*Discrete Mathematics and Theoretical Computer Science* DMTCS vol. **7**, 2005, 141–154

# *Recognizing Maximal Unfrozen Graphs with respect to Independent Sets is CO-NP-complete*

N. Abbas and J. Culberson and L. Stewart

*Department of Computing Science, University of Alberta, Edmonton, Alberta, Canada, T6G 2E8*

*received Sep 2004*, *revised Jun 2005*, *accepted Jun 2005*.

A graph is unfrozen with respect to $k$ independent set if it has an independent set of size $k$ after the addition of any edge. The problem of recognizing such graphs is known to be NP-complete. A graph is maximal if the addition of one edge means it is no longer unfrozen. We designate the problem of recognizing maximal unfrozen graphs as $\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$ and show that this problem is CO-NP-complete. This partially fills a gap in known complexity cases of maximal NP-complete problems, and raises some interesting open conjectures discussed in the conclusion.

**Keywords:** independent set, edge addition, unfrozen, extremal graph, critical graph, complexity

## 1 Introduction

In this paper we present a construction that entwines both extremal graph theory and computational complexity, with original motivations stemming from the physical notions of statistical mechanics as applied to the typical case complexity associated with random graph thresholds.

Any subset $\mathcal{P}$ of the set of all graphs[†] $\Omega$ is said to be a *property* of graphs, and a graph $\mathrm{G} \in \mathcal{P}$ is said to have the property. A property is said to be *a monotone property* (with respect to deletion of edges) if whenever $\mathrm{G} \subseteq \mathrm{G}'$ are two graphs on the same vertex set and $\mathrm{G}'$ has the property, then so does G.

Given a monotone property $\mathcal{P}$, we say that a graph G is *unfrozen* with respect to $\mathcal{P}$, written $\mathrm{G} \in \mathcal{U}(\mathcal{P})$, if $\mathrm{G} \in \mathcal{P}$ and remains in $\mathcal{P}$ under the addition of any edge. Note that $\mathcal{U}(\mathcal{P})$ is also a monotone property of graphs. Recognizing unfrozen graphs with respect to NP-complete properties is frequently, but not always, NP-complete [3, 4].

Given a non-trivial[‡] monotone property $\mathcal{P}$, a graph G is *maximal* with respect to $\mathcal{P}$, written $\mathrm{G} \in \mathrm{MAX}(\mathcal{P})$, if $\mathrm{G} \in \mathcal{P}$ but with the addition of any new edge $e$, $\mathrm{G} + e \notin \mathcal{P}$. $\mathrm{MAX}(\mathcal{P})$ is a property, but is not monotone. For most NP-complete properties $\mathcal{P}$, recognizing graphs in $\mathrm{MAX}(\mathcal{P})$ can be done in polynomial time [4]. In [4] only two potential exceptions were found, one of which was isomorphism

[†] Detailed definitions pertaining to graphs are given in Section 2.

[‡] $\mathcal{P} = \emptyset$ and $\mathcal{P} = \Omega$ are trivial properties.

complete and the other NP-hard. This is in distinct contrast to the results for extremal versions of properties in CO-NP where recognizing critical graphs for many properties (e.g. critical colorability and critical Hamiltonian path[§]) are DP-complete problems[25]. At this time we know of no proof that the maximal version of any NP-complete property is DP-complete. This paper provides another maximal property that is unlikely to be in P because in this case it is CO-NP-complete, and it is the first such result on the complexity of $\mathrm{MAX}(\mathcal{U}(\mathcal{P}))$ for any property $\mathcal{P}$ known to be NP-complete.

Here $\mathcal{P} = k$-SET, the set of graphs containing an independent set of size $k$. From [4] we know that $\mathcal{U}(k\text{-SET})$ is NP-complete. We recall that the reduction $k$-SET $\propto_m \mathcal{U}(k\text{-SET})$ was particularly simple, being the complete join of two copies of the initial graph G. Also, we recall that $\mathrm{MAX}(k\text{-SET})$ is easily seen to be in the complexity class P because such graphs must consist of a $k$-set completely joined to a clique.

Nevertheless, in this paper we show that $\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$ is CO-NP-complete. This result is also interesting in that showing NOT-$\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$ is in NP requires similar theoretical machinery as does showing it is NP-hard.

Studying unfrozen and maximal properties is related to other extremal graph research where the emphasis is on obtaining bounds on the number of vertices or edges in a graph with certain properties. We briefly review those results that seem most closely related to our results.

The intersection of all maximum independent sets has been dubbed the *core* in the literature. Various lower bounds on the core size as it relates to the size of the maximum independent set, the number of vertices, and the size of the maximum matching have been obtained or strengthened by Hammer et al. [18], Levit and Mandrescu [23], and Boros et al. [9]. Gunther et al. [17] and Zito [30] have shown that the core of a tree cannot have only one vertex. Zito [30] has shown that not all vertices of a tree are in a maximum independent set unless the tree is an even path.

For a graph to be unfrozen with respect to the maximum independent set, no more than one vertex can be in the core, as noted by Haynes et al. [21]. In [23] Levit and Mandrescu have noted that the size of the maximum independent set cannot exceed half the number of the vertices for the graph to be unfrozen. Gunther et al. [17] have constructively characterized unfrozen trees.

From a more practical viewpoint, if a graph is unfrozen it means that the property is immune to small changes in the graph structure. It might conceivably be of interest to know how far we can push such resiliency.

The original motivation for the line of enquiry in this paper began with the observed easy-hard-easy pattern in various NP-complete problems near the threshold (or phase transition in general) [10, 14, 27]. Briefly, Friedgut [14] characterized when such properties have sharp thresholds; that is, for random graphs as we vary the edge probability $p$, the probability of property $\mathcal{P}$ exhibits a sharp drop at a critical value of $p = p_c$. It has been observed that for many such problems, instances randomly generated with edge probability $p$ near the threshold are exponentially more difficult to solve for all known solvers than those generated with $p$ either less than or greater than the critical value. However, this does not hold for every monotone NP-complete property [12]. Investigation of maximal properties is motivated by the desire to understand from a complexity theoretic basis when the easy-hard-easy pattern holds and when it might not. Frozen and unfrozen properties are also thought to be related to issues of complexity near the threshold [11], and the study of maximal unfrozen properties follows naturally.

