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 "l% regime"
④ (Approximate) polymorphisms of predicates
⑤ Open questions
⑥ Bonus: More on the proof (will skip)
Polymorphisms

Suppose \( f: \{0,1\}^n \rightarrow \{0,1\} \) 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:** \( \exists \sigma \subseteq \{n\} \) s.t. \( f(x) = \bigoplus_{i \in \sigma} x_i \)

Indeed: (one direction)

\[
\begin{align*}
f(x \oplus y) &= \bigoplus_{i \in \sigma} (x_i \oplus y_i) \\
&= \bigoplus_{i \in \sigma} x_i \oplus \bigoplus_{i \in \sigma} y_i \\
&= f(x) \oplus f(y)
\end{align*}
\]
What if $f$ is a polymorphism of $g(a,b) = a \land b$, i.e., $f(xy) = f(x) \land f(y)$?

Theorem:

$f$ is constant or

$\forall s \subseteq \{y\} \text{ s.t. } f = \bigwedge_{x \in s} x$
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, $X_i, \overline{X}_i$

\[ f(\text{Maj}(X_i, y, z)) = \text{Maj}(f(x), f(y), f(z)) \]

**Minority:** $X_i, \overline{X}_i$

\[ f(\text{Min}(X_i, y, z)) = \text{Min}(f(x), f(y), f(z)) \]

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

1. $f$ depends on $\leq 1$ coordinate, or
2. $g$ depends on $\leq 1$ coordinate, or
3. $f, g$ are (possibly negated) $\oplus$ 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

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

We extend this to

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

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

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

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

[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 \( \leq 1 \) coordinate, or
2. \( g \) essentially depends on \( \leq 1 \) 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 $g(a, b) = \overline{a} \land b$:

$$
\begin{align*}
X_1 \lor \ldots \lor X_n \land \overline{y_1} \lor \ldots \lor \overline{y_n} & \quad g(\nu(x), \nu(y)) \\
= \overline{X_1} \land \ldots \land \overline{X_n} \land \overline{y_1} \land \ldots \land \overline{y_n} & \quad ||
= (\overline{X_1} \land \overline{y_1}) \land \ldots \land (\overline{X_n} \land \overline{y_n}) & \quad \wedge (g(x, y))
\end{align*}
$$

Observation: $g(\text{Ber}(\frac{1}{2}), \text{Ber}(\frac{1}{4})) \sim \text{Ber}(\frac{1}{4})$

$$
\begin{align*}
f(z_1, \ldots, z_n) &= \begin{cases} 
\nu \wedge z_n \approx 1 \quad \text{if } z_1 + \ldots + z_n \approx \frac{n}{2} \\
\nu \wedge \overline{z_n} \approx 0 \quad \text{if } z_1 + \ldots + z_n \approx \frac{n}{4} 
\end{cases}
\end{align*}
$$
Theorem: For every fixed $g$, 

If $f$ is an approx polymorphism of $g$, then 

1. $f$ essentially depends on $\leq 1$ coordinate, or 
2. $g$ depends on $\leq 1$ coordinate, or 
3. $g$ is $(\text{negated})\ XOR$ and $f$ is close to $(\text{negated})\ XOR$, or 
4. $g$ is AND and $f$ is close to AND, or 
5. $g$ is OR and $f$ is close to OR, or 
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 $\varepsilon$ is small enough,
  
  $F$ is an exact polymorphism of $g$

  (slightly cheating here: $f \approx F$ only on balanced inputs)

  Need $\varepsilon < (\frac{1}{2})^{\frac{1}{|\text{dom } F| \cdot |\text{dom } g|}}$

Step 1 is the hard part

Uses Jones' regularity lemma and "It Ain't Over Till It's Over"
"1/2 regime":

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

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

imply that $f$ is structured, that is, correlated with a junta?*

Equivalently, maximize

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

subject to $f$ not correlated with a junta

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

Guess 1: random $f$

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

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

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

$$\Pr[f(x) = f(y)] \approx \frac{3}{4}$$
Guess 3: $f(x) = \text{Maj}(X_1, \ldots, X_n)$ when $\frac{X_1 + \cdots + X_n}{n} \approx \frac{1}{2}$

$f(x) = \text{Thr}_m(X_1, \ldots, X_n)$ when $\frac{X_1 + \cdots + X_n}{n} \approx \frac{1}{4}$

where $m = \omega(1)$

and $\Theta$ is an appropriate constant

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

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

\[ \text{NAE}_3 \]

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

\[ \text{NOR} \]

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

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

Approx version
due to Kalai, Mossel

Polymorphisms =
Intersecting families

Approx version
due to Friedgut-Reger

General exact version: Post's lattice

Emilly
(Wikipedia)
Open Questions

1. Approximate polymorphisms of predicates
2. Larger alphabets
3. Improve dependence between $\epsilon$ and $\delta$

$$
\begin{align*}
\epsilon &= \Pr[f(x \land y) \neq f(x) \land f(y)] \\
\delta &= \Pr[f \neq \text{AND}]
\end{align*}
$$

(currently quasi-exponential)
**Bonus**

If \( \Pr[f(x,y)=f(x)\land f(y)] \neq 1 \)

then \( f \) is close to a junta

---

**Proof**

Average over \( y \) to get \( T_y f \approx \mathbb{E}[f] \cdot f \).

Values of \( T_y f \) around the middle slice depend on values of \( f \) around the quarter slice.

It turns out that \( \hat{T}_y f \approx (\frac{1}{\sqrt{3}}) \hat{f} \approx (s) \).

Wlog \( \mathbb{E}[f] \neq 0 \) \( \Rightarrow \hat{f} \approx \hat{\mathbb{E}} T_y f \) 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 $f_x$ 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, $f_x(x, y) = f_x(x, z) \Rightarrow f_x(x) \land f_y(y) = f_x(x) \land f_y(z) \text{ w.p.}$

If $f_x$ is far from constant, $f_x(x) = 1 \text{ w.p. } 2/3$.

"It Ain't Over Till It's Over": $f_x(y) \neq f_x(z) \text{ w.p. } 2/3$. \(\exists\)

(Requires low influences!)