Approximate Polymorphisms

Gilad Chase
Yuval Filmus
Technion

Dor Minzer
Elchanan Mossel
MIT

Nitin Saurabh
IIIT Hyderabad

Plan of Talk

1. Polymorphisms and their characterization
2. Approximate polymorphisms & their characterization
3. "1% regime"
4. (Approximate) polymorphisms of predicates
5. Open questions
6. Bonus: More on the proof (will skip)

Polymorphisms

Suppose $f: \{0,1\}^n \longrightarrow \{0,1\}$

satisfies $f(x \oplus y) = f(x) \oplus f(y)$

for all $x, y \in \{0,1\}^n$

(we say that f is a **polymorphism**

of $g(a,b) = a \oplus b$)

What can f be?

Theorem: $\exists\, S \subseteq [n]$ s.t. $f(x) = \bigoplus_{i \in S} x_i$

Indeed: (one direction)

$$f(x \oplus y) = \bigoplus_{i \in S} (x_i \oplus y_i)$$

$$= \bigoplus_{i \in S} x_i \oplus \bigoplus_{i \in S} y_i$$

$$= f(x) \oplus f(y)$$

What if f is a polymorphism of $g(a,b) = a \wedge b$, i.e., $f(x \wedge y) = f(x) \wedge f(y)$?

Theorem:

f is constant or

$\exists S \subseteq [n]$ s.t. $f = \bigwedge_{i \in S} x_i$

What if f is a polymorphism of $g(a,b,c) = a \oplus b \oplus c$, i.e., $f(x \oplus y \oplus z) = f(x) \oplus f(y) \oplus f(z)$?

XORs, constant 1, negated XORs

what if f is a polymorphism of
$g(a,b,c) = \text{Majority}(a,b,c)$ or of
$h(a,b,c) = \text{Minority}(a,b,c)$?

Majority: constants, x_i, $\overline{x_i}$

$$f(\text{Maj}(x,y,z)) = \text{Maj}(f(x), f(y), f(z))$$

Minority: x_i, $\overline{x_i}$

$$f(\text{Min}(x,y,z)) = \text{Min}(f(x), f(y), f(z))$$

what if f is a polymorphism of

$$g(a,b,c,d) = \begin{cases} 1 & \text{if } a+b+c+d=2, \\ 0 & \text{otherwise?} \end{cases}$$

0, x_i

Theorem (essentially Dokow & Holzman):

If f is a polymorphism of g then

have exact characterization

(1) f depends on ≤ 1 coordinate, or

(2) g depends on ≤ 1 coordinate, or

associativity of product

(3) f, g are (possibly negated)° XORs, or

(4) f, g are ANDs, or

(5) f, g are ORs

°subject to parity constraints

Their result is actually more general: they (essentially) find all solutions to

$$f_0(g(x_1, \dots, x_m)) = g(f_1(x_1), \dots, f_m(x_m))$$

We extend this to

$$f_0(g(x_1, \dots, x_m)) = h(f_1(x_1), \dots, f_m(x_m))$$

Dokow and Holzman, coming from a social choice perspective, put restrictions on $f_0, \dots, f_m$.

Our result holds without any restrictions.

Approximate Polymorphisms

What if f is a polymorphism of g only most of the time?

E.g. what can we say about f if

$$\Pr[f(x \oplus y) = f(x) \oplus f(y)] \geq 1 - \varepsilon?$$

(Known as Linearity testing in property testing)

Theorem (BLR):
f is $O(\varepsilon)$-close to an XOR

What if $\Pr[f(x \wedge y) = f(x) \wedge f(y)] \geq 1 - \varepsilon$?

[Filmus-Lifshitz-Minzer-Mossel 2020]

Theorem: f is close to 0 or an AND

Theorem (?)

If f is an approx polymorphism of g then

(1) f essentially depends on ≤ 1 coordinate, or

(2) g essentially depends on ≤ 1 coordinate, or

(3) f, g are close to (negated) XORs, or

(4) f, g are close to ANDs, or

(5) f, g are close to ORs

Counterexample to the conjectured theorem

Consider $g(a,b) = \overline{a} \wedge \overline{b}$:

$$
\begin{aligned}
&\overline{x_1 \vee \cdots \vee x_n} \wedge \overline{y_1 \vee \cdots \vee y_n} && g(\vee(x), \vee(y)) \\
&= \overline{x_1} \wedge \cdots \wedge \overline{x_n} \wedge \overline{y_1} \wedge \cdots \wedge \overline{y_n} && \parallel \\
&= (\overline{x_1} \wedge \overline{y_1}) \wedge \cdots \wedge (\overline{x_n} \wedge \overline{y_n}) && \wedge(g(x,y))
\end{aligned}
$$

Observation: $g(\mathrm{Ber}(\tfrac{1}{2}), \mathrm{Ber}(\tfrac{1}{2})) \sim \mathrm{Ber}(\tfrac{1}{4})$

$$
f(z_1, \ldots, z_n) = \begin{cases} z_1 \vee \cdots \vee z_n = 1 & \text{if } z_1 + \cdots + z_n \approx \frac{n}{2} \\ z_1 \wedge \cdots \wedge z_n = 0 & \text{if } z_1 + \cdots + z_n \approx \frac{n}{4} \end{cases}
$$

Theorem: For every fixed g,

If f is an approx polymorphism of g then

(1) f essentially depends on ≤ 1 coordinate, or

(2) g depends on ≤ 1 coordinate, or

(3) g is (negated) XOR and f is close to (negated) XOR, or

(4) g is AND and f is close to AND, or

(5) g is OR and f is close to OR, or

new

(6) g is NAND and f is close to AND, or

(7) g is NOR and f is close to OR

A bit on the proof:

