A Large Set of Non-Hamiltonian Graphs

Gyula Y. Katona *

Department of Computer and Information Sciences
Ibaraki University, Hitachi 316-8511 Japan

Department of Computer Science and Information Theory
Technical University of Budapest, 1521 Budapest, Hungary
kiskat@cs.bme.hu

1 Introduction

In the present paper we consider only undirected graphs with no loops or multiple edges. A graph G is Hamiltonian if it has a *Hamilton* cycle, i.e., a cycle containing all of its vertices.

Our work was motivated by [5] in which the authors give a large set of non–Hamiltonian graphs. Their motivation was to approximate the set, NH, of non–Hamiltonian graphs efficiently. This can be used to generate many different non–Hamiltonian graphs for testing the performance of algorithms experimentally. It is very easy to give a non–deterministic algorithm which generates all Hamiltonian graphs but such an algorithm cannot exist for the non-Hamiltonian graphs unless $\mathbf{NP} = \mathbf{co\text{-}NP}$. Therefore we can only aim to generate a "large" subset I of NH which can be useful if it satisfies certain properties.

The testing procedure is that we generate graphs for which we know whether they are Hamiltonian or not then see if the algorithm can give the right answer or not. If I is a polynomially recognizable set, then there are polynomial algorithms which always give the right answer to our test graphs. Hence we would like I to be **NP**-hard.

The set of not-1-tough graphs, $N1T$, is **NP**-complete ([2]) so it is a first step towards the approximation of NH since $N1T \subset NH$. (The definitions are given in the next section.)

In [5] the authors prove that the set of all non-sub-2-factor graphs, NS2F, is also **NP**-complete, $N1T \subset NS2F$ and the set $NS2F - N1T$ is $\mathbf{D^P}$-complete. The later theorem shows that $NS2F$ is essentially larger than $N1T$ which suggests that it is definitely more difficult to give an algorithm which works for $NS2F$ than an algorithm which works only for $N1T$. This would not be true if $NS2F - N1T \in \mathbf{P}$.

The main result of the present paper is to show that the set of all non-1-edge-tough graphs, $N1ET$, is an even better approximation of NH than $NS2F$. 1-edge-toughness was introduced by the author in [6]. First we observe that $NS2F \subset N1ET$. It is proved that the set of non-t-edge-tough graph is **NP**-complete for any positive rational number t (including $t = 1$). Finally we prove that $N1ET - NS2F$ is $\mathbf{D^P}$-complete, which suggests that $N1ET$ is essentially larger than $NS2F$. The proofs rely on the constructions given in [5].

*Supported by the Japan Society for the Promotion of Science, by Monbusho and by Hungarian National Foundation for Scientific Research, OTKA Grant Number T 030059.

2 Definitions

In 1973 Chvátal [3] introduced the definition of toughness. Many authors investigated the relation of toughness and Hamiltonicity since Chvátal's paper. Let $\omega(G)$ denote the number of components of a graph G. If $T \subseteq V(G)$ then the graph $G-T$ is defined as follows. $V(G-T) = V(G) - T$ and $(u,v) = e \in E(G-T)$ iff $e \in E(G)$ but neither of u and v is in T.

Definition 1 (Chvátal [3]) *A graph G is t–**tough** if $|S| \geq t\omega(G-S)$ for every subset S of the vertex set $V(G)$ with $\omega(G-S) > 1$. Hence a graph is not 1-tough if $\omega(G-S) > |S|$ and $\omega(G-S) > 1$ holds for some $S \subset V(G)$. The set of all non-1-tough graphs is denoted by $N1T$.*

In [4] Chvátal defined a larger class of non-Hamiltonian graphs.

Definition 2 (Chvátal [4]) *A graph G is a **sub-2-factor** if*

$$\omega(T) \leq |S| + \sum_{Q_i} \left\lfloor \frac{edge(Q_i, T)}{2} \right\rfloor$$

*holds for any partition of $V(G)$ into disjoint subsets R, S and T such that $T \neq V(G)$, where each Q_i denotes a connected component of the subgraph of G induced by R. $edge(Q_i, T)$ denotes the number of edges with one end in Q_i and the other in $T = V(G) - R - S$. Hence a graph is **non-sub-2-factor** if there exists $R, S \subseteq V(G)$ such that $R \cap S = \emptyset$ and*

$$\omega(G-R-S) > |S| + \sum_{Q_i} \left\lfloor \frac{edge(Q_i, V(G)-R-S)}{2} \right\rfloor.$$

The set of all non-sub-2-factor graphs is denoted by $NS2F$.

In the definition of toughness only one set of vertices is deleted, while in the above definition two different kinds of vertex subsets are deleted. In the following definition we delete a subset of the vertices and a subset of the edges.

For a graph G, let $X \subseteq V(G)$ and $Y \subseteq E(G-X)$. Let $Q_1, Q_2, \ldots, Q_n$ be the components of the edge–induced subgraph $G(Y)$ induced by Y. The vertices of $G(Y)$ are those vertices of G which are incident to an edge in Y and the edges of $G(Y)$ are those edges of G whose end vertices are both in the vertex set of $G(Y)$. Therefore, there may be edges in $G(Y)$ which are not in Y. Let H be a graph and B a subgraph of H. The boundary of B in H is defined as $\mathrm{bd}_H(B) = \{v \in V(B) | v \text{ has a neighbor outside of } B\}$ and $\mathrm{in}_H(B) = V(B) - \mathrm{bd}_H(B)$. Hence, in our case

$$\mathrm{bd}_{G-X}(Q_i) = \{v \in V(Q_i) \mid v \text{ has a neighbour outside of } Q_i \text{ and } X\}$$

and

$$\mathrm{in}_{G-X}(Q_i) = V(Q_i) - \mathrm{bd}_{G-X}(Q_i).$$

$G - X - Y - \cup_{Q_i} \mathrm{in}_{G-X}(Q_i)$ denotes the graph which is obtained by deleting the vertices of X and $\mathrm{in}_{G-X}(Q_i)$ and the edges of Y from G (but not deleting the endvertices of these edges).

Definition 3 [6] *Let t be a positive number. G is t–**edge–tough** if G is connected and*

$$\omega\left(G - X - Y - \bigcup_{Q_i} \mathrm{in}_{G-X}(Q_i)\right) \leq \frac{|X| + \sum_{Q_i} \left\lfloor \frac{|\mathrm{bd}_{G-X}(Q_i)|}{2} \right\rfloor}{t} =: \frac{cost(X,Y;G)}{t}$$

holds for every $X \subseteq V(G)$ and $Y \subseteq E(G-X)$ satisfying

$$\omega\left(G - X - Y - \cup_{Q_i} \mathrm{in}_{G-X}(Q_i)\right) > 1,$$

using the notation introduced above. The set of all non-t-edge-tough graphs is denoted by $NtET$.

