Almost all k-SAT functions are unate

József Balogh   Dingding Dong   Bernard Lidický   Nitya Mani   Yufei Zhao

AMS Sectional Meeting #1184
Mar 18, 2023

Question

Count functions

$$f : \{0,1\}^n \to \{0,1\} \qquad 2^{2^n}$$

k-SAT FUNCTION can be defined as

$$f(x_1, \ldots, x_n) = C_1 \vee C_2 \vee \cdots \vee C_m$$

$$C_i = \underbrace{z_1 \wedge z_2 \wedge \cdots \wedge z_k}_{\text{all different}} \qquad z_i \in \{x_1, \neg x_1, x_2, \neg x_2, \ldots, x_n, \neg x_n\}$$

x_i variable, C_i clause, z_i literal

example $k = 3$

$$x_1 \wedge x_2 \to (x_1 \wedge x_2 \wedge x_3) \vee (x_1 \wedge x_2 \wedge \neg x_3)$$

$$x_1 \wedge x_2 \wedge \neg x_2 \to \text{always false}$$

Every k-SAT function has a formula but the formula may not be unique.

number of $f : \{0,1\}^n \to \{0,1\}$ $\qquad 2^{2^n}$
number of k-SAT formula $2^{2^k\binom{n}{k}}$
number of k-SAT functions?

k-SAT formula is *monotone* if it uses only $x_1, x_2, \ldots, x_n$, (i.e. no $\neg x_i$ is used)

All monotone k-SAT formula give different functions

$$g \not\ni x_1 \wedge \cdots \wedge x_k \ni f \qquad f \neq g \text{ at } x_1 = \cdots = x_k = 1, x_{k+1} = \cdots = x_m = 0$$

Number of monotone k-SAT functions $2^{\binom{n}{k}}$

k-SAT formula is *unate* if it uses at most one of $\{x_i, \neg x_i\} = \{x_i, \overline{x_i}\}$
Number of unate k-SAT functions $(1+o(1))2^{n+\binom{n}{k}}$
Functions avoiding x_i counted multiple times

Conjecture (Bollobás, Brightwell, Leader 2003)

Fix $k \geq 2$, $1 - o(1)$ fraction of k-SAT functions are unate as $n \to \infty$.

- # 2-SAT functions is $2^{(1+o(1))\binom{n}{2}}$. Bollobás, Brightwell, Leader 2003
  using Szemerédi regularity lemma
- Conjecture true for $k = 2$ Allen 2007
  using Szemerédi regularity lemma
- Conjecture true for $k = 2$ Ilinca, Kahn 2009
  without Szemerédi regularity lemma
- Conjecture true for $k = 3$ Ilinca, Kahn 2012
  using hypergraph regularity lemma
- Conjecture true for $k = 4, 5$ Dong, Mani, Zhao 2022

Conjecture true for all k :-) Balogh, Dong, Lidický, Mani, Zhao 2022+

- $\{C_1, \ldots, C_m\}$ is *minimal* if deleting any C_i changes the function.
- i.e. for every C_i exists $X \in \{0,1\}^n$ s.t. only C_i is satisfied
  $\{wx, wy, x\bar{z}, \bar{y}z\}$ is not minimal

Idea: forbid non-minimal formula and transform to a Turán type problem.
k-uniform hypergraph

$$V = \{x_1, \ldots, x_n\} \qquad E = \binom{\{x_1, \ldots, x_n\}}{k}$$

Count the number of hypergraphs not containing forbidden configurations.
(forbidden configuration is a non-minimal formula)

- Trouble 1: How to reduce $\{x_1, \overline{x_1}, \ldots, x_n, \overline{x_n}\}$ to n vertices and identify forbidden configurations?
  Dong, Mani, Zhao using blow-up, saturation, container method

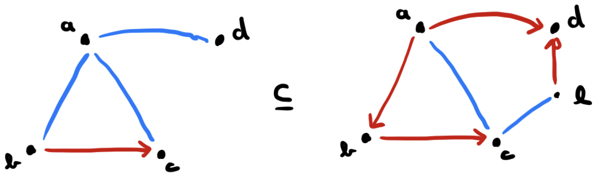
- Trouble 2: How to solve the resulting hypergraph extremal problem?
  BDLMZ: computer free flag-algebra

Directed Hypergraph Turán Problem

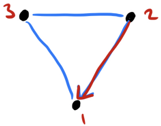
Partially directed k-graph is a k-uniform hypergraph, where every edge is

- undirected
- rooted at one vertex (directed towards one vertex)

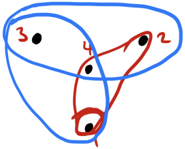
$\vec{H} \subseteq \vec{G}$ if $\vec{H}$ could be obtained from $\vec{G}$ by removing some vertices, edges, or orientations.



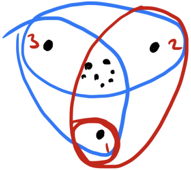
$\vec{T}_k$



- $\vec{T}_2 = \{\hat{1}2, 13, 23\}$



- $\vec{T}_3 = \{\hat{1}24, 134, 234\}$



- $\vec{T}_k = \{\hat{1}24 \cdots k+1, 134 \cdots k+1, 234 \cdots k+1\}$

Extremal problem

G is k-uniform, n-vertex, $\vec{T}_k$-free.