In the next section we present the formal definitions required for the particular results of this paper,

[§] Note that for Hamiltonian path and certain other properties, the direction of monotonicity is reversed.

and follow that in Section 3 by a partial characterization of $\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$ which can be computed in polynomial time. In Section 4 we use those results to obtain our main theorem. In the final section we observe that in some ways this result creates more questions than it answers, and pose a few of what we consider to be the most interesting open questions.

## 2 Definitions

A *graph* $\mathrm{G} = (\mathrm{V}, \mathrm{E})$ consists of a set of vertices $\mathrm{V} = \{v_1, \ldots, v_n\}$ and a set E of undirected edges, where an edge is a 2-subset of V. Any pair $e = \{u, v\} \subseteq \mathrm{V}, e \notin \mathrm{E}$ is called a non-edge in G, and the set of all non-edges is referred to as $\mathrm{E}^c$. Given a graph G, we use $\mathrm{V_G}, \mathrm{E_G}$ and $\mathrm{E}^c_{\mathrm{G}}$ to refer to the vertex, edge and non-edge sets of G. For any subset $A \subseteq \mathrm{V}$, we use the notation $\mathrm{G}[A] = (A, \mathrm{E}[A])$ for the subgraph induced by $A$.

If $\mathrm{G} = (\mathrm{V}, \mathrm{E})$ and $\mathrm{G}' = (\mathrm{V}', \mathrm{E}')$ are two graphs such that $\mathrm{V} \subseteq \mathrm{V}'$ and $\mathrm{E} \subseteq \mathrm{E}'$ then we say that G is a *subgraph* of $\mathrm{G}'$, and write $\mathrm{G} \subseteq \mathrm{G}'$. We use the notation $\mathrm{G} + e = (\mathrm{V}, \mathrm{E} \bigcup \{e\}), e \in \mathrm{E}^c$ and similarly for addition of vertices. We replace $+$ by $-$ for deletion of vertices and edges. Deletion of a vertex implies that all edges of which the vertex is a member are also deleted.

Given a graph G, an *independent set* is a subset $\mathrm{A} \subseteq \mathrm{V}$ such that for all $u, v \in \mathrm{A}, \{u, v\} \notin \mathrm{E}$. We refer to an independent set of size $k$ by the shorthand $k$-set and the set of graphs with a $k$-set as $k$-SET. It is easy to see that $k$-SET is a monotone property of graphs, and thus so is $\mathcal{U}(k\text{-SET})$. We let X be the set of isolated vertices in G, that is the vertices with no edges.

We use the notation $\Delta = \Delta(\mathrm{G})$ to represent the maximum degree of G.

The independent set problem, given a graph G and integer $k$ is $\mathrm{G} \in k$-SET, is well known to be NP-complete [15]. In [4] it was shown that a number of unfrozen versions of monotone NP-complete problems, including $\mathcal{U}(k\text{-SET})$, are also NP-complete, while others are in P.

We define $\mathcal{I}_k = \mathcal{I}_k(\mathrm{G}) = \{I : I \text{ is a } k\text{-set}\}$. We say that $\mathcal{H} \subseteq \mathcal{I}_k$ is a $k$-set *cover (or cover) of* G if $\cup_{I \in \mathcal{H}} I = \mathrm{V}$. A cover $\mathcal{H}$ is *minimal* if no proper subset of $\mathcal{H}$ covers G.

## 3 Some structural properties of $\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$

In this section we shall show that except for some special cases and vertices of degree $n-1$, the vertices of any graph in $\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$ can be partitioned into three classes which we designate as the sole vertices $\mathcal{S}$, core vertices $\mathcal{C}$ and others, $\mathcal{O}$. Each sole vertex occurs in exactly one maximum $k$-set and has degree $n-k$, the cores are closely associated with the soles, and the structure of the graph induced by $\mathcal{S} \cup \mathcal{C}$ is such that these can be identified in polynomial time. On the other hand given a graph with the appropriate sole and core structure, we can embed an arbitrary graph in $\mathcal{O}$ and whether or not the result will be in $\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$ depends on the properties of this graph. In Section 4 this is the key idea used in showing our complexity result.

First, we eliminate the special cases and vertices of too high degree, so that we can focus on the interesting properties mentioned above.

**Observation 3.1** $\mathrm{MAX}(\mathcal{U}(1\text{-SET}))$ *consists only of complete graphs (vacuously since there are no non-edges).*

**Observation 3.2** *If $k > 1$ then no vertex of degree $n-1$ can be in any* $k$-set*, thus the removal of such vertices will not alter the membership of* G *in* $\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$.

**Observation 3.3** *If there are no vertices of degree $n-1$ then the only graphs in* $\mathrm{MAX}(\mathcal{U}(2\text{-SET}))$ *are those of 3 vertices and one edge, and the square.*

**Observation 3.4** *If* G *is a* $k+1$-set *then* G *is in* $\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$ *and these are the only graphs in* $\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$ *with* $n \le k+1$.

**Observation 3.5** *If* G *is the complete bipartite graph $K_{k,k}$ or* G *consists of an isolated vertex and the complete bipartite subgraph $K_{k-1,k-1}$ then* G *is in* $\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$.

These observations justify restricting our attention to those graphs $\mathrm{G} = (\mathrm{V}, \mathrm{E}) \in \mathrm{MAX}(\mathcal{U}(k\text{-SET}))$ satisfying the following assumptions.

