Introduction to EF-games

Inexpressivity results for first-order logic

Normal forms for first-order logic

Algorithms and complexity for specific classes of structures

General complexity bounds

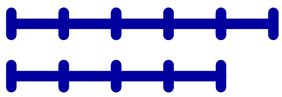
Sufficient vs. "iff" conditions

$$\mathcal{L}_k \stackrel{\text{def}}{=} (\{1, \ldots, k\}, <)$$

We know that

$$n, p \geqslant 2^m - 1 \Rightarrow \textbf{II} \text{ wins } \mathcal{G}_m(\mathcal{L}_n, \mathcal{L}_p).$$

- Given $\mathcal{L}_5$ and $\mathcal{L}_6$, does **II** win $\mathcal{G}_3(\mathcal{L}_5, \mathcal{L}_6)$?



No!

In fact,

$$n = p \text{ or } n, p \geqslant 2^m - 1 \Leftrightarrow \textbf{II} \text{ wins } \mathcal{G}_m(\mathcal{L}_n, \mathcal{L}_p)$$

- Complete characterizations are needed to exploit games algorithmically

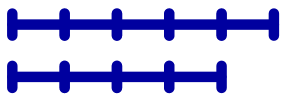
Sufficient vs. "iff" conditions

$$\mathcal{L}_k \stackrel{\text{def}}{=} (\{1, \ldots, k\}, <)$$

We know that

$$n, p \geqslant 2^m - 1 \Rightarrow \textbf{II} \text{ wins } \mathcal{G}_m(\mathcal{L}_n, \mathcal{L}_p).$$

- Given $\mathcal{L}_5$ and $\mathcal{L}_6$, does **II** win $\mathcal{G}_3(\mathcal{L}_5, \mathcal{L}_6)$?



No!

In fact,

$$n = p \text{ or } n, p \geqslant 2^m - 1 \Leftrightarrow \textbf{II} \text{ wins } \mathcal{G}_m(\mathcal{L}_n, \mathcal{L}_p)$$

- Complete characterizations are needed to exploit games algorithmically

Winning vs. optimal strategies

Winning strategy $\neq$ Optimal strategy

The distinction between winning and optimal strategies is essential in unbounded games:

- In unbounded EF-games on finite structures, **I** wins unless $\mathcal{A} \cong \mathcal{B}$
- "Play randomly" is a winning strategy for **I**
- But, how far actually is the end of a game?
- What are the *best* moves for **I** (and **II**)?

Remoteness

Optimal strategies (in combinatorial games $\mathcal{G}$) can be characterized in terms of remoteness ($rem(\mathcal{G})$):

- Current player has no legal moves from (the current configuration of) $\mathcal{G} \Rightarrow rem(\mathcal{G}) = 0$
- Current player can move to a configuration with even remoteness $\Rightarrow rem(\mathcal{G}) = 1 +$ least even remoteness

  Win Quickly!

- Current player can only move to configurations with odd remoteness $\Rightarrow rem(\mathcal{G}) = 1 +$ greatest odd remoteness

  Lose Slowly!

- The parity of the remoteness tells the winner

Win quickly, lose slowly!

Remoteness in EF-games:

- For EF-games, remoteness in terms of rounds, not moves
- Remoteness of $\mathcal{G}$: the minimum m such that **I** wins $\mathcal{G}_m$ (simplified definition under the hypothesis $\mathcal{A} \not\equiv \mathcal{B}$)
- Optimal **I** 's move: given a configuration $\mathcal{G}$, a move by **I** is *optimal* if and only if, whatever **II** replies, the remoteness of the resulting configuration is less than or equal to $rem(\mathcal{G}) - 1$.
- Optimal **II** 's move: given a configuration $\mathcal{G}$ and a move by **I** , a reply by **II** is *optimal* if and only if the remoteness of the resulting position is
  - $rem(\mathcal{G}) - 1$, if **I** 's move is optimal
  - $rem(\mathcal{G})$, otherwise

Solving Games

Example

$$n = p \text{ or } n, p \geqslant 2^m - 1 \Leftrightarrow \textbf{II} \text{ wins } \mathcal{G}_m(\mathcal{L}_n, \mathcal{L}_p)$$

How many rounds are needed to **I** to win?
Assume $n < p$. Then:

1. The remoteness can be computed as:

$$rem(\mathcal{G}(\mathcal{L}_n, \mathcal{L}_p)) = \lfloor \log(n+1) \rfloor + 1$$

2. A move of **I** from $\mathcal{G}(\mathcal{L}_n, \mathcal{L}_p)$ is optimal if and only if **I** chooses
   - an element in $[n - 2^{\lfloor \log_2(n+1) \rfloor} + 2, 2^{\lfloor \log_2(n+1) \rfloor} - 1]$ in L_n, or
   - an element in $[n - 2^{\lfloor \log_2(n+1) \rfloor} + 2, p - n + 2^{\lfloor \log_2(n+1) \rfloor} - 1]$ in L_p
3. Similarly, the set of **II**'s optimal replies can be computed

Algorithmic and complexity results

B. Khoussainov and J. Liu,
On Complexity of Ehrenfeucht-Fraïssé Games
Annals of Pure and Applied Logic, 2009

A. Montanari and A. Policriti and N. Vitacolonna,
An Algorithmic Account of Winning Strategies in Ehrenfeucht Games on Labeled Successor Structures (extended version: Ehrenfeucht Games on Labelled Successor Structures: Remoteness and Optimal Strategies, RR 02/09, Dept. of Math. and CS, University of Udine, Feb. 2009)
Proc. 12th LPAR, LNCS 3835, 2005

E. De Maria, A. Montanari, N. Vitacolonna,
Games on Strings with a Limited Order Relation
Proc. LFCS 2009, LNCS 5407

E. Pezzoli,
Computational Complexity of Ehrenfeucht-Fraïssé Games on Finite Structures
Proc. CSL 1998, LNCS 1584

EF-games on specific classes

- Equivalence relations (with/without colors)
- Embedded equivalence relations
- Trees (with level predicates)
- Labelled successor structures
- Labelled linear structures with a bounded ordering

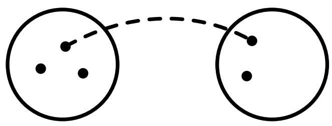
Equivalence relations: local strategy

Definition
Structures $\mathcal{A} = (A, E)$, where E is an equivalence relation on A.

Definition

- For $m, n, t \in \mathbb{N}$, $m =_t n$ iff $m = n$ or both $m, n > t$
- $(\mathcal{A}, \vec{a}, \mathcal{B}, \vec{b})$ is t-locally safe iff
  1. $\vec{a} \to \vec{b}$ is a partial isomorphism, and
  2. if $t > 0$, then $|[a_i] \setminus \{a_1, \ldots, a_k\}| =_t |[b_i] \setminus \{b_1, \ldots, b_k\}|$ for $i = 1, \ldots, k$.

When a position is t-locally safe, there is not incentive for **I** to play in a class that has already been chosen, in a game with at most t rounds.



