0$, $1 \leqslant i \leqslant r$. Equations (1) and (2) now follow

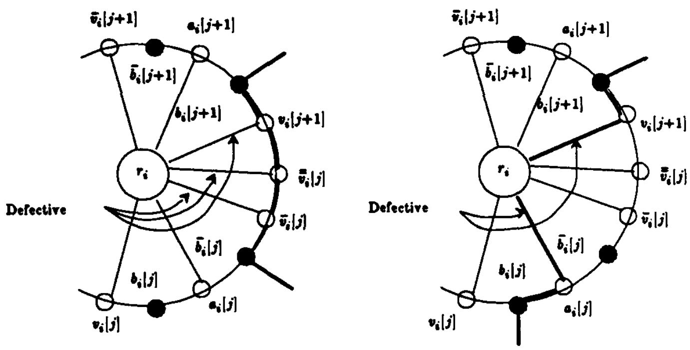


Fig. 10. Defective rays corresponding to a walk of some gadget-to-clause cycle through a gadget.

since

$$|C(GA_i \cap G')| \leqslant 2k, \qquad 1 \leqslant i \leqslant r, \text{ and } C(1, G') \leqslant k.$$

This concludes the proof. □

VI. The Complexity of a Class of I-EDP's

A significant number of well-known classes of graphs can be characterized by a set of forbidden induced subgraphs. Examples include, bipartite, claw-free, cographs [4], connected and degree constrained, domishold, indifference, interval graphs, edge graphs, strongly chordal, threshold, and transitive series–parallel directed acyclic graphs. However, little is known about the complexity status of the corresponding I-EDP in each case. Yannakakis [16] has shown that the EDP for degree constrained graphs and edge graphs are NP-hard. None of the above-mentioned cases can be handled by the result obtained in [2]. In this section, we consider the EDP's on graphs having no induced path P_l on l nodes.

For $l = 2$, any feasible solution is empty. Complications arise for $l > 2$, however. For $l = 3$, any component in any feasible solution is a clique. In this case, one can show that this problem is NP-complete using the reduction to the partitioning into triangles problem mentioned in [5]. For $l > 3$, the complexities of different $I - \text{EDP}(\{P_l\})$ do not seem to follow directly from other known results. The goal of this section is to show that these problems are NP-complete for any $l \geqslant 3$. The key to the proof is the following family of graphs, called gadgets throughout this section. Let P_l be a simple path on l nodes, say $(v[1], v[2], \cdots, v[l])$, α and β be two positive integers. Construct a graph $G[l, (\sigma[i], \sigma[j], \sigma[k]), \alpha, \beta]$ by replacing each node $v[i]$ in P_l by a clique $Q[i]$. In particular, $Q[l] = (\sigma[i], \sigma[j], \sigma[k])$, $|V(Q[l-1])| = \alpha$ and $|V(Q[i])| = \beta$, $1 \leqslant i \leqslant l-2$.

Furthermore, add the necessary edges to join every node in Q_i to every node in Q_{i+1}, $1 \leqslant i \leqslant l$. Fig. 11 illustrates this family of graphs. We need the following preliminary lemma.

Lemma 3. For any α, $\beta \geqslant 1$ and $l \geqslant 2$, the graph $G[l, (\sigma[i], \sigma[j], \sigma[k]), \alpha, \beta]$ is P_{l+1}-i-free.

Proof: To derive a contradiction, assume $H \cong P_{l+1}$, $H \subseteq G[l, (\sigma[i], \sigma[j], \sigma[k]), \alpha, \beta]$. At least two nodes of H, say $v[i]$ and $v[i+1]$, lie in the same clique, say $Q[i]$. Let v be a neighbour of $v[i]$ ($v[i+1]$) other than $v[i+1]$($v[i]$). The subgraph induced by $v, v[i]$ and $v[i+1]$ is a K_3, a contradiction. □

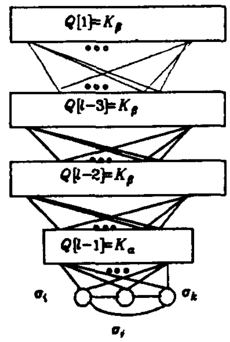


Fig. 11. $G[l, (\sigma[i], \sigma[j], \sigma[k]), \alpha, \beta]$.

