Bounds on functionality and symmetric difference

Tung Anh Vu

joint with
Pavel Dvořák, Lukáš Folwarczný, Michal Opler, Pavel Pudlák, and Robert Šámal.

Matoušek prize lecture

7th March 2024

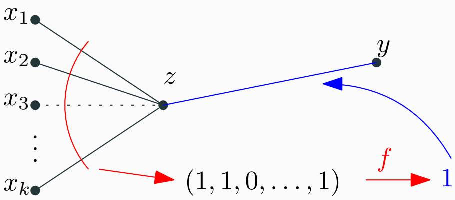
CSI

IÚUK

Graph functionality – definition

- Consider a graph G. Put $A(x,y) = 1$ iff $\{x,y\} \in E(G)$.
- Vertex $y \in V(G)$ *is a function of* $x_1, \dots, x_k \in V(G) \setminus \{y\}$ iff

$$\exists f \colon \{0,1\}^k \to \{0,1\}$$
$$\forall z \in V(G) \setminus \{y, x_1, \dots, x_k\} :$$
$$A(y,z) = f(A(x_1,z), \dots, A(x_k,z))$$



- $\mathrm{fun}_G(y) :=$ minimal k such that y is a function of some k vertices
- $\mathrm{fun}(G) := \max_{H \subseteq_{\mathrm{ind}} G} \min_{y \in V(H)} \mathrm{fun}_H(y)$
- Seed of this idea in [Atminas, Collins, Lozin, Zamaraev 2015]
- Defined in [Alecu, Atminas, Lozin 2021]

Graph functionality – motivation

- By putting f identically 0, we get $\text{fun}_G(y) \leq \deg(y)$ and thus

$$\text{fun}(G) \leq \textit{degeneracy of } G$$

- If $\text{fun}(G) \leq k$ then G can be represented using $n\big((k+1)\log(n) + 2^k\big)$ bits of information.
- So if every graph $G \in \mathcal{G}$ has bounded functionality then $\mathcal{G}$ contains at most $2^{O(n \log n)}$ graphs on n vertices.
- Kind of a reverse to the *Implicit graph conjecture*, which motivated this research (but was recently disproved [Hatami and Hatami 2022]).

Known results – general

Theorem (AAL 2021)

For any graph G:

$$\mathrm{fun}(G) \leq 2\mathrm{cwd}(G) - 1$$

Theorem (AAL 2021)

There exists a function g such that for any graph G:

$$\mathrm{vc}(G) \leq g(\mathrm{fun}(G)).$$

Known results – concrete graph classes

Theorem (AAL 2021)

- $\text{fun}(\textit{unit interval graphs}) \leq 2$.
- $\text{fun}(\textit{line graphs}) \leq 6$.
- $\text{fun}(\textit{permutation graphs}) \leq 8$.
- $\text{fun}(\textit{intersection graphs of 3-uniform hypergraphs}) \leq 462$.

Theorem (Dallard, Lozin, Milanič, Štorgel, Zamaraev 2023)

$\text{fun}(\textit{interval graphs}) \leq 8$.

Our results I

Theorem (Bounds on functionality)

There exists a graph on n vertices with functionality $\Omega(\sqrt{n})$.

$\text{fun}(G) \leq \mathcal{O}(\sqrt{n \log n})$ *for every graph G.*

Theorem (Random graphs)

W.h.p.

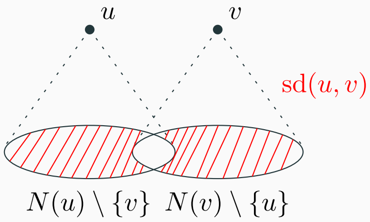
$$C_1 \log n \leq \text{fun}(G(n, 1/2)) \leq C_2 \log n$$

Theorem (Informal)

"Determining the functionality of a vertex is NP*-hard."*

Symmetric difference

Given $u, v \in V(G)$, let $\mathrm{sd}_G(u, v)$ be the number of vertices different from u and v that are adjacent to exactly one of u and v.



- It is noted by [AAL 2021] that $\mathrm{fun}(G) \leq \mathrm{sd}(G) + 1$.

Our results II

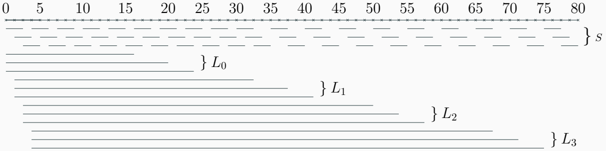
INT_n – interval graphs on n vertices

Theorem

Any interval graph $G \in \mathrm{INT}_n$ has symmetric difference at most $O(\sqrt[3]{n})$.

Theorem

There is an interval graph $G \in \mathrm{INT}_n$ of symmetric difference at least $\Omega(\sqrt[4]{n})$.



“Small” graph with “large” functionality – proof

Theorem

Consider a finite projective plane of order k and its bipartite incidence graph G. Then we have $\mathrm{fun}(G) \geq k$.

- In a finite projective plane of order k, each line has $k+1$ points and each point lies on $k+1$ lines. The total number of points, as well as lines, is k^2+k+1.
- Let ℓ be a line of the projective plane.
- Symmetry $\Rightarrow$ it is enough to show $\mathrm{fun}_G(\ell) \geq k$.

“Small” graph with “large” functionality – proof

Theorem

Consider a finite projective plane of order k and its bipartite incidence graph G. Then we have $\mathrm{fun}(G) \geq k$.

