Clique Graph Recognition Is NP-Complete*

L. Alcón[1], L. Faria[2], C.M.H. de Figueiredo[3], and M. Gutierrez[1]

[1] Departamento de Matemática, UNLP, Argentina
[2] Departamento de Matemática, FFP, UERJ, Brazil
[3] Instituto de Matemática and COPPE, UFRJ, Brazil

Abstract. A *complete set* of a graph G is a subset of V inducing a complete subgraph. A *clique* is a maximal complete set. Denote by $\mathcal{C}(G)$ the *clique family* of G. The *clique graph* of G, denoted by $K(G)$, is the intersection graph of $\mathcal{C}(G)$. Say that G is *a clique graph* if there exists a graph H such that $G = K(H)$. The clique graph recognition problem asks whether a given graph is a clique graph. A sufficient condition was given by Hamelink in 1968, and a characterization was proposed by Roberts and Spencer in 1971. We prove that the clique graph recognition problem is NP-complete.

1 Introduction

We consider finite, simple and undirected graphs. V and E denote the vertex set and the edge set of the graph G, respectively. A *complete set* of G is a subset of V inducing a complete subgraph. A *clique* is a maximal complete set.

If G is a graph, $\mathcal{C}(G)$ denotes the *clique family* of G. The *clique graph* of G, denoted by $K(G)$, is the intersection graph of $\mathcal{C}(G)$. Say that G is *a clique graph* if there exists a graph H such that $G = K(H)$. Not every graph is a clique graph. The Clique graph recognition problem can be formulated as follows.

CLIQUE GRAPH
INSTANCE: A graph $G = (V, E)$.
QUESTION: Is there a graph H such that $G = K(H)$?

A sufficient condition for a graph to be a clique graph was given in [6], and characterizations of clique graphs are given in [9] and more recently in [1]. However the time complexity of the problem of recognizing clique graphs is still open [4,8,10].

Given a *set family* $\mathcal{F} = (F_i)_{i \in I}$, the sets F_i are called *members* of the family. $F \in \mathcal{F}$ means that F is a member of $\mathcal{F}$. The family is *pairwise intersecting* if the intersection of any two members is not the empty set. The *intersection* or *total intersection* of $\mathcal{F}$ is the set $\cap\mathcal{F} = \cap_{i \in I} F_i$. The family $\mathcal{F}$ has the *Helly property*, if any pairwise intersecting subfamily has nonempty total intersection.

The edge with end vertices u and v is represented by uv. We say that the complete set C *covers* the edge uv when u and v belong to C. A *complete edge cover* of a graph G is a family of complete sets of G covering all edges of G.

* Dedicated to Alberto Santos Dumont, aviation pioneer, on the 100th anniversary of the flight of his *14 Bis* in Paris in October 1906.

F.V. Fomin (Ed.): WG 2006, LNCS 4271, pp. 269–277, 2006.

The following Theorem is a well known characterization of Clique Graphs.

Theorem 1 (Roberts and Spencer [9]). *G is a clique graph if and only if there exists a complete edge cover of G satisfying the Helly property.*

A *triangle* is a complete set with exactly 3 vertices. The set of triangles of G is denoted $T(G)$. Let $\mathcal{F}$ be a complete edge cover of G and T a triangle, $\mathcal{F}_T$ is the subfamily of $\mathcal{F}$ formed by all the members containing at least two vertices of T.

Next lemma is a characterization of a complete edge cover satisfying the Helly property, in what follows *RS-family*, which will be used in the proof of our main theorem.

Lemma 1 (Alcón and Gutierrez [2]). *Let $\mathcal{F}$ be a complete edge cover of G. The following conditions are equivalent:*
i) $\mathcal{F}$ has the Helly property.
ii) For every $T \in T(G)$, the subfamily $\mathcal{F}_T$ has the Helly property.
iii) For every $T \in T(G)$, the subfamily $\mathcal{F}_T$ has nonempty intersection, this means $\cap \mathcal{F}_T \neq \emptyset$.

