The union-closed sets conjecture

Henning Bruhn

ulm university universität ulm

joint with Oliver Schaudt

The union-closed sets conjecture

Always: $\mathcal{A}$ finite family of finite sets

- $\mathcal{A}$ union-closed: $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$.
- Example: $\emptyset$, 1, 12, 34, 134, 1234

Conjecture

Every union-closed family of at least two sets has an element that appears in at least half of the member-sets.

- power sets are union-closed
- conjecture tight for power sets!

The union-closed sets conjecture

Always: $\mathcal{A}$ finite family of finite sets

- $\mathcal{A}$ union-closed: $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$.
- Example: $\emptyset$, 1, 12, 34, 134, 1234

Conjecture

Every union-closed family of at least two sets has an element that appears in at least half of the member-sets.

- power sets are union-closed
- conjecture tight for power sets!

A union-closed family

123456

12345 12346 12356 12456 13456 23456

1234 1235 1236 1456 2456 3456

123 145 246 356 456

45 46 56

4 5 6

$\emptyset$

- union-closed
- 25 sets
- Universe: 1, 2, 3, 4, 5, 6

A union-closed family

123456

12345 12346 12356 12456 13456 23456

1234 1235 1236 1456 2456 3456

123 145 246 356 456

45 46 56

4 5 6

$\emptyset$

- union-closed
- 25 sets
- Universe: 1, 2, 3, 4, 5, 6
- 2 lies in 12 member-sets

A union-closed family

123456

12345 12346 12356 12456 13456 23456

1234 1235 1236 1456 2456 3456

123 145 246 356 456

45 46 56

4 5 6

$\emptyset$

- union-closed
- 25 sets
- Universe: $1, 2, 3, 4, 5, 6$
- 2 lies in 12 member-sets
- 4 lies in 16 member-sets $\rightarrow$ conjecture ✔

A memorable quote

Peter Winkler ’87:

> *The ‘union-closed sets conjecture’ is well known indeed, except for (1) its origin and (2) its answer!*

Some terminology

Conjecture

Every union-closed family of at least two sets has an element that appears in at least half of the member-sets.

- $\mathcal{A}$ always (finite) union-closed family
- $U := \bigcup_{A \in \mathcal{A}} A$ is the universe
- frequency: $\mathcal{A}_u := \{A \in \mathcal{A} : u \in A\}$
- u abundant if $|\mathcal{A}_u| \geq \frac{1}{2}|\mathcal{A}|$.
- n : number of member-sets
- m : number of elements

What do we know?

$\mathcal{A}$ satisfies the conjecture when

- at most 12 elements
- at most 50 member-sets
- $\mathcal{A}$ has special structure, for instance represented by cubic graph
- ...

Also

- some (weak) properties of smallest counterexample known

$\rightarrow$ conjecture wide open

Knill's argument

Knill: There's always an element appearing in $\geq \frac{n-1}{\log_2(n)}$ member-sets

		123456		
	12345	12346	12356	
	12456	13456	23456	
	1234	1235	1236	
	1456	2456	3456	
123	145	246	356	456
	45	46	56	
	4	5	6	
		$\emptyset$		

X : smallest set intersecting all members

traces $\mathcal{T} = \{A \cap X : A \in \mathcal{A}\}$

- $X = 12456$?
- smallest traces $12|4|5|6$

Knill's argument

Knill: There's always an element appearing in $\geq \frac{n-1}{\log_2(n)}$ member-sets

123456
12345 12346 12356
12456 13456 23456
1234 1235 1236
1456 2456 3456
123 145 246 356 456
45 46 56
4 5 6
$\emptyset$

X : smallest set intersecting all members
traces $\mathcal{T} = \{A \cap X : A \in \mathcal{A}\}$

- $X = 1456$
- $\mathcal{T}$ contains all singletons of X
- union-closed$\rightarrow \mathcal{T} = 2^X$
- $\rightarrow |X| = \log_2(|\mathcal{T}|) \leq \log_2(n)$
- $\rightarrow$ an element in X meets at least $(n-1)/\log_2(n)$ member-sets

Knill's argument

Knill: There's always an element appearing in $\geq \frac{n-1}{\log_2(n)}$ member-sets

		123456		
	12345	12346	12356	
	12456	13456	23456	
	1234	1235	1236	
	1456	2456	3456	
123	145	246	356	456
	45	46	56	
	4	5	6	
		$\emptyset$		