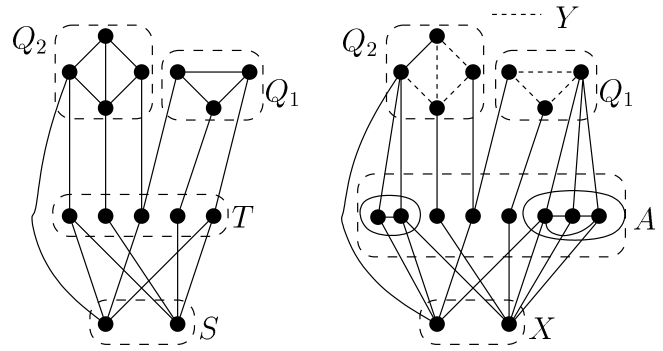

Figure 1: A non-sub-2-factor and a non-1-edge tough graph.

The first graph on Figure 1 is not a sub-2-factor graph because

$$\omega(T) = 5 > 2 + \left\lfloor \frac{3}{2} \right\rfloor + \left\lfloor \frac{3}{2} \right\rfloor = |S| + \sum_{Q_i} \left\lfloor \frac{edge(Q_i, T)}{2} \right\rfloor .$$

One can easily check that the second graph is a sub-2-factor. On the other hand it is not 1-edge-tough since

$$\begin{aligned} \omega(G - X - Y - \cup_{Q_i} \mathrm{in}_{G-X}(Q_i)) = 5 \quad & > \\ 2 + \left\lfloor \frac{3}{2} \right\rfloor + \left\lfloor \frac{3}{2} \right\rfloor \quad & = \quad |X| + \sum_{Q_i} \left\lfloor \frac{|\mathrm{bd}_{G-X}(Q_i)|}{2} \right\rfloor . \end{aligned}$$

The following Propositions describe relations of $NH, NS2F$ and $N1T$.

Proposition 4 (Chvátal [4]) *$N1T \subset NS2F \subset NH$.*

Proposition 5 [6] *$N1T \subset N1ET \subset NH$.*

Proposition 6 *$NS2F \subset N1ET \subset NH$.*

PROOF:

Lemma 7 *If $G \in NS2F$ then there exists a partition R, S, T of $V(G)$ such that for all $r \in R$ $edge(r, T) \leq 1$.*

PROOF: A vertex r is called *bad* if $edge(r, T) \geq 2$ for a $T \subset V(G)$. Since $G \in NS2F$ there is a partition R', S', T' for which

$$\omega(T') > |S'| + \sum_{Q_i'} \left\lfloor \frac{edge(Q_i', T')}{2} \right\rfloor$$

holds. Let R', S', T' be a partition which satisfies this condition and the number of bad vertices in R' is minimum. If there are no bad vertices in R' then we are done. Suppose that there is a vertex $r \in Q_i'$ for which $edge(r, T') \geq 2$. Let $R = R' - \{r\}$ and $S = S' \cup \{r\}$. It is easy to see that $T = T'$ and that

$$\left\lfloor \frac{edge(Q_i', T)}{2} - 1 \right\rfloor \geq \left\lfloor \frac{edge(Q_i, T)}{2} \right\rfloor$$

since there are at least 2 edges from Q'_i to T'. Hence

$$\omega(T) = \omega(T') > |S'| + \sum_{Q'_i} \left\lfloor \frac{edge(Q'_i, T)}{2} \right\rfloor \geq |S| - 1 + \sum_{Q_i} \left\lfloor \frac{edge(Q_i, T)}{2} \right\rfloor + 1$$

holds. This implies that the number of bad vertices decreased, which contradicts the assumption that R', S', T' is chosen such that the number of bad vertices is minimum. $\square$

Let $G \in NS2F$ and let R, S, T be a partition which satisfies the conditions of the above lemma. Let $X = S$ and $Y = E(G[R])$. It is easy to see that $G - X - Y - \bigcup_{Q_i} \mathrm{in}_{G-X}(Q_i) = T$ and $edge(Q_i, T) = \mathrm{bd}_G(Q_i)$ for every i thus

$$\begin{aligned} \omega(G - X - Y - \cup_{Q_i} \mathrm{in}_{G-X}(Q_i)) = \omega(T) \quad & > \\ |S| + \sum_{Q_i} \left\lfloor \frac{edge(Q_i, T')}{2} \right\rfloor \quad & = \quad |X| + \sum_{Q_i} \left\lfloor \frac{|\mathrm{bd}_{G-X}(Q_i)|}{2} \right\rfloor \end{aligned}$$

proving $G \in N1ET$. $\square$

The class $\mathbf{D^P}$, first defined in [7], consists of all the languages that can be expressed in the form $L_1 - L_2$ where L_1 and L_2 are in $\mathbf{NP}$. It is easy to see that $\mathbf{NP} \cup \textbf{co-NP} \subseteq \mathbf{D^P}$. A language $L \in \mathbf{D^P}$ is $\mathbf{D^P}$-complete if any language in $\mathbf{D^P}$ is reduceable to L in polynomial time. We will use the $\mathbf{D^P}$-completeness of SAT-UNSAT [7]: Given two 3-CNF predicates f and $\tilde{f}$ over disjoint variable sets, the problem asks whether it is true that f is satisfiable and $\tilde{f}$ is unsatisfiable.

3 Recognizing t-edge-tough graphs is NP-complete

Theorem 8 *Let t be a positive rational number. The problem of determining whether a given graph is in $NtET$ is* $\mathbf{NP}$*-complete.*

Proof: Let $t = a/b$ such that $a, b \geq 2$ (a and b are not necessarily relative primes). We show that the problem of deciding, given a graph G, whether or not there is a subset of vertices $X \subseteq V(G)$ and a subset of edges $Y \subseteq E(G - X)$ such that

$$a \cdot \omega(G - X - Y - \cup_{Q_i} \mathrm{in}_{G-X}(Q_i)) > b \cdot cost(X, Y; G)$$

is $\mathbf{NP}$-complete. Since we can check the condition in polynomial time for a particular X and Y the problem is in $\mathbf{NP}$. To show that the problem is $\mathbf{NP}$-hard 3-SAT is reduced to this problem.