$:= \alpha := \dfrac{e_{undirected}(G)}{\binom{n}{k}}$

$:= \beta := \dfrac{e_{directed}(G)}{\binom{n}{k}}$

Given k, θ, what is

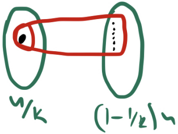
$$\max\{\alpha + \theta\beta\}?$$

Special case:

Show $\alpha + \theta\beta \leq 1$ when $1 \leq \theta \leq \left(1 - \frac{1}{k}\right)^{1-k} \approx e$

Constructions:

Complete undirected graph



Conjecture (Bollobás, Brightwell, Leader 2003)

Fix $k \geq 2$, almost all k-SAT functions are unate.

Theorem (Dong, Mani, Zhao)

If $\alpha + \theta\beta \leq 1$ when $\log_2 3 < \theta$ then almost all k-SAT functions are unate.

The above theorem is a lot of work.

Theorem (Dong, Mani, Zhao)

Conjecture true for $k \leq 5$.

Theorem (Balogh, Dong, Lidický, Mani, Zhao)

Conjecture for all k.

Containers

- few containers
- each minimal k-SAT formula is a subformula of at least one container
- Undirected edge $\{x_1, x_2, \cdots, x_k\}$ in container gives

$$x_1 x_2 \cdots x_k \quad \text{or} \quad \text{nothing}$$

- Directed edge $\{\hat{x}_1, x_2, \ldots, x_k\}$ in container gives

$$x_1 x_2 \cdots x_k \quad \text{or} \quad \overline{x}_1 x_2 \cdots x_k \quad \text{or} \quad \text{nothing}$$

- One container with $\alpha\binom{n}{2}$ undirected edges and $\beta\binom{n}{2}$ directed edges gives up to $2^{\alpha\binom{n}{2}} 3^{\beta\binom{n}{2}} = 2^{(\alpha+\beta\log_2 3)\binom{n}{2}}$ k-SAT formulas

Theorem (Füredi 1992)

$e(G^2) \geq e(G) - \lfloor \frac{n}{2} \rfloor$ *where* $E(G^2) = \{(x,y) : \exists z, xz, yz \in E(G)\}$

Theorem (Dong, Mani, Zhao)

For $k = 2$*:* $\alpha + 2\beta \leq 1 + o(1)$

Proof.

- $\vec{H}$ be $\vec{T}_2$-free graph
- G underlying graph (forget orientation) $e(G) = (\alpha + \beta)\binom{n}{2}$
- $xy \in E(G)$ and $xy \in E(G^2)$ means xy was undirected in $\vec{H}$.
- $$\binom{n}{2} \geq (\alpha + \beta)\binom{n}{2} \geq e(G^2) + \beta\binom{n}{2} \geq e(G) + \beta\binom{n}{2} - \frac{n}{2} = (\alpha + 2\beta)\binom{n}{2} - \frac{n}{2}$$

$\square$

- Averaging via link-graphs of $v \in \vec{H}$:

$$\pi\left(\vec{T}_k, \frac{(k-1)\theta+1}{k}\right) \leq \pi\left(\vec{T}_{k-1}, \theta\right)$$

- $$\pi\left(\vec{T}_3, \frac{5}{3}\right) \leq \pi\left(\vec{T}_2, 2\right)$$

- $\frac{5}{3} > \log_2 3$ implying case $k = 3$
- cases $k = 4, 5$ slightly more complicated

THEOREM (BALOGH, DONG, LIDICKÝ, MANI, ZHAO)

All values of k.

Proof for $k = 4$

$$\alpha + \theta\beta$$

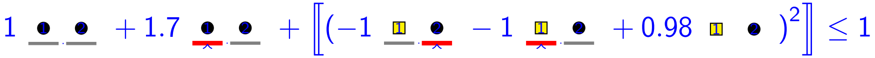
$$= \underline{1}\;\underline{2}\;\underline{3}\;\underline{4} + \theta\; \underline{1}\;\underline{2}\;\underline{3}\;\underline{4}$$

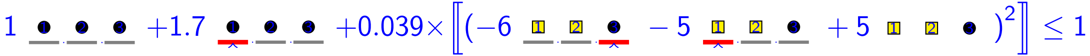
$$\leq \underline{1}\;\underline{2}\;\underline{3}\;\underline{4} + \theta\; \underline{1}\;\underline{2}\;\underline{3}\;\underline{4} + \left[\!\!\left[\left(a\; \boxed{1}\;\boxed{2}\;\boxed{3}\;4 - b\; \underline{\boxed{1}}\;\underline{\boxed{2}}\;\underline{\boxed{3}}\;\underline{4} \right)^2 \right]\!\!\right]$$

$$\leq 1$$

for $\theta = 1 + \frac{1}{\sqrt{2}} \geq 1.707 > \log_2 3 \quad a = \frac{1}{\sqrt{2}}, b = \frac{k(\theta-1)-1}{\sqrt{2}} \quad k \geq 4$

Proof for $k = 2$ and $k = 3$





Theorem (Balogh, Dong, Lidický, Mani, Zhao)

If $\vec{T}_k$ *is forbidden, then*  $+\left(1+\frac{1}{\sqrt{2}}\right)$  ≤ 1 *for all* k.

Question

If $\vec{T}_k$ *is forbidden, then*  $+\left(1-\frac{1}{k}\right)^{1-k}$  ≤ 1 *for all* k?