1-locally safe, but not 2-locally safe

Equivalence relations: "small disparity"

- $q_t^{(\mathcal{A},\vec{a})}$: number of classes of size t in $\mathcal{A}$ not containing any a_i (free classes)
- Let $\Delta^{(\mathcal{A},\vec{a})}_{(\mathcal{B},\vec{b})} = \{ t \mid q_t^{(\mathcal{A},\vec{a})} \neq q_t^{(\mathcal{B},\vec{b})} \}$
- Let $q_t = \min\{q_t^{(\mathcal{A},\vec{a})}, q_t^{(\mathcal{B},\vec{b})}\}$

Lemma
*Given $(\mathcal{A}, \vec{a}, \mathcal{B}, \vec{b})$ and $t \in \Delta^{(\mathcal{A},\vec{a})}_{(\mathcal{B},\vec{b})}$, **I** can reach a position that is not t-locally safe after $q_t + 1$ rounds.*

Corollary
***I** has a winning strategy in $\leqslant q_t + 1 + t$ rounds, with $t \in \Delta^{(\mathcal{A},\vec{a})}_{(\mathcal{B},\vec{b})}$.*

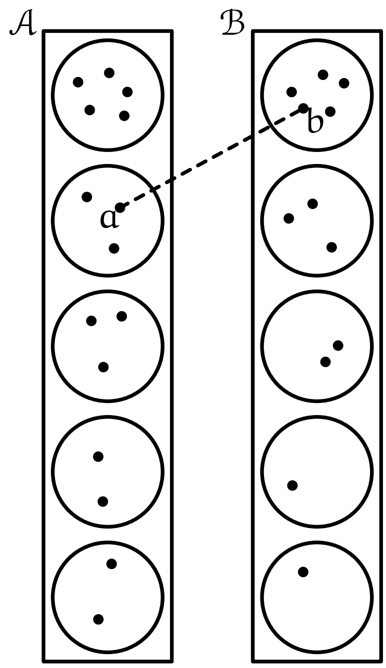
- **I** selects q_t distinct classes of size t ("global" moves)
- Then, he plays one more "global" move in a class of size t to which **II** cannot reply "appropriately"
- Then, he plays $\leqslant t$ rounds in the same class ("local" moves)

Example

- 2-locally safe, but not 3-locally safe

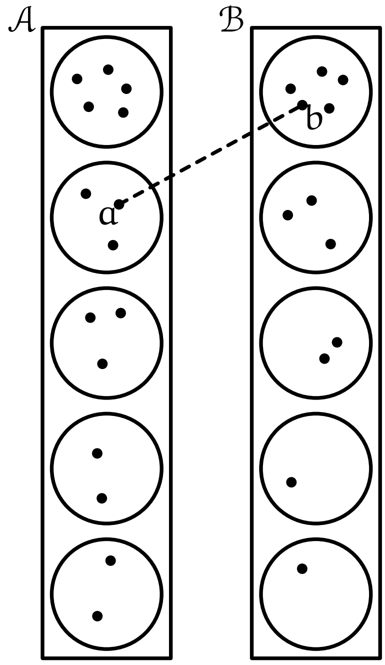
t	$q_t^{(\mathcal{A},a)}$	$q_t^{(\mathcal{B},b)}$
1	0	2
2	2	1
3	1	1
4	0	0
5	1	0

- $\Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} = \{1, 2, 5\}$



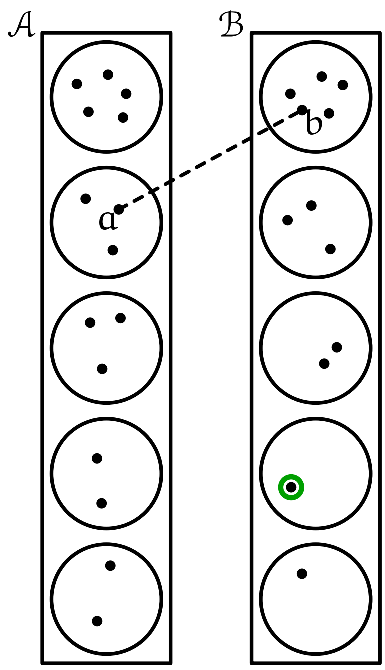
Example (cont.)

- $\Delta_{(\mathcal{B},b)}^{(\mathcal{A},a)} = \{1, 2, 5\}$
- $q_1 = 0,\ q_2 = 1,\ q_5 = 0$
- $1 \in \Delta_{(\mathcal{B},b)}^{(\mathcal{A},a)} \Rightarrow$ **I** can reach a not 1-locally safe configuration in $q_1 + 1 = 1$ round
- $2 \in \Delta_{(\mathcal{B},b)}^{(\mathcal{A},a)} \Rightarrow$ **I** can reach a not 2-locally safe configuration in $q_2 + 1 = 2$ rounds
- $5 \in \Delta_{(\mathcal{B},b)}^{(\mathcal{A},a)} \Rightarrow$ **I** can reach a not 5-locally safe configuration in $q_5 + 1 = 1$ round



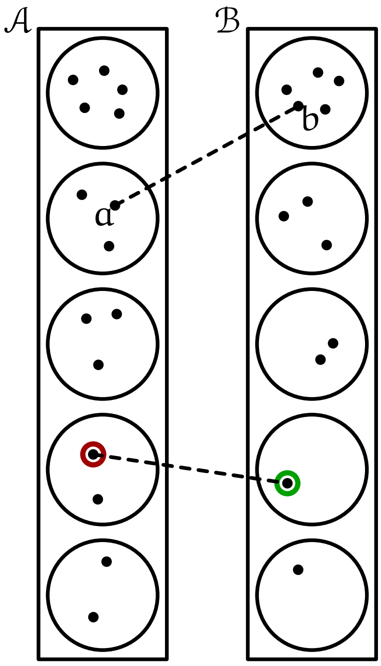
Example (cont.)

- $\Delta_{(\mathcal{B},b)}^{(\mathcal{A},a)} = \{1, 2, 5\}$
- $q_1 = 0,\ q_2 = 1,\ q_5 = 0$
- $1 \in \Delta_{(\mathcal{B},b)}^{(\mathcal{A},a)} \Rightarrow$ **I** can reach a not 1-locally safe configuration in $q_1 + 1 = 1$ round
- $2 \in \Delta_{(\mathcal{B},b)}^{(\mathcal{A},a)} \Rightarrow$ **I** can reach a not 2-locally safe configuration in $q_2 + 1 = 2$ rounds
- $5 \in \Delta_{(\mathcal{B},b)}^{(\mathcal{A},a)} \Rightarrow$ **I** can reach a not 5-locally safe configuration in $q_5 + 1 = 1$ round



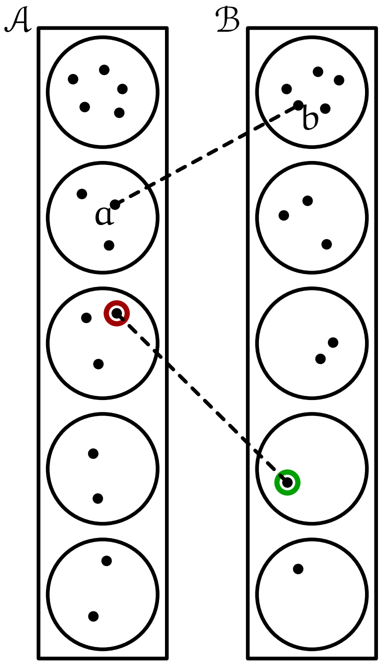
Example (cont.)

- $\Delta_{(\mathcal{B},b)}^{(\mathcal{A},a)} = \{1, 2, 5\}$
- $q_1 = 0,\ q_2 = 1,\ q_5 = 0$
- $1 \in \Delta_{(\mathcal{B},b)}^{(\mathcal{A},a)} \Rightarrow$ **I** can reach a not 1-locally safe configuration in $q_1 + 1 = 1$ round
- $2 \in \Delta_{(\mathcal{B},b)}^{(\mathcal{A},a)} \Rightarrow$ **I** can reach a not 2-locally safe configuration in $q_2 + 1 = 2$ rounds
- $5 \in \Delta_{(\mathcal{B},b)}^{(\mathcal{A},a)} \Rightarrow$ **I** can reach a not 5-locally safe configuration in $q_5 + 1 = 1$ round



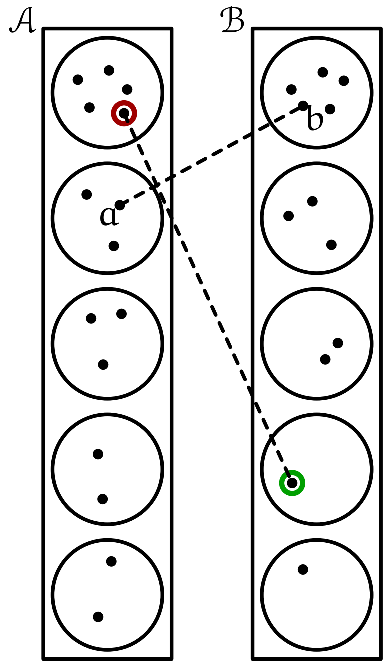
Example (cont.)

