Available online at www.sciencedirect.com

ScienceDirect

Electronic Notes in Theoretical Computer Science

Electronic Notes in Theoretical Computer Science 346 (2019) 393–400

www.elsevier.com/locate/entcs

A General Method for Forbidden Induced Subgraph Sandwich Problem NP-completeness [1]

Simone Dantas[2]

IME, Universidade Federal Fluminense, Brazil.

Celina M. H. de Figueiredo[2]

COPPE, Universidade Federal do Rio de Janeiro, Brazil.

Priscila Petito[2]

FFP, Universidade do Estado do Rio de Janeiro, Brazil.

Rafael B. Teixeira[2]

ICE, Universidade Federal Rural do Rio de Janeiro, Brazil.

Abstract

We consider the sandwich problem, a generalization of the recognition problem introduced by Golumbic and Shamir (1993), with respect to classes of graphs defined by excluding induced subgraphs. The Π graph sandwich problem asks, for a pair of graphs $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$ with $E_1 \subseteq E_2$, whether there exists a graph $G = (V, E)$ with $E_1 \subseteq E \subseteq E_2$ that satisfies property Π. We consider the property of being H-free, where H is a fixed graph. Using a new variant of the SAT problem, we present a general framework to establish the NP-completeness of the sandwich problem for several H-free graph classes which generalizes the previous strategy for the class of Hereditary clique-Helly graphs. We also provide infinite families of 3-connected special forbidden induced subgraphs for which each forbidden induced subgraph sandwich problem is NP-complete.

Keywords: algorithms and computational complexity, graph sandwich problems, satisfiability, Linear CNF-formula, 3-SAT, forbidden induced subgraph

[1] This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, CNPq and FAPERJ (Brazilian research agencies)

[2] Email: `sdantas@id.uff.br`, `celina@cos.ufrj.br`, `ppetito@uerj.br`, `rafaelbt@ufrrj.br`

https://doi.org/10.1016/j.entcs.2019.08.035

1571-0661/

1 Introduction

All graphs considered here are finite and undirected. Given a graph property Π, the Π GRAPH SANDWICH PROBLEM is defined as follows:

> *Input:* A pair (G_1, G_2) of graphs with $G_1 = (V, E_1)$, $G_2 = (V, E_2)$ and $E_1 \subseteq E_2$;
> *Question:* Is there a graph $G = (V, E)$ with $E_1 \subseteq E \subseteq E_2$ that satisfies property Π?

The graph sandwich problem was introduced by Golumbic and Shamir [6]. Remark that when $G_1 = G_2$ the problem is to decide whether G_1 satisfies property Π. So the graph sandwich problem generalizes the problem of deciding whether a graph satisfies a given property. In particular, if the decision problem is NP-complete, then the sandwich problem is also NP-complete. When the property Π is to belong to a class $\mathcal{C}$ of graphs, we call this problem the $\mathcal{C}$ *graph sandwich problem.* Golumbic, Kaplan and Shamir [7,8] proved that the interval graph, unit interval graph, permutation graph and comparability graph sandwich problems are all NP-complete; whereas the split graph, threshold graph and cograph sandwich problems are in P.

For an instance (G_1, G_2) of the graph sandwich problem with $G_1 = (V, E_1)$, $G_2 = (V, E_2)$ and $E_1 \subseteq E_2$, we say that any element of E_1 is a *forced edge*, any element of $E_2 \setminus E_1$ is an *optional edge*, and any other pair of $V \times V$ is a *forbidden edge*. Every graph $G = (V, E)$ with $E_1 \subseteq E \subseteq E_2$ is called a *sandwich graph* for the pair (G_1, G_2). In this case, E consists of all forced edges plus some (possibly zero) optional edges and no forbidden edge.