As noted by Roberts and Spencer [9], Theorem 1 yields a polynomial certificate of G being a clique graph. First, for the polynomial size of the edge cover certificate, note that if $\mathcal{F}$ has the Helly property, then every subfamily $\mathcal{F}'$ of $\mathcal{F}$ has the Helly property as well. In addition, we prove that if G admits a complete edge cover $\mathcal{F}$, then G admits a complete edge cover $\mathcal{F}'$ of size at most $|E|$ which is our considered certificate: just greedily scan the edges of E, select for $\mathcal{F}'$ one complete set of $\mathcal{F}$ covering the first edge, and for each edge e not yet covered by $\mathcal{F}'$, select for $\mathcal{F}'$ one complete set of $\mathcal{F}$ covering e. Clearly this greedy procedure labels each selected set with a corresponding scanned edge of E, yielding a subfamily $\mathcal{F}'$ of size at most $|E|$. Second, for the polynomial verification of the certificate, a result of Berge [3] says that a family of sets has the Helly property, if and only if for any triple of elements, the subfamily of sets containing at least two out of these three elements has non-empty intersection. Actually, by Lemma 1, it is enough to consider the triples of vertices a, b, c of G defining a triangle T. We consider the members of $\mathcal{F}'_T$ and check for every vertex v of V if v belongs to $\bigcap \mathcal{F}'_T$. This produces an $O(n^4 m)$ algorithm that checks if a complete edge cover $\mathcal{F}'$ of size $O(m)$ is Helly. Thus CLIQUE GRAPH belongs to NP.

In this paper we prove that CLIQUE GRAPH is NP-complete by a reduction from the following version of the 3–satisfiability problem with at most 3 occurrences per variable.

Let $U = \{u_i, 1 \leq i \leq n\}$ be a set of boolean variables. A literal is either a variable u_i or its complement $\overline{u_i}$. A clause over U is a set of literals of L. Let $C = \{c_j, 1 \leq j \leq m\}$ be a collection of clauses over U. We say that variable u_i occurs in clause c_j (and then in C) if u_i or $\overline{u_i} \in c_j$. We say that variable u_i occurs in clause c_j as literal u_i (or that literal u_i occurs in c_j) if $u_i \in c_j$, and as literal $\overline{u_i}$ (or that literal $\overline{u_i}$ occurs in c_j) if $\overline{u_i} \in c_j$.

3SAT$_{\overline{3}}$
INSTANCE: $I = (U, C)$, where $U = \{u_i, 1 \le i \le n\}$ is a set of boolean variables, and $C = \{c_j, 1 \le j \le m\}$ a set of clauses over U such that each clause has two or three variables, each variable occurs two or three times in C, each variable occurs never twice in the same clause. If variable u_i occurs twice in C, then it is once as literal u_i and once as literal $\overline{u_i}$. If variable u_i occurs three times in C, then it is once as literal u_i and twice as literal $\overline{u_i}$.
QUESTION: Is there a truth assignment for U such that each clause in C has at least one true literal?

It is a known result that 3SAT$_{\overline{3}}$ is an NP-complete problem [5,7].

In order to reduce 3SAT$_{\overline{3}}$ to CLIQUE GRAPH we need to construct in polynomial time a particular instance G of CLIQUE GRAPH from a generic instance $I = (U, C)$ of 3SAT$_{\overline{3}}$, in such a way that C is satisfiable if and only if G is a clique graph.

In Section 2 we describe the construction of instance G of CLIQUE GRAPH from instance $I = (U, C)$ of 3SAT$_{\overline{3}}$. In Section 3, we state and prove the main theorem; and in Section 4 give some conclusions and propose new related problems.

2 Construction of G from $I = (U, C)$

Let $I = (U, C)$ be any instance of 3SAT$_{\overline{3}}$,

For each variable u_i let j_i be the subindex of the unique clause where variable u_i occurs as literal u_i; and $\overline{J}_i = \{j \mid \text{ literal } \overline{u_i} \text{ occurs in } c_j\}$.

For each clause c_j with $\mid c_j \mid= 3$, let $I_j = \{i \mid \text{ variable } u_i \text{ occurs in } c_j\}$; and for each clause c_j with $\mid c_j \mid= 2$, let $I_j = \{i \mid \text{ variable } u_i \text{ occurs in } c_j\} \cup \{n+1\}$. Notice that in any case $\mid I_j \mid= 3$. Given $I_j = \{i_1, i_2, i_3\}$, with $i_1 < i_2 < i_3$, let $i_1^* = i_2$, $i_2^* = i_3$ and $i_3^* = i_1$.