- $\Delta_{(\mathcal{B},b)}^{(\mathcal{A},a)} = \{1, 2, 5\}$
- $q_1 = 0,\ q_2 = 1,\ q_5 = 0$
- $1 \in \Delta_{(\mathcal{B},b)}^{(\mathcal{A},a)} \Rightarrow$ **I** can reach a not 1-locally safe configuration in $q_1 + 1 = 1$ round
- $2 \in \Delta_{(\mathcal{B},b)}^{(\mathcal{A},a)} \Rightarrow$ **I** can reach a not 2-locally safe configuration in $q_2 + 1 = 2$ rounds
- $5 \in \Delta_{(\mathcal{B},b)}^{(\mathcal{A},a)} \Rightarrow$ **I** can reach a not 5-locally safe configuration in $q_5 + 1 = 1$ round



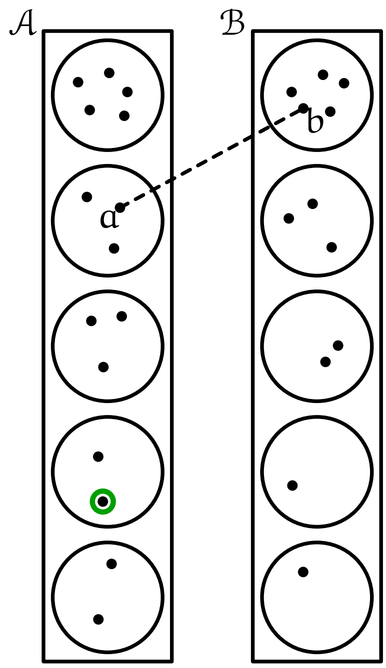
Example (cont.)

- $\Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} = \{1, 2, 5\}$
- $q_1 = 0,\ q_2 = 1,\ q_5 = 0$
- $1 \in \Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} \Rightarrow$ **I** can reach a not 1-locally safe configuration in $q_1 + 1 = 1$ round
- $2 \in \Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} \Rightarrow$ **I** can reach a not 2-locally safe configuration in $q_2 + 1 = 2$ rounds
- $5 \in \Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} \Rightarrow$ **I** can reach a not 5-locally safe configuration in $q_5 + 1 = 1$ round



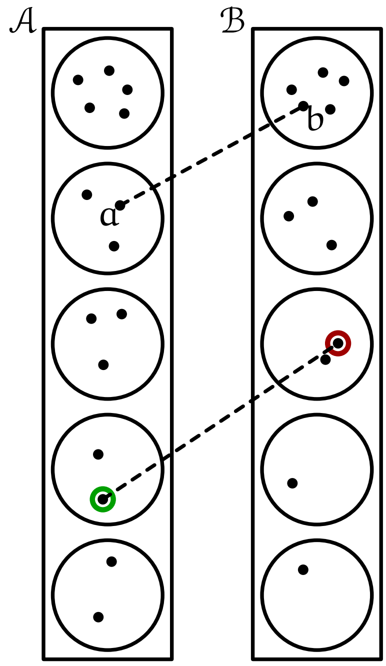
Example (cont.)

- $\Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} = \{1, 2, 5\}$
- $q_1 = 0,\ q_2 = 1,\ q_5 = 0$
- $1 \in \Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} \Rightarrow$ **I** can reach a not 1-locally safe configuration in $q_1 + 1 = 1$ round
- $2 \in \Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} \Rightarrow$ **I** can reach a not 2-locally safe configuration in $q_2 + 1 = 2$ rounds
- $5 \in \Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} \Rightarrow$ **I** can reach a not 5-locally safe configuration in $q_5 + 1 = 1$ round



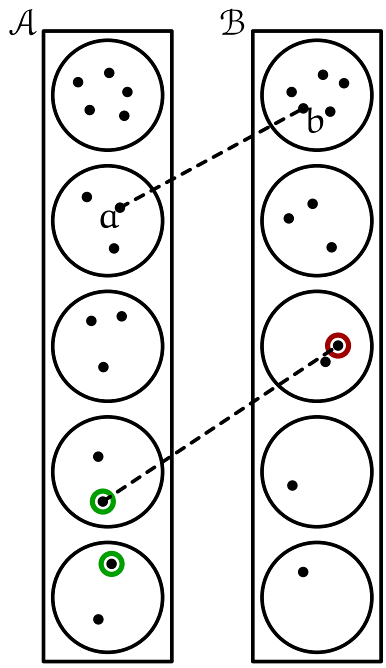
Example (cont.)

- $\Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} = \{1, 2, 5\}$
- $q_1 = 0,\ q_2 = 1,\ q_5 = 0$
- $1 \in \Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} \Rightarrow$ I can reach a not 1-locally safe configuration in $q_1 + 1 = 1$ round
- $2 \in \Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} \Rightarrow$ I can reach a not 2-locally safe configuration in $q_2 + 1 = 2$ rounds
- $5 \in \Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} \Rightarrow$ I can reach a not 5-locally safe configuration in $q_5 + 1 = 1$ round



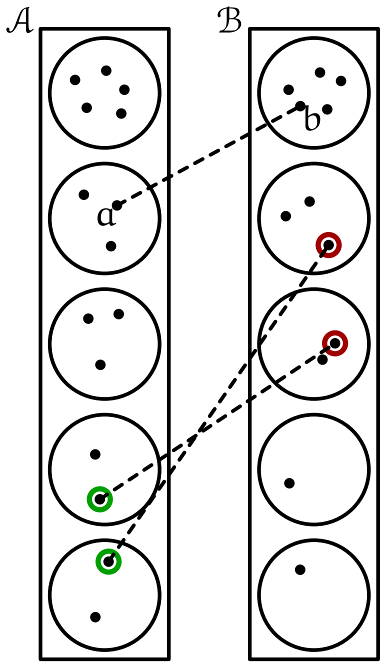
Example (cont.)

- $\Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} = \{1, 2, 5\}$
- $q_1 = 0,\ q_2 = 1,\ q_5 = 0$
- $1 \in \Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} \Rightarrow$ I can reach a not 1-locally safe configuration in $q_1 + 1 = 1$ round
- $2 \in \Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} \Rightarrow$ I can reach a not 2-locally safe configuration in $q_2 + 1 = 2$ rounds
- $5 \in \Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} \Rightarrow$ I can reach a not 5-locally safe configuration in $q_5 + 1 = 1$ round



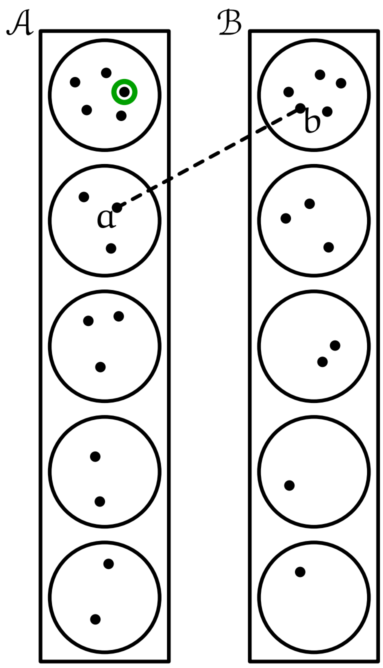
Example (cont.)