We say that a graph G *contains* a graph H if some induced subgraph of G is isomorphic to H. A graph G is H*-free* if it does not contain H. Dantas, de Figueiredo, da Silva and Teixeira [2], and Dantas, de Figueiredo, Maffray and Teixeira [3] have studied the H-free graph sandwich problems, determining the complexity status of the problem for several graphs H (paw and $(K_p \setminus e)$, for every fixed $p \geq 4$, are in P; whereas C_p, for every fixed $p \geq 4$, claw and bull are NP-complete). More recently, Couto, Faria, Gravier and Klein [1], and de Figueiredo and Spirkl [5] further investigated H-free graph sandwich problems and compared the complexities to probe problems, a variation of sandwich problems where the optional edges occur between vertices of a special subset of V.

Dourado, Petito, Teixeira and de Figueiredo [4] proved the NP-completeness of the Hereditary clique-Helly graph sandwich problem, where the class of Hereditary clique-Helly graphs is defined by a set of four forbidden induced subgraphs, the so-called ocular graphs. Here, we develop this study by providing a general method to prove NP-completeness, which generalizes their strategy, for infinite families of forbidden subgraphs.

In order to complete this task, we have introduced a suitable variant of the NP-complete Linear conjunctive normal form 3-SAT problem (LCNF 3-SAT) [9], which we call k-GIRTH LCNF 2-3-SAT.

2 k-GIRTH LCNF 2-3-SAT

We begin this section by proposing a suitable variant of the NP-complete LCNF 3-SAT [9], which we call k-girth linear conjunctive normal form 2-3-SAT problem (k-GIRTH LCNF 2-3-SAT). According to [9], in the linear conjunctive normal form 3-SAT problem (LCNF 3-SAT) the clauses have size 3 and each pair of distinct clauses have 0 or 1 variable in common.

Let $I = (X, C)$ be an instance of 3-SAT, where X denotes the set of variables, and C denotes the set of clauses. The bipartite graph of clauses and variables $B(I) = (V^I, E^I)$ is a graph constructed from a general instance $I = (X, C)$ of 3-SAT as follows. For each clause c of C, there exists a vertex c that belongs to V^I. For each variable x of X, there exists a vertex x that belongs to V^I. If a clause c contains the literal x or $\overline{x}$, then the edge cx belongs to E^I. We remark that the *linear property* of two distinct clauses having 0 or 1 variable in common implies that the girth of $B(I)$ is greater than 4. The proposed variant is stated as follows:
k-GIRTH LINEAR CNF 2-3-SAT (k-GIRTH LCNF 2-3-SAT)
Instance: set $X = \{v_1, \ldots, v_n\}$ of variables, collection $C = \{c_1, \ldots, c_m\}$ of clauses over X such that each clause $c \in C$ has size $2 \leq |c| \leq 3$ and, for all $c, c' \in C$, $c \neq c'$, $|c \cap c'| \leq 1$, and the bipartite graph of clauses and variables has girth greater than k.
Question: Is there a truth assignment for X such that each clause in C has at least one true literal?

Theorem 2.1 k-GIRTH LCNF 2-3-SAT *problem is NP-Complete.*

Proof: Given an instance (X, C) of LCNF 3-SAT, we construct an instance (X', C') of k-GIRTH LCNF 2-3-SAT as follows.

Let $X = \{v_1, \ldots, v_n\}$ be the set of variables, $C = \{c_1, \ldots c_m\}$ be a collection of clauses over X such that each clause $c \in C$ has size $|c| = 3$ and, for all $c, c' \in C$, $c \neq c'$, $|c \cap c'| \leq 1$. This last constraint ensures that the girth g of the bipartite graph of clauses and variables $B(I) = (V^I, E^I)$ is greater than 4. In order to increase the value of g we proceed as follows. Set $X' := X$ and $C' := C$. For each clause $c_j = \{\ell_1, \ell_2, \ell_3\} \in C$, $1 \leq j \leq m$, we introduce two new auxiliary variables, $X' := X' \cup \{x_{c_j}, y_{c_j}\}$, and we replace c_j by three new clauses c'_j, c''_j and c'''_j, that is, $C' := (C' \setminus c_j) \cup \{\{\ell_1, \overline{x}_{c_j}\}, \{\ell_2, x_{c_j}, y_{c_j}\}, \{\ell_3, \overline{y}_{c_j}\}\}$.