**Assumption 3.6** $k > 2$.

**Assumption 3.7** G *has no vertices of degree $n-1$.*

**Assumption 3.8** $|\mathrm{V}| > k+1$. *This eliminates the trivial case in Observation 3.4.*

**Assumption 3.9** ** G *is not one of the graphs indicated in Observation 3.5. This assumption only becomes effective after Lemma 3.17. We list it here for easy reference.*

Notice that by definition, if G is in $\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$ then for every non-edge pair of vertices there must be another pair such that adding both edges eliminates every independent set. We refer to such pairs as *destroying edges* or *destroying pairs*.

**Lemma 3.10** *The sets $I \in \mathcal{I}_k(\mathrm{G})$ are maximum independent sets.*

**Proof:** Let $I$ be a maximum independent set in G. Clearly $|I| \le k+1$, otherwise not all $k$-independent sets can be destroyed with two edges. Now suppose $|I| = k+1$ and note that $\mathrm{V} \setminus I$ is not empty by Assumption 3.8. Let $u \in \mathrm{V} \setminus I$ such that $\{u, v\} \in \mathrm{E}^c$ (such a $v$ exists by Assumption 3.6). Then $\{u, v\}$ is not part of any destroying pair, since $G + \{u, v\}$ has a $k+1$-set, namely $I$. This contradicts the hypothesis that $\mathrm{G} \in \mathrm{MAX}(\mathcal{U}(k\text{-SET}))$. □

Let $\mathrm{X} \subseteq \mathrm{V}$ be the set of isolated vertices.

**Lemma 3.11** $|\mathrm{X}| \le 1$ *and, for all* $I \in \mathcal{I}_k(\mathrm{G}), \mathrm{X} \subseteq I$.

**Proof:** As observed in [21], the intersection of all $k$-independent sets in G, $\cap_{I \in \mathcal{I}_k} I$, contains at most one vertex, otherwise there is an $e \in \mathrm{E}^c$ such that $\mathrm{G} + e$ has no $k$-independent set, violating that G is unfrozen. Trivially, every isolated vertex must appear in every maximum set, and so by Lemma 3.10 in every $k$-set. □

**Lemma 3.12** *For every pair of vertices $\{u, v\} \in \mathrm{E}^c$, there is some $I \in \mathcal{I}_k$ that contains both $u$ and $v$, and there is some $I' \in \mathcal{I}_k$ that contains at most one of $u$ and $v$.*

**Proof:** Suppose there is no $k$-set that contains both $u$ and $v$, then $\mathcal{I}_k(\mathrm{G} + e) = \mathcal{I}_k(\mathrm{G})$, where $e = \{u, v\}$. But this is contradicted by the fact that the intersection over $\mathcal{I}_k(\mathrm{G})$ has at most one vertex, while the intersection over $\mathcal{I}_k(\mathrm{G} + e)$ has at least two vertices (otherwise, there is no destroying edge for $e$ [3]).

There must be some $k$-set not containing both $u$ and $v$, or as observed in [21] G would not be unfrozen. □

**Lemma 3.13** *A cover $\mathcal{H} \subseteq \mathcal{I}_k$ exists and for any such cover, $\cap_{I\in\mathcal{H}} I = \mathrm{X}$.*

**Proof:** By Assumption 3.7 every vertex is part of at least one non-edge and so by Lemma 3.12 is in some $I \in \mathcal{I}_k$. By Lemma 3.11, $\mathrm{X} \subseteq \cap_{I\in\mathcal{H}} I$.

Consider any $v \in \mathrm{V} \setminus \mathrm{X}$ and $\{v, w\} \in \mathrm{E}$. Since $\mathcal{H}$ is a cover of V, let $I \in \mathcal{H}$ be such that $v \in I$ and $I' \in \mathcal{H}$ be such that $w \in I'$, and note that $v \notin I'$. Since this holds for any $v \in \mathrm{V} \setminus \mathrm{X}$, this implies $\cap_{I\in\mathcal{H}} I \subseteq \mathrm{X}$. □

**Lemma 3.14** *For any cover $\mathcal{H} \subseteq \mathcal{I}_k$, for every $e = \{u, v\} \in \mathrm{E}^c$, there is $I \in \mathcal{H}$ such that $e \subseteq I$.*

**Proof:** Suppose $e \in \mathrm{E}^c$ is not contained in any $I \in \mathcal{H}$. Then $\mathcal{H} \subseteq \mathcal{I}_k(G + e)$ which, by Lemma 3.13 and Lemma 3.11 implies $| \cap_{I\in\mathcal{I}_k(G+e)} I| \leq 1$, leaving no destroying edge for $e$. □

We are now in a position to discuss the sole vertices $\mathcal{S}$ mentioned at the beginning of this section. We define the sole vertices by $\mathcal{S} = \{v \in \mathrm{V} : \deg(v) = n - k\}$. We call these *sole* vertices since trivially such vertices belong to exactly one $k$-set. Lemma 3.15 below shows that $\mathcal{S}$ is not empty, and that these are the vertices of maximum degree under our assumptions.

**Lemma 3.15** *$\Delta = n - k$, and for any minimal cover $\mathcal{H} \subseteq \mathcal{I}_k$, for all $I \in \mathcal{H}$ there exists a vertex $v \in I$ with $\deg(v) = n - k$.*

**Proof:** By Assumption 3.7, $\Delta < n - 1$. Any vertex $v$ of degree $n - k < \deg(v) < n - 1$ would not be in any $k$-set, and thus the non-edges on $v$ would violate Lemma 3.12. Thus, $\Delta \leq n - k$. Let $\mathcal{H} \subseteq \mathcal{I}_k$ be a *minimal* $k$-set cover of G. Since $\mathcal{H}$ is minimal, each $I \in \mathcal{H}$ contains a vertex $v$ that does not appear in any other element of $\mathcal{H}$. Then by Lemma 3.14 for each such $v$ there must exist an edge $\{v, u\}$, for every $u \notin I$. □

Since each vertex in $\mathcal{S}$ occurs in exactly one $I \in \mathcal{I}_k$, there exists a $q$-subset $\mathcal{Z} = \{I_i, 1 \leq i \leq q\} \subseteq \mathcal{I}_k$ which partitions $\mathcal{S}$ into $S_i, 1 \leq i \leq q$. We call the $I \in \mathcal{Z}$ *sole sets*.