- $\Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} = \{1, 2, 5\}$
- $q_1 = 0,\ q_2 = 1,\ q_5 = 0$
- $1 \in \Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} \Rightarrow$ **I** can reach a not 1-locally safe configuration in $q_1 + 1 = 1$ round
- $2 \in \Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} \Rightarrow$ **I** can reach a not 2-locally safe configuration in $q_2 + 1 = 2$ rounds
- $5 \in \Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} \Rightarrow$ **I** can reach a not 5-locally safe configuration in $q_5 + 1 = 1$ round



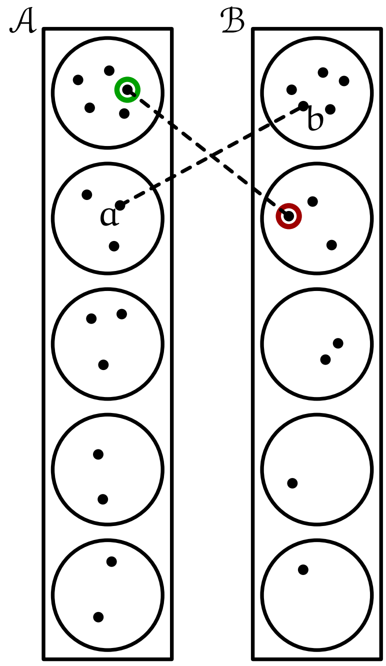
Example (cont.)

- $\Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} = \{1, 2, 5\}$
- $q_1 = 0,\ q_2 = 1,\ q_5 = 0$
- $1 \in \Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} \Rightarrow$ **I** can reach a not 1-locally safe configuration in $q_1 + 1 = 1$ round
- $2 \in \Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} \Rightarrow$ **I** can reach a not 2-locally safe configuration in $q_2 + 1 = 2$ rounds
- $5 \in \Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} \Rightarrow$ **I** can reach a not 5-locally safe configuration in $q_5 + 1 = 1$ round



Example (cont.)

- $\Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} = \{1, 2, 5\}$
- $q_1 = 0,\ q_2 = 1,\ q_5 = 0$
- $1 \in \Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} \Rightarrow$ **I** can reach a not 1-locally safe configuration in $q_1 + 1 = 1$ round
- $2 \in \Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} \Rightarrow$ **I** can reach a not 2-locally safe configuration in $q_2 + 1 = 2$ rounds
- $5 \in \Delta^{(\mathcal{A},a)}_{(\mathcal{B},b)} \Rightarrow$ **I** can reach a not 5-locally safe configuration in $q_5 + 1 = 1$ round



Equivalence relations: "large disparity"

- $q^{(\mathcal{A},\vec{a})}_{\geqslant t}$: number of free classes of size $\geqslant t$
- Let $\Gamma^{(\mathcal{A},\vec{a})}_{(\mathcal{B},\vec{b})} = \{\, t \mid q^{(\mathcal{A},\vec{a})}_{\geqslant t} \neq q^{(\mathcal{B},\vec{b})}_{\geqslant t} \,\}$
- Let $q_{\geqslant t} = \min\{q^{(\mathcal{A},\vec{a})}_{\geqslant t}, q^{(\mathcal{B},\vec{b})}_{\geqslant t}\}$

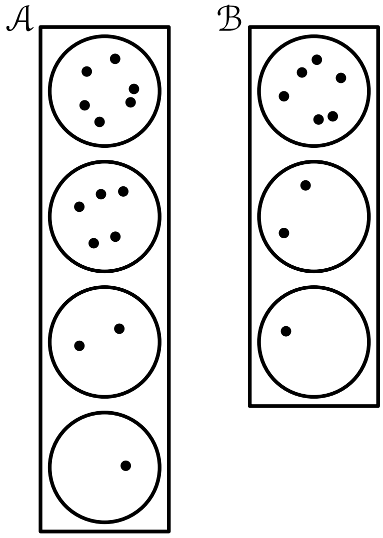
Lemma
*Given $(\mathcal{A}, \vec{a}, \mathcal{B}, \vec{b})$ and $t \in \Gamma^{(\mathcal{A},\vec{a})}_{(\mathcal{B},\vec{b})}$, **I** can reach a position that is not $(t-1)$-locally safe after $q_{\geqslant t} + 1$ rounds.*

Corollary
***I** has a winning strategy in at most $q_{\geqslant t} + t$ rounds, with $t \in \Gamma^{(\mathcal{A},\vec{a})}_{(\mathcal{B},\vec{b})}$.*

- **I** selects $q_{\geqslant t}$ distinct free classes of size $\geqslant t$ ("global" moves)
- Then, only one structure remains with a free class of size $\geqslant t$
- **I** plays t rounds in that class ("local" moves)

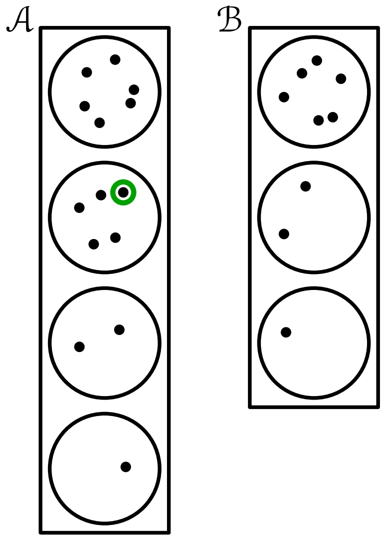
Example

- Initially, empty configuration
- Let $t = 3$
- Then $q_{\geqslant t} = 1$
- let **I** pick a free class with $\geqslant t$ elements
- **II** replies accordingly
- Now there is a free class of size $\geqslant t$ only in $\mathcal{A}$
- **II** replies with a "small" class
- **I** starts to play locally
- **II** must reply locally
- **I** wins
- $q_{\geqslant t} + t$ rounds needed



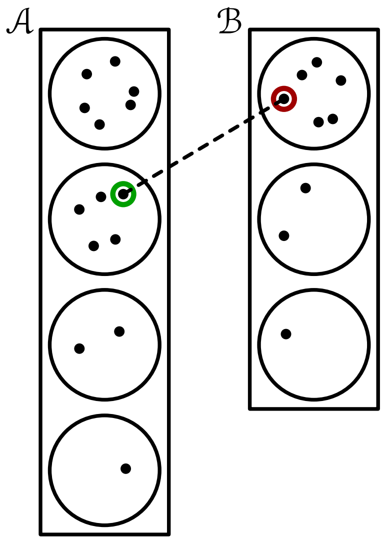
Example

- Initially, empty configuration
- Let $t = 3$
- Then $q_{\geqslant t} = 1$
- let **I** pick a free class with $\geqslant t$ elements
- **II** replies accordingly
- Now there is a free class of size $\geqslant t$ only in $\mathcal{A}$
- **II** replies with a "small" class
- **I** starts to play locally
- **II** must reply locally
- **I** wins
- $q_{\geqslant t} + t$ rounds needed



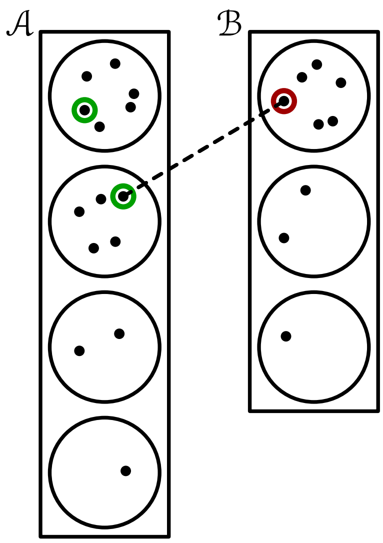
Example

- Initially, empty configuration
- Let $t = 3$
- Then $q_{\geqslant t} = 1$
- let **I** pick a free class with $\geqslant t$ elements
- **II** replies accordingly
- Now there is a free class of size $\geqslant t$ only in $\mathcal{A}$
- **II** replies with a "small" class
- **I** starts to play locally
- **II** must reply locally
- **I** wins
- $q_{\geqslant t} + t$ rounds needed