We are now ready to present the main theorem of this section.

Theorem 3. For any integer l, $l \geqslant 2$, the $I - \text{EDP}(\{P_{l+1}\})$ is NP-hard.

Proof: Let $(\sigma, \mathbf{S})$ be an arbitrary instance of the exact 3-cover problem (X3C). Where, $\sigma = \{\sigma[1], \sigma[2], \cdots, \sigma[3t]\}$ and $\mathbf{S} = \{S_i | 1 \leqslant i \leqslant r\}$ is a collection of 3-element subsets of σ. One may assume that $S_i \neq S_j$ for $i \neq j$ and that $r > t > 1$. Similarly, when no confusion arises, the nodes of the constructed graph and the elements of $(\sigma, \mathbf{S})$ share the same names. The proof is to reduce $(\sigma, \mathbf{S})$ to an instance of $I - \text{EDP}(\{P_{l+1}\})$ in polynomial time. Construct a graph $G = (V, E)$ as follows. First, create a clique of $3t$ σ-nodes, call it the σ-clique. Set constants $\alpha = \binom{3t}{2}$ and $\beta = r^2\alpha$. For every 3-element subset $S_x = \{\sigma[i], \sigma[j], \sigma[k]\}$ of $\mathbf{S}$, attach a copy, G_x, of $G[l, (\sigma[i], \sigma[j], \sigma[k]), \alpha, \beta]$ to the σ-clique by identifying $\sigma[i]$, $\sigma[j]$ and $\sigma[k]$ to their counterparts in the clique. Denote by $Q_x[i]$ the clique $Q[i]$ in G_x. Furthermore, denote its nodes by $v_x[i,1]$, $v_x[i,2], \cdots, v_x[i,\theta]$ where $1 \leqslant i \leqslant l-1$ and $\theta = \alpha$ if $i = l-1$ or β, otherwise. Finally, denote the set of edges having one node in one set, say X, and the other in another set, say Y, by E_{XY}. Thus $V = \mathbf{Q} \cup \sigma$, where, $\mathbf{Q} = \cup V(Q_x[i])$, $1 \leqslant x \leqslant r$, $1 \leqslant i \leqslant l-1$ and $E = E_{QQ} \cup E_{Q\sigma} \cup E_{\sigma\sigma}$. Thus $|\mathbf{Q}| = r(\alpha + (l-2)\beta)$ and $|\sigma| = 3t$. Fig. 12 illustrates an example of this construction. We show that $(\sigma, \mathbf{S})$ has an *X3C* if and only if G has a

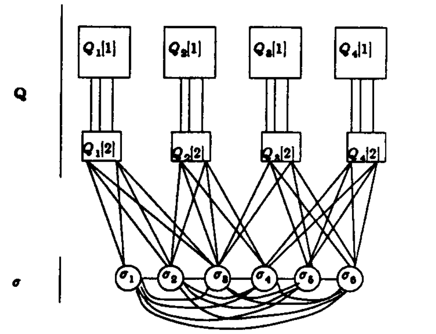


Fig. 12. A graph G corresponding to the instance: $\mathbf{S} = \{\{\sigma_1, \sigma_2, \sigma_3\}, \{\sigma_2, \sigma_3, \sigma_4\}, \{\sigma_3, \sigma_5, \sigma_6\}, \{\sigma_4, \sigma_5, \sigma_6\}\}$, and $l = 3$.