**Lemma 3.16** *$\mathcal{Z}$ is the unique minimal $k$-set cover of G.*

**Proof:** Let $\mathcal{H}$ be a minimal cover of G. Then by Lemma 3.15, every $I \in \mathcal{H}$ contains a vertex of $\mathcal{S}$. Thus, $\mathcal{S}$ induces a minimal cover of G, namely $\mathcal{Z}$, which is unique since each $v \in \mathcal{S}$ occurs in exactly one $I$. □

The above lemmas are used to define the graph property $\nabla(k)$ in Section 4.1.

We now note that $q > 1$, otherwise since $\mathcal{Z}$ is a cover we only have one set and so $\mathrm{G} \notin \mathrm{MAX}(\mathcal{U}(k\text{-SET}))$. The following takes care of another special case, when the minimal cover $\mathcal{Z}$ has $q = 2$ $k$-sets.

**Lemma 3.17** *If $q = 2$ then G is either the bipartite complete graph $K_{k,k}$ or the graph consisting of an isolated vertex and the subgraph $K_{k-1,k-1}$.*

**Proof:** If there is no isolated vertex, then every vertex is in either $I_1$ or $I_2$ by Lemma 3.16, and not in both since there is an edge on each vertex. Thus G is bipartite on $2k$ vertices. Since a complete bipartite graph is in $\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$, it follows that G must be complete. A similar argument follows for the case where one vertex is isolated. □

**Assumption:** For the remainder of this section we assume that $q \geq 3$. By Lemma 3.17 this is equivalent to Assumption 3.9, given our other assumptions on G.

Recall that $\mathrm{X} \subseteq \mathrm{V}$ is the set of isolated vertices and either X is empty or consists of a single vertex, which we label $x$.

We define the *core associated with* $S_i$ by $C_i = (\cap_{1 \leq j \leq q, j \neq i} I_j) \setminus \mathrm{X}, 1 \leq i \leq q$. We call $\mathcal{C} = (\cup_{1 \leq i \leq q} C_i) \cup \mathrm{X}$ the set of *core vertices*. We define the other vertices by $\mathcal{O} = \mathrm{V} \setminus (\mathcal{S} \cup \mathcal{C})$. We define $\mathrm{O}_i = I_i \cap \mathcal{O}, 1 \leq i \leq q$.

We have two key results on the core vertices. First, the cores are completely joined to the corresponding soles.

**Lemma 3.18** *For all* $1 \leq i \leq q$*, for all* $u \in S_i$ *and all* $v \in C_i$*,* $\{u, v\} \in \mathrm{E}$*.*

**Proof:** Since $\mathcal{Z}$ is a cover of V by Lemma 3.16, the only vertices that can occur in every set $I_j \in \mathcal{Z}$ are those in X by Lemma 3.13. So by definition $C_i \cap I_i = \emptyset$ because it excludes X. Since every vertex in $S_i$ is of degree $n - k$ they must all be adjacent to every vertex not in $I_i$ and therefore to every vertex in $C_i$. □

Second, there are no other edges incident to the core vertices.

**Lemma 3.19** *For all* $1 \leq i \leq q$*, for all* $u \in C_i$ *and* $v \in \mathrm{V} \setminus S_i$*,* $\{u, v\} \notin \mathrm{E}$*.*

**Proof:** First we note that the lemma is trivially true for $v \in \mathrm{X}$. Since we assume $q \geq 3$ then for every pair $1 \leq i, j \leq q$ there exists $h, 1 \leq h \leq q, i, j \neq h$ such that $C_i \cup C_j \subseteq I_h$, and thus there can be no edge in $\mathcal{C}$. Also, since by definition $C_i \subset I_j$ for all $j \neq i$, there can be no edges to $S_j$ or $\mathrm{O}_j$.

Since $\mathcal{Z}$ is a cover, it follows that $\cup_{1 \leq i \leq q} \mathrm{O}_i = \mathcal{O}$. This means we have only to consider pairs in $\mathrm{O}_i \times C_i$. Suppose there is an edge $\{w, t\}, w \in \mathrm{O}_i, t \in C_i$. Since for all $j \neq i$, $C_i \subseteq I_j$, we see that $w \notin I_j$; that is, $w$ occurs only in $I_i$. Then by Lemma 3.14, for all $y \notin I_i$, $\{w, y\} \in \mathrm{E}$ which implies that $w \in S_i$, a contradiction to $w \in \mathrm{O}_i$. □

Under the assumption $q \geq 3$, these lemmas lead us immediately to

**Lemma 3.20**

1. *$\mathcal{C}$ is an independent set,*
2. *for all $I \in \mathcal{I}_k \setminus \mathcal{Z}, \mathcal{C} \subseteq I$,*
3. *$\mathcal{C} \cap \mathcal{S} = \emptyset$,*
4. *for all $u \in S_i$, for all $v \in \mathcal{O} \setminus \mathrm{O}_i$, $\{u, v\} \in \mathrm{E}$.*

**Lemma 3.21** $|C_i \cup \mathrm{X}| \geq 2$*,* $|S_i| \geq |C_i|, 1 \leq i \leq q$ *and* $|\mathcal{C}| \leq k$*.*

**Proof:** Suppose $e = \{u, v\}$ where $u \in S_i$ and $v \in I_i$. Then $I_i$ is the only set in $\mathcal{I}_k$ containing $e$. Writing

$$\bigcap_{I \in \mathcal{I}_k(\mathrm{G}+e)} I = \left( \bigcap_{I \in \mathcal{Z} \setminus I_i} I \right) \bigcap \left( \bigcap_{I \in \mathcal{I}_k \setminus \mathcal{Z}} I \right)$$

we see that the term on the right contains $\mathcal{C}$ by Lemma 3.20 and thus by definition this

$$= \quad C_i \cup \mathrm{X}$$

Since there must be a destroying edge for $e$, this implies $|C_i \cup \mathrm{X}| \geq 2$.