From instance $I = (U, C)$, we construct a graph $G = (V, E)$. Please refer to Figures 1 and 2. The vertex set V is the union:

$$V = \bigcup_{1 \le i \le n} \left[\{a^i_{j_i}, c^i_{j_i}, d^i_{j_i}, e^i_{j_i}, f^i_{j_i}, g^i_{j_i}, h^i_{j_i}\} \bigcup_{j \in \overline{J}_i} \{a^i_j, c^i_j, d^i_j, e^i_j, f^i_j, g^i_j, h^i_j, z^i_j, v^i_j, w^i_j\} \right] \bigcup$$
$$\bigcup_{1 \le j \le m, |c_j|=2} \left[\{a^{n+1}_j, c^{n+1}_j, d^{n+1}_j, e^{n+1}_j, f^{n+1}_j, g^{n+1}_j, h^{n+1}_j\} \right].$$

Since $\mid \overline{J}_i \mid \le 2$, $\mid V \mid$ is bounded by $(n+1) \times 7 + n \times 2 \times 10 = 27 \times n + 7$.

The edge set E contains:

For each $j, 1 \le j \le m$, the edges of the complete graph induced by the vertex set $K_{12}(j) = \{a^i_j, d^i_j, g^i_j, h^i_j \mid i \in I_j\}$; the edges of the sets $\{c^i_j d^i_j \mid i \in I_j, i \ne n+1\}$ and $\{c^i_j a^i_j, c^i_j a^{i^*}_j, e^i_j d^i_j, e^i_j h^i_j, f^i_j g^i_j, f^i_j a^{i^*}_j \mid i \in I_j\}$.

And for each $i, 1 \le i \le n$, for each $j \in \overline{J}_i$, the edges of the complete graph induced by the vertex set $K_5(j, i) = \{h^i_{j_i}, g^i_{j_i}, v^i_j, h^i_j, g^i_j\}$; and the edges of the set $\{h^i_{j_i} w^i_j, w^i_j h^i_j, g^i_{j_i} z^i_j, z^i_j g^i_j, a^i_{j_i} v^i_j, v^i_j a^i_j\}$.

Notice that for each variable u_i, graph G contains as induced subgraph the graph depicted in Figure 1; and for each clause c_j, graph G contains as induced

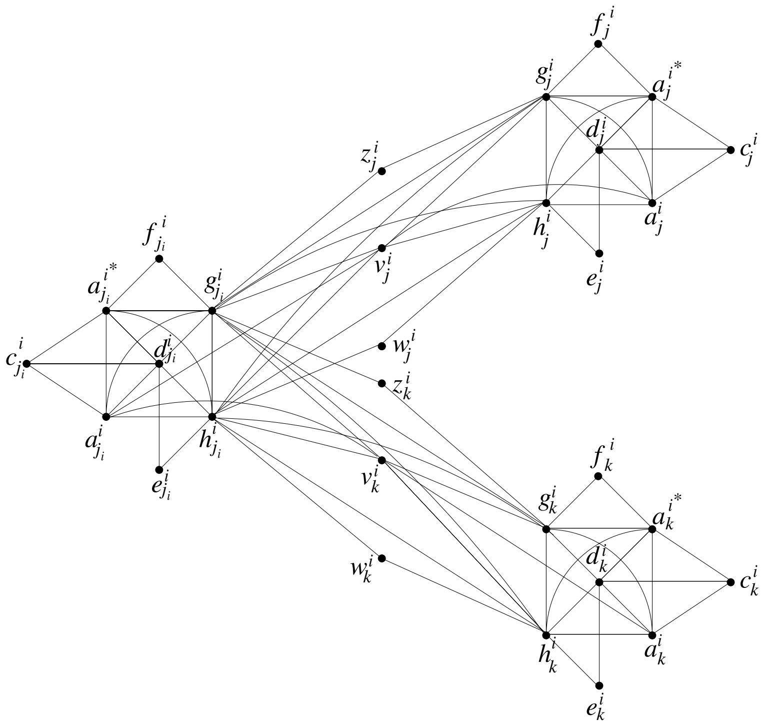

Fig. 1. Truth Setting component T_i for variable u_i with $\overline{J}_i = \{j, k\}$

