Slide 1

Counting problems in the hypercube

Guillermo Morales

Centro de Investigación y Estudios Avanzados del IPN

gmorales@cs.cinvestav.mx

November 14, 2000

CONTENTS

Introduction
Preliminaries
Some counting problems
Pairs of clauses with a common literal
k-representability
Final remarks

Slide 2

Abstract

The finite-dimensional hypercube has served as model of several processes in Electrical Engineering: from the most basic of circuit design till prototypical NP-complete problems, needless to mention its importance regarding communicactions and multi-processor computing. Any set in the hypercube can be realized as the set of satisfying assignments of a Boolean form. However, if we consider just conjunctive forms with clauses of a fixed length, the satisfying assignments form a *simple* set, i.e. a set which is the union of subspaces of a given dimension. In our approach we look towards the calculation of the cardinalities of simple sets. In the most general case, we have posed recurrence relations whose solutions are the simple sets and the sets in the hypercube that contain no a subspase of a given dimension.

Keywords: Complexity, optimization, hypercube geometry.

Slide 3

Resumen

El hipercubo de dimensión finita se ha presentado como modelo de diversos procesos: desde los más elementales tocantes al diseño de circuitos, hasta la formulación de problemas completos-NP, pasando por el anĺisis y el diseño de programas y dispositivos de cómputo en paralelo o distribuído. Cualquier conjunto en el hipercubo puede ser representado como el conjunto de asignaciones satisfactoras de alguna forma booleana. Si sólo se considera formas conjuntivas con cláusulas de longitud fija, las asignaciones satisfactoras son conjuntos *simples*, i.e., uniones de subespacios de una misma dimensión. Planteamos el problema de calcular cardinalidades de conjuntos simples. Planteamos ecuaciones de recurrencia cuyas soluciones son las cardinalidades buscadas y las cardinalidade de conjuntos en el hipercubo que no contienen a espacios de una dimensión determinada.

Palabras claves: Teoría de la complejidad, optimización, geometría del hipercubo.

Slide 4

Introduction

Satisfiability problem (SAT): To decide whether any given Boolean form is satisfiable.

#SAT: To calculate the number of satisfying assignments of an input Boolean form.

For any Boolean form $\phi(\mathbf{X})$ the set of satisfying assignments is $\phi^{-1}(1)$ and the set of falsifying assignments is $\phi^{-1}(0)$.

SAT decides whether $\phi^{-1}(1) \neq \emptyset$.

#SAT asks for the value of card$\phi^{-1}(1)$.

We pose in this paper the problem to characterize the numbers that can appear as solutions of #SAT restricted to conjunctive forms with an uniform clause length. We initiate our analysis by posing recurrent relations whose solutions are the sought values.

Slide 5

Preliminaries

Two $= \{0,1\}$: Boolean truth-values.

For any $n \geq 1$, Two^n: *n-dimensional hypercube.*

$\mathbf{X} = [X_j]_{j=1,\ldots,n}$: set of *variables*:

Literals L literal if $\exists X \in \mathbf{X}: \ (L = X) \vee (L = \neg X)$.

Phrases Literals conjunctions: $\forall \boldsymbol{\epsilon} \in \text{Three}^n$, $(\text{Three} = \text{Two} \cup \{\hat{}\})$,

$$\text{Phrase}(\boldsymbol{\epsilon}) = \bigwedge_{1 \leq j \leq n} X_j^{\epsilon_j}$$

$$\forall j = 1, \ldots, n: \ X_j^{\epsilon_j} = \begin{cases} X_j & \text{if } \epsilon_j = 1 \\ 0 & \text{if } \epsilon_j = \hat{} \\ \neg X_j & \text{if } \epsilon_j = 0 \end{cases}$$