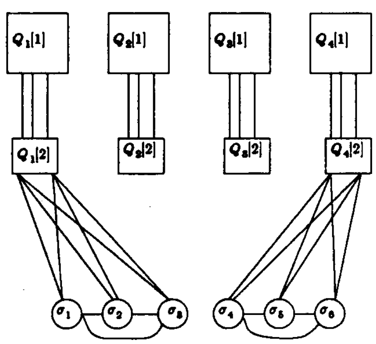


Fig. 13. A P_4-i-free subgraph corresponding to the solution $X3C = \{S_1, S_4\}$.

P_{l+1}-i-free subgraph $G' = (V, E')$ satisfying:

$$|E| - |E'| \leqslant 3(r-t)\alpha + (\alpha - 3t). \tag{1}$$

($\Rightarrow$) Suppose $(\sigma, \mathbf{S})$ has an exact 3-cover, say $S' = \{S_1, S_2, \cdots, S_t\}$. Then G has a P_{l+1}-i-free subgraph G' analogous to the one illustrated in Fig. 13. In G all of the 3α edges of any set $E_{Q_x S_x}$, $t+1 \leqslant x \leqslant r$, are missing and there are $(r-t)$ such sets. In addition, all edges of the clique induced by $E_{\sigma\sigma}$ are not in G', except $3t$ edges which appear in the t gadgets. Thus G' satisfies (1).

($\Leftarrow$) Suppose G has a P_{l+1}-i-free subgraph G' satisfying (1). We show how to construct an $X3C$ of $(\sigma, \mathbf{S})$. First, we establish the following two claims.

Claim 1. Let G_x be any gadget in G and $v_x[l-1, i_1]$, $1 \leqslant i_1 \leqslant \alpha$, be any node in the α-clique $Q_x[l-1]$. Then, there exists an induced path $P = (v_x[l-1, i_1], v_x[l-2, i_2], \cdots, v_x[1, i_{l-1}])$ isomorphic to P_{l-1} in G'.

Proof of Claim 1: To derive a contradiction, let $v_x[l-j, i_j]$, $1 \leqslant j \leqslant l-2$, be the first node in a path $P' = (v_x[l-1, i_1], \cdots, v_x[l-j, i_j])$ having no adjacent node in the β-clique $Q_x[l-j-1]$ in G'. By construction of G, $v_x[l-j, i_j]$ has $(r^2\alpha)$ adjacent nodes in the β-clique $Q_x[l-j-1]$. Thus G' satisfies (1) only if

$$\begin{aligned} r^2\alpha &\leqslant 3(r-t)\alpha + (\alpha - 3t) \\ &< \alpha(1 + 3(r-t)). \end{aligned}$$

However, for $t \geqslant 1$ the above inequality simplifies to $r^2 < 3r - 2$ which is false for all $r \geqslant 1$.

Claim 2. Let G'' be any component of G'. Then, $V(G'') \subseteq V(G_x)$ for some gadget G_x of G.

Proof of Claim 2: To derive a contradiction, assume $V(G'') \cap V(G_x) \neq \phi$ and $V(G'') \setminus V(G_x) \neq \phi$. Let $U = V(G'') \cap V(Q_x[l-1])$. In addition, let $U_1 = V(G'') \cap V(S_x)$. Moreover, let U_2 be the neighbours of U_1 in G'' other than those contained in $V(G_x)$. Thus, members of U_2 are either σ-nodes, other than S_x, or $Q_y[l-1]$-nodes of some gadget G_y, $x \neq y$. However, G'' being connected with $V(G'') \setminus V(G_x) \neq \phi$ imply that $U_2 \neq \phi$. Thus one may conclude that G'' contains a path $P' = (v_x[l-1, i], \sigma_j, z)$, where $\sigma_j \in U_1$ and $z \in U_2$. By Claim 1, there exists in G'' a path P, having $v_x[l-1, i]$ as an origin, isomorphic to P_{l-1}. In addition, P includes exactly one node from each possible β-clique Q_h, $1 \leqslant h \leqslant l-2$. Thus $P + P'$ induces a path of $l+1$ nodes in G'', a contradiction.

