On the Gap Between $ess(f)$ and $cnf\_size(f)$ (Extended Abstract)

Lisa Hellerstein and **Devorah Kletenik**
Polytechnic Institute of NYU, 6 Metrotech Center, Brooklyn, N.Y., 11201

Abstract

Given a Boolean function f, $ess(f)$ denotes the largest set of assignments that falsify f, no two of which falsify a common implicate of f. The quantity $ess(f)$ is a lower bound on $cnf\_size(f)$ (the minimum number of clauses in a CNF formula for f). Čepek et al. gave examples of functions f for which there is a small gap between $ess(f)$ and $cnf\_size(f)$. We demonstrate significantly larger gaps. We show that the gap can be exponential in n for arbitrary Boolean functions, and $\Omega(\sqrt{n})$ for Horn functions, where n is the number of variables of f.

Introduction

Determining the smallest CNF formula for a given Boolean function f is a difficult problem that has been studied for many years (cf. (Coudert 1994)). Recently, Čepek et al. introduced a combinatorial quantity, $ess(f)$, which lower bounds $cnf\_size(f)$, the minimum number of clauses in a CNF formula representing f (Čepek, Kučera, and Savický 2010). The quantity $ess(f)$ is equal to the size of the largest set of falsepoints of f, no two of which falsify the same implicate of f. [1]

For certain subclasses of Boolean functions, such as the monotone (i.e., positive) functions, $ess(f)$ is equal to $cnf\_size(f)$. However, Čepek et al. demonstrated that there can be a gap between $ess(f)$ and $cnf\_size(f)$. They constructed a Boolean function f on n variables such that there is a multiplicative gap of size $\Theta(\log n)$ between $cnf\_size(f)$ and $ess(f)$. [2] Their constructed function f is a Horn function. Their results leave open the possibility that $ess(f)$ could be a close approximation to $cnf\_size(f)$.

We show that this is not the case. We construct a Boolean function f on n variables such that there is a multiplicative gap of size $2^{\Theta(n)}$ between $cnf\_size(f)$ and $ess(f)$. Note that such a gap could not be larger than 2^{n-1}, since $cnf\_size(f) \leq 2^{n-1}$ for all functions f.

[1] This definition immediately follows from Corollary 3.2 of Čepek et al. (Čepek, Kučera, and Savický 2010).

[2] Their function is actually defined in terms of two parameters n_1 and n_2. Setting them to maximize the multiplicative gap between $ess(f)$ and $cnf\_size(f)$, as a function of the number of variables n, yields a gap of size $\Theta(\log n)$.

We also construct a Horn function f such that there is a multiplicative gap of size $\Theta(\sqrt{n})$ between $cnf\_size(f)$ and $ess(f)$. We show that no gap larger than $\Theta(n)$ is possible.

If one expresses the gaps as a function of $cnf\_size(f)$, rather than as a function of the number of variables n, then the gap we obtain with both the constructed non-Horn and Horn functions f is $cnf\_size(f)^{1/3}$. Clearly, no gap larger than $cnf\_size(f)$ is possible.

This is an extended abstract. Additional material can be found in a full version of this paper (Hellerstein and Kletenik 2011).

Preliminaries

Definitions

A variable x_i and its negation $\neg x_i$ are *literals* (positive and negative respectively). A *clause* is a disjunction ($\vee$) of literals. A *term* is a conjunction ($\wedge$) of literals. A *CNF* formula is a formula of the form $c_0 \wedge c_1 \wedge \ldots c_k$, where each c_i is a clause. A *DNF* formula is a formula of the form $t_0 \vee t_1 \vee \ldots t_k$, where each t_i is a term.

A clause c containing variables from $X_n = \{x_1, \ldots, x_n\}$ is an *implicate* of f if for all $x \in \{0,1\}^n$, if c is falsified by x then $f(x) = 0$. A term t containing variables from X_n is an *implicant* of function $f(x_1, \ldots, x_n)$ if for all $x \in \{0,1\}^n$, if t is satisfied by x then $f(x) = 1$. An implicate (implicant) is *prime* if removing any literal from it would cause it to become a non-implicate (non-implicant).

We define the *size* of a CNF formula to be the number of its clauses, and the *size* of a DNF formula to be the number of its terms.

Given a Boolean function f, $cnf\_size(f)$ is the size of the smallest CNF formula representing f, and $dnf\_size(f)$ is the size of the smallest DNF formula representing f.