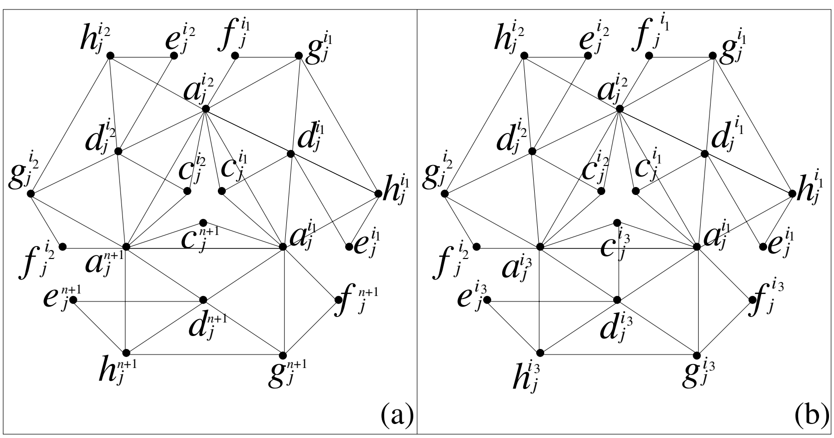

Fig. 2. Satisfaction Testing component S_j for clause c_j with: (a) 2 literals corresponding to variables u_{i_1} and u_{i_2}; and (b) 3 literals corresponding to variables u_{i_1}, u_{i_2} and u_{i_3}. Vertices $\{a_j^i, d_j^i, g_j^i, h_j^i \mid i \in I_j\}$ induce a complete graph but for simplicity some edges are not drawn.

subgraph the graph depicted either in Figure 2 (a) or (b), where some edges have been omitted. We obtain the whole graph G by superposing these subgraphs.

For the convenience of the reader we offer an example in Figure 5 of graph G obtained from the instance $I = (U, C)$, $U = \{u_1, u_2, u_3\}$, $C = \{\{u_1, u_3\}, \{\overline{u_1}, u_2, \overline{u_3}\}, \{\overline{u_1}, \overline{u_2}\}\}$.

2.1 About Graph G

The following two lemmata present properties of graph G that we will use in the proof of the main theorem. Notice that any RS-family of a graph contains the triangles of the graph that are cliques, in particular the triangles with a vertex of degree 2.

Lemma 2. (*Two Cover Lemma*) *Let $\mathcal{F}$ be an RS-family of the graph G. For each $j, 1 \leq j \leq m$, and for each $i \in I_j, i \neq n+1$, exactly one of the triangles $\{a^i_j, a^{i^*}_j, c^i_j\}$, $\{a^i_j, c^i_j, d^i_j\}$ belongs to $\mathcal{F}$.*

Proof. Please refer to Figure 1. There are three possible complete sets of G covering a^i_j, c^i_j: $\{a^i_j, a^{i^*}_j, c^i_j, d^i_j\}$, $\{a^i_j, a^{i^*}_j, c^i_j\}$, $\{a^i_j, c^i_j, d^i_j\}$.

Note that both triangles $\{d^i_j, e^i_j, h^i_j\}$ and $\{f^i_j, a^{i^*}_j, g^i_j\}$ belong to $\mathcal{F}$, as e^i_j and f^i_j are vertices of degree 2 in G.

Suppose $\{a^i_j, a^{i^*}_j, c^i_j, d^i_j\} \in \mathcal{F}$. The Helly property implies $\{a^i_j, g^i_j, v^i_j\}$ is not contained in a complete set of $\mathcal{F}$, which implies $\{a^i_j, h^i_j, v^i_j\} \in \mathcal{F}$, in order to cover edge $a^i_j v^i_j$; but then $\{a^i_j, a^{i^*}_j, c^i_j, d^i_j\}$, $\{e^i_j, h^i_j, d^i_j\}$ and $\{a^i_j, h^i_j, v^i_j\}$ violate the Helly property. Then $\{a^i_j, a^{i^*}_j, c^i_j, d^i_j\} \notin \mathcal{F}$.

