Electronic Colloquium on Computational Complexity, Report No. 10 (2000)

# Supermodels and Closed Sets

Amitabha Roy *
Dept. of Computer Science
University of Oregon
Eugene, OR 97403
aroy@cs.uoregon.edu

Christopher Wilson
Dept. of Computer Science
University of Oregon
Eugene, OR 97403
cwilson@cs.uoregon.edu

**Abstract**

A *supermodel* is a satisfying assignment of a boolean formula for which any small alteration, such as a single bit flip, can be repaired by another small alteration, yielding a nearby satisfying assignment. We study computational problems associated with super models and some generalizations thereof. For general formulas, it is NP-complete to determine if it has a supermodel. We examine 2-SAT and HORNSAT, both of which have polynomial time satisfiability tests. We see that while 2-SAT has a polynomial time test for a supermodel, testing whether a HORNSAT formula has one is NP-complete. We then look at sets of supermodels called *closed sets* - these are sets of supermodels which retain the supermodel property even after being broken and repaired. Using combinatorial methods, we examine extremal properties of closed sets. We find that they are at least linear in size. For large ones, an upper bound is trivial, but we see that the largest *minimal* closed set has size between $2^{2n/3}$ and $(4/5)2^{n-1}$. A *sparse* closed set is one in which each break of a supermodel has only a single repair. We obtain analogous, and slightly tighter, bounds on sparse closed sets, whose sizes essentially must lie between $(2e)^{n/8}$ and $2^n/n$.

## 1 Introduction

The concept of supermodels, introduced in [GPR98], formalizes a notion of fault tolerant satisfying assignments to boolean formulas. In this paper, we study the problem of identifying these supermodels and generalize this notion of fault tolerance.

The motivation for studying supermodels in the artificial intelligence/planning community was to build search algorithms finding robust solutions to problems (typically in scheduling or planning domains). These solutions have the property that if an expected resource is suddenly unavailable, then a minimal modification to the solution produces an equally acceptable alternative. Recently, this idea has been used by [BP99].

Our goal is to examine the computational and combinatorial complexity of supermodels. In the first few sections, we are concerned with the computational aspects of supermodels of boolean formulas. In the last two sections, we take a combinatorial approach to identify the structure of sets of certain kinds of supermodels.

Essentially, a *supermodel* of a boolean formula $F$ is a satisfying assignment $\alpha$ of $F$, $F(\alpha) = 1$, such that for every $i$, if we negate the $i$th bit of $\alpha$, there is another bit $j \neq i$ of $\alpha$ which we can negate to get another satisfying assignment. That is, if $\delta_i(\alpha)$ is the function which negates the $i$th bit of $\alpha$, then $(\forall i)(\exists j \neq i)F(\delta_j(\delta_i(\alpha))) = 1$. In sections 3 and 4, we study the complexity of finding supermodels for arbitrary and restricted classes of formulas. While we prove that finding supermodels in general is NP-complete, we also exhibit polynomial time algorithms for finding them in two specific types of boolean formulas: 2-SAT and Affine SAT, two classes of satisfiability where finding satisfying assignments has efficient algorithms. Intriguingly, while this (efficient satisfiability checking) is also true for HORNSAT, we show that finding supermodels for HORNSAT formulas is NP-complete.

In sections 5 and 6, we look at the concept of closed sets of supermodels. This is a extension of the notion of supermodels under which a model remains a model after an arbitrary sequence of breaks and repairs. These appear to have a rich combinatorial structure. Closed sets can also be characterized in terms of what's called $(\sigma, \rho)$-domination in [Te94, TP97].

*partially supported by NSF grant CCR9820945



ISSN 1433-8092

Due to the complex structure of closed sets, we have made certain simplifying assumptions about their behavior. Section 5 covers sparse closed sets: these are sets of supermodels such that each supermodel has only one repair for every possible break that gives another supermodel in the set. We prove that *every* sparse closed set has size at least $(2e)^{n/8}$ and at most $2^n/n$. We also give an explicit construction of sparse closed sets of nearly maximum size.

Section 6 deals with general closed sets: we see here that the smallest closed sets must have linear size (there is a simple example that achieves this bound). On the other hand, while a closed set can be as large as $2^{n-1}$, we apply a generalized "triangle-elimination" technique to bound the size of *minimal* closed sets to at most $(4/5)2^{n-1}$.

## 2 Definitions and Notations

The intended objects of study are strings over $\{0,1\}$ of some specified length $n$. The basic operations on these strings are bit flips (negations): changing a specified bit from a one to a zero or the reverse. However, for the purposes of this paper we have found it convenient to denote these objects as sets. A string $\alpha$ of length $n$ represents a subset $X$ of $\{1,2,\ldots,n\}$ as its *incidence vector* (or characteristic sequence): the $i$th bit of $\alpha$ is 1 iff $i \in X$. In this context, instead of flipping a bit of $\alpha$, we take the symmetric difference of $X$ with a singleton set, $X\triangle\{i\}$. In this manner we are able to describe a series of bit flips themselves as a set, simplifying the descriptions of our proofs.

Let $[n]$ refer to a set of $n$ elements $\{1,2\ldots,n\}$ and let $2^{[n]}$ refer to its power set. For $i \in [n]$, the operator $\delta_i : 2^{[n]} \to 2^{[n]}$ is defined as $\delta_i(X) := X\triangle\{i\}$.

We shall write $\delta_{ij}(X)$ to mean $\delta_j(\delta_i(X))(= \delta_i(\delta_j(X)))$ as shorthand (in these situations, we shall always assume that $i \neq j$ unless otherwise mentioned). Note that if $\delta_{ij}(X) = Y$ then $\delta_{ij}(Y) = X$.

The operator $\Delta_i(X)$ is defined as follows: $\Delta_i(X) = \bigcup_{j\neq i} \delta_{ij}(X)$.

For $S \subseteq [n]$, define $\delta_S(X)$ inductively: $\delta_\emptyset(X) = X$ and $\delta_S(X) = \delta_{S\setminus\{i\}}(\delta_i(X))$ where $i \in S$. Given a family of subsets $\mathcal{F}$ of $[n]$ let $\mathcal{F}_k$ ($\mathcal{F}_{\leq k}$) denote the number of sets in $\mathcal{F}$ with $k$ elements ($\leq k$ elements). Let $\binom{[n]}{k}$ ($\binom{[n]}{\leq k}$) denote the family of $k$-element ($\leq k$ element) subsets of $[n]$.

Let $F$ be a boolean formula of $n$ variables $[n]$. An assignment $X : [n] \to \{0,1\}$ is called a model if $X$ makes $F$ true. We shall also interpret $X$ as an incidence vector of a subset in $2^{[n]}$. In all our discussions on boolean formulas, we make the reasonable assumption that every variable appears in both positive and negative literals.

**Definition 2.1** *A model $X$ of $F$ is called a $(r,s)$ supermodel if $\forall R \in \binom{[n]}{\leq r}$, there exists $S \in \binom{[n]}{\leq s}$, such that $R \cap S = \emptyset$, and $\delta_{R\cup S}(X)$ is a model of $F$. We view $r$ and $s$ as fixed constants unless otherwise mentioned.*

In other words, $X$ is a $(r,s)$ supermodel iff for every bit flip (called a "break") of up to $r$ coordinates in the incidence vector of $X$, there is a disjoint set of up to $s$ bits that could be flipped to get a model of $F$. In this paper we shall be primarily concerned with $(1,1)$ supermodels, which we shall call *supermodels.* Define $\mathrm{SUP}(r,s)$ to be the set of boolean formulas which have $(r,s)$ supermodels.

A family $\mathcal{C} \subseteq 2^{[n]}$ is said to be a *closed set* if $\forall X \in \mathcal{C}\ \forall i \in [n],\ \ \Delta_i(X) \cap \mathcal{C} \neq \emptyset$.

We define $\Delta_i(X,\mathcal{C}) = \Delta_i(X) \cap \mathcal{C}$. Observe that each $Y \in \Delta_i(X,\mathcal{C})$ is of the form $Y = \delta_{ij}(X)$, for some $j \in [n]$. Define the $i$-th repair set of $X$ in $\mathcal{C}$ to be $R_i(X,\mathcal{C}) = \{j|\delta_{ij}(u) \in \mathcal{C}, j \neq i\}$ : these are the coordinates that repair the $i$-th break to $X$ in $\mathcal{C}$.

A family $\mathcal{C} \subseteq 2^{[n]}$ is said to be a *sparse closed set* if $\forall X \in \mathcal{C}\ \forall i \in [n],\ \ |\Delta_i(X,\mathcal{C})| = 1$.
Sparseness embodies the notion of needing exactly one repair. Thus, it is the number of repairs which are few (hence "sparse"), rather than the size of the sets, which can be rather large.

By *weak* we refer to a set in which breaks do not necessarily need to be repaired. Define $\Delta_i^*(X) = \Delta_i(X) \cup \delta_i(X)$. where $\mathcal{C}$ is a family of subsets of $[n]$ and $X \subseteq [n]$. Then $\mathcal{C}$ is a *weak closed set* if $\forall X \in \mathcal{C}\ \forall i \in [n],\ \ \Delta_i^*(X) \cap \mathcal{C} \neq \emptyset$. Thus weak closed sets allow for the possibility that there are supermodels at distance 1 from each other.

Given a closed set $\mathcal{C}$, an element $X \in \mathcal{C}$ is said to be redundant if $\forall i \in [n]\ |\Delta_i(X,\mathcal{C})| > 1$.

$X$ is irredundant otherwise. A family $\mathcal{C}$ is a minimal closed set if for all $\mathcal{C}' \subseteq \mathcal{C}$ and $\mathcal{C} \neq \mathcal{C}'$, $\mathcal{C}'$ is not closed. We can similarly define minimal sparse closed sets in the poset of sparse closed sets.