Denote by a Q-component any component of G' that contains at least one Q-node. The last step is to show that G' has exactly t Q-components, each one contains exactly 3 σ-nodes. By Claim 2 such a collection of t Q-components uniquely defines a solution for the $X3C$. For this purpose, let a_i, $0 \leqslant i \leqslant 3$, be the number of Q-components of G', each one contains exactly i σ-nodes. Furthermore, let b be the remaining number of σ nodes, (i.e. $b = 3t - (a_1 + 2a_2 + 3a_3)$). Our goal is to show that in G': $a_3 = t$, $a_1 = a_2 = b = 0$ and $a_0 = (r-t)$. In G' we have $|E| - |E'| \geqslant (|E_{Q\sigma}| - |E'_{Q\sigma}|) + (|E_{\sigma\sigma}| - |E'_{\sigma\sigma}|)$. However, $|E'_{Q\sigma}|$ is maximum when every σ-node that appears in any Q-component of G' is adjacent to exactly α Q-nodes. i.e. $(|E_{Q\sigma}| - |E'_{Q\sigma}|) \geqslant 3r\alpha - (3t-b)\alpha = \alpha(3(r-t) + b)$. Similarly, $|E'_{\sigma\sigma}|$ is maximum when every group of σ-nodes contained in each Q-component of G' induce a clique. In addition, the remaining b σ-nodes induces one clique. Thus, G' satisfies (1) only if

$$\alpha(3(r-t) + b) + \left(\alpha - \binom{b}{2} - a_2 - 3a_3\right) \leqslant \alpha(1 + 3(r-t)) - 3t.$$

Simplifying the above inequality gives

$$b\alpha + 3t + \frac{1}{2}b \leqslant \frac{1}{2}b^2 + a_2 + 3a_3. \tag{2}$$

In addition, the structure of G' implies

$$a_1 + 2a_2 + 3a_3 + b = 3t. \tag{3}$$

Observe that

$$a_2 + 3a_3 \leqslant 3t \text{ and } \frac{1}{2}b^2 < b\alpha + \frac{1}{2}b \text{ hold in (2).}$$

We conclude that (2) holds if and only if $b = 0$. Using $b = 0$ to simplify (2) further, we may write

$$3t \leqslant a_2 + 3a_3. \tag{4}$$

However, (3) and (4) have a unique solution when $a_2 = 0$ and $a_3 = t$. This completes the proof. $\square$

Corollary 1. The EDP's corresponding to cographs and transitive series–parallel directed acyclic graphs (TSP dags) and the completion problems to cographs and TSP dags are NP-complete.

Proof: NP-completeness of the first two problems follows from Theorem 3. It is known that cographs ($\{P_4\}$-i-free graphs) are closed under complementation. Furthermore, a graph is TSP dag if and only if its underlying undirected graph is a cograph. Thus deciding whether there exists a set of at most k edges (arcs) whose addition to a graph (dag) G will result in a cograph (TSP dag) G' amounts to solving the corresponding EDP on the complement of G. □

VII. Concluding Remarks

We observe that most of the EDP's are NP-hard. So far, versions of the edge deletion problem that are known to have efficient algorithms include computing a spanning tree ($H-\text{EDP}(\{K_3\})$), using a simple greedy algorithm, and finding a maximum matching ($H-\text{EDP}(\{P_3\})$) in a graph and some variations thereof. Therefore, it seems worthwhile to determine the complexity of the EDP's corresponding to other known classes of graphs. It also seems that the variation of the NDP that calls for finding a subset of nodes whose deletion results in a subgraph G' having a prescribed property and such that $|E'| \geqslant k$, for a given constant k, does not seem to have received an equal amount of research.

References

[1] T. Asano, "An application of duality to edge-deletion problems," *SIAM J. Comput.*, vol. 16, no. 2, pp. 312–331, Apr. 1987.