It is easy to see that this transformation is done in polynomial time. The linear property implies that the set of clauses C has a size linear in n, which implies that the resulting set of clauses C' has a size linear in n, since the preceding set of clauses is linear in n, and each clause $c_j \in C$ is replaced by three clauses with pairwise intersection of at most one variable with all clauses of the new C' (because these new clauses do not duplicate variables of c_j, and distinct additional auxiliary variables are added for each clause $c_j \in C$, $1 \leq j \leq m$).

We claim that instance (X, C) is satisfiable if, and only if, (X', C') is satisfiable because clause $\{\ell_1, \ell_2, \ell_3\}$ is logically equivalent to the set of clauses $\{\ell_1, \overline{x}_{c_j}\}$, $\{\ell_2, x_{c_j}, y_{c_j}\}$, $\{\ell_3, \overline{y}_{c_j}\}$.

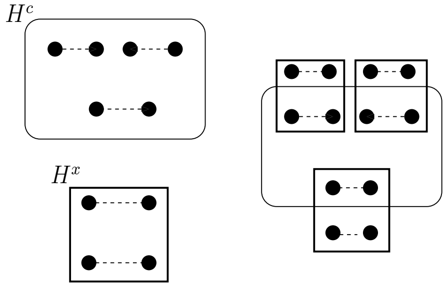

Fig. 1. Clause subgraph H^c, variable subgraph H^x, and their connection.

This procedure is reflected in the updated bipartite graph $B(I)$ as follows. Let $\mathcal{C}$ be a cycle in $B(I)$ that contains the clause vertex c_j. After applying the procedure above, the clause vertex c_j is replaced by three new clause vertices, say c'_j, c''_j and c'''_j; and we also add to V^I the two new auxiliary variable vertices x_{c_j} and y_{c_j}. Let $\ell_1 = x$ or $\ell_1 = \overline{x}$ be a literal of the clause $c_j = \{\ell_1, \ell_2, \ell_3\}$. Hence, the edge xc_j is replaced by the path x, c'_j, x_{c_j}, c''_j. So, the size of any cycle in $B(I)$ that contains c_j is increased by at least 2 vertices. Now, it is clear that by repeating the procedure at most $k/2$ times, we are guaranteed to build in polynomial time an instance I such that the girth g of $B(I)$ is greater than k, since k is a fixed number. □

3 H-free graph sandwich problem

We are interested in a special structure of the forbidden graph. The forbidden graph H is required to have a matching of size 2, say $A = \{a_1a'_1, a_2a'_2\}$, and to have an anti-matching (i.e. a matching in the complement graph $\overline{H}$) of size 3, say $B = \{b_1b'_1, b_2b'_2, b_3b'_3\}$.

Given an instance (X, C) of k-GIRTH LCNF 2-3-SAT, we construct an instance (G_1, G_2) of H-free graph sandwich problem as follows (see Figure 1). In what follows, each induced variable subgraph H^x and each induced clause subgraph H^c is a copy of H.

For each variable x of X, there exists an induced variable subgraph H^x in G_2, such that the edges $a_1^x a_1'^x$ and $a_2^x a_2'^x$ are the unique optional edges of H^x in the set $E_2 \setminus E_1$.

For each three-sized clause $c = \{l_1, l_2, l_3\}$ of C, there exists an induced clause subgraph H^c in G_1, such that for each literal l_i, $i \in \{1, 2, 3\}$, we include the additional optional edge $b_i^c b_i'^c$ in the set $E_2 \setminus E_1$.