We note the following trivial combinatorial lemma which we record below because we use it so frequently in our proofs.

**Lemma 2.1** *Let $\mathcal{F} \subseteq 2^{[n]}$ be a collection of pairwise disjoint 2-element subsets of $[n]$. Then $|\mathcal{F}| \leq n/2$.*

We shall call such families 2-partitions of $[n]$. $\mathcal{F}$ is a maximal 2-partition if $\bigcup_{X \in \mathcal{F}} X = [n]$. Clearly a maximal 2-partition has size $n/2$.

# 3 Supermodels of Boolean Formulas

Consider the following decision question:
**Problem:** SUP$(r, s)$

INSTANCE: Boolean formula $F$.
QUESTION: Does $F$ have a $(r, s)$ supermodel ?

Supermodels were defined in [GPR98], where SUP$(r, s)$ was also proved NP-complete.

**Theorem 3.1** *[GPR98]* SUP$(r, s)$ *is NP-complete*

*Proof:* Reduction from SAT. Let $F$ be an instance of SAT, where $F$ is a boolean formula over $n$ variables $\{x_1, x_2 \ldots, x_n\}$. Construct the formula $F' = F \vee x_{n+1}$ where $x_{n+1}$ is a new variable. We claim that $F$ is satisfiable iff $G$ has a $(r, s)$ supermodel. Suppose $F$ is satisfiable: let $X$ be a satisfying assignment. Extend $X$ to a satisfying assignment $X'$ of $F'$ by setting $x_{n+1}$ to 0. We claim $X'$ is a $(r, s)$ supermodel. Let us break any set of up to $r$ bits in $X$. If that break set includes the bit corresponding to $x_{n+1}$, we do not need any repairs. If it doesn't, we can repair by flipping the bit corresponding to $x_{n+1}$. Now suppose that $F'$ has a $(r, s)$ supermodel $X'$. Then $F'$ must have a model with $x_{n+1}$ set to 0: if $X'$ has $x_{n+1} = 0$ we are done, otherwise flip $x_{n+1}$, we are guaranteed a repair. The restriction of that model to $\{x_1, x, \ldots, x_n\}$ gives us a model for $F$. Hence SUP$(r, s)$ is NP-hard.

Since SUP$(r, s)$ is in NP (an NDTM needs to guess a table indexed by all possible $O(n^r)$ break sets and a repair set corresponding each break set), SUP$(r, s)$ is NP complete. **QED**

We now see interpret supermodels as generalized models of boolean formulas. Let $F$ be a boolean formula. Then $\sup^0(r, s)$ supermodels are just models of $F$:

$$\sup^0(r, s) = \{X |\ X \text{ is a model of } F\}$$

and we define $\sup^k(r, s)$ supermodels inductively:

$$\sup^k(r, s) = \{X \in \sup^{k-1}(r, s) |\ \forall R \in \binom{[n]}{\leq r}, \exists S \in \binom{[n]}{\leq s}, R \cap S = \emptyset \text{ and } \delta_{R \cup S}(X) \in \sup^{k-1}(r, s)\}$$

We define SUP$^k(a, b)$ to be the family of boolean formulas which have a $\sup^k(a, b)$ supermodel. We define SUP$^*(r, s) = \bigcap_{i=0}^{\infty}$ SUP$^i(r, s)$ and, perhaps not surprisingly, we call the corresponding models $\sup^*(r, s) = \bigcap_i \sup^i(r, s)$ , super*-models. Our primary concern will be with the case $r = s = 1$.

We had to define the concept of a weak closed set corresponding to our definition of supermodels (we have to allow the case that there might be no repairs necessary).

**Lemma 3.1** *Let $F$ be a boolean formula. Then $F \in$ SUP$^*(1, 1)$ iff there is a weak closed set consisting of models of $F$.*

In this situation, we say that $F$ has a weak closed set of models.

We shall identify SUP$^k(1, 1)$ with the decision question: given a boolean formula, does it belong to the family SUP$^k(1, 1)$, ( where the identification will be clear from the context) ?

**Lemma 3.2** SUP$^*(1, 1) \in$ *NEXP.*

*Proof:* An NDTM guesses a weak closed set (which could be of exponential size) and checks that it is indeed a closed set and that all elements in the closed set is a model of the input instance. □.

If the weak closed set had a polynomial description, then the NDTM would just use polynomial space. We wonder whether SUP$^*(1, 1)$ is complete for NEXP.

However, we can prove the following (weaker) result:

**Theorem 3.2** SUP$^*(1,1)$ *is NP-hard.*

*Proof:* We use the same reduction as in Theorem 3.1. Given instance of SAT, a boolean formula $F$ over $n$ variables $\{x_1, x_2, \ldots, x_n\}$ we construct $F' = F \vee x_{n+1}$. Suppose $F$ has a model $X$. We construct a weak closed set of models $\mathcal{C}$ of $F'$. If $Y \subseteq \{x_1, x_2, x_{n+1}\}$, let $\pi_n(Y)$ be the projection of the incidence vector of $Y$ into the first $n$ coordinates $\{x_1, x_2, \ldots, x_n\}$.

$$\mathcal{C} = \{Y \subseteq [n+1] |\ d(X, \pi_n(Y)) = 1 \text{ iff } x_{n+1} \in Y\}$$

It is trivial to see that $\mathcal{C}$ is indeed a weakly closed set of models. Also observe that if $F'$ had a super*-model, then it has a model $X$ with $x_{n+1} = 0$. Then $\pi_n(X)$ is a model of $F$. **QED**

**Corollary 3.1** SUP$^k(1,1)$ *is NP-complete.*

*Proof:* Since sup$^*(1,1)$ models are automatically sup$^k(1,1)$ supermodels, for all $k \geq 0$, Theorem 3.2 shows that SUP$^k(1,1)$ is NP hard. SUP$^k(1,1) \in NP$ by induction: SUP$^0(1,1) \in NP$ ([Pa94]) (base case). To see that SUP$^i(1,1) \in NP$, a NDTM guess a table: which contains a repair (if needed) for each break that is supposed to give a sup$^{i-1}(1,1)$ supermodel. With each such repair, it also guess a polynomial length certificate of the sup$^{i-1}(1,1)$ supermodel. It checks the certificate in polynomial time. □

# 4 Restricted Boolean Formulas

We now prove that finding supermodels (i.e. sup(1,1) supermodels) for 2-SAT formulas is in polynomial time. 2-SAT formulas are in conjunctive normal form with 2 literals per clause.

Let $\phi$ be an instance of 2-SAT. We define the graph $G(\phi)$ as follows: the vertices of the graph are the literals of $\phi$ (i.e. variables of $\phi$ along with their negations) and for each clause $\alpha \rightarrow \beta$ (where $\alpha, \beta$ are literals) we add two directed edges $(\alpha, \beta)$ and $(\neg\beta, \neg\alpha)$. Thus the edges of $G(\phi)$ capture the implications of $\phi$. The following theorem is well-known.

**Theorem 4.1** *[Pa94] $\phi$ is unsatisfiable iff there is a variable $x$ such that there is a path from $x$ to $\neg x$ and a path from $\neg x$ to $x$ in $G(\phi)$.*

If $\phi$ has a supermodel, then $G(\phi)$ has a further restriction.

**Lemma 4.1** *If $\phi$ has a supermodel, then there is no path from $u$ to $\neg u$, where $u$ is a literal in $\phi$.*

*Proof:* Suppose there was a path $u \rightarrow v_1 \rightarrow v_2 \cdots v_m \rightarrow \neg u$, where $m \geq 0$, and $v_i \notin \{u, \neg u\}$. We now claim that $u$ has to be set 0 (false) in any model of $\phi$. Suppose not: let $u$ be set true. Thus there must be some $(v_i, v_{i+1})$ such that $v_i$ is true and $v_{i+1}$ is set false (if $m = 0$ then let $v_i = u, v_{i+1} = \neg u$ in the following argument). Then the implication $v_i \Rightarrow v_{i+1}$ is not satisfied. If the value of $u$ is fixed (to either true or false) in $\phi$, then $\phi$ cannot have a supermodel. □

Define a simple path in $G(\phi)$ to be an ordered sequence $\mathcal{P} = (u_1, u_2, \ldots, u_m)$ where $u_1, u_2, \ldots, u_m$ are all distinct vertices. The length of $\mathcal{P}$, denoted by $l(\mathcal{P})$, is $m-1$. By Lemma 4.1, we know that if $\phi$ has a supermodel, then a simple path cannot include a variable and its negation. If $X$ is an 0-1 assignment to the variables of $\phi$ let $X(u)$ denote the value of the literal $u$ under $X$.

**Lemma 4.2** *Let $\phi$ have a supermodel. Let $\mathcal{P}$ be a simple path in $G(\phi)$. Then $l(\mathcal{P}) \leq 3$.*

*Proof:* Suppose $l(\mathcal{P}) \geq 4$. Then there is a simple path $\mathcal{P}' = (u_1, u_2, u_3, u_4, u_5)$ where $l(\mathcal{P}') = 4$ (take the initial 5 vertices of $\mathcal{P}$). Let $X$ be a supermodel of $\phi$. Let $F = \{u_i |\ X(u_i) = 0\}$ and $T = \{u_1, u_2 \ldots, u_5\} \setminus F$. It is easy to see that $F$ has to be the initial segment of $\mathcal{P}'$, and $T$ has to be the remaining segment. We claim that $|F| \leq 2$. Suppose not: then $X(u_1) = X(u_2) = X(u_3) = 0$. If we now break $u_1$, we will need 2 repairs, hence $X$ cannot be a supermodel. Similarly $|T| \leq 2$. However $|F| + |T| = 5$, a contradiction. □

Using arguments similar to Lemma 4.2, one can show that

**Lemma 4.3** *Let $\phi$ have a supermodel $X$.*

1. *Let $\mathcal{P} = (u_1, u_2, u_3, u_4)$ be a path in $G(\phi)$ of length 3. Then $X(u_1) = X(u_2) = 0$ and $X(u_3) = X(u_4) = 1$.*

2. *Let $\mathcal{P} = (u_1, u_2, u_3)$ be a path in $G(\phi)$ of length 2. Then $X(u_1) = 0$ and $X(u_3) = 1$.*

We need to bound the length of cycles in $G(\phi)$ as well. The proof follows the same technique as Lemma 4.2.