For the second part, if $|S_i| < |C_i|$ then $(I_i \cup C_i) \setminus S_i$ will be an independent set of size greater than $k$, violating Lemma 3.10. The bound on $\mathcal{C}$ also follows from Lemma 3.10. □

Note however that this analysis leaves open which pairs in $\mathcal{O}$ are edges. In fact, it turns out that an arbitrary graph can be embedded in $\mathcal{O}$ and this is key to showing the NP-complete results in the next section.

# 4 NOT-$\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$ is NP-complete

In this section we consider the complexity of NOT-$\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$: given a graph G and integer $k$, is $\mathrm{G} \notin \mathrm{MAX}(\mathcal{U}(k\text{-SET}))$?

**Theorem 4.1** NOT-$\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$ *is* NP-complete.

The proof of this theorem is in two parts, presented in Sections 4.1 and 4.2.

## *4.1* NOT-$\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$ *is in* NP

Analogous to the definition in Section 3, we define $\mathcal{S} = \{v \mid \deg(v) = n - k\}$. We define X to be the set of isolated vertices as in Section 3.

We define the graph property $\nabla(k)$ by $\mathrm{G} \in \nabla(k)$ if and only if

1. $n > k + 1, k > 2$.
2. $\Delta = n - k$.
3. $|\mathrm{X}| \leq 1$.
4. for all $v \in \mathcal{S}$, there exists a $k$-set $I_v \subset \mathrm{V}$ with $v \in I_v$.
5. $\cup_{v \in \mathcal{S}} I_v = \mathrm{V}$.
6. for all $e \in \mathrm{E}^c$, there exists $v \in \mathcal{S}$ such that $e \subseteq I_v$.
7. $\cap_{v \in \mathcal{S}} I_v = \mathrm{X}$.

By this definition, for $\mathrm{G} \in \nabla(k)$ we have $\{I_v \mid v \in \mathcal{S}\}$ covers V, each $v \in \mathcal{S}$ occurs only in $I_v$, and we may define $S_i, 1 \leq i \leq q$ as the partition of the elements of $\mathcal{S}$ as in Section 3. For $\mathrm{G} \in \nabla(k)$, we may then define $C_i = (\cap_{1 \leq j \leq q, j \neq i} I_j) \setminus \mathrm{X}, 1 \leq i \leq q$ and $\mathcal{C} = (\cup_{1 \leq i \leq q} C_i) \cup \mathrm{X}$. Finally as in Section 3 we define $\mathcal{O} = \mathrm{V} \setminus (\mathcal{S} \cup \mathcal{C})$.

The following graph property refines the definition $\nabla(k)$ to only include graphs also satisfying the conditions of Lemmas 3.17 through 3.19 and 3.21. We define the graph property $\nabla^c(k)$ by $\mathrm{G} \in \nabla^c(k)$ if and only if

1. $\mathrm{G} \in \nabla(k)$.

2. G is not one of the graphs in Observation 3.5.
3. for every $u \in S_i$ and every $v \in C_i$, $\{u, v\} \in \mathrm{E}$.
4. for every $u \notin S_i$ and every $v \in C_i$, $\{u, v\} \notin \mathrm{E}$.
5. for all $1 \leq i \leq q$, $|C_i \cup \mathrm{X}| \geq 2$ and $|S_i| \geq |C_i|$, and $|\mathcal{C}| \leq k$.

**Lemma 4.2** *Under the assumptions that $n > k+1$, $k > 2$, $\Delta < n-1$ and* G *is not one of the graphs in Observation 3.5,* $\mathrm{G} \notin \nabla^c(k)$ *implies* $\mathrm{G} \notin \mathrm{MAX}(\mathcal{U}(k\text{-SET}))$.

**Proof:** The definition of $\nabla^c(k)$ (including $\nabla(k)$) is just a list of properties proven to hold for any $\mathrm{G} \in \mathrm{MAX}(\mathcal{U}(k\text{-SET}))$ meeting the same assumptions in Section 3. □

**Lemma 4.3** $\mathrm{G} \in \nabla^c(k)$ *implies* $\mathrm{G} \in \mathcal{U}(k\text{-SET})$.

**Proof:** By the definition of $\nabla(k)$ items 3 and 7, for every edge $\{u, v\} \in \mathrm{E}^c$ for at least one $w \in \{u, v\}$ there exists $I_y, y \in \mathcal{S}$ such that $w \notin I_y$. Thus, $I_y$ is a $k$-set in $\mathrm{G} + \{u, v\}$. □

**Lemma 4.4** *If* $\mathrm{G} \in \nabla^c(k)$ *then for all* $e \in \mathrm{E}^c$*, there exists* $e' \in \mathrm{E}^c[\mathcal{C}]$ *such that* $\mathrm{G} + e + e'$ *contains no* $k$-set $I$ *with* $I \cap \mathcal{S} \neq \emptyset$.

**Proof:** By the definition of $\mathcal{S}$, each vertex in $\mathcal{S}$ is in exactly one $k$-set since the degree is $n-k$. By the definition of $\nabla^c(k)$, for every $e \in \mathrm{E}^c$ there is some $I_i, 1 \leq i \leq q$ such that $e \subset I_i$. Thus, $I_i$ is not a $k$-set in $\mathrm{G} + e$. Also by definition, for all $j \neq i, 1 \leq j \leq q$, $C_i \cup \mathrm{X} \subseteq I_j$ and contains at least two vertices. Thus, since these are the only sets intersecting $\mathcal{S}$, choosing any edge $e'$ from $C_i \cup \mathrm{X}$ completes the proof. □

**Lemma 4.5** *If* $\mathrm{G} \in \nabla^c(k)$ *then* $\mathrm{G} \in$ NOT-$\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$ *if and only if there exists a* $k+1$-set, $I \subseteq \mathcal{C} \cup \mathcal{O}$.