Given a predicate f a graph G is constructed with the following properties. (i) G is not 1-tough (so it is not 1-edge-tough) if f is satisfiable. (ii) G is 1-edge-tough if f is not satisfiable. The construction of the graph is very similar to the one in [5]. The global structure is different because we allow vertices with large degree. The local structure is different only because of the different choice of a and b.

Suppose that the predicate f uses n variables, $U = \{u_1, \ldots, u_n\}$, and contains m clauses, $C = \{c_1, \ldots, c_m\}$, where the jth clause includes three laterals $x_{j,1}, x_{j,2}$ and $x_{j,3}$. For every variable u_i we have a graph GC_i which is illustrated on Figure 2. It consists of two copies of a complete graph on a vertices (K_a), YC_i and NC_i, and two copies of the complement of K_b, i. e. b independent vertices, YC'_i and NC'_i. Then we make complete bipartite connections between YC_i and NC_i, and between YC'_i and YC_i, as well as between NC'_i and NC_i.

A subgraph GD_j is also defined for each clause c_j. The construction, illustrated on Figure 3, is similar to GC_i, but now we have three complete and three empty parts.

Now we add some edges between these components according to f. If the ith variable ($i \in \{1, 2, 3\}$) in the jth clause of f is u_k then we make a complete bipartite connection between YC_k

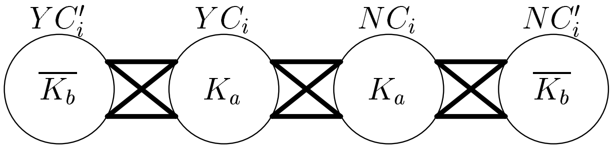

Figure 2: GC_i

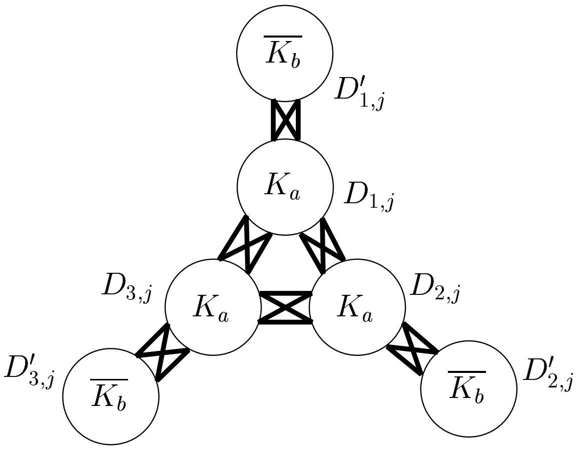

Figure 3: GD_j

and $D_{i,j}$. If it is $\overline{u_k}$ then, instead, we make the complete bipartite connection between NC_k and $D_{i,j}$. For example, if the first clause is $(\overline{u_7}, u_8, u_9)$ then there are connections between NC_7 and $D_{1,1}$, YC_8 and $D_{2,1}$, YC_9 and $D_{3,1}$. We do not add any edges from YC_i', NC_i' or $D'_{i,j}$. There are no edges between two GC_i components or between two GD_j components, either.

Next we add an empty graph B of size $3(b-1)(n+m)+1$, and a complete graph A of size $3a(n+m)$ whose vertices are connected to every other vertex of G, see Figure 4.

Lemma 9 *If f is satisfiable then G is not t-tough.*

PROOF: We show that there exists a vertex set $X \subset V(G)$ such that $t\omega(G-X) > |X|$. First of all we fix an arbitrary truth assignment of f.

Let $A \subset X$.

For each i let $YC_i \subset X$ or $NC_i \subset X$ depending of the truth assignment. If u_i is true then let $YC_i \subset X$ if u_i is false then let $NC_i \subset X$.

For each j exactly two of the vertex sets of $D_{1,j}, D_{2,j}, D_{3,j}$ will be in X. Since every clause is true in the truth assignment, every clause must contain a variable which is true. If there is more then one of these we choose an arbitrary one. The sets which corresponds to the other two variables are included in X. For example if $x_{j,1}$ is true and $x_{j,2}$ and $x_{j,3}$ are false then $D_{1,j} \not\subset X$, $D_{2,j} \cup D_{3,j} \subset X$. Note that in this way the sets, which corresponds to a false variable, are in X.

It is clear that

$$|X| = 3a(n+m) + an + 2am = 4an + 5am \tag{1}$$

Let us calculate $\omega(G-X)$. Since the vertices in B have no neighbours outside of A, they are isolated in $G-X$. If $YC_i \subset X$ then the vertices of YC_i' are isolated in $G-X$ by similar reasons. The subgraph induced by $NC_i \cup NC_i'$ will be a separate component of $G-X$, since all the edge going out from these vertices are deleted in $G-X$. Recall, that $YC_i \subset X$ implies that $\overline{u_i}$ is false

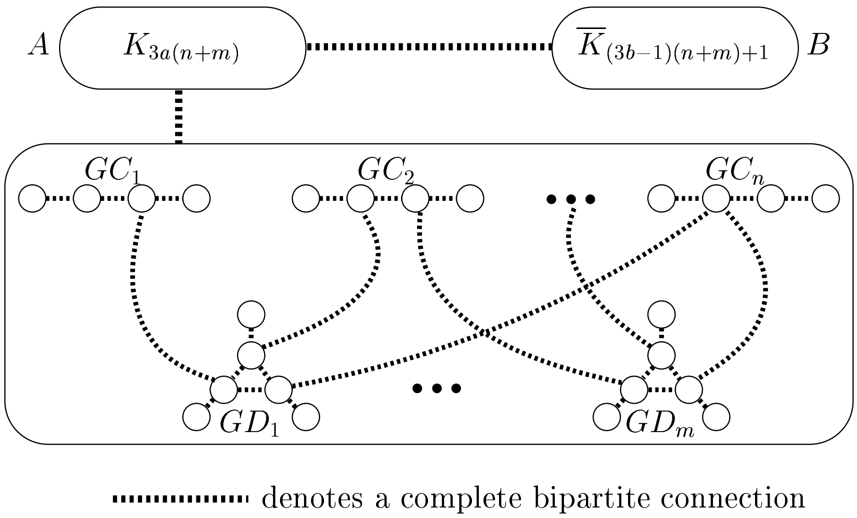

Figure 4: The construction of G

in the truth assignment. Therefore, all the neighbours of the vertices in NC_i are in X. Similar argument works if $NC_i \subset X$.

By the symmetry of GD_j, it may be supposed, that $D_{1,j} \not\subset X$, $D_{2,j} \cup D_{3,j} \subset X$. This implies that the vertices in $D'_{2,j} \cup D'_{3,j}$ are isolated vertices in $G - X$. The vertices of $D_{1,j} \cup D'_{1,j}$ induce one more component, since the vertex set YC_i or NC_i, to which $D_{1,j}$ is connected, is contained in X according to our assignment.