Slide 6

$$\epsilon_j = \begin{cases} 1 & \text{if } X_j \text{ appears in Phrase}(\boldsymbol{\epsilon}), \\ 0 & \text{if } \neg X_j \text{ appears in Phrase}(\boldsymbol{\epsilon}), \\ \hat{} & \text{if neither } X_j \text{ nor } \neg X_j \text{ appear in Phrase}(\boldsymbol{\epsilon}). \end{cases}$$

$$\forall \boldsymbol{\delta} \in Two^n: \ \text{Phrase}(\boldsymbol{\epsilon})(\boldsymbol{\delta}) = 1 \ \Leftrightarrow \ \boldsymbol{\epsilon} = \boldsymbol{\delta} \qquad (1)$$

Disjunctive forms Phrases disjunctions. $DF = [L_{ij}]_{i,j}^D$ if

$$DF = \bigvee_{1 \leq i \leq m} Ph_i = \bigvee_{1 \leq i \leq m} \bigwedge_{1 \leq j \leq n} L_{ij}.$$

A DF is True if at least one of its phrases is True.

The DF is False if each of its phrases is False. In symbols:

$$\begin{aligned} DF^{-1}(1) &= \bigvee_{i=1}^{m} Ph_i^{-1}(1) \\ DF^{-1}(0) &= \bigwedge_{i=1}^{m} Ph_i^{-1}(0) \end{aligned} \qquad (2)$$

Slide 7

Clauses Literal disjunctions: $\forall \boldsymbol{\epsilon} \in \mathrm{Three}^n$, let

$$\mathrm{Clause}(\boldsymbol{\epsilon}) = \bigvee_{1 \leq j \leq n} \overline{X_j^{\epsilon_j}}.$$

$$\epsilon_j = \begin{cases} 0 & \text{if } X_j \text{ appears in Clause}(\boldsymbol{\epsilon}), \\ 1 & \text{if } \neg X_j \text{ appears in Clause}(\boldsymbol{\epsilon}), \\ \hat{} & \text{if neither } X_j \text{ nor } \neg X_j \text{ appear in Clause}(\boldsymbol{\epsilon}). \end{cases}$$

$$\forall \boldsymbol{\delta} \in \mathit{Two}^n : \ \mathrm{Clause}(\boldsymbol{\epsilon})(\boldsymbol{\delta}) = 0 \ \Leftrightarrow \ \boldsymbol{\epsilon} = \boldsymbol{\delta} \tag{3}$$

Slide 8

Conjunctive forms Clauses conjunctions. $CF = [L_{ij}]_{i,j}^C$ if

$$CF = \bigwedge_{1 \leq i \leq m} C_i = \bigwedge_{1 \leq i \leq m} \bigvee_{1 \leq j \leq n} L_{ij}.$$

A CF is False if at least one of its clauses is False.

The CF is True if each of its clauses is True.

$$\begin{aligned} CF^{-1}(0) &= \bigvee_{i=1}^{m} C_i^{-1}(0) \\ CF^{-1}(1) &= \bigwedge_{i=1}^{m} C_i^{-1}(1) \end{aligned} \tag{4}$$

Boolean forms Any conjunctive or disjunctive form.

Each Boolean form F determines a function $F : \mathit{Two}^n \to \mathrm{Two}$. A Boolean form F is *satisfiable* if $F \neq \mathbf{0}$, *tautological* if $F = \mathbf{1}$, *refutable* if $F \neq \mathbf{1}$ and a *contradiction* if $F = \mathbf{0}$.

Slide 9

Given $A \subset Two^n$, the *disjunctive form associated to* A is

$$DF_A(\mathbf{X}) = \bigvee_{\boldsymbol{\epsilon} \in A} \text{Phrase}(\boldsymbol{\epsilon})(\mathbf{X}). \tag{5}$$

Clearly, $A = DF_A^{-1}(1)$.

Similarly, the *conjunctive form associated to* A is

$$CF_A(\mathbf{X}) = \bigwedge_{\boldsymbol{\epsilon} \in A} \text{Clause}(\boldsymbol{\epsilon})(\mathbf{X}). \tag{6}$$

Clearly, $A = CF_A^{-1}(0)$.

The following are classical problems:

Problem SAT

Instance: A CF F.

