# Finding Needles in Exponential Haystacks

Joel Spencer

Eurandom
January 6, 2014

Erdős Magic:

*If a random object has a positive probability of being good than a good object MUST exist*

Modern Erdős Magic:

*If a randomized algorithm has a positive probability of producing a good object than a good object MUST exist*

Working with Paul Erdős was like taking a walk in the hills. Every time when I thought that we had achieved our goal and deserved a rest, Paul pointed to the top of another hill and off we would go.
– Fan Chung

# PART I

The Lovász Local Lemma

# $k$-SAT

Boolean $x_1, \ldots, x_n$. $y_s$ is $x_s$ or $\overline{x_s}$

Clause $C = y_{i_1} \vee \ldots \vee y_{i_k}$

$k$-SAT instance: $\wedge_{\alpha \in I} C_\alpha$

$C_\alpha, C_\beta$ overlap if common $y_j$.

Assume: Each $C_\alpha$ overlaps $\leq d$ $C_\beta$

Assume: $ed2^{-k} \leq 1$

LLL: Then satisfiable.

# $k$-SAT

Fix $d, k$ (e.g.: $k = 5$, $d = 10$) but let $n \rightarrow \infty$

Where is the satisfying assignment?

Original Proof: Can't Find it

MOSER: I can find it!

# FIX-IT!

Moser's FIX-IT Algorithm

FIX-IT I Randomly assign $x_j \leftarrow \{t, f\}$

# FIX-IT!

Moser's FIX-IT Algorithm

FIX-IT I Randomly assign $x_j \leftarrow \{t, f\}$

FIX-IT II WHILE some clause $C_\alpha \leftarrow f$

# FIX-IT!

Moser's FIX-IT Algorithm

FIX-IT I Randomly assign $x_j \leftarrow \{t, f\}$

FIX-IT II WHILE some clause $C_\alpha \leftarrow f$

FIX-IT IIIa Select *one* bad clause $C_\alpha$

# FIX-IT!

Moser's FIX-IT Algorithm

FIX-IT I Randomly assign $x_j \leftarrow \{t, f\}$

FIX-IT II WHILE some clause $C_\alpha \leftarrow f$

FIX-IT IIIa Select *one* bad clause $C_\alpha$

FIX-IT IIIb Randomly Reassign $x_j \leftarrow \{t, f\}$ for all $x_j$ in $C_\alpha$

# The LOG

$LOG$ - Clauses reassigned in order FIX-IT IIIb applies

$TLOG$ = length of LOG. ($= \infty$ if no stop)

Modern Erdős Magic: $E[TLOG] < \infty$ implies satisfiable *and* a good [1] algorithm

Example: Variables 12345678. Clauses $A : 123$, $B : 234$, $C : 345$, $D : 456$, $E : 567$, $F : 678$. $LOG = ADCFECBF$

[1] with $E[TLOG]$ reasonable

# Let's Play Tetris!

$s = ADCFECBF$: ADCFECBF

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| - | - | - | - | - | - | - | - |
| - | - | - | - | - | - | - | - |
| - | - | - | - | - | - | - | - |
| - | - | - | - | - | - | - | - |
| - | - | - | - | - | - | - | - |
| A | A | A | - | - | - | - | - |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

# Let's Play Tetris!

$s = ADCFECBF$: ADCFECBF

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| - | - | - | - | - | - | - | - |
| - | - | - | - | - | - | - | - |
| - | - | - | - | - | - | - | - |
| - | - | - | - | - | - | - | - |
| - | - | - | - | - | - | - | - |
| A | A | A | D | D | D | - | - |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

# Let's Play Tetris!

$s = ADCFECBF$: ADCFECBF

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| - | - | - | - | - | - | - | - |
| - | - | - | - | - | - | - | - |
| - | - | - | - | - | - | - | - |
| - | - | - | - | - | - | - | - |
| - | - | C | C | C | - | - | - |
| A | A | A | D | D | D | - | - |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