X : smallest set intersecting all members
traces $\mathcal{T} = \{A \cap X : A \in \mathcal{A}\}$

- $X = 1456$
- $\mathcal{T}$ contains all singletons of X
- union-closed$\rightarrow \mathcal{T} = 2^X$
- $\rightarrow |X| = \log_2(|\mathcal{T}|) \leq \log_2(n)$
- $\rightarrow$ an element in X meets at least $(n-1)/\log_2(n)$ member-sets

- constant factor improved by Wójcik

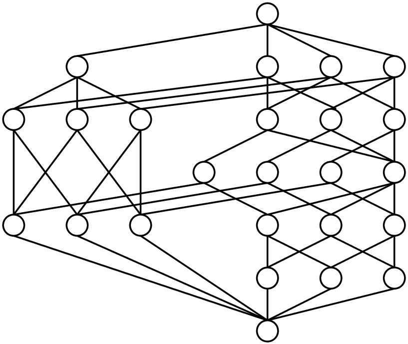
Many faces

Equivalent reformulations

- in terms of lattices
- in terms of graphs
- in terms of “very full sets”

The lattice formulation

Lattice: Finite poset $(L, <)$ so that

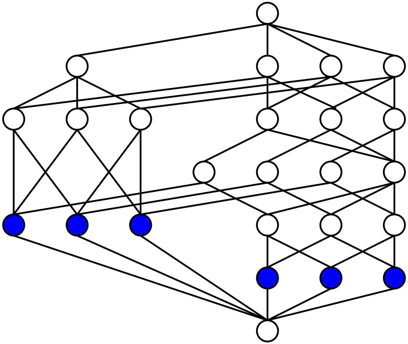


- any two $x, y \in L$ have unique greatest lower bound $x \wedge y$
- any two $x, y \in L$ have unique smallest upper bound $x \vee y$

non-zero $x \in L$ is join-irreducible if $x = y \vee z$ implies $x = y$ or $x = z$.

The lattice formulation

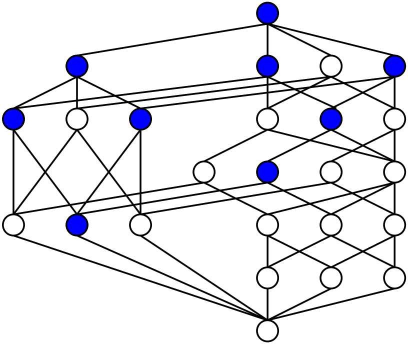
Lattice: Finite poset $(L, <)$ so that



- any two $x, y \in L$ have unique greatest lower bound $x \wedge y$
- any two $x, y \in L$ have unique smallest upper bound $x \vee y$

non-zero $x \in L$ is join-irreducible if $x = y \vee z$ implies $x = y$ or $x = z$.

The lattice formulation II



Conjecture

Let L be a lattice with $|L| \geq 2$. Then there is join-irreducible $x \in L$ so that

$$|\{y : x \leq y\}| \leq \tfrac{1}{2}|L|.$$

- true for lower semimodular lattices

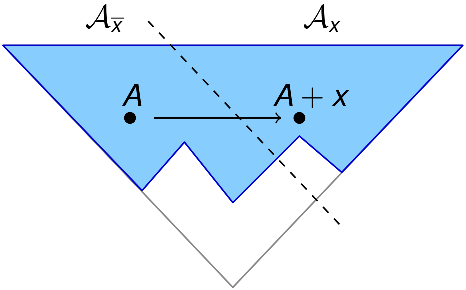
Main techniques

- Injections
- Local configurations
- Averaging

Injections

Up-set: If $A \in \mathcal{A}$ and $B \supseteq A$ then $B \in \mathcal{A}$

- Up-sets satisfy the conjecture



Proof:

- Injection $\mathcal{A}_{\overline{x}} \to \mathcal{A}_x$
- $\Rightarrow 2|\mathcal{A}_x| \geq |\mathcal{A}_{\overline{x}}| + |\mathcal{A}_x| = |\mathcal{A}|$.
- x abundant!