An assignment $x \in \{0,1\}^n$ is a *falsepoint* of f if $f(x) = 0$, and is a *truepoint* of f if $f(x) = 1$. We say that a clause c *covers* a falsepoint x of f if x falsifies c. A term t *covers* a truepoint x of f if x satisfies t.

A CNF formula ϕ representing a function f forms a *cover* of the falsepoints of f, in that each falsepoint of f must be covered by at least one clause of ϕ. Further, if x is a truepoint of f, then no clause of ϕ covers x. Similarly, a DNF formula ϕ representing a function f forms a *cover* of the truepoints of f, in that each truepoint of f must be covered by at least one term of ϕ. Further, if x is a falsepoint of f, then no term of ϕ covers x.

Given two assignments $x, y \in \{0,1\}^n$, we write $x \leq y$ if $\forall i, x_i \leq y_i$. An assignment r *separates* two assignments p and q if $\forall i, p_i = r_i$ or $q_i = r_i$.

A *partial* function f maps $\{0,1\}^n$ to $\{0,1,*\}$, where $*$ indicates that the value of f is undefined. A Boolean formula ϕ is *consistent* with a partial function f if $\phi(a) = f(a)$ for all $a \in \{0,1\}^n$ where $f(a) \neq *$. If f is a partial Boolean function, then $cnf\_size(f)$ and $dnf\_size(f)$ are the size of the smallest CNF and DNF formulas consistent with f, respectively.

A Boolean function $f(x_1, \ldots, x_n)$ is *monotone* if for all $x, y \in \{0,1\}^n$, if $x \leq y$ then $f(x) \leq f(y)$. Function f is *anti-monotone* if for all $x, y \in \{0,1\}^n$, if $x \geq y$ then $f(x) \leq f(y)$.

A DNF or CNF formula is *monotone* if it contains no negations. A CNF formula is a Horn-CNF if each clause contains at most one variable without a negation. If each clause contains exactly one variable without a negation it is a *pure* Horn-CNF. A *Horn function* is a Boolean function that can be represented by a Horn-CNF. It is a *pure Horn function* if it can be represented by a pure Horn-CNF. Horn functions are a generalization of anti-monotone functions, and have applications in artficial intelligence (see e.g. (Russell and Norvig 2003)).

We say that two falsepoints, x and y, of a function f are *independent* if no implicate of f covers both x and y. Similarly, we say that two truepoints x and y of a function f are *independent* if no implicant of f covers both x and y. We say that a set S of falsepoints (truepoints) of f is independent if all pairs of falsepoints (truepoints) in S are independent.

The *set covering problem* is as follows: Given a ground set $A = \{e_1, \ldots, e_m\}$ of elements, a set $\mathcal{S} = \{S_1, \ldots, S_n\}$ of subsets of A, and a positive integer k, does there exist $\mathcal{S}' \subseteq \mathcal{S}$ such that $\bigcup_{S_i \in \mathcal{S}'} = \mathcal{S}$ and $|\mathcal{S}'| \leq k$? Each set $S_i \in \mathcal{S}$ is said to cover the elements it contains. Thus the set covering problem asks whether A has a "cover" of size at most k.

A set covering instance is r-uniform, for some $r > 0$, if all subsets $S_i \in \mathcal{S}$ have size r.

Given an instance of the set covering problem, we say that a subset A' of ground set A is *independent* if no two elements of A' are contained in a common subset S_i of $\mathcal{S}$.

The quantity $ess(f)$

We begin by restating the definition of $ess(f)$ in terms of independent falsepoints. We also introduce an analogous quantity for truepoints. (The notation ess^d refers to the fact that this is a dual definition.)

Definition: Let f be a Boolean function. The quantity $ess(f)$ denotes the size of the largest independent set of falsepoints of f. The quantity $ess^d(f)$ denotes the largest independent set of truepoints of f.

The fact that $ess(f) \leq cnf\_size(f)$ follows easily from the above definitions, and from the following facts: (1) if ϕ is a CNF formula representing f, then every falsepoint of f must be covered by some clause of ϕ, and (2) each clause of ϕ must be an implicate of f.

Let f' denote the function that is the complement of f, i.e. $f'(a) = \neg f(a)$ for all assignments a. Since, by duality, $ess(f') = ess^d(f)$ and $cnf\_size(f') = dnf\_size(f)$, it follows that $ess(f') \leq dnf\_size(f)$. We use the following property (cf. (Čepek, Kučera, and Savický 2010)).