**Lemma 4.4** *Let $\phi$ have a supermodel. Let $(u_1, u_2, \ldots, u_{m+1} = u_1)$ be a cycle of length $m$. Then $m \leq 2$.*

Let $X$ be a partial assignment of the variables in $\phi$. We now show an algorithm that takes $X$ and makes forced choices (but only with regard to vertices that take part in cycles) and checks to see whether $X$ can be extended to a supermodel.

**Extend**$(\phi, X)$

1. For each cycle $(u_1, u_2), (u_2, u_1)$ in $G(\phi)$, such that exactly one of $X(u_1)$ and $X(u_2)$ is defined, set $X(u_1) = X(u_2)$. If there is a conflict, because of a vertex taking part in more than 1 cycle, then abort. Let $X'$ be the new (partial) assignment.

2. For each edge $(u_1, u_2)$ in $G(\phi)$ such that both $X(u_1)$ and $X(u_2)$ are defined, check to see whether the implication $u_1 \rightarrow u_2$ is satisfied by $X$. If not, abort.

3. For each triple of assigned vertices $u, v, w$ such that $X(u) = X(v) = X(w) = 0$, check if $(u, v), (u, w)$ are edges in $G(\phi)$. If so, abort.

4. For each triple of assigned vertices $u, v, w$ if $X(u) = X(v) = X(w) = 1$, check if $(v, u), (w, u)$ are edges in $G(\phi)$. If so, abort.

5. Return $X'$.

It is not difficult to see that if $X$ can be extended to a supermodel for $\phi$, then Extend$(X, \phi)$ returns $X'$ which can also be extended to a supermodel. The assignments in step 1 are forced: so if there are conflicts then there cannot be any supermodel. The (forcibly extended) supermodel now has to pass step 2 as well, to make sure it is a model. If there were vertices $u, v, w$ all assigned false by $X'$ with $(u, w), (u, v)$ edges in $G(\phi)$, then $X'$ cannot be extended to a supermodel: a break to $u$ requires 2 repairs whatever the extension. Hence $X'$ has to pass through step 3, and similarly, through step 4.

Now we are ready to describe our algorithm **Supermodel**$(\phi)$ to find supermodels for 2-SAT theories where the input instance is the 2-SAT formula $\phi$.

**Supermodel**$(\phi)$

1. Construct $G(\phi)$. Set initial partial assignment $X = \emptyset$.

2. Check to see whether there is any vertex $u$ such that there is a directed path from $u$ to $\neg u$. If so, abort.

3. Check to see whether there is any path of length 5. If so, abort.

4. For every simple path of length 3 and every simple path of length 2, construct a partial assignment $X$ as prescribed by Lemma 4.3. If there is a conflict in assigning a value to a vertex, abort.

5. Run Extend$(\phi, X)$ which either aborts or returns a (possibly new) partial assignment $X'$.

6. For each isolated cycle $(u, v), (v, u)$ (where both $u, v$ have both in-degree and out-degree 1) such that both $X'(u)$ and $X'(v)$ are undefined, set both $X'(u) = X'(v) = 1$.

7. Let $U$ be the set of literals left unassigned by $X'$. Construct a 2-SAT formula $\beta$ as follows:

   (a) Initially set $\beta$ to the trivial (empty) formula.

   (b) for each pair of unassigned literals $u \in U$ and $v \in U$ such that there is a vertex $w$ in $G(\phi)$ with $X'(w) = 0$, and $(w, u)$ and $(w, v)$ are edges in $G(\phi)$, set $\beta = \beta \wedge (u \vee v)$.

   (c) for each pair of unassigned literals $u \in U$ and $v \in U$ such that there is a vertex $w$ in $G(\phi)$ with $X'(w) = 1$, and $(u, w)$ and $(v, w)$ are edges in $G(\phi)$, set $\beta = \beta \wedge (\neg u \vee \neg v)$.

(d) For each pair of literals $u$ and $v \in U$ such that there is a vertex $w$ in $G(\phi)$ with $X'(w) = X'(u) = 0$, and $X(v)$ unassigned with $(w, u)$ and $(w, v)$ as edges in $G(\phi)$, set $\beta = \beta \wedge (v)$.

(e) For each pair of literals $u$ and $v \in U$ such that there is a vertex $w$ in $G(\phi)$ with $X'(w) = X'(u) = 1$, and $X(v)$ unassigned with $(u, w)$ and $(v, w)$ as edges in $G(\phi)$, set $\beta = \beta \wedge (\neg v)$.

If $\beta$ is unsatisfiable, then abort else find a model for $\beta$ and combine with $X'$ to get an assignment $M$.

If **Supermodel**($\phi$) does not abort, the returned assignment $M$ is a supermodel. We first observe that it is a model since each step we make sure that if $(u, v)$ is an edge in $G(\phi)$ then $X'(u) \leq X'(v)$. Let $X$ be the partial assignment before executing step 7. Let $u$ be a vertex in $G(\phi)$ set false by $X$. Let $N(u) = \{v \in G(\phi) | (u, v)$ is an edge in $G(\phi)\}$. If $X$ was extensible to a supermodel, then we can have at most one vertex in $N(u)$ set to false in that extension and this happens exactly when $\beta$ is true. The argument is symmetric when $v$ is set false by $X$. Also breaks to the variables in $U$ do not need repairs as they are intermediate vertices in a chain of length 2.

Each step in supermodel($\phi$) is easily seen to be in polynomial time. Hence

**Theorem 4.2** *In polynomial time, one can determine if a 2-SAT theory has a* $(1, 1)$ *supermodel and find one if it exists.*

In fact, one can show that each step can be done in NL.

**Theorem 4.3** SUP$(1, 1)$–2-SAT $\in$ *NL.*

Surprisingly, the situation completely alters when we consider sup$(1, b)$ supermodels for $b > 1$.

**Theorem 4.4** SUP$(1, b)$–2-SAT *is NP-complete for* $b > 1$.

*Proof:* Reduction from $(b+1)$-SAT. Let $T = C_1 \wedge C_2 \ldots \wedge C_m$ be an instance of $(b+1)$-SAT where each clause $C_i$ is a disjunction of $b + 1$ literals $l_1^i \cup l_2^i \ldots \cup l_{b+1}^i$. We construct an instance $T'$ of SUP$(1, b)$–2-SAT as follows:

$$T' = \bigwedge_{1 \leq i \leq m} F(i)$$

where $F(i)$ is a 2-SAT theory defined for each clause $C_i$ as follows:

$$\begin{aligned} F(i) \quad = \quad & \bigwedge_{1 \leq j \leq (b+1)} (c_i \Rightarrow l_j^i) \\ & \bigwedge_{1 \leq j \leq (b+1)} (l_j^i \Rightarrow a^i[1, j]) \\ & \bigwedge_{1 \leq j \leq b-1} \bigwedge_{1 \leq k \leq (b+1)} (a^i[j, k] \Rightarrow a^i[j + 1, k]) \end{aligned}$$

where we have introduced $1 + b(b + 1)$ new variables $c_i$ and $a^i[j, k]$ for $1 \leq j \leq b, 1 \leq k \leq (b + 1)$ to define the gadget $F(i)$.

Now suppose $T$ had a model $X$. Extend that to a model of $T'$ by setting $c_i = 0$ for all $1 \leq i \leq m$ and $a^i[j, k] = 1$ for all $1 \leq i \leq m, 1 \leq j \leq b, 1 \leq k \leq (b + 1)$. We claim that this is a sup$(1, b)$ model of $T'$. Flip any variable $v$. Now we do a case analysis of how many repairs are needed:

- $[v = c_i]$ Since $l_1^i \cup l_2^i \ldots \cup l_{b+1}^i = 1$ (since $X$ is a model, there is at least one literal in $\{l_1^i, l_2^i, \ldots, l_{b+1}^i\}$ which is already set to 1: so we need to flip at most $b$ literals in $\{l_1^i, \ldots, l_{b+1}^i\}$.

- $[v = a^i[j, k]]$ Need to flip $a^i[i, k]$ where $1 \leq i < j$ and we might need to flip $l_k^i$ if that was set to true by $X$. Hence we flip at most $j \leq b - 1 + 1 = b$ variables.

- $[v = l_j^i]$ No repairs are necessary.

Now suppose $T'$ has a $\sup(1, b)$ supermodel. Note that in such a model $c_i = 0$ for all $i$ (otherwise we will need more than $b$ repairs when we flip the value of $c_i$). Now all literals $\{l_1^i, l_2^i, \ldots, l_{b+1}^i\}$ cannot be set to 0, since a break to $c_i$ would again necessitate $b+1$ repairs. Hence at least one of the literals in $\{l_1^i, l_2^i, \ldots, l_{b+1}^i\}$ is set to 1. In other words, the literal $C_i$ is true. Since $c_i = 0$ for all $i$, $T$ must have a model. **QED**

We can also show that finding super*-models for 2-SAT is in polynomial time.