Problem with this technique:
Need to know where to find an abundant element

Local configurations

Early observation:

Singleton $\{x\} \in \mathcal{A}$	$\longrightarrow$	x abundant!
2-set $\{x, y\} \in \mathcal{A}$	$\longrightarrow$	x or y abundant!
3-set $\{x, y, z\} \in \mathcal{A}$	$\longrightarrow$	x, y or z abundant?

Local configurations

Early observation:

Singleton $\{x\} \in \mathcal{A}$	$\longrightarrow$	x abundant!
2-set $\{x, y\} \in \mathcal{A}$	$\longrightarrow$	x or y abundant!
3-set $\{x, y, z\} \in \mathcal{A}$	$\longrightarrow$	x, y or z abundant?

NO!

123456

12345 12346 12356 12456 13456 23456

1234 1235 1236 1456 2456 3456

123 145 246 356 456

45 46 56

4 5 6

$\emptyset$

Local configurations II

However:

- $123, 124, 134 \in \mathcal{A}$ then one of $1, 2, 3, 4$ is abundant
- Family $\mathcal{L}$ is Frankl-complete if any union-closed $\mathcal{A}$ that contains $\mathcal{L}$ has abundant element among the elements of $\mathcal{L}$
- all Frankl-complete families known on five elements (Morris)
- General characterisation due to Poonen

Averaging

Strategy: determine average frequency $\frac{1}{|U|}\sum_{u\in U}|\mathcal{A}_u|$

		123456		
	12345	12346	12356	
	12456	13456	23456	
	1234	1235	1236	
	1456	2456	3456	
123	145	246	356	456
	45	46	56	
	4	5	6	
		$\emptyset$		

- $1, 2, 3 \rightarrow$ each 12 times
- $4, 5, 6 \rightarrow$ each 16 times
- average frequency $\frac{1}{6}(3 \cdot 12 + 3 \cdot 16) = 15$
- $\rightarrow$ there is element of frequency ≥ 15
- 25 member-sets $\rightarrow$ ✔

Averaging

Strategy: determine average frequency $\frac{1}{|U|}\sum_{u\in U}|\mathcal{A}_u|$

123456
12345 12346 12356
12456 13456 23456
1234 1235 1236
1456 2456 3456
123 145 246 356 456
45 46 56
4 5 6
$\emptyset$

- $1, 2, 3 \rightarrow$ each 12 times
- $4, 5, 6 \rightarrow$ each 16 times
- average frequency $\frac{1}{6}(3 \cdot 12 + 3 \cdot 16) = 15$
- $\rightarrow$ there is element of frequency ≥ 15
- 25 member-sets $\rightarrow$ ✔

complete rubbish approach!

Average set size

Double counting:

$$\sum_{A \in \mathcal{A}} |A| = \sum_{u \in U} |\mathcal{A}_u|$$

Usually, total set size easier to control!
Thus, if

$$\frac{1}{|U|} \sum_{A \in \mathcal{A}} |A| \geq \frac{1}{2} |\mathcal{A}|$$

then $\mathcal{A}$ satisfies the conjecture.

Average set size

Double counting:

$$\sum_{A \in \mathcal{A}} |A| = \sum_{u \in U} |\mathcal{A}_u|$$

Usually, total set size easier to control!
Thus, if average set size

$$\frac{1}{|\mathcal{A}|} \sum_{A \in \mathcal{A}} |A| \geq \frac{1}{2}|U|$$

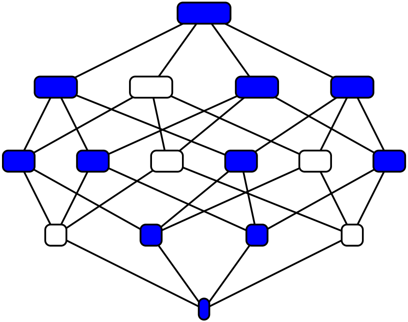
then $\mathcal{A}$ satisfies the conjecture.

- Advantage: Don't need to know where to look for abundant element
- Drawback: Averaging does not always work!