Property 1: Two falsepoints of f, x and y, are independent iff there exists a truepoint a of f that separates x and y.

Consider the following decision problem, which we will call *ESS*: "Given a CNF formula representing a Boolean function f, and a number k, is $ess(f) \leq k$?" Using Property 1, this problem is easily shown to be in co-NP (Čepek, Kučera, and Savický 2010). We can combine the fact that *ESS* is in co-NP with results on the hardness of approximating CNF-minimization, to get the following preliminary result, based on a complexity-theoretic assumption.

Proposition 1. *If co-NP $\neq \Sigma_2^P$, then for some $\gamma > 0$, there exists an infinite set of Boolean functions f such that $ess(f)n^\gamma < cnf\_size(f)$, where n is the number of variables of f.*

The proof of the proposition follows from a result of Umans, which states that it is Σ_2^P-complete to approximate the minimum-size CNF formula equivalent to a given CNF formula f to within a factor of n^γ. Here γ is a positive constant, and n is the number of variables of f (Umans 1999). We omit the details here.

The assumption that $\Sigma_2^P \neq$ co-NP is not unreasonable, so we have grounds to believe that there is an infinite set of functions for which the gap between $ess(f)$ and $cnf\_size(f)$ is greater than n^γ (or $cnf\_size(f)^\gamma$) for some γ. Below, we will explicitly construct such sets with larger gaps than that of Proposition 1, and with no complexity theoretic assumptions.

We can also prove a proposition similar to Proposition 1 for Horn functions, using a different complexity theoretic assumption. (Since the statement of the proposition includes a complexity class parameterized by the standard input-size parameter n, we use N instead of n to denote the number of inputs to a Boolean function.)

Proposition 2. *If NP $\not\subseteq$ co-NTIME($n^{polylog(n)}$), then for some ϵ such that $0 < \epsilon < 1$, there exists an infinite set of Horn functions f such that $\frac{cnf\_size(f)}{ess(f)} \geq 2^{\log^{1-\epsilon} N}$, where N is the number of input variables of f.*

The proof follows from a non-approximability result of Bhattacharya et al. (Bhattacharya et al. 2010) for the problem of minimizing Horn formulas.

Constructions of functions with large gaps between $ess(f)$ and $cnf\_size(f)$

We will begin by constructing a function f, such that $\frac{cnf\_size(f)}{ess(f)} = \Theta(n)$. This is already a larger gap than the multiplicative gap of $\log(n)$ achieved by the construction of (Čepek, Kučera, and Savický 2010), and the gap of n^γ in Proposition 1. We describe the construction of f, prove bounds on $cnf\_size(f)$ and $ess(f)$, and then prove that the ratio $\frac{cnf\_size(f)}{ess(f)} = \Theta(n)$.

We will then show how to modify this construction to give a function f such that $\frac{cnf\_size(f)}{ess(f)} = 2^{\Theta(n)}$, thus increasing the gap to be exponential in n.

Finally, in Section , we give our Horn function constructions.

Constructing a function with a linear gap

Theorem 1. *There exists a function $f(x_1, \ldots, x_n)$ such that $\frac{cnf\_size(f)}{ess(f)} = \Theta(n)$.*

Proof. We construct a function f such that $\frac{dnf\_size(f)}{ess^d(f)} = \Theta(n)$. Theorem 1 then follows immediately by duality.

Our construction relies heavily on a reduction of Gimpel from the 1960's (Gimpel 1965), which reduces a generic instance of the set covering problem to a DNF-minimization problem. See (Czort 1999) or (Allender et al. 2008) for more recent discussions of this reduction.

Gimpel's reduction is as follows. Let $A = \{e_1, \ldots, e_m\}$ be the ground set of the set covering instance, and let $\mathcal{S}$ be the set of subsets A from which the cover must be formed. With each element e_i in A, associate a Boolean input variable x_i. For each $S \in \mathcal{S}$, let x_S denote the assignment in $\{0,1\}^m$ where $x_i = 0$ iff $e_i \in S$. Define the partial function $f(x_1, \ldots, x_m)$ as follows:

$$f(x) = \begin{cases} 1 & \text{if } x \text{ contains exactly } m-1 \text{ ones} \\ * & \text{if } x \geq x_S \text{ for some } S \in \mathcal{S} \\ 0 & \text{otherwise} \end{cases}$$

There is a DNF formula of size at most k that is consistent with this partial function if and only if the elements e_i of the set covering instance A can be covered using at most k subsets in $\mathcal{S}$ (cf. (Czort 1999)).

