A path of reductions from Clique to 3Sat

Larcheri Paolo

April 19, 2003

1 Introduction

Generally 3Sat is used to demonstrate another problem to be NP-complete. Despite this common habit, in this document we will show a reduction (made up of many reduction steps) of Clique to 3Sat in order to strengthen our conviction that every NP-complete problem can be reduced to another NP-complete problem (we will also call NPC the set of all NP-complete problems). We will firstly reduce Clique to Isomorphic Subgraph problem, Isomorphic Subgraph to Sat, Sat to 3Sat and then 3SAT to Vertex Cover to close our path of reductions. Figure 1 illustrates the way of reductions we will cover within NPC.

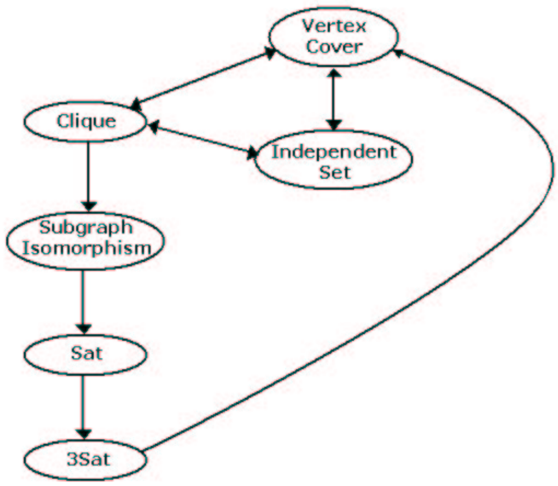

Figure 1: Reduction steps we'll follow in this document.

This document is organized as follows: in Section 2 Clique, Node Cover and Independent Set problems are formally defined and their equivalence is briefly discussed; the reduction of Clique to Isomorphic Subgraph is shown in Section 3; in Section 4 it is proved that $SAT \in NPC$ by demonstrating SAT reducibility to Isomorphic Subgraph; in Section 5, it is shown how to reduce SAT to 3SAT; at last, in in Section 6 3SAT is reduced to Vertex Cover (VC).

2 Clique, Node Cover and Independent Set

It isn't difficult to see that Clique, Vertex Cover and Independent Set are equivalent. To formally define these problems we adopt notation used in [Gar79]:

CLIQUE (CQ)
INSTANCE:
A graph $G = (V, E)$ and a positive integer $k \leq |V|$.
QUESTION:
Does there exist a subset $V' \subseteq V$ such that $|V'| \geq k$ and every two distinct vertices in V' are joined by an edge $e \in E$?

VERTEX COVER (VC)
INSTANCE:
A graph $G = (V, E)$ and a positive integer $k \leq |V|$.
QUESTION:
Does there exist a subset $V' \subseteq V$ such that $|V'| \leq k$ and, for each edge $\{u, v\} \in E$, at least one of u and v belongs to V'?

INDEPENDENT SET (IS)
INSTANCE:
A graph $G = (V, E)$ and a positive integer $k \leq |V|$.
QUESTION:
Does there exist a subset $V' \subseteq V$ such that $|V'| \geq k$ and, for all two distinct vertices $u, v \in V'$, $uv \notin E$?

The following lemma shows that these three problems can be considered simply as "different versions" of the same one.

Lemma 1 *For any graph $G = (V, E)$ and subset $V' \subseteq V$, the following statements are equivalent:*

- *V' is a vertex cover for G;*
- *$V - V'$ is an independent set for G;*
- *$V - V'$ is a clique in the* complement *G^c of G, where $G^c = (V, E^c)$ with $E^c = \{uv : u, v \in V \text{ and } uv \notin E\}$.*

Thus, $< G, j > \in VC \quad \Leftrightarrow \quad < G, n - j > \in IS \quad \Leftrightarrow \quad < G^c, n - j > \in CQ$.