**Proof:** If there exists such a set in $\mathcal{C} \cup \mathcal{O}$, then for any edge $e$ such that at least one end is in $\mathcal{S}$, the $k+1$-set will remain independent in $\mathrm{G} + e$, and so no other edge can destroy all the subsets of it.

Otherwise we first note that by the definition of $\nabla^c(k)$, for every maximum independent set $I \subseteq \mathcal{C} \cup \mathcal{O}$, $\mathcal{C} \subseteq I$. Thus, since $|I| \leq k$ then for every $e' \in \mathrm{E}^c[\mathcal{C}]$, $\mathrm{G}[\mathcal{C} \cup \mathcal{O}] + e'$ contains no $k$-set. It then follows by Lemmas 4.3 and 4.4 that $\mathrm{G} \in \mathrm{MAX}(\mathcal{U}(k\text{-SET}))$. □

**Lemma 4.6** $\mathrm{G} \in \nabla^c(k)$ *can be determined in polynomial time.*

**Proof:** (Outline) It is easy to identify the vertices of $\mathcal{S}$ if they exist by degree criteria, and then the $I_v$ are uniquely forced. Once these are identified, the partitioning into sole sets $S_i$ and identification of cores is a matter of computing a polynomial number of set intersections and unions over subsets of V. Verifying the remaining conditions is merely a matter of testing for appropriate edges and requires no search. □

**Theorem 4.7** NOT-$\mathrm{MAX}(\mathcal{U}(k\text{-SET})) \in$ NP.

**Proof:** The proof when $k = 1$ is trivial. For $k > 1$ recursively delete all vertices of degree $n - 1$. Check the special cases eliminated by the assumptions, and if none apply determine whether $\mathrm{G} \in \nabla^c(k)$. This is all polynomial, the last step by Lemma 4.6. If $\mathrm{G} \notin \nabla^c(k)$ we are done by Lemma 4.2. Otherwise by Lemma 4.5 if $\mathrm{G} \in$ NOT-$\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$ then there is a $k + 1$-set contained in $\mathcal{C} \cup \mathcal{O}$. We non-deterministically choose a set of $k+1$ vertices from $\mathcal{C} \cup \mathcal{O}$ and verify it is an independent set in polynomial time. □

## 4.2 NOT-$\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$ *is* NP*-hard*

We will show that NOT-$\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$ is NP-hard by a reduction from independent set.

Let $\langle \mathrm{G}, k\rangle$ be an instance of the independent set problem. We will construct an instance $\langle \mathrm{G}', h\rangle$ of NOT-$\mathrm{MAX}(\mathcal{U}(h\text{-SET}))$. We assume without loss of generality that $|\mathrm{V_G}| \geq k \geq 4$. The outline of the idea is that we construct a new graph that has all the properties required by the lemmas in Section 3 and embed G in $\mathcal{O}$.

First, define a graph $\mathcal{O} = (\mathrm{V_G} \bigcup \{v_{n+1}, v_{n+2}\}, \mathrm{E_G})$ which will be a subgraph of $\mathrm{G}'$ and consists of G plus two independent vertices. Next for each non-edge $e \in \mathrm{E}^c_{\mathcal{O}}$ we create two independent sets of vertices, $S_e$ and $C_e$. Each $C$ has two vertices. Each $S$ has $k + 1$ vertices.

We create additional edges for $\mathrm{E_{G'}}$ as follows. For each non-edge $e = \{u, v\} \in \mathrm{E}^c_{\mathcal{O}}$, we add the edges in $S_e \times C_e$, $S_e \times (\mathrm{V}_{\mathcal{O}} - \{u, v\})$ and $S_e \times S_{e'}$, for all $e' \neq e, e' \in \mathrm{E}^c_{\mathcal{O}}$.

Let $m = |\mathrm{E}^c_{\mathcal{O}}|$. Then define $h = 2m + k + 1$.

This completes the construction of $\langle \mathrm{G}', h\rangle$. The following lemmas all pertain to this construction and the terms defined above.

**Lemma 4.8** *For all $e \in \mathrm{E}^c_{\mathcal{O}}, S_e$ is contained in a unique maximal $h$-set which is $I_e = S_e \bigcup e \bigcup_{e' \neq e} C_{e'}$.*

**Proof:** We see by the construction that the only vertices not adjacent to any vertex in $S_e$, are those in $S_e$, those in $e$ and those in $C_{e'}$, for each $e' \neq e$. The $C$'s are all independent of each other, and of $\mathrm{V}_{\mathcal{O}}$, so this is an independent set. The size of this set is $k + 1 + 2 + 2(m - 1) = h$. □

Define $I_e$ to be the unique maximal $h$-set containing $S_e$ for each $e \in \mathrm{E}^c_{\mathcal{O}}$ as in Lemma 4.8.