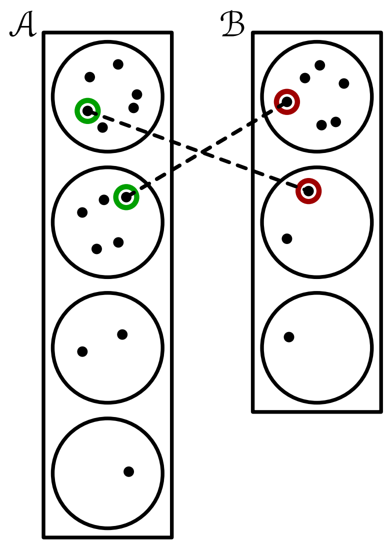
Example

- Initially, empty configuration
- Let $t = 3$
- Then $q_{\geqslant t} = 1$
- let **I** pick a free class with $\geqslant t$ elements
- **II** replies accordingly
- Now there is a free class of size $\geqslant t$ only in $\mathcal{A}$
- **II** replies with a "small" class
- **I** starts to play locally
- **II** must reply locally
- **I** wins
- $q_{\geqslant t} + t$ rounds needed



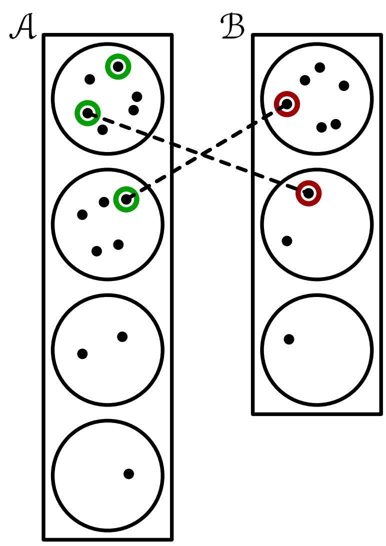
Example

- Initially, empty configuration
- Let $t = 3$
- Then $q_{\geqslant t} = 1$
- let **I** pick a free class with $\geqslant t$ elements
- **II** replies accordingly
- Now there is a free class of size $\geqslant t$ only in $\mathcal{A}$
- **II** replies with a "small" class
- **I** starts to play locally
- **II** must reply locally
- **I** wins
- $q_{\geqslant t} + t$ rounds needed



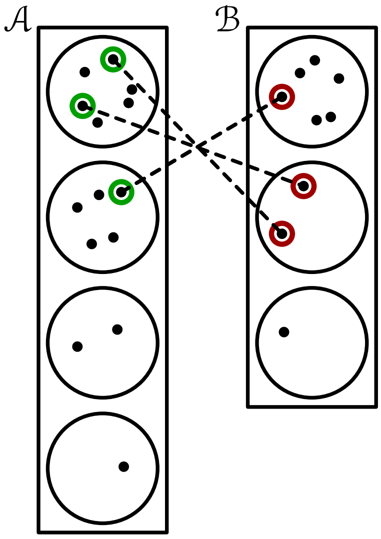
Example

- Initially, empty configuration
- Let $t = 3$
- Then $q_{\geqslant t} = 1$
- let **I** pick a free class with $\geqslant t$ elements
- **II** replies accordingly
- Now there is a free class of size $\geqslant t$ only in $\mathcal{A}$
- **II** replies with a "small" class
- **I** starts to play locally
- **II** must reply locally
- **I** wins
- $q_{\geqslant t} + t$ rounds needed



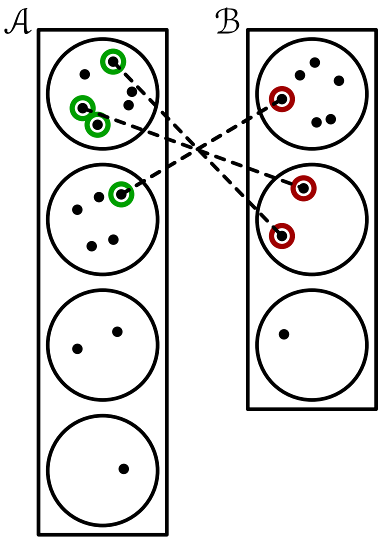
Example

- Initially, empty configuration
- Let $t = 3$
- Then $q_{\geqslant t} = 1$
- let **I** pick a free class with $\geqslant t$ elements
- **II** replies accordingly
- Now there is a free class of size $\geqslant t$ only in $\mathcal{A}$
- **II** replies with a "small" class
- **I** starts to play locally
- **II** must reply locally
- **I** wins
- $q_{\geqslant t} + t$ rounds needed



Example

- Initially, empty configuration
- Let $t = 3$
- Then $q_{\geqslant t} = 1$
- let **I** pick a free class with $\geqslant t$ elements
- **II** replies accordingly
- Now there is a free class of size $\geqslant t$ only in $\mathcal{A}$
- **II** replies with a "small" class
- **I** starts to play locally
- **II** must reply locally
- **I** wins
- $q_{\geqslant t} + t$ rounds needed



Equivalence relations: characterization

Definition

Given $(\mathcal{A}, \vec{a}, \mathcal{B}, \vec{b})$ and $m \in \mathbb{N}$, $(\mathcal{A}, \vec{a}, \mathcal{B}, \vec{b})$ is m-globally safe iff

- $q_t > m - t - 1$ for all $t \in \Delta^{(\mathcal{A},\vec{a})}_{(\mathcal{B},\vec{b})}$
- $q_{\geqslant t} > m - t$ for all $t \in \Gamma^{(\mathcal{A},\vec{a})}_{(\mathcal{B},\vec{b})}$

Theorem

***II** wins $\mathcal{G}_m(\mathcal{A}, \vec{a}, \mathcal{B}, \vec{b})$ iff $(\mathcal{A}, \vec{a}, \mathcal{B}, \vec{b})$ is m-locally safe and m-globally safe.*

Corollary

The remoteness of $\mathcal{G}(\mathcal{A}, \vec{a}, \mathcal{B}, \vec{b})$ is the minimum between the minimum m such that $(\mathcal{A}, \vec{a}, \mathcal{B}, \vec{b})$ is not m-locally safe and

$$\min\left\{\min\{t + q_{\geqslant t} \mid t \in \Gamma^{(\mathcal{A},\vec{a})}_{(\mathcal{B},\vec{b})}\}, 1 + \min\{t + q_t \mid t \in \Delta^{(\mathcal{A},\vec{a})}_{(\mathcal{B},\vec{b})}\}\right\}.$$

The remoteness can be computed in $O(|A| + |B|)$ time and space.

Sketch of the proof

Theorem

***II** wins $\mathcal{G}_m(\mathcal{A}, \vec{a}, \mathcal{B}, \vec{b})$ iff $(\mathcal{A}, \vec{a}, \mathcal{B}, \vec{b})$ is m-locally safe and m-globally safe.*

- If a position is m-locally safe and **I** play a local move, then **II** can reach a position $(m-1)$-locally safe
- If a position is m-globally safe, then **II** can reach a position $(m-1)$-globally safe
  - The only tricky case is when **I** chooses an element in a free class of size $t \in \Delta^{(\mathcal{A},\vec{a})}_{(\mathcal{B},\vec{b})}$ or $t \in \Gamma^{(\mathcal{A},\vec{a})}_{(\mathcal{B},\vec{b})}$
  - But, m-global safety allows **II** to reply properly

The result easily extends to structures colored homogeneously, i.e., if $E(x, y)$ then $x \in P \Leftrightarrow y \in P$, for all $x, y \in A$ and unary predicate P

Equivalence structures with one color

Definition

Structures $\mathcal{A} = (A, E, P)$, where E is an equivalence relation on A and P is a unary predicate.