We apply this reduction to the simple, 2-uniform, set covering instance over m elements where $\mathcal{S}$ consists of all subsets containing exactly two of those m elements. The smallest set cover for this instance is clearly $\lceil m/2 \rceil$. The largest independent set of elements is only of size 1, since every pair of elements is contained in a common subset of $\mathcal{S}$. Note that this gives a ratio of minimal set cover to largest independent set of $\Theta(m)$.

Applying Gimpel's reduction to this simple set covering instance, we get the following partial function $\hat{f}$:

$$\hat{f}(x) = \begin{cases} 1 & \text{if } x \text{ contains exactly } m-1 \text{ ones} \\ * & \text{if } x \text{ contains exactly } m-2 \text{ ones} \\ * & \text{if } x \text{ contains exactly } m \text{ ones} \\ 0 & \text{otherwise} \end{cases}$$

Since the smallest set cover for the instance has size $\lceil m/2 \rceil$,

$$dnf\_size(\hat{f}) = \lceil m/2 \rceil.$$

Allender et al. extended the reduction of Gimpel by converting the partial function f to a total function g. The conversion is as follows:

Let $t = m + 1$ and let s be the number of $*$'s in $f(x)$. Let y_1 and y_2 be two additional Boolean variables, and let $z = z_1 \ldots z_t$ be a vector of t more Boolean variables. Let $S \subseteq \{0,1\}^t$ be a collection of s vectors, each containing an odd number of 1's (since $s \leq 2^m$, such a collection exists). Let χ be the function such that $\chi(x) = 0$ if the parity of x is even and $\chi(x) = 1$ otherwise.

The total function g is defined as follows:

$$g(x, y_1, y_2, z) = \begin{cases} 1 & \text{if } f(x) = 1 \text{ and } y_1 = y_2 = 1 \text{ and } z \in S \\ 1 & \text{if } f(x) = * \text{ and } y_1 = y_2 = 1 \\ 1 & \text{if } f(x) = *, y_1 = \chi(x), \text{ and } y_2 = \neg\chi(x) \\ 0 & \text{otherwise} \end{cases}$$

Allender et al. proved that this total function g obeys the following property:

$$dnf\_size(g) = s(dnf\_size(f) + 1).$$

Let $\hat{g}$ be the total function obtained by setting $f = \hat{f}$ in the above definition of g.

We can now compute $dnf\_size(\hat{g})$. Let n be the number of input variables of $\hat{f}$. The total function $\hat{g}$ is defined on $n = 2m + 3$ variables. Since $dnf\_size(\hat{f}) = \lceil m/2 \rceil$, we have

$$dnf\_size(\hat{g}) = s\left(\lceil \frac{m}{2} \rceil + 1\right) \geq s\left(\frac{n-3}{4} + 1\right)$$

where s is the number of assignments x for which $\hat{f}(x) = *$.

We will upper bound $ess^d(\hat{g})$ by dividing the truepoints of $\hat{g}$ into two disjoint sets and upper-bounding the size of a maximum independent set of truepoints in each. (Recall that two truepoints of $\hat{g}$ are independent if they do not satisfy a common implicant of $\hat{g}$.)

Set 1: The set of all truepoints of $\hat{g}$ whose x component has the property $f(x) = *$.
Let a_1 be a maximum independent set of truepoints of $\hat{g}$ consisting only of points in this set. Consider two truepoints p and q in this set that have the same x value. It follows that they share the same values for y_1 and y_2. Let t be the term containing all variables x_i, and exactly one of the two y_j variables, such that each x_i appears without negation if it set to 1 by p and q, and with negation otherwise, and y_j is set to 1 by both p and q. Clearly, t is an implicant of $\hat{g}$ by definiton of $\hat{g}$, and clearly t covers both p and q. It follows that p and q are not independent.
Because any two truepoints in this set with the same x value are not independent, $|a_1|$ cannot exceed the number of different x assignments. There are s assignments such that $\hat{f}(x) = *$, so $|a_1| \leq s$.