**Lemma 4.9** $\mathrm{G}' \in \mathcal{U}(h\text{-SET})$.

**Proof:** We need to show that, for each $e \in \mathrm{E}^c_{\mathrm{G}'}$, $\mathrm{G}' + e \in h$-SET. If $e \in \mathrm{E}^c_{\mathcal{O}}$ then it follows from Lemma 4.8 that there exists $e' \in \mathrm{E}^c_{\mathcal{O}}$ such that $e \not\subseteq I_{e'}$. If $e = \{u, v\}$ and $u \in S_{e'}$, then by Lemma 4.8 choose $e'' \neq e', e'' \in \mathrm{E}^c_{\mathcal{O}}$ and $I_{e''}$ is a suitable set. Finally, if $u \in C_{e'}$ then $I_{e'}$ is a suitable set, again by Lemma 4.8. □

**Lemma 4.10** *For all $v \in \mathrm{V}_{\mathcal{O}}$, there exist $e, e' \in \mathrm{E}^c_{\mathcal{O}}$ such that $e \neq e'$ and $v \in I_e \bigcap I_{e'}$.*

**Proof:** Since $\mathcal{O}$ contains two independent vertices and $|\mathrm{V_G}| \geq 4$, there are at least two non-edges incident on every vertex in $\mathcal{O}$. □

**Lemma 4.11** *For all $e' \in \mathrm{E}^c_{\mathrm{G}'}$ there exists $e \in \mathrm{E}^c_{\mathcal{O}}$ such that $e' \subset I_e$.*

**Proof:** We provide a case analysis on the possible $e'$, giving an $e$ for each case.

**case** $e' \in \mathrm{E}^c_{\mathcal{O}}$**:** Let $e = e'$.

**case** $e' \subset S_{e''}$**:** Let $e = e''$.

**case** $e' = C_{e''}$**:** Any $e \neq e'', e \in \mathrm{E}^c_{\mathcal{O}}$. There is always such a non-edge in $\mathcal{O}$ by Lemma 4.10.

**case** $e' = \{u, v\}, u \in S_{e''}, v \in \mathrm{V}_{\mathcal{O}}$**:** Let $e = e''$.

**case** $e' = \{u, v\}, u \in S_{e''}, v \in C_{e'''}$**:** Necessarily by construction $e''' \neq e''$. Then by Lemma 4.8 we can let $e = e''$.

**case** $e' = \{u, v\}, u \in C_{e''}, v \in C_{e'''}$**:** By Lemma 4.10 and our construction, there are more than 3 non-edges in $\mathrm{E}^c_{\mathcal{O}}$. By Lemma 4.8 we can choose any $e$ such that $e \neq e''$ and $e \neq e'''$.

**case** $e' = \{u, v\}, u \in C_{e''}, v \in \mathrm{V}_{\mathcal{O}}$**:** By Lemma 4.8 choose any $e = \{v, w\}, e \neq e''$.

By construction there are no other non-edges in $\mathrm{G}'$. □

**Lemma 4.12** $\mathrm{G}' \notin \mathrm{MAX}(\mathcal{U}(h\text{-SET}))$ *if and only if* $\mathrm{G} \in k$-SET.

**Proof:** By Lemma 4.9 we already know that $\mathrm{G}' \in \mathcal{U}(h\text{-SET})$. Thus, we only need to determine whether or not for every $e \in \mathrm{E}^c_{\mathrm{G}'}$ there exists $e' \in \mathrm{E}^c_{\mathrm{G}'}$ such that $\mathrm{G}' + e + e' \notin h$-SET.

Let $j$ be the size of a maximum independent set in G. We consider two cases:

**Claim 1:** $j < k$ *implies* $\mathrm{G}' \in \mathrm{MAX}(\mathcal{U}(h\text{-SET}))$.

Let $C^* = \bigcup_{e \in \mathrm{E}^c_{\mathcal{O}}} C_e$ and let $A \subseteq \mathrm{V}_{\mathcal{O}}$ be any maximal independent set in $\mathcal{O}$. Note that since $A$ is maximal, it corresponds to a maximal set in G, plus the two independent vertices added to $\mathcal{O}$. Thus, if $j < k - 1$ then $|C^* \bigcup A| < h$, and it follows that the only $h$-sets in $\mathrm{G}'$ are the $I_e$. If $j = k - 1$ then when $A$ is a maximum set, $C^* \bigcup A$ is an $h$-set. Thus, we have such a set for every maximum set of size $k - 1$ in G. However, all of these non-$I_e$ $h$-sets contain all vertices in $C^*$, and thus adding any edge to $C^*$ destroys them all.

Let $e \in \mathrm{E}^c_{\mathrm{G}'}$ be any non-edge. Then by Lemma 4.11 there is an $e'$ such that $e \subset I_{e'}$. Then choose a second edge $e'' \in C_{e'}$. By Lemma 4.8 this edge is in all remaining $I$ and together with the above observation this implies $\mathrm{G}' + e + e''$ has no $h$-set.

**Claim 2:** $j \geq k$ *implies* $\mathrm{G}' \notin \mathrm{MAX}(\mathcal{U}(h\text{-SET}))$.

Defining $C^*$ and $A$ as in the previous claim, we see that $|C^* \bigcup A| > h$ when $A$ is maximum, since $|A| \geq k + 2$. If we let $e \in S_{e'}$ for some $e'$, then $|C^* \bigcup A|$ is still an independent set of size greater than $h$, and so no other edge can destroy all $h$-sets in it.

This completes the proof of the lemma. □

**Theorem 4.13** NOT-$\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$ *is* NP-*hard.*

**Proof:** The reduction is correct by Lemma 4.12. We see that $h$ is at most quadratic in $|\mathrm{V}_{\mathrm{G}}|$ by our assumptions on G and $k$. The number of vertices in $\mathrm{G}'$ is at most cubic in $|\mathrm{V}_{\mathrm{G}}|$ and the construction is a straight-forward plug in of components plus additional edges. Thus, we have a polynomial reduction from an NP-complete problem. □

## 5 Conclusions and Open Problems

We have shown that NOT-$\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$ is NP-complete, or equivalently that $\mathrm{MAX}(\mathcal{U}(k\text{-SET}))$ is CO-NP-complete. It seems a slightly curious twist that we start with an NP-complete property and on considering the maximal version we obtain a problem in CO-NP-complete. We ask, is there an NP-complete property $\mathcal{P}$ such that $\mathrm{MAX}(\mathcal{P})$ is NP-complete?