Definition

- Let $P[a_i]$ be the set of elements $a_j \in [a_i]$ such that $P(a_j)$ holds ("a_j is colored")
- Let $\bar{P}[a_i]$ be the set of elements $a_j \in [a_i]$ such that $\neg P(a_j)$ holds ("a_j is non-colored")
- Type of an equivalence class X of $\mathcal{A}$: $tp(X) = (i, j)$, where i is the number of elements e in X such that $P(e)$ holds and j is the number of elements e in X such that $\neg P(e)$ holds

Colored and non-colored equivalences

Remind: for $m, n, t \in \mathbb{N}$, $m =_t n$ iff $m = n$ or both $m, n > t$

Definition

Two types (i, j) and (i', j') are *colored* n*-equivalent*, denoted by $(i, j) \equiv_n^C (i', j')$ if, and only if,

1. $i =_{n-1} i'$
2. $j =_{n-2} j'$

Definition

Two types (i, j) and (i', j') are *non-colored* n*-equivalent*, denoted by $(i, j) \equiv_n^N (i', j')$ if, and only if,

1. $i =_{n-2} i'$
2. $j =_{n-1} j'$

Lemma

If either $(i, j) \equiv_n^C (i', j')$ *or* $(i, j) \equiv_n^N (i', j')$, *then* $(i, j) \equiv_{n-1}^C (i', j')$ *and* $(i, j) \equiv_{n-1}^N (i', j')$

Counting up to colored equivalence

- For structures $\mathcal{A}$ and $\mathcal{B}$, type (i,j) and $k \geqslant 1$,

$$C^{\mathcal{A}}_{(i,j),k} \stackrel{\text{def}}{=} \{X \mid X \text{ is an equivalence class of } \mathcal{A} \text{ and } tp(X) \equiv^{C}_{k} (i,j)\}$$

- Let $q^{\mathcal{A},C}_{(i,j),k} \stackrel{\text{def}}{=} |C^{\mathcal{A}}_{(i,j),k}|$
- Let $q^{C}_{(i,j),k} \stackrel{\text{def}}{=} \min(q^{\mathcal{A},C}_{(i,j),k}, q^{\mathcal{B},C}_{(i,j),k})$
- Let $\mathcal{A}^{C}((i,j),k)$ be the structure obtained from $\mathcal{A}$ by removing $q^{C}_{(i,j),k}$ equivalence classes in $C^{\mathcal{A}}_{(i,j),k}$
- $N^{\mathcal{A}}_{(i,j),k}$, $q^{\mathcal{A},N}_{(i,j),k}$, $q^{N}_{(i,j),k}$ and $\mathcal{A}^{N}((i,j),k)$ are defined similarly w.r.t. $\equiv^{N}_{k}$

Colored and non-colored disparity

Definition

We say that a *colored disparity* occurs in a game $\mathcal{G}_n(\mathcal{A},\mathcal{B})$ if there exists a type (i,j) and $n > k \geqslant 0$ such that the following holds:

1. $k = q^{C}_{(i,j),n-k}$
2. In one of $\mathcal{A}^{C}((i,j),n-k)$ and $\mathcal{B}^{C}((i,j),n-k)$, there is an equivalence class whose type is colored $(n-k)$-equivalent to (i,j), and no such equivalence class exists in the other structure

Non-colored disparity is defined in a similar way.

Theorem

***II** has a winning strategy in $\mathcal{G}_n(\mathcal{A},\mathcal{B})$ if and only if neither colored disparity nor non-colored disparity occurs.*

Proof's idea

(⇒)

- Assume that colored disparity occurs for some (i, j) and k
- W.l.o.g, suppose that in $\mathcal{A}^C((i, j), n-k)$ there is an equivalence class whose type is colored $(n-k)$-equivalent to (i, j), and no such class exists in $\mathcal{B}^C((i, j), n-k)$
- First, **I** chooses $k = q^C_{(i,j),n-k}$ mutually non-equivalent elements in $C^{\mathcal{A}}_{(i,j),n-k}$
- Then, **I** selects a colored element in a class X with $tp(X) \equiv^C_{n-k} (i, j)$ and he plays the rest of the game inside it

(⇐)

- Describe a winning strategy for **II** that maintains the following invariant: at round k, for $1 \leqslant l, m \leqslant k$,
  1. a_l is colored iff b_l is colored
  2. $E(a_l, a_m)$ iff $E(b_l, b_m)$
  3. $(i_l, j_l) \equiv^C_{n-l} (i'_l, j'_l)$ and $(i_l, j_l) \equiv^N_{n-l} (i'_l, j'_l)$
  4. $\mathcal{G}_{n-k}$ has neither colored disparity nor non-colored disparity

Embedded equivalence structures: local strategy

Definition
Structures $\mathcal{A} = (A, E_1, \ldots, E_h)$, where each E_i is an equivalence relation on A and $E_i \subseteq E_j$ for $i < j$.

- We consider the case $h = 2$
- Let $\mathcal{A} = (A, E_1, E_2)$ and $\mathcal{B} = (B, E_1, E_2)$

Definition
A local game on $(\mathcal{A}, \vec{a}, \mathcal{B}, \vec{b})$ is a game played only within non-free equivalence classes, i.e., classes containing some $a_i \in \vec{a}$ or $b_i \in \vec{b}$.

Definition
$(\mathcal{A}, \vec{a}, \mathcal{B}, \vec{b})$ is t-locally safe iff **II** has a winning strategy in the t-round local game on $(\mathcal{A}, \vec{a}, \mathcal{B}, \vec{b})$.

- t-round local games are characterized as in "flat" equivalence games

Embedded equivalence structures: global strategy

Definition

- Type of an E_2-class X of $\mathcal{A}$: $\mathrm{tp}(X) = (q_1, \dots, q_t)$, if the largest E_1-equivalence class in X has size t and, for all $1 \leqslant i \leqslant t$, q_i is the number of E_1-classes of size i in X
- $\mathrm{tp}(X) \equiv_t \mathrm{tp}(Y)$ iff **II** wins $\mathcal{G}_t((X, E_1 \restriction X), (Y, E_1 \restriction Y))$
- (Free) t-multiplicity of type σ in $(\mathcal{A}, \vec{a})$:

$$q_{\sigma,t}^{(\mathcal{A},\vec{a})} \stackrel{\text{def}}{=} \left|\{ Y \mid Y \text{ is a free } E_2\text{-class of } (\mathcal{A}, \vec{a}) \wedge \mathrm{tp}(Y) \equiv_t \sigma \}\right|$$

- $\Delta_{(\mathcal{B},\vec{b})}^{(\mathcal{A},\vec{a})} = \{ (\sigma, t) \mid q_{\sigma,t}^{(\mathcal{A},\vec{a})} \neq q_{\sigma,t}^{(\mathcal{B},\vec{b})} \}$

Lemma

Given $(\mathcal{A}, \vec{a}, \mathcal{B}, \vec{b})$ *and* $(\sigma, t) \in \Delta_{(\mathcal{B},\vec{b})}^{(\mathcal{A},\vec{a})}$, ***I*** *has a winning strategy in* $\min\{ q_{\sigma,t}^{(\mathcal{A},\vec{a})}, q_{\sigma,t}^{(\mathcal{B},\vec{b})} \} + 1 + t$ *rounds.*

- A complete characterization can be given

Trees with height h

Definition

A tree $\mathcal{T}$ is a pair $(T, \preceq)$ where

1. $\preceq$ is a partial ordering with a unique minimum
2. for all $x \in T$, $\{ y \mid y \preceq x \}$ is finite and linearly ordered
3. maximal elements are leaves
4. Level of a node: distance from the root
5. Height of $\mathcal{T}$: number of levels -1

- $\mathcal{K}_h$: class of trees of height h
- $x \preceq y$ iff x is an ancestor of y
- The idea of Khoussainov and Liu's paper is to map $\mathcal{K}_h$ into the class of embedded equivalence relations of height h
- Sounds nice!
- Unfortunately, it does not work (without a level predicate)