**Theorem 4.5** SUP*$(1,1)$–2-*SAT* $\in P$.

*Proof:* Let $\phi$ be the input 2-SAT formula over $n$ variables $[n]$. We construct the graph $G(\phi)$ as described before.

Since a super*-model is by definition also a $(1,1)$ supermodel, we must have the same path restrictions set forth by Lemma 4.1 and Lemma 4.2. However, if $\phi$ has a super*-model then we can show that any simple path in $G(\phi)$ can have length at most 1. Suppose not: let $(u, v, w)$ be a simple path of length 2. Let $X$ be a super*-model of $\phi$. Because of Lemma 4.3, we know that $X(u) = 0, X(w) = 1$. Suppose $X(v) = 1$ (the argument for $X(v) = 0$ is similar). If we break the value of $X(v)$, then we have 2 alternatives: no repair or set $X(w) = 0$. In the latter case, if we now break $X(u)$, we need 2 repairs. In the former case, we now break $X(w)$, then we face the same problem once we break $X(u)$. Hence the length of a simple path in $G(\phi)$ can have length at most 1. Note $G(\phi)$ may have cycles $(u, v), (v, u)$, however in that situation $\{u, v\}$ must form one connected component. We can assign either 0 or 1 to both $u, v$. So wlog, assume that $G(\phi)$ has no cycles. In that case, the simple path length restriction means that $G(\phi)$ is a bipartite graph.

Let $G(\phi) = R \cup B$ where $R, B$ are disjoint vertex sets and all edges in $G(\phi)$ are between vertices in $R$ and vertices in $B$. Let $R$ could be the vertices with in-degree 0 and $B$ be the vertices with out-degree 0. Observe that a vertex cannot have positive in-degree and positive out-degree.

**Claim 1** *If* $(u, v)$ *is an edge in* $G(\phi)$*, then the out-degree of* $\neg u$ *is* 0.

*Proof:* Otherwise, there would be a path of length 2 or a cycle, both of which we have excluded. □

Hence if $u \in R$ iff $\neg u \in B$. We also observe that there are no isolated points in $G(\phi)$ since every clause is a disjunction of distinct literals. Now let $X$ be an assignment that sets every literal in $R$ false (0) and (that automatically sets) every literal in $B$ true.

**Claim 2** $X$ *is a super*-model.*

*Proof:* We exhibit a weakly closed set that contains only models of $\phi$ and that contains $X$. Let $Y_B$ ($Y_R$) denote the restriction of any assignment $Y$ onto the literals in $B$ ($R$). Consider the set of assignments $\mathcal{C} = \{Y |\ Y_B$ contains at most one positive literal $\}$. Note that if $B$ contains at most one positive literal under $Y$, then $Y_R$ contains at most one negative literal. It is clear that $\mathcal{C}$ is a closed set. □

An instance of HORNSAT is a boolean formula in CNF where each clause contains at most 1 positive literal. As in 2-SAT, there is a polynomial time algorithm to find a model of a Horn formula. However, *unlike* the situation in 2-SAT, finding supermodels for Horn formulas is NP complete.

**Theorem 4.6** SUP$(1,1)$-*HORNSAT is NP-complete.*

*Proof:* SUP$(1,1)$-HORNSAT is clearly in NP.To prove that it is NP-hard, we reduce from 3-SAT. Let $T = C_1 \wedge C_2 \cdots \wedge C_m$ be an instance of 3-SAT. We assume without loss of generality, that there are no pure literals in $T$.

For ease of description, we first apply an intermediate transformation to $T$ by replacing any positive literal (say $x$) in $C_i$ by a new negative literal $(\neg a_x)$. But then we add clauses to $T$ to signify that $\neg a_x \Leftrightarrow x \equiv (((\neg a_x) \vee x) \wedge (a_x \vee x))$. Thus we obtain

$$T' = \bigwedge_{1 \le i \le m} C_i' \bigwedge_x (\neg a_x \Leftrightarrow x)$$

where $x$ refers to a variable (positive literal) in $T$ and $C_i'$ refers to the pure Horn clause (no positive literal) by replacing all the positive literals in $C_i$ as described above. Note that since we assume that there are no pure literals in $T$, we add clauses for $((\neg a_x) \Leftrightarrow x)$ for all variables $x$. Observe that $T'$ is *almost* Horn, the bad clauses are only the clauses $(a_x \vee x)$. Clearly $T'$ has a model iff $T$ has a model.

Now we produce an instance of SUP$(1,1)$-HORNSAT from $T'$. We first introduce two new variables $A, B$. For each clause $C'_i = \neg l_i[1] \vee \neg l_i[2] \vee \neg l_i[3]$ of $T'$ (note that $l_i[1], l_i[2], l_i[3]$ are variables) define the clause

$$\begin{aligned} C''_i &= (\neg n_i \vee \neg n_i(l_i[1]) \vee \neg n_i(l_i[2]) \vee \neg n_i(l_i[3])) \\ &\wedge (n_i \Rightarrow A) \wedge (A \Rightarrow B) \\ &\wedge (l_i[1] \Rightarrow n_i(l_i[1])) \wedge (l_i[2] \Rightarrow n_i(l_i[2])) \wedge (l_i[3] \Rightarrow n_i(l_i[3])) \end{aligned}$$

where $n_i, n_i(l_i[1]), n_i(l_i[2]), n_i(l_i[3])$ are all newly introduced variables. Note that $C'$ is Horn. For the clauses that represent $(a_x \Leftrightarrow x)$ we construct the gadget

$$D(x) = (h_x \Rightarrow a_x) \wedge (a_x \Rightarrow t_x) \wedge (h_x \Rightarrow x) \wedge (x \Rightarrow t_x)$$

where $h_x$ and $t_x$ are new variables, $x$ refers to a variable in the original theory $T$. Note that $D(x)$ is also Horn. Our instance of SUP$(1,1)$-HORNSAT is

$$T' = \bigwedge_{1 \le i \le m} C''_i \bigwedge_x D(x)$$

It is easy to that each $n_i(l_i[1])$ is always at the end of a chain of implications of length 2 (some $u \Rightarrow v \Rightarrow n_i(l_i[1])$).

Suppose $T'$ had a model $X'$. Extend that to a model $X''$ of $T''$ by setting $n_i = 0$ for all $i \in \{1, \ldots, m\}$, $A = 1, B = 1$ and $n_i(l_i[j]) = 1$ for all $1 \le i \le m$ and for all $1 \le j \le 3$ and setting $h_x = 0, t_x = 1$ for all $x$.

We now show that that $X''$ is a supermodel. Suppose some variable in $T'_1 s$ domain is flipped, then we need no repairs. If $A$ is flipped, we need no repairs. If $B$ is flipped, we need one repair$(A)$. If $h_x$ or $t_x$ are flipped, we need one repair: because of $T'$, $(x \vee a_x) \wedge (\neg x \vee \neg a_x)$. If $n_i$ is flipped, we do not need to repair $A, B$. But we do need to repair either $n_i(l_i[1]), n_i(l_i[2])$ or $n_i(l_i[3])$ where $C'_i = \neg l_i[1] \vee \neg l_i[2] \vee \neg l_i[3]$. To be able to set one of these to 0 we require one of $l_i[1], l_i[2], l_i[3]$ to be set to 0, which is true since $X'$ is a model of $T'$. Hence $X''$ is a supermodel.

Now suppose $T''$ has a supermodel $X''$. We claim that in such a model $n_i(l_i[1]) = 0$ (because it is at the end of path of length 2 in a 2 SAT sub formula, see Lemma 4.3) and similarly $h_x = 0, t_x = 1$ (also by Lemma 4.3). We also claim that in such a supermodel, $A = 1$. Suppose not, then $A = 0$. This implies that $n_i = 0$ for all $i$. If $n_i$ is now flipped, we need to repair at least one of the literals $n_i(l_i[1])$ appearing in $C''_i$ *and* repair $A$. SO 2 repair will be needed. Hence $A = 1$. Since $h_x = 0, t_x = 1$, both $a_x$ and $x$ cannot be set to the same value. Hence $a_x \Leftrightarrow x$ is true. We now need to show that $X''$ also makes $C'_i = l_i[1] \vee l_i[2] \vee l_i[3]$ true. If I flip $n_i$, one needs to repair with one of $n_i(l_i[1]), n_i(l_i[2]), n_i(l_i[3])$. But this is possible only if one of $l_i[1], l_i[2], l_i[3]$ is set to 0. **QED**

Another class of boolean formulas that have polynomial time satisfiability checkers is Affine SAT: these are formulas which are a conjunction of clauses, where each clause is an exclusive-or (denoted by $\oplus$). One can find a satisfying assignment for a formula in affine form by a variant of gaussian elimination. We now prove that finding supermodels for affine formulas is also in polynomial time.

**Theorem 4.7** SUP$(1,1)$–*Affine-SAT* $\in$ *P.*

*Proof:* Let $\phi = C_1 \wedge C_2 \ldots \wedge C_m$ be a boolean formula in affine form over the set of variables $X = \{x_1, x_2, \ldots, x_n\}$.

For each variable $x$, define $I(x) = \{i |\ 1 \le i \le m, x \text{ appears in } C_i\}$. For $i \in I(x)$, let $N_i(x) = \{y \in X |\ y \text{ appears in } C_i, y \ne x\}$ denote the set of variables that appear with $x$ in clause $C_i$. With a slight abuse of notation, let $I(Y) = \cap_{y \in Y} I(y)$ denote the set of clauses where all variables in $Y$ appear together where $Y \subseteq X$. We similarly define $N_i(Y)$ for $i \in I(Y)$. For $1 \le i \le m$ let $Y \cap C_i$ denote the set of variables in $Y$ that appear in clause $C_i$.