Assume $\{a^i_j, a^{i^*}_j, c^i_j\}$ and $\{a^i_j, c^i_j, d^i_j\}$ belong to $\mathcal{F}$. Since $\{a^i_j, a^{i^*}_j, c^i_j\} \in \mathcal{F}$, then $\{a^i_j, v^i_j, g^i_j\}$ is not contained in a complete of $\mathcal{F}$, thus $\{a^i_j, v^i_j, h^i_j\} \in \mathcal{F}$, but in this case $\{a^i_j, v^i_j, h^i_j\}$, $\{a^i_j, c^i_j, d^i_j\}$ and $\{e^i_j, h^i_j, d^i_j\}$ violate the Helly property. □

Lemma 3. (*Literal Communication Lemma*) *Let $\mathcal{F}$ be an RS-family of the graph G. For each $i, 1 \leq i \leq n$, and for each $j \in \overline{J}_i$, if $\{a^i_j, c^i_j, d^i_j\} \in \mathcal{F}$ then $\{a^i_{j_i}, a^{i^*}_{j_i}, c^i_{j_i}\} \in \mathcal{F}$.*

Proof. Please refer to Figure 1. Since $\{a^i_j, c^i_j, d^i_j\} \in \mathcal{F}$ and $\{e^i_j, h^i_j, d^i_j\} \in \mathcal{F}$, we have that $\{a^i_j, h^i_j, v^i_j\}$ and $\{a^i_j, h^i_j, g^i_j, v^i_j\}$ do not belong to $\mathcal{F}$, because this violates the Helly property. Thus $\{a^i_j, g^i_j, v^i_j\} \in \mathcal{F}$, in order to cover edge $a^i_j v^i_j$.

Triangles $\{a^i_j, g^i_j, v^i_j\}$, $\{g^i_{j_i}, g^i_j, z^i_j\}$ belong to $\mathcal{F}$ implies $\{a^i_{j_i}, g^i_{j_i}, v^i_j\}$ and $\{a^i_{j_i}, h^i_{j_i}, g^i_{j_i}, v^i_j\}$ do not belong to $\mathcal{F}$. Thus $\{a^i_{j_i}, h^i_{j_i}, v^i_j\} \in \mathcal{F}$.

Since $\{a^i_{j_i}, h^i_{j_i}, v^i_j\}$ and $\{d^i_{j_i}, h^i_{j_i}, e^i_{j_i}\}$ belong to $\mathcal{F}$, then $\{a^i_{j_i}, c^i_{j_i}, d^i_{j_i}\} \notin \mathcal{F}$. By the <u>Two Cover Lemma</u> $\{a^i_{j_i}, a^{i^*}_{j_i}, c^i_{j_i}\} \in \mathcal{F}$. □

These two lemmata are the basis of the proof of the main theorem. It follows that given any RS-family of G, and any variable u_i, by looking if one triangle of the satisfaction testing subgraph S_{j_i} belongs or not to the RS-family it is possible to know whether one triangle of the satisfaction testing subgraph S_j, with $j \in \overline{J}_i$, belongs or not to the RS-family. The two possible cases are shown in Figures 3 and 4.

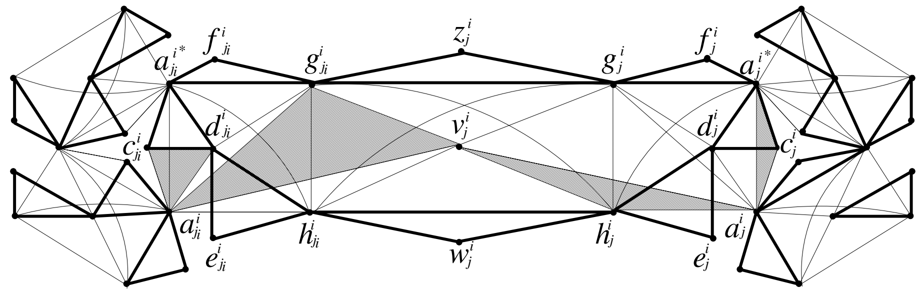

Fig. 3. Truth value of u_i equals to True

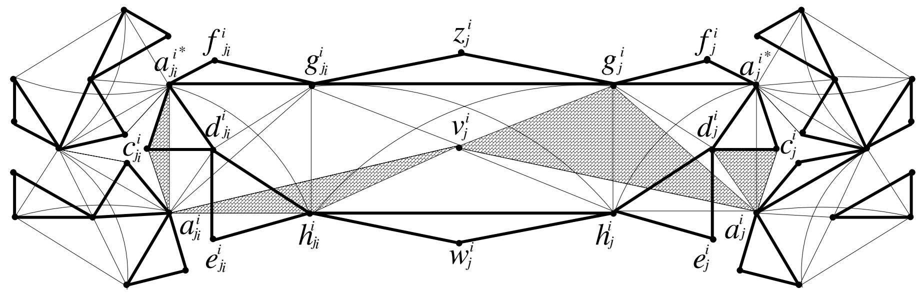

Fig. 4. Truth value of u_i equals to False

3 Main Theorem

Theorem 2. CLIQUE GRAPH *is NP-complete.*

Proof. As shown in the Introduction, CLIQUE GRAPH belongs to NP.