For each two-sized clause $c = \{l_1, l_2\}$ of C, there exists an induced clause subgraph H^c in G_1, such that for each literal l_i, $i \in \{1, 2\}$, we include the additional optional edge $b_i^c b_i'^c$ in the set $E_2 \setminus E_1$. Note that, in this case, the edge $b_3^c b_3'^c$ is forbidden, i.e., $b_3^c b_3'^c \notin E_2$.

Whenever a variable x occurs as positive (resp. negative) literal l_i in clause c, then the edge $a_1^x a_1'^x$ (resp. $a_2^x a_2'^x$) is equivalent to the edge $b_i^c b_i'^c$, by identifying $a_1^x = b_i^c$ and $a_1'^x = b_i'^c$ (resp. $a_2^x = b_i^c$ and $a_2'^x = b_i'^c$), $i \in \{1, 2, 3\}$ (resp. $i \in \{1, 2\}$ in case of a two-sized clause).

This concludes the construction of the particular instance (G_1, G_2) of H-free graph sandwich problem.

We claim that this construction gives sufficient conditions to analyze the NP-completeness of H-free graph sandwich problems by studying some properties of graph H and of the structure of problem k-GIRTH LCNF 2-3-SAT.

Theorem 3.1 *Let H be a graph, containing a matching of size 2 and an anti-matching of size 3. If the particular instance (G_1, G_2) constructed above admits an H-free sandwich graph G, then there exists a truth assignment that satisfies instance (X, C) for the k-GIRTH LCNF 2-3-SAT.*

Proof: Suppose G is an H-free sandwich graph. So, every H^c clause subgraph of G_1 is destroyed by using at least one optional edge of set B. However, no H^x variable subgraph is created by adding both edges of A.

We now define the truth assignment for (X, C): if an edge of B belongs to $G \setminus G_1$ then set the truth value of the corresponding literal to true. Suppose that two edges are corresponding to the positive and negative literals of the same variable x. This generates an H induced subgraph in G, corresponding to the H^x variable subgraph for x, a contradiction. □

The converse theorem is not so straightforward and it requires a deeper study of the structure of the graph H and of the bipartite graph of clauses and variables $B(I)$.

Every rule we use in order to construct a sandwich graph from a truth assignment of k-GIRTH LCNF 2-3-SAT needs to ensure that no side-effect H subgraph is generated.

It is quite easy to give a rule that destroys all H^c clause subgraphs of G_1 without creating an H^x variable subgraph of G_2.

Our attempt consists in giving a simple rule (every optional edge corresponding to a true literal is added to G_1 to form a sandwich graph G), and then search for a side-effect H sandwich subgraph. We remark that every vertex of the constructed instance (G_1, G_2) belongs to an H^c clause subgraph or to an H^x variable subgraph, and possibly to both an H^c clause subgraph and an H^x variable subgraph. Although a vertex may belong to at most one H^x variable subgraph, possibly a vertex may belong to several H^c clause subgraphs. A *side-effect* H is an induced subgraph isomorphic to H of some sandwich graph G, such that H is neither associated to an H^c clause subgraph nor to an H^x variable subgraph.

In the special structure of the forbidden graph H, in order to have the converse theorem, we require further the forbidden graph H to be 3-connected.

Theorem 3.2 *Let H be a 3-connected graph, containing a matching of size 2 and an anti-matching of size 3. If there exists a truth assignment that satisfies instance (X, C) of the k-GIRTH LCNF 2-3-SAT, then the particular instance (G_1, G_2) constructed above admits an H-free sandwich graph G.*

Proof: Let H be a 3-connected graph, containing a matching of size 2 and an anti-matching of size 3, and let ℓ be the size of the largest chordless cycle of H. Let

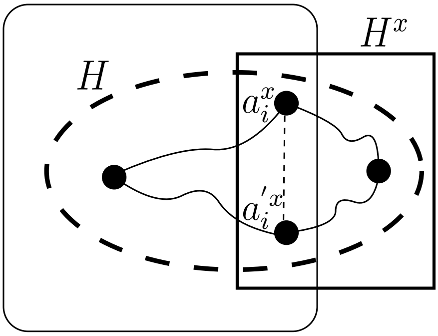

Fig. 2. The removal of vertices a_i^x and $a_i'^x$ disconnects H, a contradiction.

parameter $k = 2\ell$.