**Lemma 4.5** *$\phi$ has a supermodel iff for all $x \in X$, there exists $y = y(x) \in X$, such that $y \in \bigcap_{i \in I(x)} N_i(x)$ and $x \in \bigcap_{i \in I(y)} N_i(y)$.*

*Proof:* It is easy to see that $\{x, y(x)\}$ are a break-repair pair. □

Since the conditions in Lemma 4.5 are easily checkable in polynomial time, we have a polynomial time algorithm for SUP(1,1)–Affine SAT. **QED**

We can in fact, prove the following stronger theorem

**Theorem 4.8** SUP$(r,s)$–*Affine-SAT* $\in$ *P.*

*Proof Sketch:* Let $r = 1$ and apply induction on $s$. We first check whether the input formula $\phi$ (having $n$ variables $\{x_1, x_2, \ldots, x_n\}$ has a sup$(1,j)$ supermodel (for $1 \leq j \leq s-1$). If so, we are done. Otherwise construct the set $L = X_s = \{S \subseteq X| \; |S| = s\}$. Note that $|L| \in \mathrm{O}(n^s)$ and $L$ contains all possible repair sets of size $s$ for each break. However not all sets in $L$ can be repair sets: we might break a variable $x$, repair by some set $l \in L$ but those $s$ repairs might themselves need further repairs. The following lemma characterizes valid repair sets.

**Lemma 4.6** *The variable $x \in X$ be repaired by the set $S \in L$ in a* sup$(1,s)$ *supermodel of $\phi$ iff for each $i \in I(x)$, $|S \cap C_i|$ is odd and for each $i \in \{1, \ldots, m\} \setminus I(x)$, $|S \cap C_i|$ is even.*

Since the conditions of Lemma 4.6 can be checked in polynomial time (there are $2^s$ subsets to check), there is a polynomial time test for sup$(1,s)$ supermodels for affine formulas (recall that $s$ is a fixed constant and not a part of the input).

For arbitrary $r$, the above description can be easily modified to provide an algorithm for sup$(r,s)$ as well. **QED**

## 5 Sparse Closed Sets

In this section, we give a lower bound to the size of sparse closed sets. We first make some easy observations about sparse closed sets. In the following discussion, let $\mathcal{C}$ refer to a sparse closed set.

**Lemma 5.1** *If $Y = \delta_{ij}(X)$ and $Z = \delta_{kl}(X)$ are distinct, where $X, Y, Z \in \mathcal{C}$, then $\{i,j\} \cap \{k,l\} = \emptyset$.*

*Proof:* Wlog assume $i = k$. Then $\Delta_i(X, \mathcal{C}) = \{Y, Z\}$ contradicting the fact that $|\Delta_i(X, \mathcal{C})| = 1$. □

If we have three sets $X, Y, Z$ with $Y = \delta_{ij}(X), Z = \delta_{ik}(X)$, then the incidence vectors of $X, Y$ and $Z$ form an equilateral triangle with sides of length 2, the metric being the Hamming distance. We shall refer to this as a 2-triangle. Thus Lemma 5.1 says that a sparse closed set is 2-triangle free.

Given $X \in \mathcal{C}$, let $N_2(X) = \bigcup_{i \in [n]} \Delta_i(X, \mathcal{C})$.

**Lemma 5.2** $|N_2(X)| = n/2$.

*Proof:* Consider the set $N'(X) = \{\{i,j\}| \; \delta_{ij}(X) \in N_2(X)\}$. Because of Lemma 5.1, we know that $N'$ is a 2-partition of $[n]$ and since $\mathcal{C}$ is closed, it is a maximal 2-partition. Hence $|N'(X)| = n/2$ using Lemma 2.1. Also since the map $\phi : N_2(X) \to N'(X)$ defined by $\phi(\delta_{ij}(X)) = \{i,j\}$ is a bijection, we have $|N'(X)| = |N_2(X)|$. □

Observe that Lemma 5.2 implies that if $\mathcal{C}$ is a sparse closed subset of $2^{[n]}$ then $n$ is even. This also follows from observing that the elements of $[n]$ belong to a unique break-repair pair (otherwise there would be more than one repair for a particular break).

We can define the (undirected) graph $G = G(\mathcal{C})$ as follows: the vertices of $G$ are the elements of $\mathcal{C}$ and the edges are $\{u, v\}$ where $v = \delta_{ij}(u)$ for some $i \neq j \in [n]$.

$\mathcal{C}$ is said to be connected if $G(\mathcal{C})$ is connected.

**Lemma 5.3** *Let $\mathcal{C}$ be a sparse closed set. Then $\mathcal{C}$ is minimal iff $\mathcal{C}$ is connected.*

*Proof:* Let $H$ be a connected component of $G(\mathcal{C})$. Let $u \in H$. By definition, for each $i \in [n]$, $u$ in $G(\mathcal{C})$ is connected to $v$ where $v \in \Delta_i(u, \mathcal{C})$. Hence $\Delta_i(u, H) = \Delta_i(u, \mathcal{C})$. Hence $H$ is closed (it is obviously sparse). Thus $\mathcal{C}$ is minimal iff $H = \mathcal{C}$. □

If $\mathcal{C}$ is sparse closed, wlog we can assume that $\emptyset \in \mathcal{C}$ (We can relabel 0's and 1's in the incidence vectors in $\mathcal{C}$ appropriately). Let $\mathcal{C}_k = \{A \in \mathcal{C}| \; |A| = k\}$. We note the following two obvious facts about minimal sparse closed sets.

**Lemma 5.4** *If $\mathcal{C}$ is minimal and $\emptyset \in \mathcal{C}$ then*

1. *$\mathcal{C}_k = \emptyset$ for all odd $k$.*

2. *If $\mathcal{C}_k$ be the highest non-empty level (i.e. $|\mathcal{C}_k| > 0$ and $\mathcal{C}_i = \emptyset$ for all $i > k$), then $k \geq n/2$.*

*Proof:* (i) We observe that if $u = \emptyset$ and $\{u, v\}$ is an edge in $G(\mathcal{C})$, then $|u| \equiv 0 (\bmod 2)$. Using Lemma 5.3 and the fact that $|X| + |\delta_{ij}(X)| \equiv 0(\bmod 2)$, the result follows by induction.
(ii) Suppose $k < n/2$. Let $\mathcal{C}_k$ be the highest non-empty level and let $u \in \mathcal{C}_k$. Since $\mathcal{C}_k$ is non-empty and highest, $\Delta_i(u, \mathcal{C}) \subseteq \mathcal{C}_{k-2} \cup \mathcal{C}_k$ for all $i \in [n]$. Consider the set $R(u) = \bigcup_{i \in u} R_i(u, \mathcal{C})$. Clearly $|R(u) \cup u| \leq 2k < n$ (as $|R_i(u, \mathcal{C})| = 1$). Consider $R' = [n] \setminus (R(u) \cup u)$. There must be some $l, m \in R'$ such that $\delta_{lm}(u) \in \mathcal{C}$. But $|\delta_{lm}(u)| = k + 2$ (since $l, m \notin u$), a contradiction. □

The idea behind lemma 5.4 (ii) is that if there are more 0's than 1's in the incidence vector, there must be two 0's which form a break-repair pair.

**Lemma 5.5** *Let $u \in \mathcal{C}_k$, where $2 \leq k < n/2$ is even. Let $P(u) = \{v \in \mathcal{C}_{k-2} |\ u = \delta_{ij}(v)$ for some $i$ and $j$ $\}$. Then $|P(u)| \leq k/2$.*

*Proof:* Consider the set $P'(u) = \{\{i, j\} |\ u = \delta_{ij}(v), v \in \mathcal{C}_{k-2}\}$ so that $|P(u)| = |P'(u)|$. Since $u \in \mathcal{C}_k$ and $v = \delta_{ij}(u) \in \mathcal{C}_{k-2}$, this means that $i, j \notin v$ and $i, j \in u$. Lemma 5.1 implies that for distinct elements $\{i, j\} \in P'(u)$ and $\{k, l\} \in P'(u)$, $\{i, j\} \cap \{k, l\} = \emptyset$. This means that $P'(u)$ forms a partition of the elements in $u$ into 2 element subsets, hence $|P'(u)| \leq k/2$ by Lemma 2.1. Thus $|P(u)| \leq k/2$. □.