Solution: $\begin{cases} \mathit{Yes} & \text{if } F \text{ is satisfiable,} \\ \mathit{No} & \text{otherwise.} \end{cases}$

Slide 10

Problem #SAT

Instance: A CF F.

Solution: $\text{card}\{\boldsymbol{\epsilon} \in Two^n | F(\boldsymbol{\epsilon}) = 1\}$

Problem MaxSAT

Instance: A CF F.

Solution: An assignment $\boldsymbol{\epsilon}_0 \in Two^n$ such that $\forall \boldsymbol{\epsilon} \in Two^n$:

$$\text{card}\{c \in F | c(\boldsymbol{\epsilon}_0) = 1\} \geq \text{card}\{c \in F | c(\boldsymbol{\epsilon}) = 1\}.$$

Problem co-SAT

Instance: A DF F.

Solution: $\begin{cases} \mathit{Yes} & \text{if } F \text{ is a tautology,} \\ \mathit{No} & \text{otherwise.} \end{cases}$

Slide 11

For $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n) \in \mathrm{Three}^n$ the *instantiated variables by* $\boldsymbol{\epsilon}$ are:

$$\mathrm{Def}(\boldsymbol{\epsilon}) = [\![1, n]\!] - \boldsymbol{\epsilon}^{-1}(\hat{}) = \{i \in [\![1, n]\!] | \epsilon_i \in \mathrm{Two}\}$$

If $k = \mathrm{card}\ \mathrm{Def}(\boldsymbol{\epsilon})$ then $\boldsymbol{\epsilon}$ determines a hypercube

$$H(\boldsymbol{\epsilon}) = \{\boldsymbol{\delta} \in \mathit{Two}^n | \forall i \in \mathrm{Def}(\boldsymbol{\epsilon})\, [\delta_i = \epsilon_1]\},$$

and it is a subspace whose dimension is $n - k$. In this case, $\boldsymbol{\epsilon} \in \mathrm{Three}^n$ is a *k-assignment*. Evidently

$$\mathrm{Phrase}(\boldsymbol{\epsilon})^{-1}(1) = H(\boldsymbol{\epsilon}) = \mathrm{Clause}(\boldsymbol{\epsilon})^{-1}(0).$$

A *k-phrase* (resp. *k-clause*) is of the form $\mathrm{Phrase}(\boldsymbol{\epsilon})$ (resp. $\mathrm{Clause}(\boldsymbol{\epsilon})$), where $\boldsymbol{\epsilon}$ is a k-assignment. In Circuit Design, a k-phrase is called also a *k-miniterm*.

A k-DF (resp.k-CF), $k \geq 1$, is a DF (resp. CF) consisting just of k-phrases (resp. k-clauses).

Slide 12

kSAT, #kSAT, Max$k$SAT and k-co-SAT are the corresponding problems restricted to k-forms.

Thus the sets of assignments satisfying a k-DF (resp. falsifying a k-CF) is the union of $(n-k)$-dimensional subspaces.

In this paper we will deal with the problem to count the number of sets that can be expressed as the union of k_1-dimensional subspaces, for any given $k_1 \in [\![1, n]\!]$, or, stated equivalently, are the satisfying sets of k_2-DF's, or falsifying sets of k_2-CF's, with $k_1 + k_2 = n$.

Slide 13

Some counting problems

n: number of variables,

$k \leq n$: clauses length, and

$C_n^{(k)}$: set of k-clauses over n variables.

There are $c_{nk} = \binom{n}{k} 2^k$ elements in $C_n^{(k)}$.

Pairs of clauses with common literals

Problem 1 *Given $l \in [\![0, k]\!]$, calculate the number p_{lkn} of pairs in $\left(C_n^{(k)}\right)^{(2)}$ whose components share exactly l literals.* In other words: *Count the pairs of clauses sharing exactly l literals.*

Slide 14

Let us consider first the case where $k \in [\![1, \frac{n}{2}[\![$.