3 Clique to Isomorphic Subgraph

ISub (Isomorphic Subgraph) can be defined as follows:

ISOMORPHIC SUBGRAPH (ISub)
INSTANCE:
Two graphs $G = (V, E)$ and $H = (U, F)$.

QUESTION:
Is H an Isomorphic Subgraph of G? The question can be reformulated as: does there exist a total injective function $s : U \to V$ such that, for all couples of distinct vertices $u_1, u_2 \in U$, $s(u_1)s(u_2) \in E$ *if and only if* $u_1u_2 \in F$?

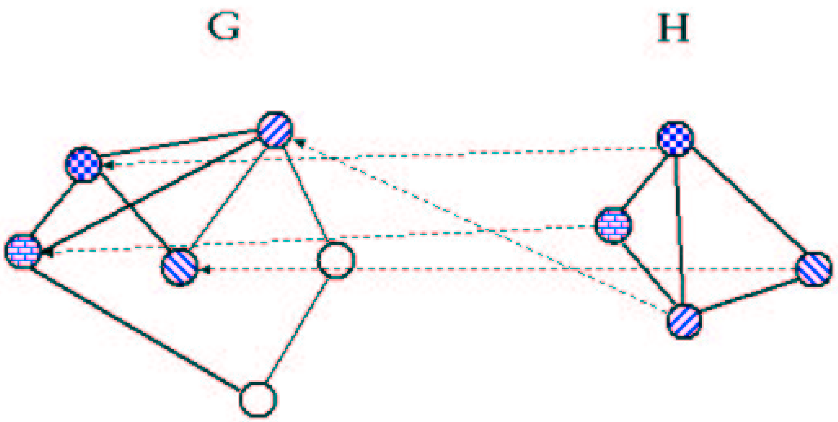

Figure 2: An example of isomorphic subgraph. Dotted arrows indicate vertex associations between H and G.

Figure 2 illustrates an example isomorphic subgraph. After having described the problem, we have to prove that it is NP-complete, starting from the assumption that Clique is NP-complete.

Theorem 2 *CQ reduces to ISub.*

Proof: We transform Clique to Isomorphic Subgraph. Let $I =< G, j >$ be any instance of Clique. Note that $< G, j > \in CQ$ *if and only if* $G, K_k \in ISub$ where K_j denotes complete graph on j nodes. It should be clear that $< G, k > \in CQ \quad \Leftrightarrow \quad < G, K_k > \in ISub$

4 Isomorphic Subgraph to SAT

SAT is the following problem:

SAT
INSTANCE:
Let X be a set of n boolean variables $x_1 \dots x_n$. An instance of SAT is a conjunction $\bigwedge_{i=0}^{m} C_i$ of m clauses on X.
QUESTION:
Is there a truth assignment for X that satisfies $\bigwedge_{i=0}^{m} C_i$?

Theorem 3 *ISub reduces to SAT.*

Proof: Suppose to have an instance of Isomorphic Subgraph composed by two graphs $G = (V, E)$ and $H = (U, F)$, where $|V| = n$, $|U| = m$ and $m \leq n$. We must construct in polynomial time an instance $< X, C >$ of SAT that is satisfiable if and only if there exists a total and injective function $s : U \rightarrow V$ such that it's true that for all couples of distinct vertices $u_1, u_2 \in U$, $s(u_1)s(u_2) \in E$ *if and only if* $u_1u_2 \in F$.