**Lemma 5.6** *Let $\mathcal{C}$ be minimal sparse, $\emptyset \in \mathcal{C}$ and $2 \leq k \leq n/2$. Then $|\mathcal{C}_{k+2}| \geq |\mathcal{C}_k|(n - 2k)/(k + 2)$.*

*Proof:* Let $u \in \mathcal{C}_k$. Define $R(u)$ as in Lemma 5.4. Let $R' = [n] \setminus (R(u) \cup u)$. Since $k \leq n/2$, and $|R'| \geq n - 2k$. For each $i \in R'$, there is a unique $j \in R'$ such that $\delta_{ij}(u) \in \mathcal{C}$. Observe that $|\delta_{ij}(u)| = k + 2$. The set $R'' = \{\{i, j\} |\ i, j \in R', \delta_{ij}(u) \in \mathcal{C}\}$ is a partition of $R'$ into disjoint 2-element subsets. Hence $|R''| \geq (n - 2k)/2$. Thus each $u \in \mathcal{C}_k$ is connected to at least $(n - 2k)/2$ vertices $v \in \mathcal{C}_{k+2}$. By Lemma 5.5, $|P(v)| \leq (k + 2)/2$ which bounds the over count. Hence there are at least $\frac{|\mathcal{C}_k|(n-2k)/2}{(k+2)/2}$ elements in $\mathcal{C}_{k+2}$. □

If $\mathcal{C}$ is minimal sparse with $\phi \in \mathcal{C}$, then we know that $|\mathcal{C}_2| = n/2$ by Lemma 5.2. Hence by Lemma 5.4, (assume for clarity that $8|n$)

$$
\begin{aligned}
|\mathcal{C}| &\geq \sum_{i=0,\ i \text{ even}}^{i<n/2} |\mathcal{C}_i| \\
&\geq 1 + \frac{n}{2} + \frac{n}{2}\frac{(n-4)}{2 \cdot 2} + \frac{n}{2}\frac{(n-4)}{2 \cdot 2}\frac{(n-8)}{3 \cdot 2} + \cdots \\
&\geq \frac{1}{(n/8)!}\sum_{k=0}^{n/8} \frac{(n/8)!}{k!}(n/4)^k \\
&\geq \frac{1}{(n/8)!}\sum_{k=0}^{n/8} \frac{(n/8)!}{k!(n/8-k)!}(n/4)^k \\
&= \frac{1}{(n/8)!}(1 + n/4)^{n/8}
\end{aligned}
$$

Estimating

$$
\frac{(n/4)^{n/8}}{(n/8)!} \approx \frac{(n/4)^{n/8}}{(n/8e)^{n/8}} = (2e)^{n/8} \approx 1.235^n
$$

We have proved the following theorem.

**Theorem 5.1** *Every sparse closed subset of $2^{[n]}$ has size at least $(2e)^{n/8}$.*

We now give some examples of sparse closed sets. Let $\mathcal{B}$ denote the family of subsets of $[n]$ whose incidence vectors satisfy the following boolean formula

$$\mathcal{B} = \bigwedge_{1 \leq i \leq n/2} (x_{2i-1} = x_{2i}) \tag{1}$$

where $x_j$ refers to the $j$th bit of the incidence vector.

It is easy to see that $\mathcal{B}$ is a sparse closed set (since it is connected) of size $2^{n/2}$ and that it is minimal. We do not know examples of sparse closed sets of size smaller than $2^{n/2}$.

**Conjecture 5.1** *Every sparse closed subset of* $2^{[n]}$ *has size at least* $2^{n/2}$.

We now turn to the problem how large sparse closed sets can be. Since we know that sparse closed sets are 2-triangle free, this allows us to derive an upper bound. In the following discussion, let $d(x, y)$ denote the Hamming distance between the $n$-vectors $x$ and $y$.

**Theorem 5.2** *If* $\mathcal{F} \subseteq 2^{[n]}$ *is 2-triangle free, then* $|\mathcal{F}| \leq \mathrm{O}(2^n/n)$.

*Proof:* Let $x$ be the incidence vector of a set in $\mathcal{F}$. Let $N_i(x) = \{y \in 2^{[n]} | d(x, y) = i\}$.

It is easy to see that $|N_1(x) \cap \mathcal{F}| \leq 2$: if there were 3 elements $a, b, c \in N_1(x) \cap \mathcal{F}$, then they form a 2-triangle. Thus $|N_1(x) \setminus \mathcal{F}| > n - 2$.

Define

$$N_i(\mathcal{F}) = \left( \bigcup_{x \in \mathcal{F}} N_i(x) \right) \setminus \mathcal{F}.$$

We will find a bound for $|N_1(\mathcal{F})|$. One can easily show that for each $\alpha \in N_1(\mathcal{F})$, $|N_1(\alpha) \cap \mathcal{F}| \leq 2$. Note that if $x \in N_1(\alpha) \cap \mathcal{F}$, then $\alpha \in N_1(x)$. This means that for at most two different $x \in \mathcal{F}$ can $\alpha$ be in $N_1(x)$. Hence $|N_1(\mathcal{F})| \geq |\mathcal{F}|(n-2)/2$. Since $\mathcal{F} \cup N_1(\mathcal{F}) \subseteq 2^{[n]}$ and $\mathcal{F} \cap N_1(\mathcal{F}) = \emptyset$, we have $|\mathcal{F}| \leq 2^n/(n/2)$. **QED**

**Corollary 5.1** *If* $\mathcal{C} \subseteq 2^{[n]}$ *is sparse closed, then* $|\mathcal{C}| \leq \mathrm{O}(2^n/n)$.

The best explicit construction of sparse closed sets we can give are of size $2^n/n^2$ which we describe below. Constructions of large sparse sets (not necessarily closed) are considerably easier: an easy probabilistic construction shows that there are sparse (i.e. 2-triangle free) sets of size $2^n/n^{1.5}$.

We now describe a construction of sparse closed sets of size $2^n/n^2$. We slightly modify the notation of Equation 1 to define the following family of boolean functions. Let $S \subseteq [n/2]$. Define

$$\mathcal{B}_S = \bigwedge_{i \in S} (x_{2i-1} \neq x_{2i}) \bigwedge_{i \notin S} (x_{2i-1} = x_{2i}).$$

So $\mathcal{B}_\emptyset = \mathcal{B}$ of Equation 1. It is clear that each $\mathcal{B}_S$ defines a sparse closed set: we shall refer to an element of this sparse closed set as some $x \in \mathcal{B}_S$ to mean that $x$ satisfies $\mathcal{B}_S$. The following lemma allows us to build large (non-minimal) sparse closed sets.

**Lemma 5.7** *If* $S, T \subseteq [n/2]$ *such that* $d(S, T) \geq 3$ *then* $\mathcal{B}_S \vee \mathcal{B}_T$ *defines a sparse closed set.*

*Proof:* We prove that if $x \in \mathcal{B}_S$ and $y \in \mathcal{B}_T$, then $d(x, y) \geq 3$, which implies that $\mathcal{B}_S \vee \mathcal{B}_T$ is 2-triangle free. Let $D = S \triangle T$ thus $|D| \geq 3$. Observe that any vector that satisfies $x_{2i-1} = x_{2i}$ is at least distance 1 away from any vector that satisfies $x_{2i-1} \neq x_{2i}$ where $i \in D$. Hence $\mathcal{B}_S \vee \mathcal{B}_T$ is 2-triangle free. □

Thus if $\mathcal{F}$ is a family of subsets of $[n/2]$ such that $S, T \in \mathcal{F}$, $S \neq T \Rightarrow d(S, T) \geq 3$, then $\bigvee_{S \in \mathcal{F}} \mathcal{B}_S$ will define a sparse closed set. We quote the following standard result from coding theory [Va92, Ro92].

**Theorem 5.3 (Gilbert-Varshamov Inequality)** *There exists a family* $\mathcal{F} \subseteq 2^{[n]}$ *satisfying the condition* $S, T \in \mathcal{F}$, $S \neq T \Rightarrow d(S, T) \geq 3$ *such that* $|\mathcal{F}| \geq \frac{2^n}{V_2(n)}$ *where* $V_2(n) = |\binom{[n]}{\leq 2}| = \binom{n}{0} + \binom{n}{1} + \binom{n}{2}$.

Using a family of size $\Omega(\frac{2^{n/2}}{n^2})$ guaranteed by Theorem 5.3 as our "index" set, we get that many disjoint balls each of size $2^{n/2}$ with a 2-triangle free union.

**Corollary 5.2** *There exist (non-minimal) sparse closed sets of size* $\Omega(2^n/n^2)$.

We suspect that minimal sparse closed sets cannot achieve this bound. A bound on the largest known one is as follows:

**Theorem 5.4** *There exists a minimal sparse subset of* $2^{[n]}$ *of size* $80^{n/8} \approx 1.69^n$.

The proof follows from the existence of a minimal sparse closed set of 80 subsets for $n = 8$. We can use this to construct a minimal sparse set of size $80^{n/8}$ by taking direct products. We do not have succint description of this set. There is another (equally mysterious) example: a minimal sparse closed set of size 10 consisting of subsets of [6]. We include this example in the appendix. Their odd structure seems to be due to the fact that the break-repair pairing changes from element to element in the closed set, unlike in the $2^{n/2}$ example.

# 6 Closed Sets

In this section, we study the extremal properties of general closed sets. The first theorem concerns the minimum size of a closed set. Again without any loss of generality, we can assume that if $\mathcal{C}$ is closed, then $\emptyset \in \mathcal{C}$. In the following discussion, let $\mathcal{C}$ refer to a closed set such that $\emptyset \in \mathcal{C}$.

**Theorem 6.1** $|\mathcal{C}| \geq n$.

*Proof:* Following Lemma 5.2 we can easily show that $|N_2(x)| \geq n/2$ for each $x \in \mathcal{C}$. Let $a = \emptyset \in \mathcal{C}$ and assume that $b = \delta_{st}(a) = \{s,t\}$ where $s,t \in [n]$. Define $I = R_s(b,\mathcal{C}) \cup R_t(b,\mathcal{C})$.

Observe that if $i \in I$, then $\delta_{si}(b)$(or $\delta_{ti}$) $\in N_2(a)$. Hence $|N_2(a) \cap N_2(b)| \geq |I|$.