- Proof by contradiction: Let $p_1, \ldots, p_a$ and $\ell_1, \ldots, \ell_b$ be points and lines such that ℓ is a function of these and $a + b \leq k - 1$.
- We claim that then there exist points q_1 and q_2 satisfying:
  1. q_1 and q_2 are distinct from all $p_1, \ldots, p_a$.
  2. Neither q_1 nor q_2 is incident to any of the lines $\ell_1, \ldots, \ell_b$.
  3. q_1 is incident to ℓ, but q_2 is not.
- The existence of these two points then implies that ℓ is not a function of $p_1, \ldots, p_a$ and $\ell_1, \ldots, \ell_b$ which is the desired contradiction.

“Small” graph with “large” functionality – proof

Theorem

Consider a finite projective plane of order k and its bipartite incidence graph G. Then we have $\text{fun}(G) \geq k$.

- There are $k+1$ points on each line and every two lines intersect in one point.
- Therefore there exists a point on ℓ which is distinct from each p_i and it is not incident with any ℓ_i. This is the point q_1.
- The total number of points is k^2+k+1. The total number of points incident to $\ell, \ell_1, \ldots, \ell_b$ is at most $(b+1)(k+1)$ and we have a points $p_1, \ldots, p_a$. This is the total of $(b+1)(k+1)+a \leq (a+b+1)(k+1) \leq k(k+1) < k^2+k+1$. This implies the existence of the point q_2.

Determining the functionality of a vertex is hard – proof

"Equivalent" view of functionality:

- Suppose that y is a function of vertices $x_1, \ldots, x_k$.
- For vertices $a, b \in V \setminus \{y, x_1, \ldots, x_k\}$ define relation $\sim$ as follows:

$$a \sim b \Leftrightarrow A(a, x_i) = A(b, x_i) \qquad \forall i \in [k]$$

- Then $\sim$ has at most 2^k equivalence classes.
- And all vertices in a single equivalence class have the same adjacency with y.

Determining the functionality of a vertex is hard – proof

- We reduce from DOMINATING SET to “determining whether a vertex v is a function of at most k vertices.”
- Given (G, k) an instance of dominating set, we add
  - a vertex v connected to all vertices of $V(G)$, and
  - a set I of $k + 1$ isolated vertices
- resulting in graph H, where $V(H) = V(G) \cup \{v\} \cup I$.

$\Rightarrow$

- Any two vertices of $V(G)$ or I have the same adjacency with v.

Determining the functionality of a vertex is hard – proof

$\Leftarrow$

- Suppose that v is a function of some vertices $u_1, \dots, u_k \in V(H)$.
- We claim that $\{u_1, \dots, u_k\} \cap V(G)$ is a dominating set in G.
- Assume it is not for contradiction.
- Then there exists $x \in V(G)$ such that it is not adjacent to any $u_1, \dots, u_k$.
- At least one $y \in I$ is not adjacent to v.
- x and y have the same neighbourhood with $u_1, \dots, u_k$ but $A(x, v) \neq A(y, v)$.

Upper bound – Proof

Theorem

Let G be a graph on n vertices. Then $\mathrm{fun}(G) = \mathcal{O}(\sqrt{n \log n})$.

- Fix a constant $c > 2$.
- We show that there is a $v \in V(G)$: $\mathrm{fun}_G(v) \leq d(n) := \sqrt{cn \log n}$.
- As $d(n)$ is increasing, this suffices for the proof of the theorem.
- We write $d = d(n)$ and $PN(u,v) = \big(N(u) \,\Delta\, N(v)\big) \setminus \{u,v\}$.
- Case 1: There exist $u \neq v$ s.t. $sd(u,v) = |PN(u,v)| < d$.
  - Then the set $PN(u,v) \cup \{u\}$ suffices to certify the neighborhood of v.
  - We are done with this case.

Upper bound – Proof continued

- Case 2: All sets $PN(u,v)$ have size at least d.
  - Choose $v \in V(G)$ arbitrarily.
  - Choose a random set $S \subseteq V(G)$ by independently putting each vertex of $G - v$ to S with probability $p = d/n$.
  - The probability that adjacency to v is not certified by S is equal to

$$\Pr(\exists u_1 \in N(v) \setminus S, u_2 \notin N(v) \cup S \cup \{v\} : S \cap PN(u_1,u_2) = \emptyset).$$

  - We estimate this using the union bound by

$$\sum_{u_1,u_2} (1-p)^{|PN(u_1,u_2)|-1} \leq n^2(1-p)^{d-1}$$

  - The probability that S is “bad for v” is at most $n^2 e^{-p(d-1)}$, which is $o(1)$ whenever $c > 2$.

Upper bound – Proof continued

- The expected size of S is $p \cdot (n-1) = \frac{n-1}{n} d$.
- By Markov inequality: $P(|S| > d) \leq \frac{n-1}{n}$.
- This means that with positive probability $|S| \leq d$ and S certifies adjacency to v.

Random graphs – Rough sketch of proofs

- $G = G(n, 1/2)$
- Lower bound: directly estimate $P(\text{fun}(v) < \frac{1}{2}\log n) < \frac{1}{n^{\log n}}$, then use union bound to get $P(\text{fun}(G) < \frac{1}{2}\log n) = o(1)$.
- Upper bound: we in fact show, that w.h.p. every induced $H \subset G(n, 1/2)$ has a distinguishing subset of size $C \log n$, from this the result about functionality follows directly. But having to prove this for 2^n subgraphs requires us to get much tighter bounds for each of the subgraphs. We use *Poisson approximation* for "balls-into-bins" together with *custom Chernoff-type bound*.

Open problems

- What is the maximum functionality of a graph on n vertices?
  (Are projective planes the worst?)
- What is the complexity of computing functionality?
  (Should be hard, put unclear how to prove it.)
- What is $\max \operatorname{sd}(G)$ for $G \in \mathrm{INT}_n$?
  ($n^{1/4}$ or $n^{1/3}$?)