Large families

Nishimura & Takahashi '96:

- If $|\mathcal{A}| > 2^m - \sqrt{2^m}$, where $m = |U|$ then $\mathcal{A}$ ✔



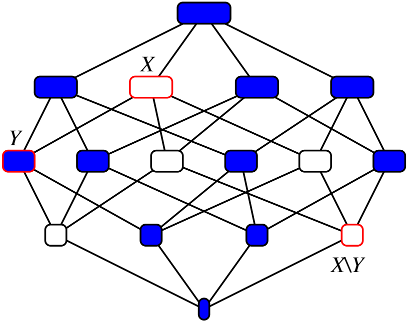
Proof:

Subfamily of power set on $1, 2, 3, 4$

Large families

Nishimura & Takahashi '96:

- If $|\mathcal{A}| > 2^m - \sqrt{2^m}$, where $m = |U|$ then $\mathcal{A}$ ✔



Subfamily of power set on $1, 2, 3, 4$

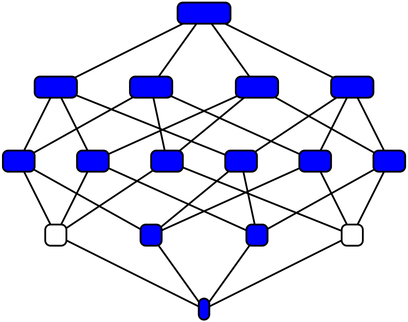
Proof:

- Assume set $X \notin \mathcal{A}$ with $|X| \geq \frac{m}{2}$
- If $Y \subseteq X$ in $\mathcal{A} \Rightarrow Y \setminus X \notin \mathcal{A}$
- Thus: $\frac{1}{2}2^{|X|}$ sets missing in $\mathcal{A}$
- $|\mathcal{A}| \leq 2^m - 2^{\frac{m}{2}}$, contradiction!

Large families

Nishimura & Takahashi ’96:

- If $|\mathcal{A}| > 2^m - \sqrt{2^m}$, where $m = |U|$ then $\mathcal{A}$ ✔



Subfamily of power set on $1, 2, 3, 4$

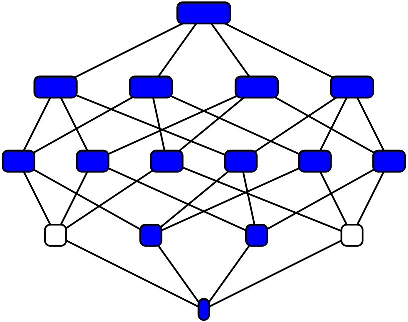
Proof:

- Assume set $X \notin \mathcal{A}$ with $|X| \geq \frac{m}{2}$
- If $Y \subseteq X$ in $\mathcal{A} \Rightarrow Y \setminus X \notin \mathcal{A}$
- Thus: $\frac{1}{2}2^{|X|}$ sets missing in $\mathcal{A}$
- $|\mathcal{A}| \leq 2^m - 2^{\frac{m}{2}}$, contradiction!
- $\Rightarrow \mathcal{A}$ contains all large sets
- Average set size $\geq \frac{m}{2}$

Large families

Nishimura & Takahashi '96:

- If $|\mathcal{A}| > 2^m - \sqrt{2^m}$, where $m = |U|$ then $\mathcal{A}$ ✔



Subfamily of power set on $1, 2, 3, 4$

Proof:

- Assume set $X \notin \mathcal{A}$ with $|X| \geq \frac{m}{2}$
- If $Y \subseteq X$ in $\mathcal{A} \Rightarrow Y \setminus X \notin \mathcal{A}$
- Thus: $\frac{1}{2}2^{|X|}$ sets missing in $\mathcal{A}$
- $|\mathcal{A}| \leq 2^m - 2^{\frac{m}{2}}$, contradiction!
- $\Rightarrow \mathcal{A}$ contains all large sets
- Average set size $\geq \frac{m}{2}$

$\rightarrow \mathcal{A}$ satisfies the conjecture!

Large families II

Balla, Bollobás & Eccles ’13:

- If $|\mathcal{A}| \geq \frac{2}{3}2^m$, where $m = |U|$ then $\mathcal{A}$ ✔

Result...

- is best possible
- builds on Kruskal-Katona theorem
- and on approach of Reimer

Reimer ’03:

- Average set size always

$$\frac{1}{|\mathcal{A}|}\sum_{A\in\mathcal{A}}|A| \geq \frac{1}{2}\log_2|\mathcal{A}|$$

Hungarian family

Let $A, B \subset \mathbb{N}$ finite
Set $A < B$ if

1. largest element: $\max A < \max B$
2. reverse colex: $\max(A \Delta B) \in A$

Initial segment:

$$\emptyset < 1 < 12 < 2 < 123 < 23 < 13 < 3 < 1234 < 234$$
$$< 134 < 34 < 124 < 24 < 14 < 4 < 12345 < \dots$$

Czédli, Maróti and Schmidt '09:

- $\mathcal{H}(m)$: initial segment of length $\lfloor \frac{2}{3} 2^m \rfloor$
- Average too low for $\mathcal{H}(m)$!