Now each $j \in ([n] \setminus (I \cup \{s,t\}))$ in $b$ is repaired by (at least one) $r \in [n] \setminus \{s,t\}$. Also $\delta_{jr}(b) \in N_2(b) \setminus N_2(a)$. Hence

$$|N_2(b) \setminus N_2(a)| \geq \frac{n - |I| - 2}{2}.$$

One can similarly show

$$|N_2(a) \setminus N_2(b)| \geq \frac{n - |I| - 2}{2}.$$

Hence

$$\begin{aligned} |N_2(a) \cup N_2(b)| &= |N_2(a) \setminus N_2(b)| + |N_2(b) \setminus N_2(a)| + |N_2(a) \cap N_2(b)| \\ &\geq \frac{n - |I| - 2}{2} + \frac{n - |I| - 2}{2} + |I| \\ &= n - 2 \end{aligned}$$

Counting $a, b$ along with $N_2(a) \cup N_2(b)$, we have the desired result. **QED**

We note that while our lower bound for sparse closed sets (Theorem 5.1) is not known to be achieved, the lower bound in Theorem 6.1 is achieved. The boolean formula $E_1(x_1, x_2, \ldots, x_n)$ defines a closed set of size $n$, where $E_1(x_1, x_2, \ldots, x_n)$ is true iff exactly one variable in $\{x_1, x_2, \ldots, x_n\}$ is true.

As in the previous section, we define the undirected graph $G(\mathcal{C})$ with vertices as elements of $\mathcal{C}$ and edges $\{u, \delta_{ij}(u)\}$ for some $i, j \in [n]$. We can also easily see that a weak version of Lemma 5.3 holds for general closed sets.

**Lemma 6.1** *Let* $\mathcal{C}$ *be a minimal closed set. Then* $\mathcal{C}$ *is connected.*

Because minimal closed sets are connected and $\emptyset \in \mathcal{C}$, all sets in $\mathcal{C}$ have even parity (analogous to Lemma 5.4). It is clear that minimal closed sets can have size at most $2^{n-1}$. We can improve this result to a constant fraction of $2^{n-1}$.

Let $\mathcal{F} = \{\{X_1, Y_1\}, \ldots, \{X_m, Y_m\}\}$ be a family of unordered tuples of subsets of $[n]$ such that $X_i \neq Y_i$ for all $i$. An independent set of $\mathcal{F}$ is a sub-family $\mathcal{I} \subseteq \mathcal{F}$ such that if $\{X,Y\}, \{R,S\} \in \mathcal{I}$ and $\{R,S\} \neq \{X,Y\}$ then $\{X,Y\} \cap \{R,S\} = \emptyset$. Let $m(\mathcal{F}) = |\{X | \exists Y, \{X,Y\} \in \mathcal{F}\}|$ be the number of distinct sets which appear in any tuple of $\mathcal{F}$. If $X \subseteq [n]$, define the degree of $X$ as $deg(X, \mathcal{F}) = |\{Y|\ \{X,Y\} \in \mathcal{F}\}|$.

**Lemma 6.2** *Let $\mathcal{F} = \{\{X_1, Y_1\}, \ldots, \{X_m, Y_m\}\}$ be a family of subsets of $[n]$ such that $X_i \neq Y_i$ and $deg(X) \leq 2$ for all $X \subseteq [n]$. Let $\mathcal{I}$ be a maximal independent subset of $\mathcal{F}$. Then $m(\mathcal{F}) \leq 2m(\mathcal{I})$.*

*Proof:* If $\mathcal{F} = \mathcal{I}$ we are done. Else let $\overline{\mathcal{I}} = \mathcal{F} \setminus \mathcal{I}$. If $\{X, Y\} \in \overline{\mathcal{I}}$, then either $\deg(X, \mathcal{I}) = 1$ or $\deg(Y, \mathcal{I}) = 1$ (observe that neither can be $> 1$ as $\mathcal{I}$ is an independent set): otherwise $\mathcal{I} \cup \{\{X, Y\}\}$ would be an independent set, contradicting maximality of $\mathcal{I}$. Hence, $|\{X| \deg(X, \mathcal{I}) = 0\}| \leq |\overline{\mathcal{I}}|$. Observe also that $\deg(X, \overline{\mathcal{I}}) \leq 1$ for each $X$ with $\deg(X, \mathcal{I}) = 1$. Hence $|\overline{\mathcal{I}}| \leq m(\mathcal{I})$. Thus

$$m(\mathcal{F}) = m(\mathcal{I}) + |\{X| \deg(X, \mathcal{I}) = 0\}| \leq 2m(\mathcal{I}). \quad \textbf{QED}$$

The following lemma is useful in characterizing closed sets.

**Lemma 6.3** *If $\mathcal{C}$ is a minimal closed set, then all $X \in \mathcal{C}$ are irredundant.*

*Proof:* Suppose there is some $X \in \mathcal{C}$ which is redundant. We claim that $\mathcal{C} \setminus \{X\}$ is a closed set. Suppose not. Then there is some $Y \in \mathcal{C} \setminus \{X\}$ and an $i \in [n]$ such that $\Delta_i(Y, \mathcal{C}) = \{X\}$. Then $X = \delta_{ij}(Y)$ for some $j \neq i$. We claim now that $|\Delta_j(X, \mathcal{C})| = 1$ contradicting the hypothesis. Clearly, $Y = \delta_{ji}(X) \in \Delta_j(X, \mathcal{C})$. Suppose there is some $k \neq i$ such that $Z = \delta_{jk}(X) \in \mathcal{C}$. Then $\delta_{ik}(Y) = Z$, contradicting our assumption that $\Delta_i(Y, \mathcal{C}) = \{X\}$. □

Thus if $\mathcal{C}$ is minimal, then for each $X \in \mathcal{C}$, there is some $i \in [n]$ such that $Y = \delta_{ij}(X)$ is the unique element in $\Delta_i(X, \mathcal{C})$. We shall write $Y = \delta^*_{(i,j)}(X)$ in that case. Observe that the order $(i, j)$ is important. For $X, Y \in \mathcal{C}$, $Z \subseteq [n]$, we say that the pair $\{X, Y\}$ *excludes* $Z$ if $Y = \delta^*_{(i,j)}(X)$, $Z = \delta_{ik}(X)$ for some distinct $i, j, k \in [n]$, and $|\Delta_i(X, \mathcal{C})| = 1$ (we note that we can write $\{X, Y\}$ as unordered pair because if $Y = \delta^*_{(i,j)}(X)$ then $X = \delta^*_{(j,i)}(Y)$ and $Z$ is still excluded). Thus if $Y = \delta^*_{(i,j)}(X)$, then $\{X, Y\}$ excludes $n - 2$ elements $Z = \delta_{ik}(X) \in \Delta_i(X) \setminus \Delta_i(X, \mathcal{C})$ where $k \neq j$.

**Lemma 6.4** *Let $\mathcal{C}$ be a minimal closed set and let $X \in \mathcal{C}, Z \subseteq [n]$. Then*

$$|\{Y \in \mathcal{C}| \ \{X, Y\} \text{ excludes } Z\}| \leq 2.$$

*Proof:* Wlog assume $Z = \emptyset$. Then $X = \{i, j\}$ for some $i, j \in [n]$. Any $Y$ such that $\{X, Y\}$ excludes $Z$ has to have $|Y| = 2$ and $Y \cap X \neq \emptyset$. Hence there can be at most 2 such sets. □

For $Z \subseteq [n]$, define $e(Z, \mathcal{C}) = \{\{X, Y\}| \ \{X, Y\}$ excludes $Z$, $X, Y \in \mathcal{C}\}$. Lemma 6.4 implies that $\deg(X, e(Z)) \leq 2$.

**Lemma 6.5** *Let $e(Z) \neq \emptyset$. Let $\mathcal{I} \subseteq e(Z, \mathcal{C})$ be a maximal independent set. Then $m(\mathcal{I}) \leq 2n$.*

*Proof:* Assume wlog $Z = \emptyset$. Then $e(Z)$ consists of unordered tuples of 2-element subsets of $[n]$. Consider the family $\mathcal{F} = \{X \cap Y| \ \{X, Y\} \in \mathcal{I}\}$. It is clear that $\mathcal{F}$ consists of one element subsets of $[n]$. We claim that $|\mathcal{F}| = |\mathcal{I}|$. Suppose not. Then there are two pairs $\{X, Y\}, \{R, S\} \in \mathcal{I}$ such that $X \cap Y = R \cap S$. Let $X = \{a, b\}, Y = \{a, c\}$ so that $Y = \delta^*_{(b,c)}(X)$ and let $R = \{a, d\}, S = \{a, e\}$ so that $\delta^*_{(d,e)}(R) = S$. Note that since $\mathcal{I}$ is an independent set $d \notin \{b, c\}$ and $e \notin \{b, c\}$. Since $\delta_{(d,c)}(R) = Y$, we cannot have $\delta^*_{(d,e)}(R) = S$, a contradiction. Since $|\mathcal{F}| \leq n$, this implies that $|\mathcal{I}| \leq n$. Since $\mathcal{I}$ is independent, $m(\mathcal{I}) \leq 2n$. □

Now Theorem 6.2 implies that $m(e(Z)) \leq 2(2n) = 4n$. So each $X \in \mathcal{C}$ excludes at least $n - 2$ sets and each set is excluded by at most $4n$ elements in $\mathcal{C}$. Hence

$$2^{n-1} \geq |\mathcal{C}| + \frac{|\mathcal{C}|(n-2)}{4n}$$

which implies that $|\mathcal{C}| \leq \frac{4n}{5n-2} 2^{n-1} \approx (4/5)2^{n-1}$. Thus we have proved

**Theorem 6.2** *If $\mathcal{C}$ is a minimal closed subset of $2^{[n]}$, then $|\mathcal{C}| \leq (4/5)2^{n-1}$.*

As in the situation for sparse closed sets, we do not know of constructions of large minimal closed sets which achieve the above bound. However we have the following:

**Theorem 6.3** *There exists a minimal closed subset of $2^{[n]}$ of size $2^{2n/3}$.*

The proof of the theorem relies on the existence of a minimal closed set of size 16 for $n = 6$ (displayed in the appendix), which was found by exhaustive search. A direct product of these yields a minimal set of size $16^{n/6}$, thus proving the theorem. Similar searches found the minimal set for $n = 5$ is 8 and for $n = 4$ is 4. This suggests the following conjecture.