Set 2: The set of all truepoints of $\hat{g}$ whose x component has the property $\hat{f}(x) = 1$.
Let a_2 be a maximum independent set consisting only of points in this set. Consider any two truepoints p and q in this set that contain the same assignment for z. We can construct a term t of the form $wy_1y_2\tilde{z}$ such that w contains exactly $m - 2$ x_i's that are set to 1 by both p and q, and

all z_is that are set to 1 by p and q appear in $\widetilde{z}$ without negation, and all other z_is appear with negation. It is clear that t is an implicant of $\hat{g}$ and that t covers both p and q. Once again, it follows that p and q are not independent truepoints of g.

Because any two truepoints in this set with the same z value are not independent, $|a_2|$ cannot exceed the number of different z assignments. There are s assignments to z such that $z \in S$, so $|a_2| \leq s$.

Since a maximum independent set of truepoints of $\hat{g}$ can be partitioned into an independent set of points from the first set, and an independent set of points from the second set, it immediately follows that [3]

$$ess^d(\hat{g}) \leq |a_1| + |a_2| \leq s + s = 2s.$$

Hence, the ratio between the DNF size and $ess(g)$ size is:

$$\frac{s(\frac{n-3}{4}+1)}{2s} \geq \frac{n+1}{8} = \Theta(n)$$

□

Note that the above construction gives a class of functions satisfying the conditions of Proposition 1, for $\gamma = 1$. The construction also yields the following corollary.

Corollary 1. *There exists a function f such that $\frac{cnf\_size(f)}{ess(f)} \geq cnf\_size(f)^{\epsilon}$ for an $\epsilon \geq 0$.*

Constructing a function with an exponential gap

Theorem 2. *There exists a function f on n variables such that $\frac{cnf\_size(f)}{ess(f)} \geq 2^{\Theta(n)}$.*

Proof. As before, we will reduce a set covering instance to a DNF-minimization problem involving a partial Boolean function f. However, here we will rely on a more general version of Gimpel's reduction, due to Allender et al., described in the following lemma.

Lemma 1. *(Allender et al. 2008) Let $\mathcal{S} = \{S_1, \ldots, S_p\}$ be a set of subsets of ground set $A = \{e_1, \ldots, e_m\}$. Let $t > 0$ and let $V = \{v^i : i \in \{1, \ldots, m\}\}$ and $W = \{w^j : j \in \{1, \ldots, p\}\}$ be sets of vectors from $\{0,1\}^t$ such that for all $j \in \{1, \ldots, p\}$ and $i \in \{1, \ldots, m\}$,*

$$e_i \in S_j \text{ iff } v^i \geq w^j$$

Let $f : \{0,1\}^t \rightarrow \{0,1,*\}$ be the partial function such that

$$f(x) = \begin{cases} 1 & \text{if } x \in V \\ * & \text{if } x \geq w \text{ for some } w \in W \text{ and } x \notin V \\ 0 & \text{otherwise} \end{cases}$$

Then $\mathcal{S}$ has a minimum cover of size k iff $dnf\_size(f) = k$.

[3] It can actually be proved that in fact, $ess^d(\hat{g}) = 2s$, but details of this proof are omitted.

(Note that the construction in the above lemma is equivalent to Gimpel's if we take $t = m$, $V = \{v \in \{0,1\}^m | v$ contains exactly $m-1$ 1's $\}$, and $W = \{x_S | S \in \mathcal{S}\}$, where x_S denotes the assignment in $\{0,1\}^m$ where $x_i = 0$ iff $e_i \in S$.)

As before, we use the simple 2-uniform set covering instance over m elements where $\mathcal{S}$ consists of all subsets of two of those elements. The next step is to construct sets V and W satisfying the properties in the above lemma for this set covering instance. To do this, we use a randomized construction of Allender et al. that generates sets V and W from an r-uniform set-covering instance, for any $r > 0$. This randomized construction appears in the appendix of (Allender et al. 2008), and is described in the following lemma.