This gives that $\omega(G-X) = (3b-1)(n+m)+1+(b+1)n+(2b+1)m = 4bn+5bm+1$ holds, so by (1)

$$t\omega(G-X) = \frac{a}{b}(4bn+5bm+1) > 4an+5am = |X|.$$

□

Lemma 10 *If f is not satisfiable then G is t-edge-tough.*

Proof: Suppose indirectly that G is not t-edge-tough. Therefore, there exist $X \subseteq V(G)$ and $Y \subset (G-X)$ such that

$$\omega(G - X - Y - \cup_{Q_i}\mathrm{in}_{G-X}(Q_i)) > 1$$

and

$$a \cdot \omega(G - X - Y - \cup_{Q_i}\mathrm{in}_{G-X}(Q_i)) > b \cdot cost(X, Y; G).$$

Choose X and Y such

(i) $a \cdot \omega(G - X - Y - \cup_{Q_i}\mathrm{in}_{G-X}(Q_i)) - b \cdot cost(X, Y; G)$ is maximal;

(ii) X is maximal, subject to (i);

(iii) Y is minimal, subject to (i) and (ii);

Let $diff(G)$ denote the maximum value taken in (i). It follows from the definition of t-edge-toughness that $A \subseteq X$. If $v \in B \cap X$ then by choosing $X' = X - \{v\}, Y' = Y$ we obtain a contradiction with (i), so $B \cap X = \emptyset$. $A \subseteq X$ implies $B \cap G(Y) = \emptyset$, because $Y \subseteq E(G-X)$ by definition.

Proposition 11 *If the closed neighbourhood of vertices x and y is the same, then $x \in X$ implies $y \in X$. Therefore, either all vertices of a subgraph $YC_i, NC_i, D_{1,j}, D_{2,j}$ or $D_{3,j}$ are contained in X or none of them.*

PROOF: If $x \in X$ and $y \notin X$ then we can leave out x from X. In this way the number of components does not change. If two vertices are connected by a path which contains x then there is a path between them which does not contain x but it contains y, since y has the same neighbourhood. On the other hand, the cost decreases by leaving out x, which contradicts condition (i). $\square$

Proposition 12 *If an edge $e \in Y$ is between two subgraphs among $YC_i, NC_i, D_{1,j}, D_{2,j}$ and $D_{3,j}$ than $G(Y)$ contains all the edges between the two subgraphs. This implies that either all of the edges between two such subgraphs are contained in $G(Y)$ or none of them.*

PROOF: We will prove the claim, when $e = \{x, y\}$, $x \in YC_i$ and $y \in NC_i$. The proof works in the same way for the other cases.

It follows from Proposition 11 that $X \cap YC_i = \emptyset$ and $X \cap NC_i = \emptyset$.

Suppose that there is an edge $e' = \{x', y'\}$ between YC_i and NC_i which is not in $G(Y)$. This implies that at least one of x' and y' is not incident to any edge of $G(Y)$. Assume that x' has this property. Note, that x' is adjacent to all vertices in $YC_i \cup NC_i$. Therefore, leaving out e from Y does not increase the number of components. If two vertices are connected by a path which contains e then there is a path between them which does not contain e but it contains x'. On the other hand, either the cost increases, so we obtain a contradiction with condition (i) or it does not change, then a contradiction is reached with condition (iii). $\square$

Proposition 13 *None of the vertices of the $YC'_i, NC'_i, D'_{1,j}, D'_{2,j}$ and $D'_{3,j}$ are in X and none of the edges incident to these vertices are in $G(Y)$.*

PROOF: Both claims are clear from the definitions and the previous Propositions. $\square$

It is shown next that $diff(G') \geq diff(G)$ if G' is a spanning subgraph of G. It is enough to prove that for the same X and Y, for which the maximum is achieved,

$$a \cdot \omega(G' - X - Y - \cup_{Q_i} \mathrm{in}_{G-X}(Q_i)) - b \cdot cost(X, Y; G') \geq diff(G)$$

holds. Since G' is a spanning subgraph of G it is obvious that

$$\omega(G' - X - Y - \cup_{Q_i} \mathrm{in}_{G-X}(Q_i)) \geq \omega(G - X - Y - \cup_{Q_i} \mathrm{in}_{G-X}(Q_i)).$$

$cost(X, Y; G') \geq cost(X, Y; G)$ would be only possible if some of interior vertices of the Q_i components became border vertices, when we delete some edges of G to get G'. However, this is not possible by the definitions. Note, that although some of the Q_i components may break up into smaller components during the deletion process, the cost cannot increase.

Using the above observation, it would be enough to prove for a suitable spanning subgraph G' that $diff(G') \leq 0$ if f is not satisfiable, but we have to use a bit complicated method. Let G' be the spanning subgraph of G which arises by deleting all the edges between the GC_i and GD_j components. In a series of lemmas we will show that $diff(G') \leq a$ but a positive value can only be reached by a very special symmetric choice of X and Y. Then it is proved that $diff(G') - a \geq diff(G)$ if f is not satisfiable.

Let H be a GC_i or GD_j component of G'. Let

$$diff(H) = a \cdot \omega(H - X - Y - \cup_{Q_i} \mathrm{in}_{G-X}(Q_i)) - b \cdot cost(X \cap H, Y \cap H; H).$$

Note that $diff(H)$ is not necessarily the maximum of the function taken over all the choices of X' and Y' in H. The maximum is an overall maximum on G.

Sub-lemma 14 *$diff(GC_i) \le a$ holds for every $1 \le i \le n$ and $diff(GC_i) > 0$ holds if and only if $Y \cap E(GC_i) = \emptyset$ and $X \cap V(GC_i) = V(YC_i)$, $V(NC_i)$ or $\emptyset$.*

PROOF: It follows from Propositions 11-13 that there are only a few cases.

If $X \cap V(GC_i) = \emptyset$ and $Y \cap E(GC_i) = \emptyset$ then $diff(GC_i) = a$.

If $X \cap V(GC_i) = V(YC_i)$ or $X \cap V(GC_i) = V(NC_i)$ then $Y \cap E(GC_i) = \emptyset$ follows and $diff(GC_i) = a$ holds.

If $X \cap V(GC_i) = V(YC_i) \cup V(NC_i)$ then $Y \cap E(GC_i) = \emptyset$ and $diff(GC_i) = 0$.