Suppose there exists a truth assignment that satisfies instance $I = (X, C)$ of the k-GIRTH LCNF 2-3-SAT, and consider the simple rule that adds to G_1 every optional edge corresponding to a true literal in order to define a sandwich graph G. Since the simple rule destroys all H^c clause subgraphs of G_1 without creating an H^x variable subgraph of G_2, it remains to prove that the sandwich graph G contains no side-effect induced subgraph H, associated to neither an H^c clause subgraph nor to an H^x variable subgraph. Assume to get a contradiction that G contains such a side-effect induced subgraph H. By construction, there exists a variable x such that the side-effect subgraph H must contain at least one among the two vertices a_i^x and $a_i'^x$, endvertices of the optional edge $a_i^x a_i'^x$. The two vertices a_i^x and $a_i'^x$ are the only connection between an H^c clause subgraph and an H^x variable subgraph, or between two clause subgraphs H^{c_j} and $H^{c_{j'}}$, by the linear property. Note that the optional edge $a_i^x a_i'^x$ belongs to a unique variable subgraph H^x. Note further that the optional edge $a_i^x a_i'^x$ must be associated to at least one clause subgraph H^c. Suppose first that the removal of the two vertices a_i^x and $a_i'^x$ disconnects the side-effect induced subgraph H, this gives a contradiction since H is a 3-connected graph (see Figure 2).

Hence, the side-effect subgraph H must contain a chordless cycle S that contains at least one of a_i^x or $a_i'^x$ (see Figure 3). Note that the size of S is less than or equal to the size of the largest chordless cycle of H, that is, $|S| \leq \ell$. Since the removal of the two vertices a_i^x and $a_i'^x$ does not disconnect the side-effect induced subgraph H, cycle S has vertices of at least two different variable subgraphs. We claim that the induced subgraph in $B(I)$ constructed from S by taking the corresponding vertices of variable and clause subgraphs of S, has a cycle R_I. Otherwise, this subgraph is a tree in $B(I)$ and there exists a variable vertex x_i with two adjacent clause vertices c_j and $c_{j'}$, such that vertices of the corresponding H^{c_j} and $H^{c_{j'}}$ in S would be disconnected by the removal of the two vertices a_i^x and $a_i'^x$, again a contradiction. We observe that the cycle R_I has size at most 2ℓ because, in the worst case, every edge of S is incident to two different variable subgraphs.

In this case, R_I has ℓ vertices corresponding to H^x variable subgraphs and ℓ vertices corresponding to H^c clause subgraphs. This contradicts the fact that, by definition, the girth of $B(I)$ is greater than 2ℓ, which concludes the proof. $\square$

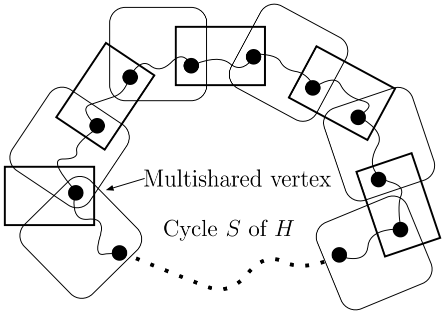

Fig. 3. An example of a cycle S of H. Note the possible shared vertex by two clause subgraphs.

We present next two applications by considering two particular graphs H. Denote by K_p the complete graph on p vertices, denote by $3K_2$ the graph consisting of an induced matching with three edges, and call *p-wheel* a graph consisting of a chordless cycle on p vertices and an additional vertex u adjacent to all p vertices on the cycle. Denote by $E[3K_2]$ the edge-set of the $3K_2$ graph.