Separating families

Consider $\mathcal{A} = \emptyset, 1, 12, 234, 1234$

- 4 does not add anything!
- delete 4 from every member
- $\mathcal{A}' = \emptyset, 1, 12, 23, 123$

$\rightarrow$ may assume that $\mathcal{A}$ separating:
for every $x, y \in U$ there is $A \in \mathcal{A}$ containing exaxtly one of x, y

Small families

Let $\mathcal{A}$ be separating, $U = \{x_1, \ldots, x_m\}$.

x_1	x_2	x_3	x_4	x_5	x_6
✔	✔	✔	✔	✔	✔

Assume $x_1, \ldots, x_m$ ordered by increasing frequency

- let X_0 be universe
- for $i < j$ there exists X_{ij} with $x_i \notin X_{ij} \ni x_j$
- set $X_i := \bigcup_j X_{ij}$
- all $X_0, \ldots, X_{m-1}$ distinct
- all contain x_m

$\rightarrow$ If $|\mathcal{A}| \leq 2m$ then $\mathcal{A}$ satisfies conjecture (Falgas-Ravry '11)

Small families

Let $\mathcal{A}$ be separating, $U = \{x_1, \ldots, x_m\}$.

x_1	x_2	x_3	x_4	x_5	x_6
✔	✔	✔	✔	✔	✔
✖	✔	✔	✔	✔	✔

Assume $x_1, \ldots, x_m$ ordered by increasing frequency

- let X_0 be universe
- for $i < j$ there exists X_{ij} with $x_i \notin X_{ij} \ni x_j$
- set $X_i := \bigcup_j X_{ij}$
- all $X_0, \ldots, X_{m-1}$ distinct
- all contain x_m

$\rightarrow$ If $|\mathcal{A}| \leq 2m$ then $\mathcal{A}$ satisfies conjecture (Falgas-Ravry '11)

Small families

Let $\mathcal{A}$ be separating, $U = \{x_1, \ldots, x_m\}$.

x_1	x_2	x_3	x_4	x_5	x_6
✔	✔	✔	✔	✔	✔
✖	✔	✔	✔	✔	✔
?	✖	✔	✔	✔	✔

Assume $x_1, \ldots, x_m$ ordered by increasing frequency

- let X_0 be universe
- for $i < j$ there exists X_{ij} with $x_i \notin X_{ij} \ni x_j$
- set $X_i := \bigcup_j X_{ij}$
- all $X_0, \ldots, X_{m-1}$ distinct
- all contain x_m

$\rightarrow$ If $|\mathcal{A}| \leq 2m$ then $\mathcal{A}$ satisfies conjecture (Falgas-Ravry '11)

Small families

Let $\mathcal{A}$ be separating, $U = \{x_1, \dots, x_m\}$.

x_1	x_2	x_3	x_4	x_5	x_6
✔	✔	✔	✔	✔	✔
✖	✔	✔	✔	✔	✔
?	✖	✔	✔	✔	✔
?	?	✖	✔	✔	✔
?	?	?	✖	✔	✔
?	?	?	?	✖	✔

Assume $x_1, \dots, x_m$ ordered by increasing frequency

- let X_0 be universe
- for $i < j$ there exists X_{ij} with $x_i \notin X_{ij} \ni x_j$
- set $X_i := \bigcup_j X_{ij}$
- all $X_0, \dots, X_{m-1}$ distinct
- all contain x_m

$\rightarrow$ If $|\mathcal{A}| \leq 2m$ then $\mathcal{A}$ satisfies conjecture (Falgas-Ravry '11)

Future directions?

- What families on n member-sets have lowest max frequency?
- for $n = 2^m \to$ power sets
- in between powers of two?

Future directions?

- What families on n member-sets have lowest max frequency?
- for $n = 2^m \to$ power sets
- in between powers of two?

Details, bibliography and more:
The journey of the union-closed sets conjecture,
Henning Bruhn and Oliver Schaudt