If $X \cap V(GC_i) = \emptyset$ and $G(Y)$ contains all the edges between YC_i and NC_i then $diff(GC_i) = a(2-b) \le 0$, since $b \ge 2$. □

Sub-lemma 15 *$diff(GD_j) \le a$ holds for every $1 \le i \le m$ and $diff(GD_i) > 0$ holds if and only if $Y \cap E(GD_j) = \emptyset$ and $X \cap V(GD_i) = \emptyset$ or $X \cap V(GD_j) = V(D_{k,j})$ or $X \cap V(GD_j) = V(D_{k,j}) \cup V(D_{l,j})$ $(k, l \in \{1,2,3\})$.*

PROOF: It follows from Propositions 11-13 that there are only a few cases.

If $X \cap V(GC_i) = \emptyset$ and $Y \cap E(GC_i) = \emptyset$ then $diff(GC_i) = a$.

If $X \cap V(GC_i) = V(D_{k,j})$ and $Y \cap E(GC_i) = \emptyset$ then $diff(GC_i) = a$.

If $X \cap V(GC_i) = V(D_{k,j})$ then the only other possibility is that $G(Y)$ contains all edges between the other two $D_{l,j}$ ($l \in \{1,2,3\} - \{k\}$) subgraphs. Then $diff(GC_i) = a(b+2) - 2ba = a(2-b) \le 0$, since $b \ge 2$.

If $X \cap V(GD_j) = V(D_{k,j}) \cup V(D_{l,j})$ then $Y \cap E(GC_i) = \emptyset$ and $diff(GC_i) = a$ must hold.

If $X \cap V(GD_j) = V(D_{1,j}) \cup V(D_{2,j}) \cup V(D_{3,j})$ then $Y \cap E(GC_i) = \emptyset$ and $diff(GC_i) = 0$.

If $X \cap V(GD_j) = \emptyset$ then the only case left is when $G(Y)$ contains all edges between the three $D_{k,j}$ sets. In this case $diff(GC_i) = 3ab - b\left\lfloor \frac{3a}{2} \right\rfloor$ which is ≤ 0 if a is even or $b \ge 3$, but $= 1$ if a is odd and $b = 2$. However, this choice of X and Y violate condition (i). Remove the edges in $E(DG_i)$ from Y and add $V(D_{1,j})$ and $V(D_{2,j})$ to X. In this way the number of components is increased by 2, while the cost increases by $2a - \left\lfloor \frac{3a}{2} \right\rfloor$. Therefore, $diff(G)$ increases by

$$2a - b\left(2a - \left\lfloor \frac{3a}{2} \right\rfloor\right) = 2\left\lfloor \frac{3a}{2} \right\rfloor - 2a \ge 1.$$

□

Using the above observations and Sub-lemmas we obtain

$$\begin{aligned} diff(G') = \;& \textstyle\sum_{i=1}^{n} diff(GC_i) + \sum_{i=1}^{m} diff(GD_i) + a|B| - b|A| \le \\ & an + am + a((3b-1)(n+m)+1) - 3ab(n+m) = a. \end{aligned}$$

Moreover, if X and Y is not chosen as it is mentioned in the sub-lemmas in at least one of the subgraphs, then $diff(G') \le 0$ and we are done. Hence we may suppose, that it is chosen that way and $diff(G') = a$.

This choice of X and Y means that Y has no edge in the GC_i and GD_j components and X does not contain at least one of $V(YC_i)$ and $V(NC_i)$ for each i and at least one of $V(D_{1,j}), V(D_{2,j}), V(D_{3,j})$ for each j.

Now we put back the edges between the GC_i and GD_j components which were removed to obtain G'. These edges were constructed according to f. Recalling the ideas in the proof of Lemma 9 it is easy to see that if f is not satisfiable then $G - X$ contains a YC_i or NC_i subgraph and a $D_{k,j}$ subgraph such that there is an edge between them.

Suppose the contrary. If $YC_i \subset X$ then set $u_i = true$, if $NC_i \subset X$ then $u_i = false$. We claim, that f is true with these settings. Assume that the jth clause is false and that $D_{k,j} \not\subset X$. If the kth variable of the jth clause is u_i then $D_{k,j}$ is connected to YC_i, but since this is false $YC_i \not\subset X$, a contradiction. If the kth variable of the jth clause is $\overline{u_i}$, then $D_{k,j}$ is connected to NC_i, but since u_i is true, $NC_i \not\subset X$, a contradiction. So there must be at least one edge, which connects two components of $G - X$.

If there is such and edge, which is not in $G(Y)$ then the number of components is decreased by one compared to G', hence, $diff(G) \leq 0$ and we are done.

To finish the proof it is shown that if all such edges are in $G(Y)$ then the conditions (i) is violated. Suppose that the edges between YC_i and $D_{1,j}$ are all in $G(Y)$ which means that $V(YC_i) \cap X = \emptyset$ and $V(D_{1,j}) \cap X = \emptyset$. Leave out the edges between YC_i and $D_{1,j}$ from Y and add $V(YC_i)$ to X. In this way, it is clear that the number of components increase, while the cost does not change, since $V(YC_i) \cup V(D_{1,j})$ were all boundary vertices. $\square$ $\square$

4 $N1ET$ is essentially larger than $NS2F$

Theorem 16 *The problem of determining whether a given graph is in $N1ET - NS2F$ is $\mathbf{D^P}$-complete.*

Proof: Since $N1ET$ and $NS2F$ are both in **NP**, $N1ET - NS2F \in \mathbf{D^P}$. To show the completeness the following SAT-UNSAT problem is reduced to our problem: Given two 3-CNF predicates over a disjoint variable set. Is it true that f is satisfiable and $\tilde{f}$ is not satisfiable? We may suppose, without loss of generality, that f and $\tilde{f}$ both uses n variables and m clauses. We construct a graph G a similar way as in the proof of Theorem 8. The construction is also similar to the one in [5], however the differences are crucial in order to obtain the different result.

Using the notations of the proof of Theorem 8 take a graph GC_i for every variable in f and a graph GD_j for every clause in f. GC_i and GD_j are exactly the same as in the proof of Theorem 8. Next we take a graph $\widetilde{GC_i}$ for every variable in $\tilde{f}$ and a graph $\widetilde{GD_j}$ for every clause in $\tilde{f}$. These graphs are illustrated on Figure 5 and 6. They are obtained from GC_i and GD_j by taking $a = b = 2$ and making some modifications.