Lemma 2. *Let $r > 0$ and let $\mathcal{S} = \{S_1, \ldots, S_p\}$ be a set of subsets of $\{e_1, \ldots, e_m\}$, where each S_i contains exactly r elements. Let $t \geq 3r(1+\ln(pm))$. Let $V = \{v^1, \ldots, v^m\}$ be a set of m vectors of length t, where each $v^i \in V$ is produced by randomly and independently setting each bit of v^i to 0 with probability $1/r$. Let $W = \{w^1, \ldots, w^p\}$, where each w^j = the bitwise AND of all v^i such that $e_i \in S_j$. Then, the following holds with probability greater than 1/2: For all $j \in \{1, \ldots, p\}$ and $i \in \{1, \ldots, m\}$, $e_i \in S_j$ iff $v^i \geq w^j$.*

By Lemma 2, there exist sets V and W, each consisting of vectors of length $6(1 + \ln(m^2(m-2)/2)) = O(\log m)$, satisfying the conditions of Lemma 1 for our simple 2-uniform set covering instance. Let $\tilde{f}$ be the partial function on $O(\log m)$ variables obtained by using these V and W in the definition of f in Lemma 1,

The DNF-size of $\tilde{f}$ is the size of the smallest set cover, which is $\lceil m/2 \rceil$, and the number of variables $n = \Theta(\log m)$; hence the DNF size is $2^{\Theta(n)}$.

We can convert the partial function $\tilde{f}(x)$ to a total function $\tilde{g}(x)$ just as done in the previous section. The arguments regarding DNF-size and $ess^d(\tilde{g})$ remain the same. Hence, the DNF-size is now $s\left(2^{\Theta(n)} + 1\right)$, and $ess^d(\tilde{g})$ is again at most $2s$.

The ratio between the DNF-size and $ess^d(\tilde{g})$ is therefore at least $2^{\Theta(n)}$. Once again, the CNF result follows. □

Size of the gap for Horn Functions

Because Horn-CNFs contain at most one unnegated variable per clause, they can be expressed as implications; eg. $\bar{a} \vee b$ is equivalent to $a \rightarrow b$. Moreover, a conjunction of several clauses that have the same antecedent can be represented as a single *meta-clause*, where the antecedent is the antecedent common to all the clauses and the consequent is comprised of a conjunction of all the consequents, eg. $(a \rightarrow b) \wedge (a \rightarrow c)$ can be represented as $a \rightarrow (b \wedge c)$.

Bounds on the ratio between $cnf\_size(f)$ and $ess(f)$

Angluin, Frazier and Pitt (Angluin, Frazier, and Pitt 1992) presented an algorithm (henceforth:the AFP algorithm) to learn Horn-CNFs, where the output is a series of meta-clauses. It can be proven (Arias and Balcázar 2008; 2011) that the output of the algorithm is of minimum implication size (henceforth: $min\_imp(f)$) – that is, it contains the fewest number of meta-clauses needed to represent function f. Each meta-clause can be a conjunction of at most n clauses; hence, each implication is equivalent to the conjunction of at most n clauses. Therefore,

$$cnf\_size(f) \leq n \times min\_imp(f).$$

The learning algorithm maintains a list of negative and positive examples (falsepoints and truepoints of the Horn function, respectively), containing at most $min\_imp(f)$ examples of each.

Lemma 3. *The set of negative examples maintained by the AFP algorithm is an independent set.*

Proof. The proof of this lemma relies heavily on (Arias and Balcázar 2008); see that paper for further details.

Let us consider any two negative examples, n_i and n_j, maintained by the algorithm. There are two possibilities:

1. $n_i \leq n_j$ or $n_j \leq n_i$. (These two examples are comparable points; one is below the other on the Boolean lattice.)
2. n_i and n_j are incomparable points (Neither is below the other on the lattice).

Let us consider the first type of points: Without loss of generality, assume that $n_i \leq n_j$. Arias et al. define a positive example n_i^* for each negative example n_i. This example n_i^* has several unique properties; amongst them, that $n_i < n_i^*$ for all negative examples n_i (Section 3 in (Arias and Balcázar 2008)). They further prove (Lemma 6 in (Arias and Balcázar 2008)) that if $n_i \leq n_j$, then $n_i^* \leq n_j$ as well. Hence, any attempt to falsify both falsepoints, n_i and n_j, with a common implicate of the Horn function would falsify the positive example (n_i^*) that lies between them as well. Therefore, these two points are independent.