# Let's Play Tetris!

$s = ADCFECBF$: ADCFECBF

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| - | - | - | - | - | - | - | - |
| - | - | - | - | - | - | - | - |
| - | - | - | - | - | - | - | - |
| - | - | - | - | - | - | - | - |
| - | - | C | C | C | F | F | F |
| A | A | A | D | D | D | - | - |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

# Let's Play Tetris!

$s = ADCFECBF$: ADCFECBF

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| - | - | - | - | - | - | - | - |
| - | - | - | - | - | - | - | - |
| - | - | - | - | - | - | - | - |
| - | - | - | - | E | E | E | - |
| - | - | C | C | C | F | F | F |
| A | A | A | D | D | D | - | - |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

# Let's Play Tetris!

$s = ADCFECBF$: ADCFECBF

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| - | - | - | - | - | - | - | - |
| - | - | - | - | - | - | - | - |
| - | - | C | C | C | - | - | - |
| - | - | - | - | E | E | E | - |
| - | - | C | C | C | F | F | F |
| A | A | A | D | D | D | - | - |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

# Let's Play Tetris!

$s = ADCFECBF$: ADCFECBF

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| - | - | - | - | - | - | - | - |
| - | B | B | B | - | - | - | - |
| - | - | C | C | C | - | - | - |
| - | - | - | - | E | E | E | - |
| - | - | C | C | C | F | F | F |
| A | A | A | D | D | D | - | - |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

# Let's Play Tetris!

$s = ADCFECBF$: ADCFECBF

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| - | - | - | - | - | F | F | F |
| - | B | B | B | - | - | - | - |
| - | - | C | C | C | - | - | - |
| - | - | - | - | E | E | E | - |
| - | - | C | C | C | F | F | F |
| A | A | A | D | D | D | - | - |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

# Let's Play Tetris!

$s = ADCFECBF$

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| - | - | - | - | - | F | F | F |
| - | B | B | B | - | - | - | - |
| - | - | C | C | C | - | - | - |
| - | - | - | - | E | E | E | - |
| - | - | C | C | C | F | F | F |
| A | A | A | D | D | D | - | - |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

The Pyramid $Pyr(s) = ADCFEF$ is support of last:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| - | - | - | - | - | F | F | F |
| - | - | - | - | E | E | E | - |
| - | - | C | C | C | F | F | F |
| A | A | A | D | D | D | - | - |

# Analysis

Pyramids of prefixes of LOG distinct.

$ADCFECBF$: $A$; $D$; $ADC$;
$DF$;$ADCFE$;$ADCFEC$;$ADCFECB$;$ADCFEF$

$E[TLOG] = \sum_s \Pr[s \text{ pyramid of prefix }]$

# Preprocess Randomness

Each $x_j$ chooses countably many assignments.

Probability $X_1 \cdots X_t$ is pyramid of prefix is $\leq (2^{-k})^t$

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| - | - | C | C | C | - | - | - |
| A | A | A | - | - | - | - | - |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

$A, C$ false with different coinflips!

# An Interesting Algebra

$X, Y$ commute if no overlap.
Tetris same if and only if equal in algebra
$ADC = DAC$

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| - | - | C | C | C | - | - | - |
| A | A | A | D | D | D | - | - |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

$ADC, DAC$ use same coinflips.

# Algebraic Combinatorics

Property KNUTH: $\sum_s (2^{-k})^{length(s)} < \infty$
(sum over pyramids $s$ in algebra)
$E[TLOG] \leq \sum_s (2^{-k})^{length(s)}$
Modern Erdős Magic: [KNUTH] implies satisfiability *and* FIX-IT takes "time" at most sum.
Algebraic Combinatorics: When does [KNUTH] hold?
Partial Answer: if $ed2^{-k} \leq 1$

# A Prescient Adversary

FIX-IT I Randomly assign $x_j \leftarrow \{t, f\}$

FIX-IT II WHILE some clause $C_\alpha \leftarrow f$

FIX-IT IIIa Select *one* bad clause $C_\alpha$

FIX-IT IIIb Randomly Reassign $x_j \leftarrow \{t, f\}$ for all $x_j$ in $C_\alpha$

Each $x_j$ selects countably many $t, f$.
Adversary knows coinflips in advance
Still can't stop FIX-IT from halting!