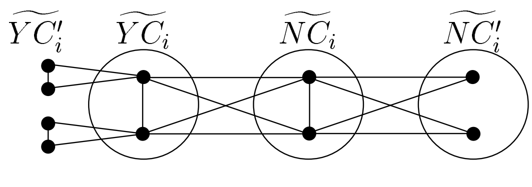

Figure 5: $\widetilde{GC_i}$

We add the edges between GC_i and GD_j according to the truth assignment of f in the same way, as in the construction in Theorem 8 . Next, edges are added between $\widetilde{GC_i}$ and $\widetilde{GD_j}$ in the same way, according to the truth assignment of $\tilde{f}$. Finally, a component B is added which consists of $10(n+m)+1$ independent vertices and every vertex is connected to all vertices of a complete graph A on $12(n+m)$ vertices (see Figure 7).

In series of lemmas it is proved that if f is satisfiable and $\tilde{f}$ is not satisfiable then G is a sub-2-factor graph which is not 1-edge-tough. This implies that $G \in N1ET - NS2F$ in this case.

Applying the method used in [5] it is easy to prove that if f and $\tilde{f}$ are both satisfiable then $G \in N1T \subset NS2F$.

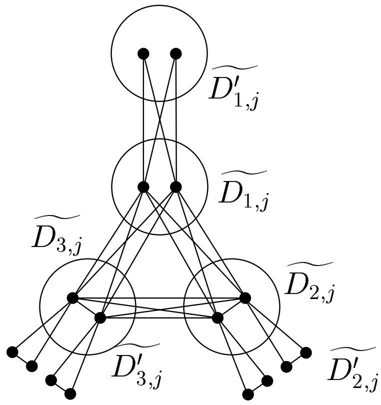

Figure 6: $\widetilde{GD_i}$

Finally, using the same method and some lemmas that are used in the other proofs it is easy to prove that if f is not satisfiable then G is 1-edge-tough. Hence, we have shown that $G \in N1ET - NS2F$ if and only if f is satisfiable and $\tilde{f}$ is not satisfiable, which completes the proof.

Lemma 17 *If f is satisfiable and $\tilde{f}$ is not, then $G \in N1ET$ and $G \notin NS2F$ ($G \in S2F$).*

PROOF: We prove first that $G \in N1ET$, namely that there exist a pair X, Y such that $\omega(G - X - Y) > cost(X, Y; G)$. (There are no interior vertices in our choice.)

First of all, we fix an arbitrary truth assignment of f. Let X be defined exactly the same way as in Lemma 9. It is the union of A, one of YC_i and NC_i for each i and two of $D_{1,j}, D_{2,j}, D_{3,j}$ for each j, according to the truth assignment of f.

Let Y consists of all edges between any two of the $\widetilde{YC_i}, \widetilde{NC_i}, \widetilde{D_{1,j}}, \widetilde{D_{2,j}}, \widetilde{D_{3,j}}$ subgraphs for all i and j.

Let us calculate $\omega(G-X-Y)$. It is easy to see that the vertices of B are isolated in $G-X-Y$, each GC_i breaks up into 3 components and each GD_j into 5 components. In the other part, each $\widetilde{GC_i}$ will contribute also 3 components and each $\widetilde{GD_j}$ 5 components. Hence,

$$\omega(G - X - Y) = 10(n + m) + 1 + 3n + 5m + 3n + 5m = 16n + 21m + 1.$$

To calculate the cost, note that all vertices of $\widetilde{YC_i}, \widetilde{NC_i}, \widetilde{D_{1,j}}, \widetilde{D_{2,j}}, \widetilde{D_{3,j}}$ are boundary vertices and that

$$\sum_{Q_i} \left\lfloor \frac{|\mathrm{bd}_{G-X}(Q_i)|}{2} \right\rfloor \leq \frac{|V(G(Y))|}{2}.$$

Therefore,

$$cost(X, Y; G) \leq 12(n + m) + 2n + 4m + \frac{4n + 6m}{2} = 16n + 19m < \omega(G - X - Y)$$

holds, which implies $G \in N1ET$.

To prove the other claim of the Lemma, suppose, that $G \in NS2F$. By Lemma 7 this means that there exist a partition of $V(G)$ into R, S, T such that

$$\omega(T) > cost(R, S) = |S| + \sum_{Q_i} \left\lfloor \frac{edge(Q_i, V(G) - R - S)}{2} \right\rfloor$$

and for all $r \in R$ $edge(r, T) \leq 1$.

Choose R and S such that

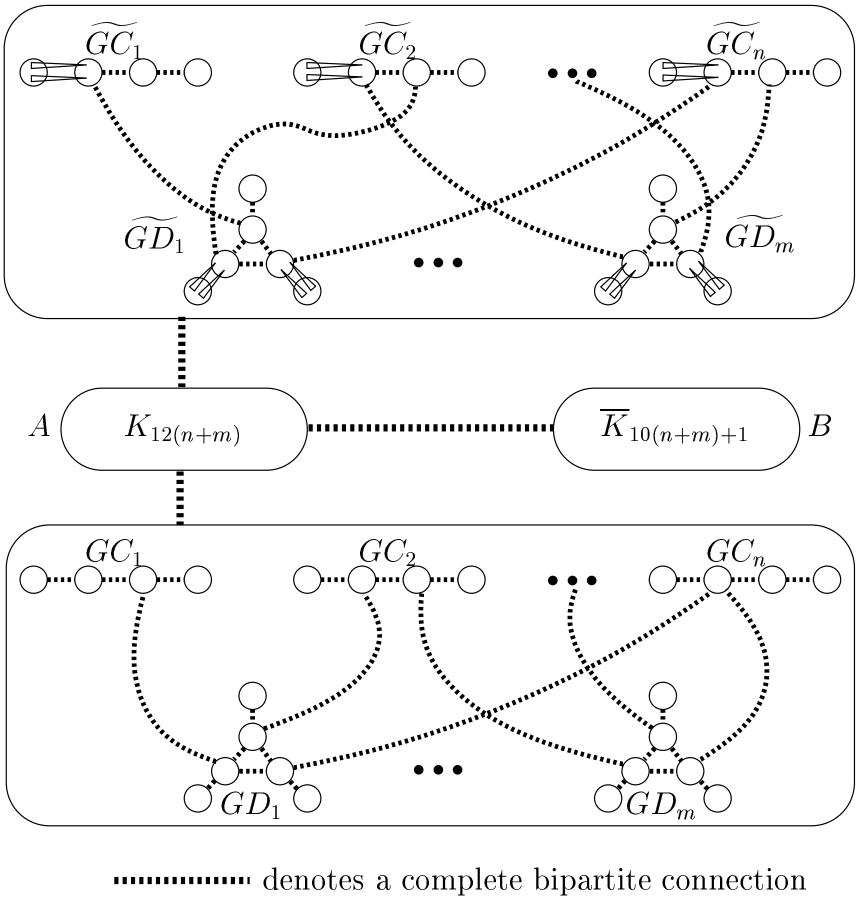

Figure 7: The construction of G

(i) $\omega(T) - cost(R, S)$ is maximal;

(ii) S is maximal, subject to (i);

(iii) R is minimal, subject to (i) and (ii);

Let $diff(G)$ denote the maximum value taken in (i).