Let G be the graph obtained by Section 2 process from an instance $I = (U, C)$ of $3\text{SAT}_{\overline{3}}$. Suppose G is a clique graph, we will exhibit a truth assignment for U such that C is satisfied.

Let $\mathcal{F}$ be an RS-family for G. Let $u_i \in U$ be a variable. Set u_i equal to true if and only if $\{a^i_{j_i}, c^i_{j_i}, d^i_{j_i}\} \in \mathcal{F}$.

To see that this truth assignment for U satisfies C consider a clause c_j.

The Helly property on $\mathcal{F}$ implies there exists $i \in I_j$ such that the triangle $\{a^i_j, a^{i^*}_j, c^i_j\}$ is not a member of $\mathcal{F}$. Notice that $i \neq n+1$.

By the Two Cover Lemma, $\{a^i_j, a^{i^*}_j, c^i_j\} \notin \mathcal{F}$ implies $\{a^i_j, c^i_j, d^i_j\} \in \mathcal{F}$.

If $j = j_i$ then variable u_i is true and clause c_j is satisfied.

If $j \neq j_i$, then $j \in \overline{J}_i$, by Literal Communication Lemma, $\{a^i_j, c^i_j, d^i_j\} \in \mathcal{F}$ implies $\{a^i_{j_i}, a^{i^*}_{j_i}, c^i_{j_i}\} \in \mathcal{F}$, thus by Two Cover Lemma, $\{a^i_{j_i}, c^i_{j_i}, d^i_{j_i}\} \notin \mathcal{F}$. It follows that u_i is false, then c_j is satisfied.

Conversely, given a truth assignment of U that satisfies C, we exhibit a complete edge cover $\mathcal{F}$ of G.