# PART II

Eliminating Outliers

# Six Standard Deviations Suffice

$S_1, \ldots, S_n \subseteq \{1, \ldots, n\}$

$\chi : \{1, \ldots, n\} \to \{-1 + 1\} = \{red, blue\}$

$\chi(S) := \sum_{j \in S} \chi(j)$, $\texttt{disc}(\texttt{S}) = |\chi(\texttt{S})| = |\texttt{red} - \texttt{blue}|$

**Theorem (JS/1985):** There *exists* $\chi$

$$\texttt{disc}(\texttt{S}_\texttt{i}) \leq 6\sqrt{\texttt{n}} \text{ for all } 1 \leq \texttt{i} \leq \texttt{n}$$

# Six Standard Deviations Suffice

$S_1, \ldots, S_n \subseteq \{1, \ldots, n\}$

$\chi : \{1, \ldots, n\} \to \{-1+1\} = \{red, blue\}$

$\chi(S) := \sum_{j \in S} \chi(j)$, $\mathtt{disc}(\mathtt{S}) = |\chi(\mathtt{S})| = |\mathtt{red} - \mathtt{blue}|$

**Theorem (JS/1985):** There *exists* $\chi$

$$\mathtt{disc}(\mathtt{S_i}) \leq 6\sqrt{\mathtt{n}} \text{ for all } 1 \leq \mathtt{i} \leq \mathtt{n}$$

**Conjecture (JS/1986-2009)** You can't find $\chi$ in polynomial time.

# Six Standard Deviations Suffice

$S_1, \ldots, S_n \subseteq \{1, \ldots, n\}$

$\chi : \{1, \ldots, n\} \to \{-1 + 1\} = \{\textit{red}, \textit{blue}\}$

$\chi(S) := \sum_{j \in S} \chi(j)$, $\texttt{disc}(\texttt{S}) = |\chi(\texttt{S})| = |\texttt{red} - \texttt{blue}|$

**Theorem (JS/1985):** There *exists* $\chi$

$$\texttt{disc}(\texttt{S}_\texttt{i}) \leq 6\sqrt{\texttt{n}} \text{ for all } 1 \leq \texttt{i} \leq \texttt{n}$$

**Conjecture (JS/1986-2009)** You can't find $\chi$ in polynomial time.

**Theorem (Bansal/2010):** Yes I can!

**Theorem (Lovett, Meka/2012):** We can too!

# A Vector Formulation

$\vec{r}_i \in R^n$, $1 \leq i \leq n$, $|\vec{r}_i|_\infty \leq 1$

Initial $\vec{z} \in [-1, +1]^n$ (e.g.: $\vec{z} = \vec{0}$.)

**Theorem:** There exists $\vec{x} \in \{-1, +1\}^n$ with

$$|\vec{r}_i \cdot (\vec{x} - \vec{z})| \leq K\sqrt{n}$$

for *all* $1 \leq i \leq n$.

# A Vector Formulation

$\vec{r_i} \in R^n$, $1 \le i \le n$, $|\vec{r_i}|_\infty \le 1$

Initial $\vec{z} \in [-1, +1]^n$ (e.g.: $\vec{z} = \vec{0}$.)

**Theorem:** There exists $\vec{x} \in \{-1, +1\}^n$ with

$$|\vec{r_i} \cdot (\vec{x} - \vec{z})| \le K\sqrt{n}$$

for *all* $1 \le i \le n$.

$\vec{z} = \vec{0}$, $\vec{x}$ random. Problem: OUTLIERS!

# Phase I

Find $\vec{x} \in [-1, +1]^n$ with all least $\frac{n}{2}$ at $\pm 1$.

Idea: Start $\vec{x} \leftarrow \vec{z}$. Move $\vec{x}$ in a Controlled Brownian Motion.