From [4] the reason that most maximal properties are in P seems to relate to the idea that features of the graph that prevent a graph being in the class are in some sense complete, for example in $\mathrm{MAX}(k\text{-SET})$ all pairs not in the independent set are edges, and in the maximal version of $k$-coloring the graph is $k$-partite complete. This completeness is usually polynomial time checkable. On the other hand the problem $\mathrm{MAX}(3\text{-coloring and maximum degree} = 4)$ is seen to be NP-hard, and the reduction in [4] indicates this is because the maximum degree restriction limits the completeness structure. Note that the degree restriction is an easy condition to check in polynomial time.

In the current result we again get a restriction on the completeness in that there is a polynomial time checkable super structure, but then the remainder of the graph, $\mathcal{O}$ in the constructions, does not need to be complete. In this case the polynomial structure is induced by the unfrozen condition on the property, rather than being an explicit condition of the property. Note that unfrozen can be seen as a polynomial composition of the property; that is, it means there may be $O(n^2)$ possible independent sets, one for each possible addition of an edge. A careful examination of the isomorphism complete result in [4] also seems to exhibit a type of limit on completeness. On the other hand, combining two NP-complete properties did not sufficiently limit completeness to move the resulting maximal properties out of P.

So, the question is can we somehow generalize and make precise these observations and thus predict into which complexity classes different modified properties will fall?

## Acknowledgements

The authors thank the Natural Sciences and Engineering Research Council of Canada for financial support.

# References


[1] A.V. Aho, J.E. Hopcroft, and J.D. Ullman. *The Design and analysis of Computer Algorithms*. Addison-Wesley, 1974.

[2] D. Bauer, F. Harary, J. Nieminen, and C. Suffel. Domination alteration sets in graphs. *Discrete Mathematics*, 47:153–161, 1983.

[3] A. Beacham. The complexity of problems without backbones. Master's thesis, The University of Alberta, 2000.

[4] A. Beacham and J. Culberson. On the complexity of unfrozen problems. *Discrete Applied Mathematics*, 2004. To appear.

[5] C. Berge. *Graphs and Hypergraphs*. North-Holland-American Elsevier, 1973.

[6] B. Bollobás. *Graph Theory*. Springer-Verlag New York Inc., 1979.

[7] B. Bollobás, C. Borgs, J.T. Chayes, J.H. Kim, and D.B. Wilson. The scaling window of the 2-SAT transition. Technical Report MSR-TR-99-41, Microsoft Research MSR-TR-99-41, 1999.

[8] J.A. Bondy and U.S.R. Murty. *Graph Theory with Applications*. McMillan Press, New York, 1976.

[9] E. Boros, M.C. Golumbic, and V.E. Levit. On the number of vertices belonging to all maximum stable sets of a graph. *Discrete Applied Mathematics*, 124:17–25, 2002.

[10] P. Cheeseman, B. Kanefsky, and W. M. Taylor. Where the *really* hard problems are. In *International Joint Conference on Artificial Intelligence*, pages 331–337, 1991.

[11] J. Culberson and I. Gent. Frozen development in graph coloring. *Theoretical Computer Science*, 265(1):227–264, 2001.

[12] J. Culberson and B. Vandergriend. The $G_{n,m}$ phase transition is not hard for the hamiltonian cycle problem. *Journal of Artificial Intelligence Research*, 9:219–245, 1998.

[13] R.D. Dutton and R.C. Brigham. An extremal problem for edge domination insensitive graphs. *Discrete Applied Mathematics*, 20:113–125, 1988.

[14] E. Friedgut. Sharp thresholds of graph properties, and the $k$-sat problem. *Journal of the American Mathematical Society*, 12:1017–1054, 1999.

[15] M.R. Garey and D.S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-completeness*. W.H. Freeman, San Francisco, 1979.

[16] I. Gent, E. MacIntyre, P. Prosser, and T. Walsh. The constrainedness of search. In *Proceedings of AAAI-96*, pages 246–252, 1996.

[17] G. Gunther, B. Hartnell, and D.F. Rall. Graphs whose vertex independence number is unaffected by single edge addition or deletion. *Discrete Applied Mathematics*, 46:167–172, 1993.

[18] P.L. Hammer, P. Hansen, and B. Simeone. Vertices belonging to all or to no maximum stable sets of a graph. *SIAM J. Algebraic Discrete Methods*, 3:511–522, 1982.

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

[20] F. Harary. Changing and unchanging invariants for graphs. *Bull. Malaysian Math. Soc.*, 5:73–78, 1982.

[21] T.W. Haynes, L.M. Lawson, R.C. Brigham, and R.D. Dutton. Changing and unchanging of the graphical invariants: minimum and maximum degree, maximum clique size, node independence number and edge independence number. *Congressus Numerantium*, 72:239–252, 1990.

[22] G. Hopkins and W. Staton. Graphs with unique maximum independent sets. *Discrete Mathematics*, 57:245–251, 1985.

[23] V.E. Levit and E. Mandrescu. Combinatorial properties of the family of maximum stable sets of a graph. *Discrete Applied Mathematics*, 117:149–161, 2002.

[24] D. L. Mammen and T. Hogg. A new look at the easy-hard-easy pattern of combinatorial search difficulty. *Journal of Artificial Intelligence Research*, 7:47–66, 1997.

[25] C. H. Papadimitriou. *Computational Complexity*. Addison-Wesley, 1994.

[26] A. Parkes. Clustering at the phase transition. In *Proceedings of AAAI-97*, pages 340–345, 1997.

[27] B. Selman, H. Levesque, and D. Mitchell. A new method for solving hard satisfiability problems. In *Proceedings of the Tenth National Conference on Artificial Intelligence (AAAI-92), San Jose, CA*, pages 440–446, 1992.

[28] J. Singer, I.P. Gent, and A. Smaill. Backbone fragility and the local search cost peak. *Journal of Artificial Intelligence*, 12:235–270, 2000.

[29] D.P. Sumner and P. Blitch. Domination critical graphs. *Journal of Combinatorial Theory, Series B*, 34:65–76, 1983.

[30] J. Zito. The structure and maximum number of maximum independent sets in trees. *J. Graph Theory*, 15:207–221, 1991.