If $\{\mathbf{c}_1, \mathbf{c}_2\} \in \left(C_n^{(k)}\right)^{(2)}$ is such a pair let us denote by $\mathbf{d}$ the list of the common literals. Then we may write

$$\begin{aligned} \mathbf{c}_1 &= \mathbf{d} \vee \mathbf{c}_1' \\ \mathbf{c}_2 &= \mathbf{d} \vee \mathbf{c}_2' \end{aligned}$$

where $\mathbf{c}_1'$ and $\mathbf{c}_2'$ are $(k-l)$-clauses. Clearly,

1. there are $\binom{n}{l} 2^l$ possibilities to select the common l-clause $\mathbf{d}$,
2. there are $\binom{n-l}{k-l} 2^{k-l}$ possibilities to select the first remaining $(k-l)$-clause $\mathbf{c}_1'$,
3. there are $\binom{n-k}{k-l} 2^{k-l}$ possibilities to select the second remaining $(k-l)$-clause $\mathbf{c}_2'$, and
4. for $l < k$ each pair of the remaining $(k-l)$-clauses $\{\mathbf{c}_1', \mathbf{c}_2'\}$ is counted exactly twice by the above counting process.

Slide 15

Thus, let

$$\begin{aligned} q_{lkn} &= \binom{n}{l}2^l \cdot \binom{n-l}{k-l}2^{k-l} \cdot \binom{n-k}{k-l}2^{k-l} \\ &= \binom{n}{l}\binom{n-l}{k-l}\binom{n-k}{k-l}2^{2k-l} \end{aligned}$$

then

$$p_{lkn} = \begin{cases} \frac{1}{2}q_{lkn} & \text{if } l < k \\ q_{kkn} & \text{if } l = k \end{cases}$$

indeed, $q_{kkn} = c_{nk}$.

In next tables we present some values of q_{lkn} with $0 \leq q \leq k$ and $1 \leq k < \frac{n}{2}$.

Slide 16

$n = 2$:

$k\backslash l$	0	1
1	4	4

$n = 3$:

$k\backslash l$	0	1
1	12	6

$n = 4$:

$k\backslash l$	0	1	2
1	24	8	
2	48	96	24

$n = 5$:

$k\backslash l$	0	1	2
1	40	10	
2	240	240	40

$n = 6$:

$k\backslash l$	0	1	2	3
1	60	12		
2	720	480	60	
3	640	2880	1440	160

$n = 7$:

$k\backslash l$	0	1	2	3
1	84	14		
2	1680	840	84	
3	4480	10080	3360	280

Slide 17

$n = 8$:

$k \backslash l$	0	1	2	3	4
1	112	16			
2	3360	1344	112		
3	17920	26880	6720	448	
4	8960	71680	80640	17920	1120

$n = 9$:

$k \backslash l$	0	1	2	3	4
1	144	18			
2	6048	2016	144		
3	53760	60480	12096	672	
4	80640	322560	241920	40320	2016

$n = 10$:

$k \backslash l$	0	1	2	3	4	5
1	180	20				
2	10080	2880	180			
3	134400	120960	20160	960		
4	403200	1075200	604800	80640	3360	
5	129024	1612800	3225600	1612800	201600	8064

Slide 18

Now, if $k \in [\![\frac{n}{2}, n]\!]$ then p_{lkn} may be computed as before whenever $n - k \geq k - l$. Thus, for the current case one must have

$$p_{lkn} = \begin{cases} 0 & \text{if } 0 \leq l < 2k - n \\ \frac{1}{2} q_{lkn} & \text{if } 2k - n \leq l < k \\ q_{kkn} & \text{if } l = k \end{cases}$$

Slide 19

k-representability

Basic geometrical notions

Let $0 \leq k \leq n$. Let us say that a set $A \subset \mathit{Two}^n$

- is *k-representable* if there is a collection $S_1, \ldots, S_m$ of k-subspaces such that $A = \bigcup_{i=1}^{m} S_i$, and
- is *k-thin* if no k-subspace is contained in A.