[2] T. Asano and T. Hirata, "Edge-deletion and edge-contraction problems," in *Proc. Fourteenth Ann. ACM Symp. on Theory of Computing*, pp. 245–254, 1982.

[3] J. A. Bondy and U. S. R. Murty, *Graph Theory with Applications*. New York: MacMillan, 1976.

[4] D. G. Corneil, H. Lerchs, and L. B. Stewart, "Complement reducible graphs," *Disc. Applied Math.*, vol. 3, pp. 163–174, 1981.

[5] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. San Francisco, CA: Freeman, 1979.

[6] F. Harary, *Graph Theory*. Reading MA: Addison-Wesley, 1970.

[7] D. S. Johnson, "The NP-completeness column: An ongoing guide," *J. Algorithms*, vol. 6, pp. 434–451, 1984.

[8] M. S. Krishnamoorthy and N. Deo, "Node-deletion NP-complete problems," *SIAM J. Comput.*, vol. 8, pp. 619–625, 1979.

[9] E. L. Lawler, *Combinatorial Optimization: Networks and Matroids*. New York: Holt, Reinhart and Winston, 1976.

[10] J. M. Lewis and M. Yannakakis, "The node-deletion problem for hereditary properties is NP-complete," *J. Comput. Systems Sci.*, vol. 20, pp. 219–230, 1980.

[11] J. M. Lewis, D. P. Dobkin, and R. J. Lipton, "Graph properties defined by a forbidden subgraph," in *Proc. 1977 Conf. on Info. Sci. and Systems*, Johns Hopkins Univ., Baltimore, MD., pp. 108–112.

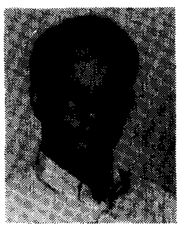
[12] P. C. Liu and R. C. Geldmacher, "On the deletion of nonplanar edges of a graph," Tech. Report, Stevens Institute of Technology, 1976.

[13] T. Watanabe, T. Ae, and A. Nakamura, "On the Removal of Forbidden graphs by Edge-Deletion or by Edge-Contraction" *Discrete Applied Math.*, vol. 3, pp. 151–153, 1981.

[14] ——, "On the NP-hardness of Edge-Deletion and -Contraction Problems" *Discrete Applied Math.*, vol. 6, pp. 63–78, 1983.

[15] M. Yannakakis, "Node- and edge-deletion NP-complete problems," in *Proc. Tenth ACM Symp. on Theory of Computing*, 1978.

[16] M. Yannakakis, "Edge-deletion problems," *SIAM J. Comput.*, vol. 10, no. 2, pp. 297–309, 1981.



Ehab S. El-Mallah received the B.E. degree from the University of Alexandria, Egypt, in 1978 in computer engineering and automatic control, the M.Sc. degree from the University of Saskatchewan, Saskatchewan, in 1984), and the Ph.D. degree from the University of Waterloo, Ont., in 1987 in computer science.

He is presently an assistant professor in the Department of Computing Science at the University of Alberta. His research interests include computational complexity and the design and analysis of combinatorial algorithms for problems in computer networks, digital circuits design and distributed processing.

Charles J. Colbourn received the B.Sc. degree from the University of Toronto, in 1976, the M. Math. degree from the University of Waterloo, in 1978, and Ph.D. degree from the University of Toronto, 1980 degrees in computer science.

From 1980 until 1983, he was assistant, and later associate, professor in computational science at the University of Saskatchewan. Since 1984, he has been an associate professor in combinatorics and optimization, and computer science, at the University of Waterloo. Dr. Colbourn is the author of *The Combinatorics of Network Reliability* and co-edited *Algorithms in Combinatorial Design Theory* and *Combinatorial Design Theory*. He is the author of more than eighty refereed journal papers in computer networks, graph algorithms, and combinatorial design theory. He also serves as an associate editor for *Networks* and the *Journal of Combinatorial Mathematics and Combinatorial Computing*.