It is shown first that $A \subseteq S$. Suppose that $T \cap A \neq \emptyset$. This implies that $\omega(T) = 1$, therefore, $cost(R, S) = 0$, hence $|S| = 0$ and $e(Q_i, T) = 1$ must hold for each i. Let $\{x, y\}(x \in T, y \in Q_i)$ such an edge. If x is moved to S from T then $w(T) = 2$ and $cost(R, S) = 1$, so $diff(G)$ does not change, but (ii) is violated, so we have a contradiction. If there exists a vertex $x \in A \cap R$ then $|T| = 1$ and $G - T \subset S \cup R$ follows from the fact that for all $r \in R$ $edge(r, T) \leq 1$. This is obviously a contradiction.

We prove now that $B \subseteq T$. If $b \in B$ is in R or S then moving b to T, increases the number of the components, by the previous claim, but cannot increase the cost. Hence, (i) is violated and a contradiction is reached.

Note that two vertices of T which are in different components of $G - S - R$ cannot be connected.

Proposition 18 *None of the vertices of the* $YC'_i, NC'_i, D'_{1,j}, D'_{2,j}, D'_{3,j}, \widetilde{YC'_i}, \widetilde{NC'_i}, \widetilde{D'_{1,j}}, \widetilde{D'_{2,j}}, \widetilde{D'_{3,j}}$ *are in* $S \cup R$.

PROOF: Suppose the contrary, that there is a vertex $x \in S \cup R$ in one of these subgraphs. Move x to T. The number of components cannot decrease in this way, because any two neighbour of such a vertex is connected by and edge, thus it cannot be in different components. On the other hand, the cost decreases if $x \in S$. If $x \in R$ then the cost may not decrease, but R becomes smaller, violating (iii). $\square$

Proposition 19 $R = \emptyset$.

PROOF: By the previous Proposition all the vertices mentioned there are in T. The leftover vertices all have at least two neighbours among these vertices, therefore, they cannot be in R, since for all $r \in R$ $edge(r, T) \leq 1$. $\square$

Proposition 20 *If the closed neighbourhood of vertices x and y is the same then $x \in S$ implies $y \in S$. Therefore, either all vertices of a subgraph $YC_i, NC_i, D_{1,j}, D_{2,j}, D_{3,j}, \widetilde{NC_i}$ or $\widetilde{D_{1,j}}$ are contained in S or none of them.*

PROOF: If $x \in S$ and $y \notin S$ then we can leave out x from S. In this way the number of components does not change. If two vertices are connected by a path which contains x then there is a path between them which does not contain x but it contains y, since y has the same neighbourhood. On the other hand, the cost increases by leaving out x. This contradicts condition (i). $\square$

Sub-lemma 21 $\omega(T \cap \widetilde{GC_i}) - cost(R \cap V(\widetilde{GC_i}), S \cap V(\widetilde{GC_i})) \leq 1$ *holds for every* $1 \leq i \leq n$ *and equality holds if and only if* $T \cap (\widetilde{YC_i} \cup \widetilde{NC_i}) \neq \emptyset$.

PROOF: By the previous propositions and by the symmetry of the graph there are only a few cases which we have to check. In each case $\omega(\widetilde{GC_i} - S)$ and the cost is calculated for the actual choice of $V(\widetilde{GC_i}) \cap S$, see Figure 8. $\square$

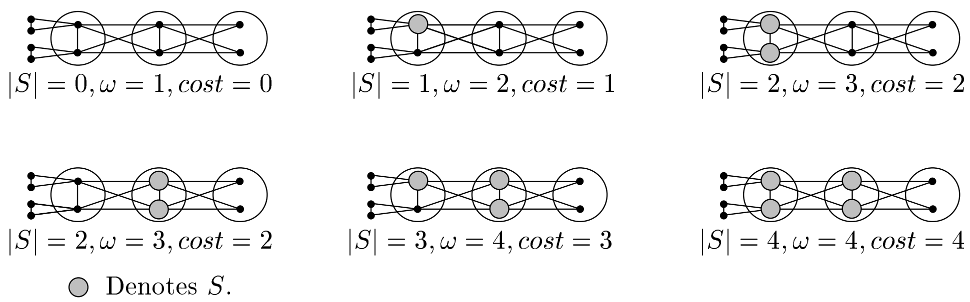

Figure 8: The remaining cases

Sub-lemma 22 $\omega(T \cap GC_i) - cost(R \cap V(GC_i), S \cap V(GC_i)) \leq 1$ *holds for every* $1 \leq i \leq n$ *and equality holds if and only if* $T \cap (YC_i \cup NC_i) \neq \emptyset$.

PROOF: The number of cases is even less than the previous case, it is left to the reader to check them. $\square$

Sub-lemma 23 $\omega(T \cap \widetilde{GD_j}) - cost((R \cap V(\widetilde{GD_j}), S \cap V(\widetilde{GD_j})) \leq 1$ *holds for every* $1 \leq j \leq m$ *and equality holds if and only if* $T \cap (\widetilde{D_{1,j}} \cup \widetilde{D_{2,j}} \cup \widetilde{D_{3,j}}) \neq \emptyset$.