Now let us assume that n_i and n_j are incomparable. Any implicate that falsifies both points is composed of variables on which the two points agree. Clearly, this implicate would likewise cover a point that is the componentwise intersection of n_i and n_j. However, Arias et al. prove (Lemma 7 in (Arias and Balcázar 2008)) that $n_i \wedge n_j$ is a positive point if n_i and n_j are incomparable. Hence, any implicate that falsifies both n_i and n_j would likewise falsify the truepoint $n_i \wedge n_j$ that lies between them. Therefore, these two points cannot be falsified by the same implicate and they are independent. □

Theorem 3. *For any Horn function* f, $\frac{cnf\_size(f)}{ess(f)} \leq n$

Proof. For any Horn function f, there exists a set of negative examples of size at most $min\_imp(f)$, and these examples are all independent. Hence, $ess(f) \geq min\_imp(f)$. We have already stated that $min\_imp(f)$ is at most a factor of n times larger than the minimum CNF size for this function.

Hence, $cnf\_size(f) \leq n \times ess(f)$.

Moreover, since Lemma 3 holds for general Horn functions in addition to pure Horn (Arias and Balcázar 2011), this bound holds for all Horn functions. □

Constructing a Horn function with a large gap between $ess(f)$ and $cnf\_size(f)$

Theorem 4. *There exists a definite Horn function* f *on* n *variables such that* $\frac{cnf\_size(f)}{ess(f)} \geq \Theta(\sqrt{n})$.

To prove this theorem, we construct f by embedding the 2-uniform set-covering instance consisting of all subsets of two elements into a definite Horn function. The construction uses techniques of (Crama and Hammer 2011), with modifications based on (Bhattacharya et al. 2010). Details are in the full version of the paper.

We earlier posited that if $\Sigma_p^2 \neq$ *co-NP*, there exists an infinite set of functions for which $\frac{cnf\_size(f)}{ess(f)} \geq cnf\_size(f)^\gamma$ for some $\gamma > 0$. The construction in the proof of the previous theorem yields a stronger result:

Theorem 5. *There exists an infinite set of Horn functions* f *for which* $\frac{cnf\_size(f)}{ess(f)} \geq cnf\_size(f)^\gamma$.

Acknowledgements

This work was partially supported by the US Department of Education GAANN grant P200A090157, and by NSF Grant CCF-0917153.

References

Allender, E.; Hellerstein, L.; McCabe, P.; Pitassi, T.; and Saks, M. E. 2008. Minimizing disjunctive normal form formulas and AC^0 circuits given a truth table. *SIAM J. Comput.* 38(1):63–84.

Angluin, D.; Frazier, M.; and Pitt, L. 1992. Learning conjunctions of horn clauses. *Machine Learning* 9:147–164.

Arias, M., and Balcázar, J. L. 2008. Query learning and certificates in lattices. In Freund, Y.; Györfi, L.; Turán, G.; and Zeugmann, T., eds., *Algorithmic Learning Theory*, volume 5254 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg. 303–315.

Arias, M., and Balcázar, J. 2011. Construction and learnability of canonical horn formulas. *Machine Learning* 1–25. 10.1007/s10994-011-5248-5.

Bhattacharya, A.; DasGupta, B.; Mubayi, D.; and Turán, G. 2010. On approximate horn formula minimization. In Abramsky, S.; Gavoille, C.; Kirchner, C.; Meyer auf der Heide, F.; and Spirakis, P., eds., *Automata, Languages and Programming*, volume 6198 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg. 438–450.

Coudert, O. 1994. Two-level logic minimization: an overview. *Integration, the VLSI Journal*.

Crama, Y., and Hammer, P. L., eds. 2011. *Boolean Functions: Theory, Algorithms, and Applications*. Cambridge University Press.

Czort, S. L. A. 1999. The complexity of minimizing disjunctive normal form formulas. Master's thesis, University of Aarhus, Aarhus, Denmark.

Gimpel, J. 1965. Method of producing a boolean function having an arbitrarily presribed prime impliant table. *IEEE Trans. Computers*.

Hellerstein, L., and Kletenik, D. 2011. On the gap between ess(f) and cnf_size(f). *CoRR* abs/1106.4247.

Russell, S. J., and Norvig, P. 2003. *Artificial Intelligence: A Modern Approach*. Pearson Education.

Čepek, O.; Kučera, P.; and Savický, P. 2010. Boolean functions with a simple certificate for CNF complexity. Technical report, Rutgers Center for Operations Research.

Umans, C. 1999. Hardness of approximating Σ_2^p minimization problems. In *Proc. IEEE Symposium on Foundations of Computer Science*, 465–474.