The following relations are self-evident:

1. The empty set $\emptyset$ is k-representable, since it is represented as an empty union of k-subspaces, and it is also k-thin.
2. The whole space Two^n is k-representable.
3. Any non-k-representable set is the disjoint union of a k-representable set and a non-empty k-thin set.

Slide 20

4. If A is k-representable then $\mathrm{card}(A) \geq 2^k$.
5. Since Two^n can be partitioned as the union of 2^{n-k} pairwise independent k-subspaces $S_0^{(k)}, \ldots, S_{2^{n-k}-1}^{(k)}$, any k-thin set A can be partitioned as a collection of 2^{n-k} pairwise disjoint sets, thus $\exists \{l_i\}_{i=1}^{2^{n-k}-1}$:

$$\left[0 \leq l_0 \leq l_1 \leq \cdots \leq l_{2^{n-k}-1} < 2^k\right] \;\&\; \left[\mathrm{card}(A) = \sum_{i=1}^{2^{n-k}-1} l_i\right]. \tag{7}$$

Slide 21

Let us define

$$\begin{aligned}
\mathcal{R}_{kn} &= \{A \subset Two^n | A \text{ is } k\text{-representable}\} \\
\rho_{kn} &= \mathrm{card}(\mathcal{R}_{kn}), \\
\mathcal{N}_{kn} &= \{A \subset Two^n | A \text{ is not } k\text{-representable}\} \\
\eta_{kn} &= \mathrm{card}(\mathcal{N}_{kn}), \\
\mathcal{T}_{kn} &= \{A \subset Two^n | A \text{ is } k\text{-thin}\} \\
\tau_{kn} &= \mathrm{card}(\mathcal{T}_{kn}).
\end{aligned}$$

For $k = 0$ the following relations hold:

$$\begin{aligned}
\mathcal{R}_{0n} &= \mathcal{P}(Two^n), \\
\mathcal{T}_{0n} &= \emptyset, \\
\mathcal{T}_{nn} &= \mathcal{P}(Two^n) - \{Two^n\},
\end{aligned}$$

Slide 22

thus

$$\begin{aligned}
\rho_{0n} &= 2^{2^n}, \\
\tau_{0n} &= 0, \\
\tau_{nn} &= 2^{2^n} - 1.
\end{aligned}$$

Besides, after 7 we get an estimation of τ_{kn}:

$$\tau_{kn} = \sum_{a=1}^{2^n-1} \sum_{\substack{l_0+\cdots+l_{2^{n-k}-1}=a \\ 0 \le l_i < 2^k}} \binom{2^k}{l_0} \cdots \binom{2^k}{l_{2^{n-k}-1}}. \qquad (8)$$

Slide 23

$$M_{akn} = \left\{ (l_0, \ldots, l_{2^{n-k}-1}) \in \mathbb{N}^{2^{n-k}} \mid \sum_{j=0}^{2^{n-k}-1} l_j = a \right\},$$

$$L_{akn} = \left\{ (l_0, \ldots, l_{2^{n-k}-1}) \in M_{akn} \mid \forall j \leq (2^{n-k}-1) : l_j < 2^k \right\}.$$

1. If $a > (2^k - 1) \cdot 2^{n-k}$ then $L_{akn} = \emptyset$. Thus 8 can be written as

$$\tau_{kn} = \sum_{a=1}^{(2^k-1)\cdot 2^{n-k}} \sum_{\substack{l_0+\cdots+l_{2^{n-k}-1}=a \\ 0 \leq l_i < 2^k}} \binom{2^k}{l_0} \cdots \binom{2^k}{l_{2^{n-k}-1}}.$$

2. For $k = 1$, $(l_0, \ldots, l_{2^{n-1}-1})$ is in L_{a1n} iff it is the characteristic function of a subset with a elements in $\{0, \ldots, 2^{n-1} - 1\}$. Thus, $\tau_{1n} = 3^{2^{n-1}} - 1$.
3. If $1 \leq a \leq (2^k - 1)$ then $L_{akn} = M_{akn}$.