Corollary 3.3 *If H is $K_p \setminus E[3K_2]$, for $p \geq 6$, then the H-free graph sandwich problem is NP-complete.*

Proof: Any two non incident edges of $K_p \setminus E[3K_2]$ is a matching of size 2. The missing $3K_2$ is an anti-matching of size 3. Finally, $K_p \setminus E[3K_2]$ is 3-connected. Thus, Theorems 3.1 and 3.2 can be applied. □

Corollary 3.4 *If H is p-wheel, for $p \geq 6$, then H-free graph sandwich problem is NP-complete.*

Proof: Any two non incident edges of the p-cycle of a p-wheel is a matching of size 2. Any induced p-cycle, $p \geq 6$, contains an anti-matching of size 3. Finally, a p-wheel, $p \geq 6$, is 3-connected. Thus, Theorems 3.1 and 3.2 can be applied. □

4 Concluding remarks

In the present work, we provide a new variant of SAT called k-GIRTH LCNF 2-3-SAT, which yields the classification of the graph sandwich problem for H-free graph classes such that the forbidden graph H is 3-connected, H contains an anti-matching (i.e. $\overline{H}$ contains a matching) of size 3 and H contains a matching of size 2.

In particular, we prove that, when H is $K_p \setminus E[3K_2]$ or a p-wheel, for $p \geq 6$, the H-free graph sandwich problem is NP-complete.

Furthermore, our results establish an interesting dichotomy: for every fixed $p \geq 6$, the $(K_p \setminus e)$-free graph sandwich problem is in P [2], whereas both the $(K_p \setminus E[3K_2])$-free graph sandwich problem (from Corollary 3.3) and the $(K_p \setminus E[2K_2])$-free graph sandwich problem are NP-complete. The NP-completeness of the $(K_p \setminus E[2K_2])$-free graph sandwich problem is implied by the NP-completeness of the C_4-free graph sandwich problem [2].

The strategy used to prove these results provides a tool to classify as NP-complete the graph sandwich problem for several families of graphs defined by forbidden induced subgraphs. For instance, we remark that $K_6 \setminus E[3K_2]$ is isomorphic to the power of cycle C_6^2, and that the H-free graph sandwich problem is NP-complete when H is the power of cycle C_n^p, for $n \geq 6$ and $p < \lfloor n/2 \rfloor$.

References

[1] F. Couto, L. Faria, S. Gravier, S. Klein. On the forbidden induced subgraph probe and sandwich problems. Discrete Appl. Math., 234 (2018) 56–66.

[2] S. Dantas, C.M.H. de Figueiredo, M.V.G. da Silva, R.B. Teixeira. On the forbidden induced subgraph sandwich problem. Discrete Appl. Math., 159 (2011) 1717–1725.

[3] S. Dantas, C.M.H. de Figueiredo, F. Maffray, R.B. Teixeira. The complexity of forbidden subgraph sandwich problems and the skew partition sandwich problem. Discrete Appl. Math., 182 (2015) 15–24.

[4] M.C. Dourado, P. Petito, R.B. Teixeira, C.M.H. de Figueiredo. Helly property, clique graphs, complementary graph classes, and sandwich problems. J. Braz. Comput. Soc., 14 (2008) 45–52.

[5] C.M.H. de Figueiredo, S. Spirkl. Sandwich and probe problems for excluding paths. Discrete Appl. Math., 251 (2018) 146–154.

[6] M.C. Golumbic, R. Shamir. Complexity and algorithms for reasoning about time: a graph-theoretic approach. J. Assoc. Comp. Mach., 40 (1993) 1108–1133.

[7] M.C. Golumbic, H. Kaplan, R. Shamir. On the complexity of DNA physical mapping. Adv. Appl. Math., 15 (1994) 251–261.

[8] M.C. Golumbic, H. Kaplan, R. Shamir. Graph sandwich problems. J. Algorithms, 19 (1995) 449–473.

[9] S. Porschen, E. Speckenmeyer, X. Zhao. Linear CNF formulas and satisfiability. Discrete Appl. Math., 157 (2009) 1046–1068.