For each possible couple of vertices (u, v) where $u \in U$ and $v \in V$, let $x_{u,v}$ be a boolean variable in X. We will assign *true* to $x_{u,v}$ if $s(u) = v$, *false* otherwise.
We construct the clauses belonging to SAT instance:

1. s is a function. This means that if $v_i = s(u)$ and $v_j = s(u)$ then $i = j$. Consequently we can write:
$$\forall\ 0 \leq i, j \leq n,\ \forall\ 0 \leq k \leq m\ \left(\overline{x}_{u_k,v_i} \vee \overline{x}_{u_k,v_j}\right)$$

2. s is total. This means that for each $u \in U$ there exists an element $v \in V$ such that $s(u) = v$. This is expressible as:
$$\forall\ 0 \leq i \leq m,\ \left(x_{u_i,v_1} \vee x_{u_i,v_2} \vee\ \ldots\ \vee x_{u_i,v_n}\right)$$

3. s is injective. This means that if $s(u_i) = v$ and $s(u_j) = v$ then $i = j$. Consequently we can write:
$$\forall\ 0 \leq i, j \leq m,\ i \neq j,\ \forall\ 0 \leq k \leq n\ \left(\overline{x}_{u_i,v_k} \vee \overline{x}_{u_j,v_k}\right)$$

4. s is an isomorphism function. This means that if $s(u_1) = v_1$ and $s(u_2) = v_2$ then $(u_1, u_2) \in U$ *if and only if* $(v_1, v_2) \in V$. Consequently we can write:

   $\forall\ 0 \leq i, j \leq m,\ i \neq j$
   $\forall\ 0 \leq k, l \leq n,\ k \neq l$
   such that $u_iu_j \in F$ *and* $v_kv_l \notin E$
$$\left(\overline{x}_{u_i,v_k} \vee \overline{x}_{u_j,v_l}\right)$$

   $\forall\ 0 \leq i, j \leq m,\ i \neq j$
   $\forall\ 0 \leq k, l \leq n,\ k \neq l$
   such that $u_iu_j \notin F$ *and* $v_kv_l \in E$
$$\left(\overline{x}_{u_i,v_k} \vee \overline{x}_{u_j,v_l}\right)$$

Note that we have introduced a polynomial number of variables and clauses. We claim that $< G, H > \in ISub$ *if and only if* $< X, C > \in SAT$. Firstly assume $< G, H > \in ISub$. Let s be the total injective function that associate each vertex in U to another in V. We have to verify that $< X, C > \in SAT$. Consider the following truth assignment: $x_{u,v} = true$ if $s(u) = v$, $x_{u,v} = false$ otherwise. Let us check all clauses are satisfied: those of type 1 are satisfied since if $v_i = s(u)$ and $v_j = s(u)$ then $i = j$ being s a function; clauses of type 2 are satisfied because s is a total function, that is for each $u \in U$ there exists an element $v \in V$ such that $s(u) = v$; those of type 3 are satisfied since s is injective, that is if $s(u_i) = v$ and

$s(u_j) = v$ then $i = j$; clauses of type 4 are satisfied because s maps H into an isomorphic subgraph of G.

Conversely, suppose $< X, C >\in SAT$ is satisfiable. Let Φ be a satisfying truth assignment for $< X, C >$. Consider the map $s : U \rightarrow V$ such that $s(u) = v$ *if and only if* $\Phi(x_{u,v}) = TRUE$. Analogously we can say that satisfiability of clauses 1 implies that s is a function, satisfiability of clauses 2 means that s is a total function and of clauses 3 implies that s is an injective function; at last satisfiability of clauses 4 implies that all constraints of isomorphism are respected.

5 SAT to 3SAT

3SAT is the following problem:

3SAT
INSTANCE:
Let X be a set of n boolean variables $x_1 \ldots x_n$. An instance of 3SAT is a conjunction $\bigwedge_{i=0}^{m} C_i$ of m clauses on X where each clause has exactly three variables.
QUESTION:
Is there a truth assignment for X such that satisfies $\bigwedge_{i=0}^{m} C_i$?