Slide 24

4. For any a, k, n the set M_{akn} can be expressed as the disjoint union of

$$\binom{a-1}{s-1} : \begin{array}{l}\text{number of possibilities to express } a \text{ as the sum of } s \\ \text{non-nul summands,}\end{array}$$

sets, each consisting of

$$\binom{2^{n-k}}{s} : \begin{array}{l}\text{number of possibilities to put } s \text{ non-nul summands into a sum of } 2^{n-k} \text{ non-negative summands.}\end{array}$$

elements. s varies from 1 to a.

Slide 25

Orders: For any $A \subset Two^n$ let $o(A) = \mathrm{Min}\{k | A \in \mathcal{T}_{kn}\} - 1$. Let

$$\begin{aligned} \mathcal{U}_{kn} &= \{A \subset Two^n | o(A) = k\} \\ \omega_{kn} &= \mathrm{card}(\mathcal{U}_{kn}), \\ \forall l \leq 2^n : \mathcal{V}_{lkn} &= \{A \in \mathcal{U}_{kn} | \mathrm{card}(A) = l\} \\ \upsilon_{lkn} &= \mathrm{card}(\mathcal{V}_{lkn}). \end{aligned}$$

The following relations hold:

1. $\{\mathcal{V}_{lkn}\}_{1 \leq l \leq 2^n}$ is a partition of $\mathcal{U}_{kn}$, thus $\omega_{kn} = \sum_{i=1}^{2^n} \upsilon_{lkn}$.

Slide 26

2. For $k = 0$,

$\upsilon_{10n} = 2^n$: Any singleton has order 0.

$\upsilon_{2^{n-1}0n} = 2$: The only way to choose 2^{n-1} points that form no an edge is by selecting opposite points along the main diagonal on 2-dimensional cubes.

$l > 2^{n-1} \Rightarrow \upsilon_{l0n} = 0$: Any set with more than 2^{n-1} points has at least an edge.

Slide 27

3. For any k with $0 < k \leq n$,

$l < 2^k \Rightarrow \upsilon_{lkn} = 0$: Any set of order k should have, at least, 2^k points.

$\upsilon_{2^k kn} = 2^{n-k}\binom{n}{k}$: This is the number of k-dimensional cubes in the n-dimensional cube.

$l > 2^{n-k}(2^k - 1) \Rightarrow \upsilon_{lkn} = 0$: Any set with more than $2^{n-k}(2^k - 1)$ points has at least a k-dimensional cube.

4. For any $k < n$, $\upsilon_{2^{n-k}(2^k-1),k,n} = 2^k$.

Slide 28

Final remarks

In tis paper we have posed several problems related to counting representable sets as well as sets of a given order.

Just a few results are presented here. Indeed the whole calculation of the defined parameters has not been completed.

Such calculation may provide an approach to solving #SAT decision problems: Given a CF ϕ and a $N \in \mathbb{N}$ we could design a procedure to decide whether $N \leq M$, where M is the number of assignments falsifying ϕ, by looking at the geometric properties of this last set.

There is still a long way to follow-up in order to complete the calculation of the parameters introduced here.

Slide 29

References

[1] Creignou N., Hermann M., *Complexity of Generalized Satisfiability Counting Problems*, Information and Computation 125, 1-12, 1996.

[2] Dubois Olivier, *Counting the number of solutions for instances of satisfiability*, Theor. Comp. Sc. 81, 49-64, 1991.

[3] Kahn J., Linial N., Samorodnitsky A., *Inclusion-exclusion: exact and approximate*, Combinatorica 16(4), 465-477, 1996.

[4] Linial N., Nissan N., *Approximate inclusion-exclusion*, Combinatorica 10(4), 349-365, 1990.

[5] Lozinskii E.L., *Counting propositional models*, Information Processing Letters 41, 327-332, 1992.

[6] Toda S., *On the computational power of PP and* $\bigoplus P$, Proc. 30th IEEE Ann. Symp. on Foundations of Computer Science, 514-519, 1989.