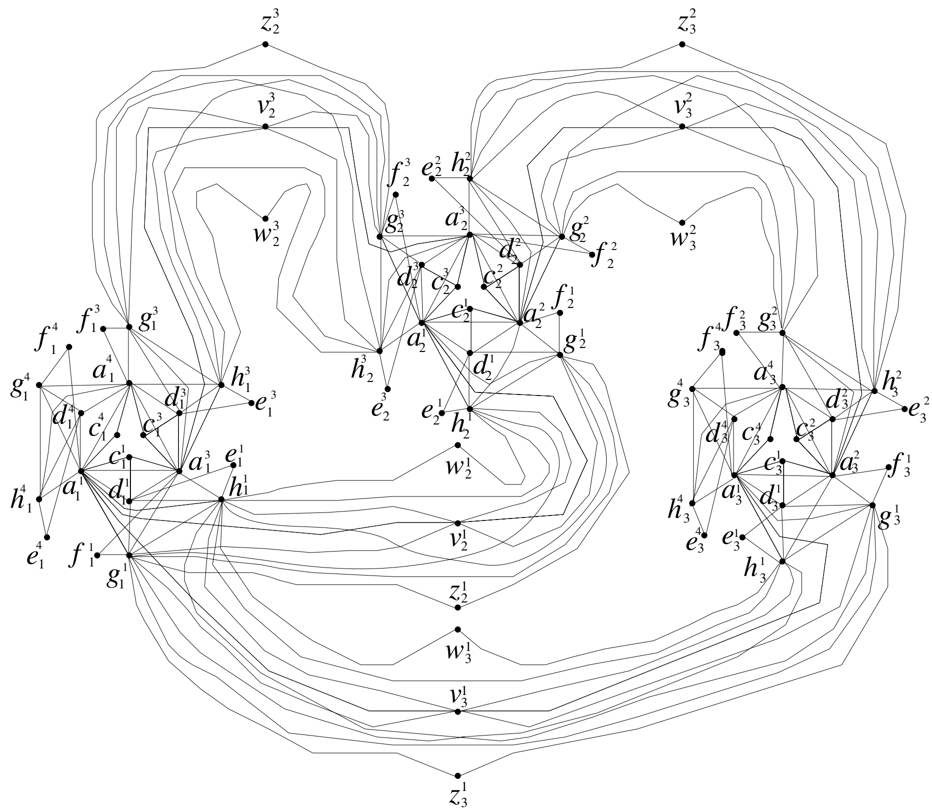

Fig. 5. Graph G obtained from the $3\text{SAT}_{\overline{3}}$ instance $U = \{u_1, u_2, u_3\}$, $C = \{\{u_1, u_3\}, \{\overline{u_1}, u_2, \overline{u_3}\}, \{\overline{u_1}, \overline{u_2}\}\}$.

For each $j, 1 \leq j \leq m$, complete set $K_{12}(j) = \{a_j^i, d_j^i, g_j^i, h_j^i \mid i \in I_j\}$;

For each $j, 1 \leq j \leq m$, for each $i \in I_j$, the triangles $\{f_j^i, a_j^{i^*}, g_j^i\}$, $\{e_j^i, d_j^i, h_j^i\}$.

For each $j, 1 \leq j \leq m$, for each $i \in I_j, i \neq n+1$, $\{c_j^i, a_j^{i^*}, d_j^i\}$; and for $i = n+1$, $\{c_j^{n+1}, a_j^{n+1^*}, a_j^{n+1}\}$.

For each $i, 1 \leq i \leq n$, for each $j \in \overline{J}_i$, the complete set $K_5(j, i) = \{h_{j_i}^i, g_{j_i}^i, v_j^i, h_j^i, g_j^i\}$;

For each $i, 1 \leq i \leq n$, for each $j \in \overline{J}_i$, $\{z_j^i, g_{j_i}^i, g_j^i\}$, $\{w_j^i, h_{j_i}^i, h_j^i\}$.

For each $i, 1 \leq i \leq n$, such that variable u_i is true, $\{c_{j_i}^i, d_{j_i}^i, a_{j_i}^i\}$; and for each $j \in \overline{J}_i$, $\{a_{j_i}^i, g_{j_i}^i, v_j^i\}$, $\{v_j^i, h_j^i, a_j^i\}$, $\{a_j^i, a_j^{i^*}, c_j^i\}$.

For each $i, 1 \leq i \leq n$, such that variable u_i is false, $\{c_{j_i}^i, a_{j_i}^i, a_{j_i}^{i^*}\}$; and for each $j \in \overline{J}_i$, $\{a_{j_i}^i, h_{j_i}^i, v_j^i\}$, $\{v_j^i, g_j^i, a_j^i\}$, $\{a_j^i, d_j^i, c_j^i\}$.

The proof is completed by showing that the complete edge cover $\mathcal{F}$ of G has the Helly property. By Lemma 1, it is enough to show that for each triangle $T \in T(G)$, $\cap \mathcal{F}_T \neq \emptyset$.