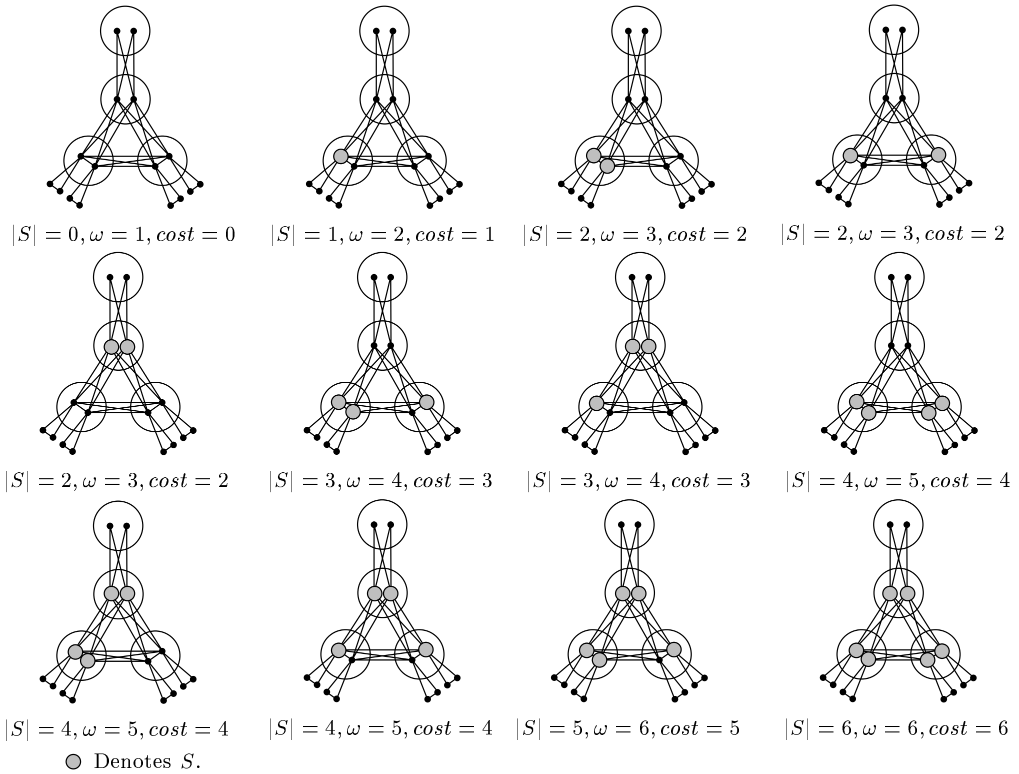

Figure 9: The remaining cases

PROOF: The number of leftover cases is small again, so we simply check them one by one, see Figure 9. $\square$

Sub-lemma 24 *$\omega(T \cap GD_j) - cost((R \cap V(GD_j), S \cap V(GD_j)) \leq 1$ holds for every $1 \leq j \leq m$ and equality holds if and only if $T \cap (D_{1,j} \cup D_{2,j} \cup D_{3,j}) \neq \emptyset$.*

PROOF: It is left to the reader to check the few remaining cases. $\square$

Note, that $w(T)$ cannot be larger then the sum of the components in for the $GC_i, GD_j, \widetilde{GC}_i, \widetilde{GD}_j$ subgraphs, since adding the edges between these subgraphs can only decrease the number of components. On the other hand, the overall cost obviously cannot be smaller than the sum of the costs for each component, because the number of $R - T$ edges can only increase. Therefore, if equality does not hold for at least one of the $GC_i, GD_j, \widetilde{GC}_i, \widetilde{GD}_j$ components then we obtain

$$\omega(T) - cost(S, R) \leq 10(n + m) + 1 - 12(n + m) - 2n + 2m + (-1) \leq 0$$

and hence, the proof of the Lemma is finished. If equality holds in each case then the edges between the remaining parts of $GC_i, GD_j, \widetilde{GC}_i, \widetilde{GD}_j$ must be considered. $\omega(T) - cost(S, R) = 1$ would imply that no two components of the $GC_i, GD_j, \widetilde{GC}_i, \widetilde{GD}_j$ subgraphs are connected by an edge after deleting S and R. Since these cross edges are constructed according to f and $\widetilde{f}$, the argument at the end of the proof of Theorem 8 implies that both f and $\widetilde{f}$ are satisfiable, a contradiction. $\square$

Lemma 25 *If f and $\widetilde{f}$ are both satisfiable then $G \in N1T$ and therefore, $G \in NS2F$.*

PROOF: We give a set $X \subset V(G)$ for which $\omega(G-X) > |X|$. The construction of X starts in the same way as in the proof of Lemma 9. Let X be the union of A, one of YC_i and NC_i for each i and two of $D_{1,j}, D_{2,j}, D_{3,j}$ for each j according to the truth assignment of f. Now repeat the same process for the other part of the graph, so X also contains one of $\widetilde{YC_i}$ and $\widetilde{NC_i}$ for each i and two of $\widetilde{D_{1,j}}, \widetilde{D_{2,j}}, \widetilde{D_{3,j}}$ for each j according to the truth assignment of $\widetilde{f}$. It is easy to see that

$$\omega(G-X) = 10(n+m) + 1 + 3n + 5m + 3n + 5m = 16n + 20m + 1$$

and

$$|X| = 12(n+m) + 2n + 4m + 2n + 4m = 16n + 20m < \omega(G-X)$$

holds. $\square$

Lemma 26 *If f is not satisfiable then G is 1-edge-tough and therefore, $G \notin N1ET$.*

PROOF: This proof is almost the same as the proof of Lemma 10 with $a = b = 2$, so it is left to the reader to check the details. Of course, the graph is different there, but there are no difference in the methods. The only new part is that we have to prove the analogue of Sub-lemma 14 and 15 for $\widetilde{GC_i}$ and $\widetilde{GD_j}$, but it is not difficult to do. $\square$ $\square$

5 Conclusions

As it was explained in the introduction Theorem 8 and 16 together shows that the set of not 1-edge-tough graphs is large and complicated, moreover, it is essentially larger than the previously known sets of non-Hamiltonian graphs. The obvious question is that are there larger sets of non-Hamiltonian graphs. The graphs given in [1] are known to be 1-edge-tough and non-Hamiltonian. However, it is not clear if it is possible to obtain an essentially larger class of graphs using these graphs.

References

[1] D. Bauer, H. J. Broersma, and H. J. Veldman, Not every 2-tough graph is Hamiltonian, to appear in *Discrete Applied Mathematics*

[2] D. Bauer, S. L. Hakimi, E. Schmeichel, Recognizing tough graphs is NP-hard, *Discrete Applied Mathematics*, **28** (1990) 191-195.

[3] V. Chvátal, Tough graphs and Hamiltonian circuits, *Discrete Mathematics* **5** (1973) 215-228.

[4] V. Chvátal, Hamiltonian cycles, *in: The Traveling Salesman Problem,* Wiley, New York (1985) 403-429.

[5] K. Iwama, E. Miyano, Better approximations of non-Hamiltonian graphs, *Discrete Applied Mathematics*, **81** (1998) 239-261.

[6] G. Y. Katona, Toughness and edge–toughness, *Discrete Mathematics* **164** (1997) 187-196.

[7] C. H. Papadimitriu, M. Yannakakis, The complexity of facets (and some facets of complexity). *J. Comput. Systems Sci.* **28** (1984) 244-259.