Mapping trees onto embedded equivalences

- $T' \stackrel{\text{def}}{=} T \cup \{ a_x \mid x \text{ is a leaf of } \mathcal{T} \}$
- E_1: minimal equivalence containing $\{ (x, a_x) \mid x \text{ is a leaf of } \mathcal{T} \}$
- E_{i+1}: minimal equivalence containing $E_i \cup (T_1 \times T_1) \cup \cdots \cup (T_k \times T_k)$, where $T_1, \ldots, T_k$ are the subtrees rooted at nodes of level $h - i + 1$
- $E_i \subseteq E_{i+1}$ (E_i is finer than E_{i+1})
- Embedded equivalence structure induced by $\mathcal{T}$:

$$\mathcal{A}(\mathcal{T}) \stackrel{\text{def}}{=} (T', E_1, \ldots, E_h)$$

Claim

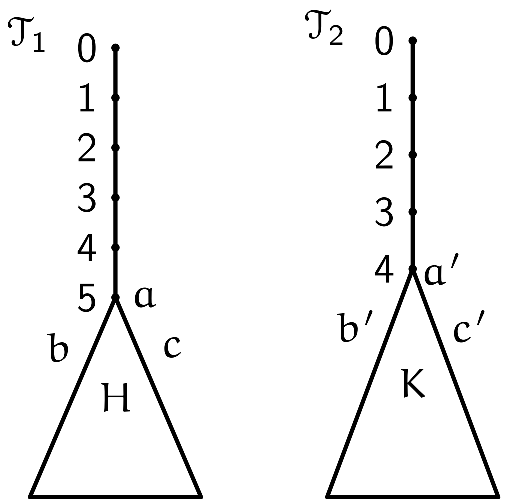
1. *$\mathcal{T}_1 \cong \mathcal{T}_2$ iff $\mathcal{A}(\mathcal{T}_1) \cong \mathcal{A}(\mathcal{T}_2)$ (ok!)*
2. ***II** wins $\mathcal{G}_m(\mathcal{T}_1, \mathcal{T}_2)$ iff **II** wins $\mathcal{G}_m(\mathcal{A}(\mathcal{T}_1), \mathcal{A}(\mathcal{T}_2))$ (wrong!)*

Why it does not work

Claim (wrong)

***II** wins $\mathcal{G}_m(\mathcal{T}_1, \mathcal{T}_2)$ iff **II** wins $\mathcal{G}_m(\mathcal{A}(\mathcal{T}_1), \mathcal{A}(\mathcal{T}_2))$.*

- Observe that $x \preceq y$ iff x has level t, y has level $s \geqslant t$ and $E_{h-t+1}(x, y)$
- Every winning strategy for **II** in $\mathcal{G}_m(\mathcal{A}(\mathcal{T}_1), \mathcal{A}(\mathcal{T}_2))$ must map elements at level k in $\mathcal{A}(\mathcal{T}_1)$ to elements at level k in $\mathcal{A}(\mathcal{T}_2)$
- How to fix the correspondence? Enrich the tree structure with a level predicate



Labelled successor structures (LSS)

- Let Σ be a finite alphabet
- Let $u \in \Sigma^*$ be a word on Σ
- Let $u[i]$ be the ith letter of u

Definition

A (labelled) successor structure is a pair $(u, \vec{i})$, where the elements of $\vec{i}$ are distinguished indices of u.
Successor structures $(u, \vec{i})$ interpret FO-formulas $\phi(\vec{x})$ in the vocabulary $(=, s, (P_a)_{a \in \Sigma})$ according to the following rules:

$$\begin{aligned} (u, \vec{i}) &\models x_h = x_l && \text{if } i_h = i_l; \\ (u, \vec{i}) &\models s(x_h, x_l) && \text{if } i_l = i_h + 1; \\ (u, \vec{i}) &\models P_a(x_h) && \text{if } u[i_h] = a. \end{aligned}$$

Local conditions

$$\eta_d(i, j) = \begin{cases} j - i & \text{if } |i - j| \leqslant d; \\ \infty & \text{otherwise.} \end{cases}$$

Definition

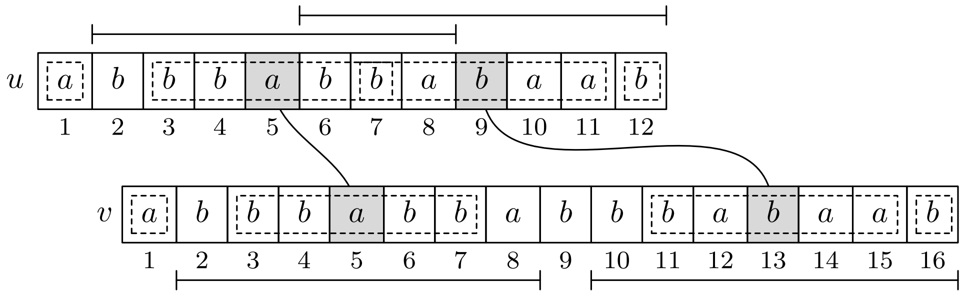
A configuration $(u, \vec{i}, v, \vec{j})$ is t-locally safe iff, for all $i_h, i_l \in \vec{i}$,

1. $\eta_{2^t}(i_h, i_l) = \eta_{2^t}(j_h, j_l)$
2. $\mathcal{N}^u_{2^t-1}(i_h) = \mathcal{N}^v_{2^t-1}(j_h)$

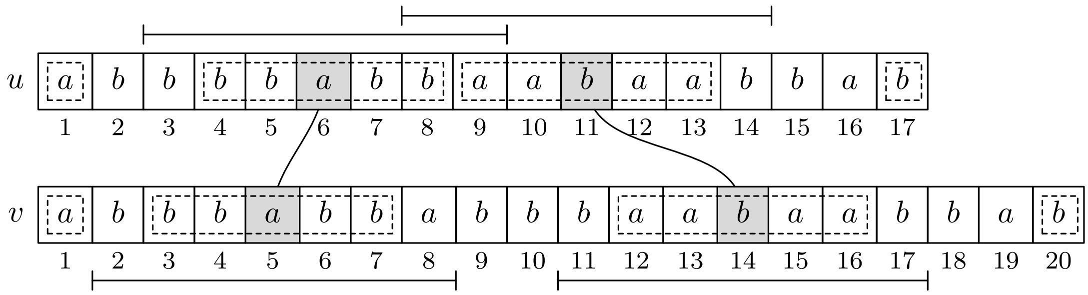
- If a configuration is not t-locally safe, **I** has a "local" winning strategy in t rounds
- **II** can turn a t-locally safe configuration into a $(t-1)$-locally safe configuration if **I** plays "locally"

Local safety: an example

Not 2-locally safe:



2-locally safe:

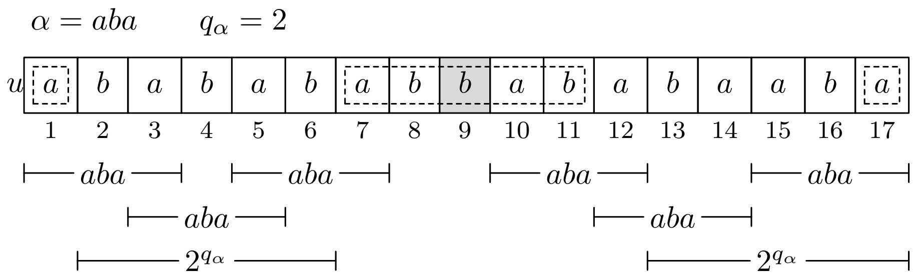


Free factors

Definition

- Let α be a word of length $2^t - 1$
- An occurrence of α centered at index k in $(u, \vec{i})$ is free iff $|k - \vec{i}| > 2^{t-1}$
- (Free) multiplicity of α in $(u, \vec{i})$: number of free occurrences of α in $(u, \vec{i})$
- Scattering of α in $(u, \vec{i})$: cardinality of a maximal 2^t-scattered subset of the free occurrences of α in $(u, \vec{i})$
- (A set $X \in \mathbb{N}$ is d-scattered iff $|x - y| > d$ for all $x, y \in X$)