# Phase I

Find $\vec{x} \in [-1, +1]^n$ with all least $\frac{n}{2}$ at $\pm 1$.

Idea: Start $\vec{x} \leftarrow \vec{z}$. Move $\vec{x}$ in a Controlled Brownian Motion.

Technical: $i$ frozen if $|x_i| \geq 1 - \epsilon$. Each step of distance $\delta$

# Phase I

Find $\vec{x} \in [-1, +1]^n$ with all least $\frac{n}{2}$ at $\pm 1$.

Idea: Start $\vec{x} \leftarrow \vec{z}$. Move $\vec{x}$ in a Controlled Brownian Motion.

Technical: $i$ frozen if $|x_i| \geq 1 - \epsilon$. Each step of distance $\delta$

Set $L_j = [n^{-1/2}\vec{r_j}] \cdot [\vec{x} - \vec{z}]$

WANT: ALL $|L_j| \leq K$

Space $V$ of allowable moves $\vec{y} = (y_1, \ldots, y_n)$

Space $V$ of allowable moves $\vec{y} = (y_1, \ldots, y_n)$

$i$ frozen $\Rightarrow y_i = 0$.

Space $V$ of allowable moves $\vec{y} = (y_1, \ldots, y_n)$

$i$ frozen $\Rightarrow y_i = 0$.

$\vec{y}$ orthogonal to current $\vec{x}$

Space $V$ of allowable moves $\vec{y} = (y_1, \ldots, y_n)$

$i$ frozen $\Rightarrow y_i = 0$.

$\vec{y}$ orthogonal to current $\vec{x}$

**KEY:** $\vec{y}$ orthogonal to $\vec{r_j}$ for $j$ with top $\frac{n}{4}$ $|L_j|$

# The Random Move

$d = \dim(V) \geq \frac{n}{4} - 1 \sim \frac{n}{4}$.

Gaussian $\vec{g} = d^{-1/2}[n_1 \vec{b_1} + \ldots + n_d \vec{b_d}]$, orthonormal $\vec{b_s}$

Move $\vec{x} \leftarrow \vec{x} + \delta \vec{g}$

# Analysis

$|\vec{x}|^2 \leftarrow |\vec{x}|^2 + \delta^2$ so $T \leq n\delta^{-2}$

$L_j$ moves Gaussian, Variance $\leq \delta^2 \frac{1}{d} \leq \delta^2 \frac{4}{n}$

Total Variance $\leq 4$. Martingale

$\Pr[|L_j| \geq K] \leq 2e^{-K^2/8} \leq 0.1$

SUCCESS: Fewer than $\frac{n}{5}$ $j$ with $|L_j| \geq K$.

Positive Probability of SUCCESS

# Analysis

$|\vec{x}|^2 \leftarrow |\vec{x}|^2 + \delta^2$ so $T \leq n\delta^{-2}$

$L_j$ moves Gaussian, Variance $\leq \delta^2 \frac{1}{d} \leq \delta^2 \frac{4}{n}$

Total Variance $\leq 4$. Martingale

$\Pr[|L_j| \geq K] \leq 2e^{-K^2/8} \leq 0.1$

SUCCESS: Fewer than $\frac{n}{5}$ $j$ with $|L_j| \geq K$.

Positive Probability of SUCCESS

SUCCESS implies that ALL $|L_j| \leq K$

# Phase $s$

$m = 2^{1-s}n$. Start $\vec{z}$ with $\leq m$ coordinates frozen. End $\vec{x}$ with $\leq \frac{m}{2}$ coordinates frozen.

Effectively $|\vec{r_j}| \leq \sqrt{m}$

Would get $K\sqrt{m}$ but still have $n = m2^{s-1}$ vectors.

Actually get: $K\sqrt{m}\sqrt{s} = K\sqrt{n}\sqrt{s}2^{(1-s)/2}$

Converges!

# Thank You!

*It is six in the morning.*
*The house is asleep.*
*Nice music is playing.*
*I prove and conjecture.*
*– Paul Erdős, in letter to Vera Sós*