Theorem 4 *SAT reduces to 3SAT.*

Proof: Let $U = \{u_1, u_2, \ldots, u_n\}$ be a set of boolean variables and $C = \{c_1, c_2, \ldots, c_m\}$ be an arbitrary set of clauses. ¡U,C¿ is a SAT instance. We construct a set C' of clauses, each made up of three literals; we also construct a new collection U' of boolean variables obtained from U and a set of additional variables U'_j used "ad hoc" in C'_j:

$$U' = U \cup \left[\bigcup_{j=1}^{m} U'_j \right]$$

and

$$C' = \bigcup_{j=1}^{m} C'_j$$

Let c_j be given by $\{x_1, x_2, \ldots, x_k\}$, literals in U used in the clause c_j. How C'_j and U'_j are constructed depends on k, the number of distinct variables used in c_j:
$-- \ k = 1 \ --$
$U'_j = \{y_j^1, y_j^2\}$
$C'_j = \{\{z_1, y_j^1, y_j^2\}, \{z_1, y_j^1, \overline{y}_j^2\}, \{z_1, \overline{y}_j^1, y_j^2\}, \{z_1, \overline{y}_j^1, \overline{y}_j^2\}\}$
$-- \ k = 2 \ --$
$U'_j = \phi$
$C'_j = \{\{c_j\}\}$
$-- \ k = 3 \ --$
$U'_j = \{y_j^1\}$
$C'_j = \{\{z_1, z_2, y_j^1\}, \{z_1, z_2, \overline{y}_j^1\}\}$

$-- \ k > 3 \ --$
$U'_j = \{y^i_j : 1 \leq i \leq k-3\}$
$C'_j = \{\{x_1, x_2, y^1_j\}\} \cup \{\{\overline{y}^i_j, x_{i+2}, y^{i+1}_j\} : 1 \leq i \leq k-4\} \cup \{\{\overline{y}^{k-3}_j, x_{k-1}, x_k\}\}$

Clearly this construction can be accomplished in polynomial time. To prove the validity of this transformation we must show that C' is satisfiable *if and only if* C is. First suppose to have a satisfying truth assignment for C called t. We show that t can be extended to obtain a satisfying truth assignment for C'.

Since t is satisfying truth assignment for C, there must be at least one integer l such that $z_l = true$. If either $l = 1$ or $l = 2$ then we set $t'(y^i_j) = false$ for $1 \leq i \leq k-3$. If either $l = k-1$ or $l = k$ then we set $t'(y^i_j) = true$ for $1 \leq i \leq k-3$. Otherwise we set $t'(y^i_j) = false$ for $1 \leq i \leq l-2$ and $t'(y^i_j) = true$ for $l-1 \leq i \leq k-3$.

Conversely, if t' is a satisfying truth assignment for C', it easy to show that the restriction of assignment to U is a satisfying truth assignment for C.

6 3SAT to VC

Vertex Cover (VC) problem is already been described in Section 2.

Theorem 5 *3SAT reduces to VC.*

Proof: Let $< U, C >$ be an instance of 3SAT, where $U = \{u_1, u_2 \dots u_n\}$ and $C = \{c_1, c_2 \dots c_m\}$. We must construct a graph $G = (V, E)$ and a positive integer $k \leq |V|$ such that G has a vertex cover of size at most K if and only if $< U, C > \in 3SAT$. For each variable $u_i \in U$, we consider a component $T_i = (V_i, E_i)$, where:
$V_i = \{u_i, \overline{u}_i\}$ $E_i = \{\{u_i, \overline{u}_i\}\}$.
For each clause $c_j \in C$, we consider a component $S_j = (V'_j, E'_j)$ composed by three vertices, connected each other by an edge:
$V'_j = \{a_1[j], a_2[j], a_3[j]\}$
$E'_j = \{a_1[j]a_2[j], a_1[j]a_3[j], a_2[j]a_3[j]\}$
Note that any vertex cover will have to contain al least one edge in V_i for each component T_i and at least two nodes in V'_j for each component S_j.