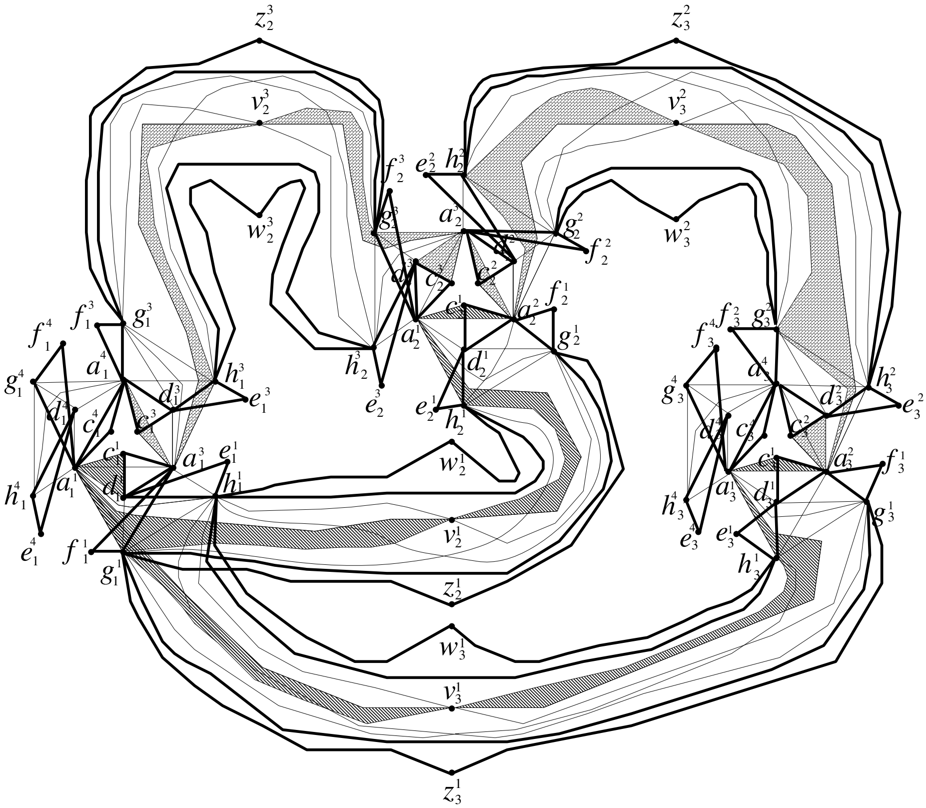

Fig. 6. RS-cover $\mathcal{F}$ for graph G of Figure 5, obtained from the satisfying truth assignment where u_1 is true, and u_2 and u_3 are false. Bold edges depict forced triangles of $\mathcal{F}$, dashed connected regions depict triangles of $\mathcal{F}$ which depend on the truth assignment for $I = (U, C)$, complete sets $K_5(j, i)$ and $K_{12}(j)$ are not depicted in order to make simpler the drawing.

If a triangle T contains an edge e for which any complete set of $\mathcal{F}$ covering e contains also T, then $\cap\mathcal{F}_T \neq \emptyset$. We call such a triangle an *easy triangle* and use this tool in order to accomplish the proof. We classify the triangles of G into types according to either they are, or they are not contained in a $K_{12}(j)$ or in a $K_5(j, i)$. Details are omitted in the extended abstract. $\square$

Figure 6 exhibits the RS-cover $\mathcal{F}$ defined from Theorem 2 for graph G of Figure 5.

4 Final Remarks

We have proved that deciding whether a given graph is a clique graph is an NP-complete problem. From the same proof, it follows that the problem remains NP-complete even for bounded degree graphs and for graphs with bounded clique size. However the problem is polynomial when restricted to graphs with maximum degree less than 5 and also when restricted to graphs with clique size less than 4 [10]. This fact suggests the search of the best bounds both for the maximum degree and for the clique size for which the problem is polynomial. Notice that the problem of recognizing clique graphs restricted to Planar Graphs of maximum clique size 4 was left open in [2]. It seems not to be trivial.

Acknowledgement. This research was partially supported by CAPES, CNPq, FAPERJ, PROCIÊNCIA-UERJ Project, CNPq–Project PROSUL- Proc. no. 490333/04-4.

References

1. Alcón, L., Gutierrez, M.: A new characterization of Clique Graphs. Matemática Contemporânea **25** (2003) 1–7
2. Alcón, L., Gutierrez, M.: Cliques and Extended Triangles. A necessary condition to be Clique Planar Graph. Discrete Applied Mathematics **141/1-3** (2004) 3–17
3. Berge, C.: "Hypergraphes". Gauthier-Villars Paris (1987)
4. Brandstädt, A., Le, V. B., Spinrad, J. P.: Graph Classes: A survey. SIAM Monographs on Discrete Mathematics and Applications (1999)
5. Garey, M. R., Johnson, D. S.: Computers and Intractability: a guide to the theory of NP-completeness. W. H. Freeman (1979)
6. Hamelink, R. C.: A partial characterization of clique graphs. Journal of Combinatorial Theory B **5** (1968), 192–197
7. Papadimitriou, C. M.: Computational Complexity. Addison-Wesley (1994)
8. Prisner, E.: A common generalization of Line Graphs and Clique Graphs. Journal of Graph Theory **18** (1994) 301–313
9. Roberts, F. S., Spencer, J. H.: A characterization of clique graphs. Journal of Combinatorial Theory B **10** (1971) 102–108
10. Szwarcfiter, J. L.: A survey on Clique Graphs. Recent Advances in Algorithmic Combinatorics, C. Linhares-Sales and B. Reed eds., Springer-Verlag (2002)