- Step 1: Show that f is close to a **junta** F
- Step 2: When ε is small enough,
  F is an **exact** polymorphism of g
  (slightly cheating here: $f \approx F$ only on balanced inputs)
  Need $\varepsilon < \left(\frac{1}{2}\right)^{|\mathrm{dom}\, F| \cdot |\mathrm{dom}\, g|}$

Step 1 is the hard part

Uses Jones' regularity lemma and "It Ain't Over Till It's Over"

"1% regime":

For which $c \in (0,1)$ does

$$\Pr[f(x \wedge y) = f(x) \wedge f(y)] \geq c$$

imply that f is **structured**,

that is, correlated with a junta?*

Equivalently, maximize

$$\Pr[f(x \wedge y) = f(x) \wedge f(y)]$$

subject to f **not** correlated with a junta

* If $\Pr[f(x \oplus y) = f(x) \oplus f(y)] \geq \frac{1}{2} + \varepsilon$ then f is correlated with XOR

Guess 1: random f

$$\Pr[f(x \wedge y) = f(x) \wedge f(y)] \approx \frac{1}{2}$$

Guess 2: $f(x)$ random for $\frac{x_1 + \cdots + x_n}{n} \approx \frac{1}{2}$

$f(x) = 0$ for $\frac{x_1 + \cdots + x_n}{n} \approx \frac{1}{4}$

$$\Pr[f(x \wedge y) = f(x) \wedge f(y)] \approx \frac{3}{4}$$

Guess 3: $f(x) = \mathrm{Maj}(x_1, \ldots, x_m)$ when $\frac{x_1 + \cdots + x_n}{n} \approx \frac{1}{2}$

$f(x) = \mathrm{Thr}_{\theta m}(x_1, \ldots, x_m)$ when $\frac{x_1 + \cdots + x_n}{n} \approx \frac{1}{4}$

where $m = \omega(1)$

and θ is an appropriate constant

Gives $\Pr[f(x \wedge y) = f(x) \wedge f(y)] \approx 0.815$

Optimal! Proof uses the invariance principle and a generalization of Borell's isoperimetric inequality

(A reduction from Boolean cube to "Gaussian space", which is $\mathbb{R}^n$ wrt Gaussian measure)

Approximate Polymorphisms of Predicates

NAE_3

0 0 1
0 1 0
0 1 1
1 0 0
1 0 1
1 1 0

Arrow's theorem:
only polymorphisms
are (anti) dictators

Approx version
due to Kalai, Mossel

NOR

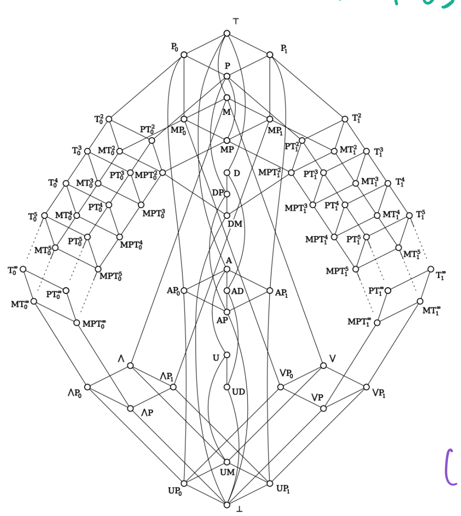
0 0
0 1
1 0

0 1
1 0
0 0
0 1
0 1
$f\downarrow\ \downarrow f$
× 1 1

Polymorphisms =
Intersecting
families

Approx version
due to Friedgut-Regev

General exact version: Post's lattice



Emily
(Wikipedia)

Open Questions

1. Approximate polymorphisms of predicates
2. Larger alphabets
3. Improve dependence between ε and δ

   $\varepsilon = \Pr[f(x \wedge y) \neq f(x) \wedge f(y)]$

   $\delta = \Pr[f \neq \mathrm{AND}]$

   (currently quasi-exponential)

Bonus — If $\Pr[f(x \wedge y) = f(x) \wedge f(y)] \approx 1$ then f is close to a junta

Proof 1

Average over y to get $T_\downarrow f \approx \mathbb{E}[f] \cdot f$.

Values of $T_\downarrow f$ around the middle slice depend on values of f around the quarter slice.

It turns out that $\widehat{T_\downarrow f}(S) = \left(\frac{1}{\sqrt{3}}\right)^{|S|} \tilde{f}(S)$.

$(\mu_{\frac{1}{2}})$ $(\mu_{\frac{1}{4}})$

Wlog $\mathbb{E}[f] \gg 0 \Rightarrow f \approx \frac{1}{\mathbb{E} f} T_\downarrow f$ has small Fourier tails

$\Rightarrow f$ is close to a junta by Bourgain's theorem.

Proof 2

Jones' regularity lemma expresses f as a uniform decision tree whose leaves are low-influence functions.

We will show that the leaves f_ℓ are nearly constant.

Sample x, y, and let $z_i = \begin{cases} y_i & \text{if } x_i = 1, \\ \text{random} & \text{if } x_i = 0. \end{cases}$

By construction, $f_\ell(x \wedge y) = f_\ell(x \wedge z) \Rightarrow f_\ell(x) \wedge f_\ell(y) = f_\ell(x) \wedge f_\ell(z)$ whp.

If f_ℓ is far from constant, $f_\ell(x) = 1$ w.p. $\Omega(1)$.

"It Ain't Over Till It's Over": $f_\ell(y) \neq f_\ell(z)$ w.p. $\Omega(1)$. $\lightning$

(Requires low influences!)