Multiplicity and scattering: an example



- Let $\alpha = aba$ $(t = 2)$
- Centers of free occurrences of aba in $(u, 9)$: $\{2, 4, 6, 13, 16\}$
- Multiplicity: 5
- Scattering: 2 $(\{2, 4, 6\}, \{13, 16\})$

LSS: Characterization

- Let $p_\alpha^{(u,\vec{i})}$ denote the free multiplicity
- Let $q_\alpha^{(u,\vec{i})}$ denote the scattering
- Let $\Delta_{(v,\vec{j})}^{(u,\vec{i})} = \{\, \alpha \mid p_\alpha^{(u,\vec{i})} \neq p_\alpha^{(v,\vec{j})} \vee q_\alpha^{(u,\vec{i})} \neq q_\alpha^{(v,\vec{j})} \,\}$
- $\Delta_{(v,\vec{j})}^{(u,\vec{i})}$ is the set of words that **I** can potentially exploit in order to win
- Let $q_\alpha = \min\{\, q_\alpha^{(u,\vec{i})}, q_\alpha^{(v,\vec{j})} \,\}$

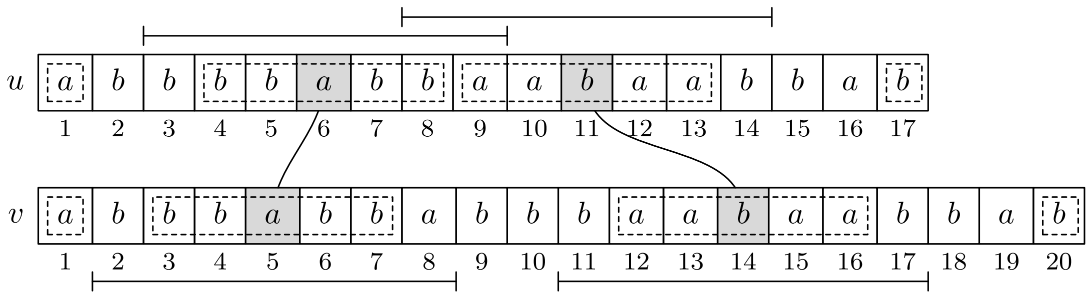
Definition

A configuration $(u, \vec{i}, v, \vec{j})$ is m-globally safe iff $q_\alpha > m - \log_2(|\alpha| + 1)$ for all words $\alpha \in \Delta_{(v,\vec{j})}^{(u,\vec{i})}$.

Theorem

II *has a winning strategy in* $\mathcal{G} = \mathcal{G}_m(u, \vec{i}, v, \vec{j})$ *iff* $\mathcal{G}$ *is* m*-locally safe and* m*-globally safe.*

Example



	α	$p_\alpha^{(u,6,11)}$	$p_\alpha^{(v,5,14)}$	$q_\alpha^{(u,6,11)}$	$q_\alpha^{(v,5,14)}$
$q = 1$	a	4	5	4	5
	b	7	9	4	5
$q = 2$	abb	2	3	2	3
	bab	1	2	1	2
	bba	1	2	1	2
	bbb	1	1	1	1

It is also 2-globally safe!

Definability and m-equivalence

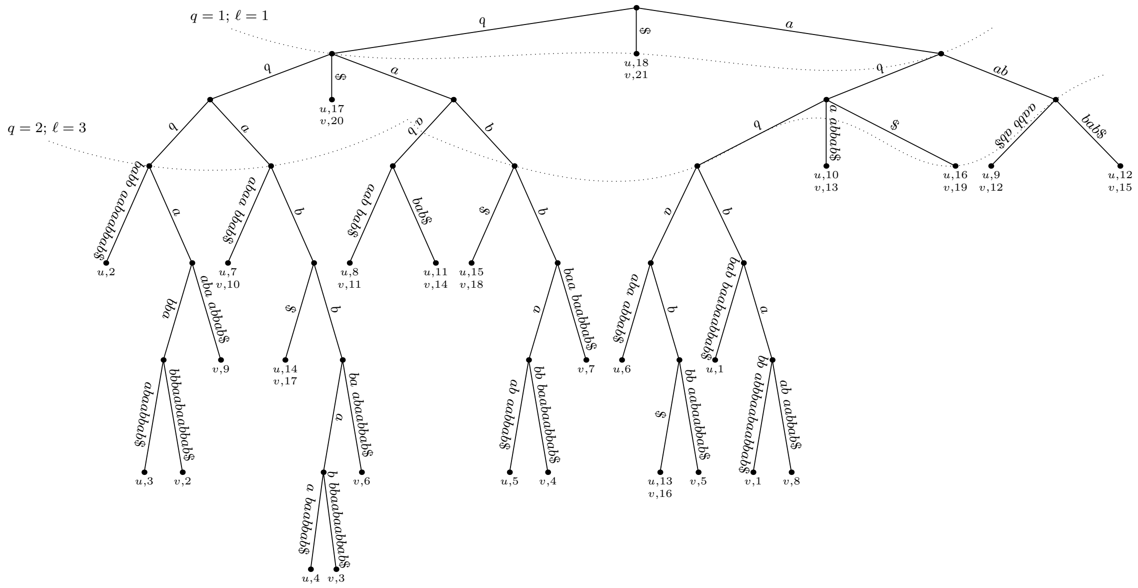
$\mathcal{L}$	Definable class	m-equivalence
FO(s)	threshold locally testable	Previous theorem

- From FO(s) to FO$(<)$:

$$\text{FO}(<_p), \text{ where } x <_p y \Leftrightarrow 0 < y - x \leqslant p.$$

$\mathcal{L}$	Definable class	m-equivalence
FO$(<)$	$*$-free	$<_p$, with $p \to \infty$

Testing $\equiv_m$ with generalized suffix trees



- Let $n = |u| + |v|$
- Remoteness of $\mathcal{G}(u, v)$: $O(n \log n)$ time and space
- **I**'s optimal moves: $O(n^2 \log n)$ time, $O(n \log n)$ space
- **II**'s optimal moves: $O(n)$ time and space (if the remoteness is known)

An emerging pattern

Let $\mathcal{A}$ and $\mathcal{B}$ arbitrary structures.

Definition

A t-round local game on $(\mathcal{A}, \vec{a}, \mathcal{B}, \vec{b})$ is a game played on $\mathcal{N}^{\mathcal{A}}_{2^t-1}(\vec{a})$ and $\mathcal{N}^{\mathcal{B}}_{2^t-1}(\vec{b})$ such that, at round $t - k + 1$, with $1 \leqslant k \leqslant t$, **I** must choose an element at distance at most 2^{k-1} from $\vec{a}$ or from $\vec{b}$.

Definition

A configuration $(\mathcal{A}, \vec{a}, \mathcal{B}, \vec{b})$ is t-locally safe if **II** has a winning strategy in the t-round local game on $(\mathcal{A}, \vec{a}, \mathcal{B}, \vec{b})$.

- We write $(\mathcal{A}, \vec{a}) \equiv^{loc}_t (\mathcal{B}, \vec{b})$
- **II** can play t rounds provided that **I** plays "near" distinguished elements (nearer and nearer after each round)

How to count neighborhoods?

- The analysis of equivalence structures shows that we need to count up to isomorphism and up to $\equiv_t^{loc}$-equivalence (in equivalence structures, neighborhoods coincide with equivalence classes; two equivalence classes are isomorphic iff they have the same number of elements and they are $\equiv_t^{loc}$-equivalent iff they both have at least t elements)
- The analysis of labelled successor structures shows that we need to count both the (free) multiplicity and the scattering of neighborhoods (for equivalence structures, the two notions collapse into one)

Conjecture

Counting the multiplicity and scattering of "small" neighborhoods up to isomorphism and up to $\equiv_t^{loc}$-equivalence is enough for characterizing the "global" winning strategy for arbitrary structures.