While the sets of components just defined depend only on U and $|C|$, the following part depends also on which literals occurs in which clauses. For each clause $c_j \in C$ consider its three literals denoted by x_j, y_j, z_j and add the following edge set:
$E''_j = \{a_1[j]x_j, a_2[j]y_j, a_3[j]z_j\}$
The last thing to do is to set $K = n + 2m$ and $G = (V, E)$ where:

$$V = \left[\bigcup_{i=1}^{n} V_j\right] \cup \left[\bigcup_{i=1}^{m} V'_j\right]$$

and

$$E = \left[\bigcup_{i=1}^{n} E_i\right] \cup \left[\bigcup_{i=1}^{m} E'_j\right] \cup \left[\bigcup_{i=1}^{m} E''_j\right]$$

Figure 3 shows an example of graph obtained with $U = \{u_1, u_2, u_3, u_4\}$ and $C = \{\{\overline{u}_1, \overline{u}_3, u_4\}, \{\overline{u}_1, \overline{u}_2, u_4\}\}$.

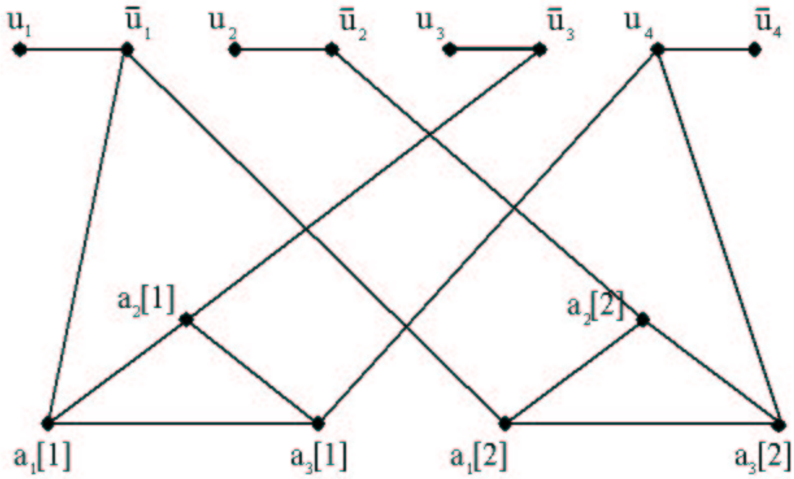

Figure 3: VC instance resulting from 3SAT instance in which $U = \{u_1, u_2, u_3, u_4\}$ and $C = \{\{\overline{u}_1, \overline{u}_3, u_4\}, \{\overline{u}_1, \overline{u}_2, u_4\}\}$.

It's easy to see that this construction can be made in polynomial time. We have to show that $< U, C > \in 3SAT$ *if and only if* $< G, k > \in VC$. Let $V' \subseteq V$ be a vertex cover for G where $|V| \leq k$. As said before V' must contain at least one vertex from each T_i and at least two vertex from each S_j. Since this gives a total of $2n + m$ and we must have that $|V'| \leq k$ and we have set $k = 2n + m$, then we consider exactly one vertex from each T_i and two vertex from each S_j. For each $1 \leq i \leq n$ assign u_i `true` if $u_i \in V'$, `false` otherwise. For each clause $c_j \in C$ consider the three edges in E''_j: only two of these can be covered by vertexes in $V'_j \cap V'$; this means that the not-covered one is covered by a vertex in V_i and, thus, implies that the j-esime clause is satisfied.
Conversely, suppose that there is a satisfying truth assignment for C. The vertex cover is made up of one vertex from each T_i and two vertex from each S_j. The vertex from T_i is u_i if the correspondent assigned value in 3SAT is true, is $\overline{u}_i$ otherwise. This implies that each set E''_j has at least one edge covered by vertex not in c_j is satisfied. Now we only include in V' the endpoints of such edge, obtaining the vertex cover.

References

[Gar79] Garey, Johnson. *Computers and Intractability, A Guide to the Theory of NP-Completeness.* W.H.Freeman and Company, 1979.