**Conjecture 6.1** *The largest minimal closed subset of* $2^{[n]}$ *is of size* $2^{n-2}$.

We now turn to the following algorithmic question : given a closed set as input, is it minimal ? Observe that a brute force algorithm that checks each subset of the closed set will run in exponential time in the size of the input. The size of the input could itself be exponential with respect to $n$, the length of the strings. Our goal is to find an algorithm that runs in polynomial time in the size of the input (not necessarily polynomial time in $n$).

**Theorem 6.4** *In polynomial time, one can test if a closed set is minimal.*

*Proof:* Let $\mathcal{C}$ denote the input closed set. For each vertex $u \in \mathcal{C}$ the algorithm runs the procedure expand($\{u\}$). The procedure expand($X$) executes the following steps in sequence:

1. If $X = \mathcal{C}$ return true.
2. If $\exists u \in \mathcal{C} \setminus X$, such that $\Delta_i(u, \mathcal{C}) \subseteq X$ for some $i \in [n]$, set $X = X \cup \{u\}$.
3. If no such $u$ exists and $X \subset \mathcal{C}$ then return false ($\mathcal{C}$ is not minimal). Else go to step 1.

If expand($u$) returns true for each $u \in \mathcal{C}$ then $\mathcal{C}$ is minimal. If expand($u$) is false for any $u \in \mathcal{C}$ then $\mathcal{C}$ is not minimal.

To see correctness, observe that if $\mathcal{C}$ was minimal, then no node $u$ can be removed so expand($u$) will return true for each invocation. If $\mathcal{C}$ was not minimal, there would be some node $u$ that can be removed. The procedure expand would detect which node it is. **QED**

# 7 Conclusions and future work

The principal results of sections 3 and 4 are summarized in the first table.

| | SAT | SUP(1,1) | SUP(1,2) | SUP*(1,1) |
|---|---|---|---|---|
| general | NP-complete | NP-complete | NP-complete | NEXP, NP-hard |
| 2-SAT | P | P | NP-complete | P |
| HORNSAT | P | NP-complete | open | open |
| Affine SAT | P | P | P | P |

The complexity of SUP($r, s$) where $r$ and $s$ are part of the input as opposed to being fixed constants remains an interesting open question. It is not hard to see that SUP($r, s$) is in $\Sigma_3^P$: it is not known whether it is complete for that class. The status of this problem for restricted cases such as 2-SAT is similarly open. Another interesting question is whether we can improve on SUP*(1,1) $\in$ NEXP (e.g. to PSPACE or even NP). This improvement seems to rely on finding suitable small certificates for closed sets. Finally, a practical modification of supermodels involves weakening the condition to allow only a high percentage of the breaks to be repaired. We wonder how this would affect the complexity issues.

The second table summarizes the results from sections 5 and 6.

| | closed | closed minimal | sparse closed | sparse closed minimal |
|---|---|---|---|---|
| largest | $2^{n-1}$ | $\leq (4/5)2^{n-1}$<br>$\geq 2^{2n/3}$ | $\leq O(2^n/n)$,<br>$\geq \Omega(2^n/n^2)$ | $\leq O(2^n/n)$<br>$\geq 80^{n/8}$ |
| smallest | | $n$ | | $\leq 2^{n/2}$<br>$\geq (2e)^{n/8}$ |

The most interesting questions here involve tightening the bounds in the table above and understanding the structure of the minimal sets. The sparse minimal sets in particular seem to have a rich combinatorial structure.

## References


[BP99] Olivier Bailleux and Pierre Marquis, Distance-SAT: Complexity and Algorithms, *Proceedings of the Sixteenth National Conference on Artificial Intelligence (AAAI–99)*, 1999, pp 642–647.

[GPR98] Matthew Ginsberg, Andrew Parkes, Amitabha Roy, Supermodels and Robustness, *AAAI-98/IAAI-98 Proceedings* , Madison, WI, 1998, pp 334-339.

[Pa94] Christos H. Papadimitriou, *Computational Complexity*, Addison-Wesley, 1994.

[Ro92] Steven Roman, *Coding and Information Theory*, Springer-Verlag, 1992.

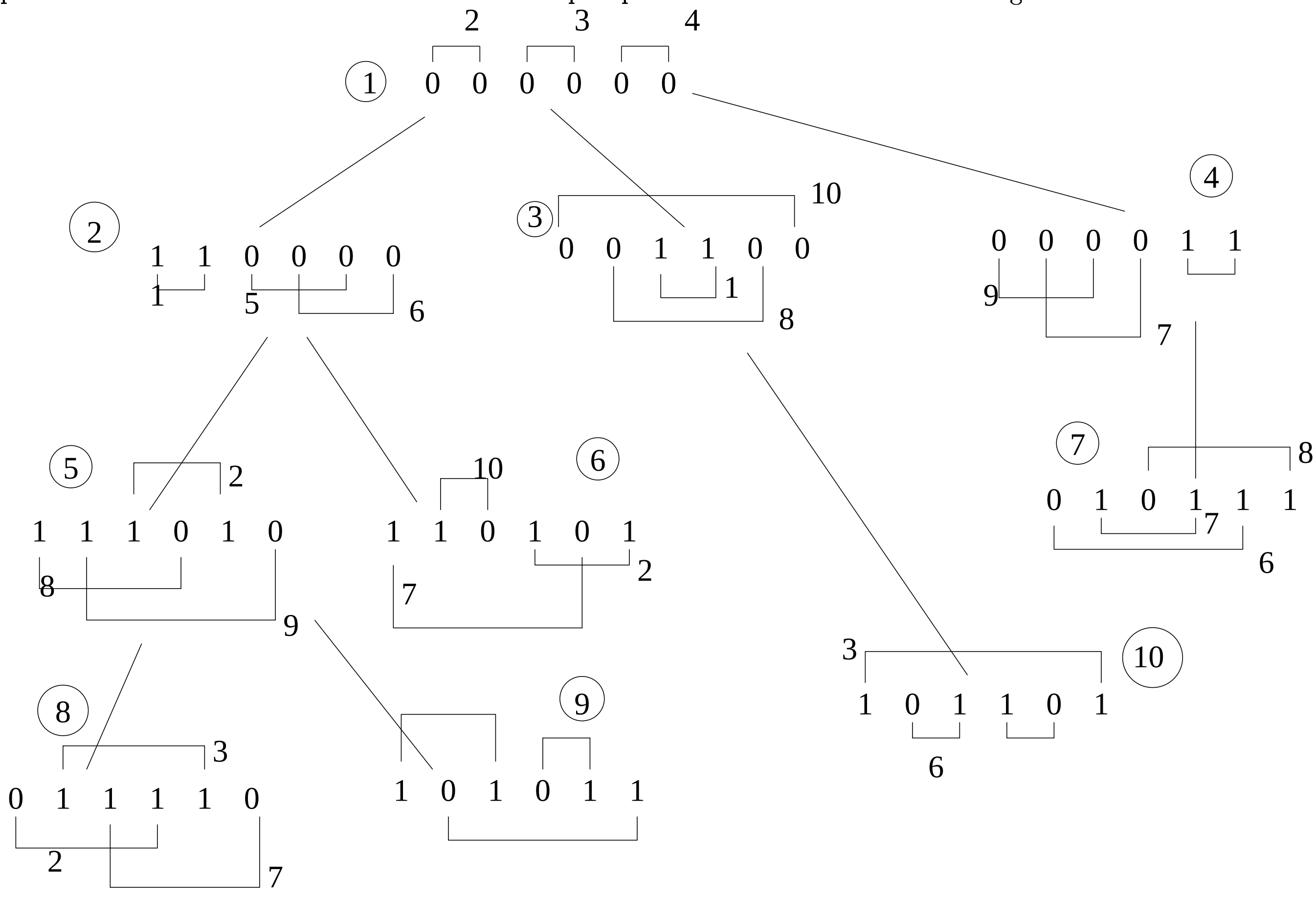
[TP97] J. A. Telle and A. Proskurowski, Algorithms for Vertex Partitioning Problems on Partial k-Trees, *SIAM Journal on Discrete Mathematics*, vol. 10, no. 4, pp 529-550, November 1997.

[Te94] J. A. Telle, Complexity of domination type problems in graphs, *Nordic Journal of Computing* 1 (1994), pp 157-171.

[Va92] J. H. van Lint, *Introduction to Coding Theory, Second Edition* Springer-Verlag, 1992.

# Appendix

We conclude with some examples of minimal sparse closed set and minimal closed sets. The figure below shows a sparse closed set of size 10 for $n = 6$. The break repair pairs are also shown for each string in the set.



We also give an example below of a sparse closed set of size 24 for $n = 8$ :

$$\{00000000, 11000000, 00110000, 00001100,\\ 00000011, 11100001, 11011000, 11000110,\\ 01101001, 10110001, 11100111, 00101101,\\ 01001011, 01111000, 01111110, 10111101,\\ 00100111, 10010011, 11111111, 10011100,\\ 11011011, 01001110, 10010110, 00110110\}$$

An example of a minimal set of size 16 for $n = 6$ used to prove Theorem 6.3 is below:

$$\{100100, 010100, 001100, 111100\\ 010010, 111010, 000110, 110110\\ 101110, 011110, 100001, 001001\\ 101101, 011101, 000011, 110011\}$$



ECCC ISSN 1433-8092
http://www.eccc.uni-trier.de/eccc
ftp://ftp.eccc.uni-trier.de/pub/eccc
ftpmail@ftp.eccc.uni-trier.de, subject 'help eccc'