MathLogAps PhD training — summer school Aussois, June 2007

# Game Based Methods
# and the Model Theory of Fragments of FO
# over Special Classes of (Finite) Structures

Martin Otto, Technische Universität Darmstadt
www.mathematik.tu-darmstadt.de/~otto

# Overview

## Part I: Ingredients

### Part I A: Games and Ehrenfeucht–Fraïssé Techniques

- Model checking games
- Back & Forth games, FO Ehrenfeucht–Fraïssé
- Modularity and Locality: Hanf, Gaifman
- Variations

### Part I B: Some Fragments of First-Order Logic

and some extensions, too

- Universal, existential and finite-variable fragments
- The modal fragment and bisimulation
- MSO and fixed points as a frame of reference

# Part II: Two Model Theoretic Themes

## Part II A: Preservation and Expressive Completeness

- Expressive completeness issues: classical and elsewhere
- Game based model constructions vs. classical arguments
- Limited variants of classical theorems

## Part II B: Relational Recursion

- Fixed point recursion
- Boundedness and related algorithmic issues

## I A: Games and Ehrenfeucht–Fraïssé Techniques

**Q1: Is $\mathfrak{A} \models \varphi$ ?**

model checking problem MC($L$):

given (finite) $\mathfrak{A}$ and $\varphi \in L$,

decide whether $\mathfrak{A} \models \varphi$

**Q2: What can be expressed in $L$ ?**

definability, expressive power, measured against, e.g.,

- other logics
- semantic criteria
- complexity criteria

$\longrightarrow$ development of *model checking games* and
model theoretic *comparison games*

later link the two via bisimulation

## the model checking game for $\mathsf{FO}^k$

as a general proviso: all vocabularies finite & relational

$\mathsf{FO}^k$: FO with variables $x_1, \ldots, x_k$ only
[every formula defines a $k$-ary predicate]

## the model checking game $\mathrm{MC}^{\boldsymbol{k}}(\mathfrak{A})$

**players**: $\mathbf{I}/\mathbf{II}$ with roles as *verifier* vs. *falsifier*

**positions**: $(\boldsymbol{a}, \varphi, \wp) \in A^k \times \mathsf{FO}^k \times \{\mathbf{I}, \mathbf{II}\}$
$\boldsymbol{a}$: assignment to $\boldsymbol{x} = (x_1, \ldots, x_k)$
$\wp$: verifier claiming $\mathfrak{A} \models \varphi[\boldsymbol{a}]$
$\overline{\wp}$: falsifier claiming $\mathfrak{A} \not\models \varphi[\boldsymbol{a}]$

**moves**: depending on $\varphi$ and $\wp$,
$\wp$ or $\overline{\wp}$ chooses successor position

**end**: in positions $(\boldsymbol{a}, \varphi, \wp)$ with atomic $\varphi$:
$\wp$ wins if $\mathfrak{A} \models \varphi[\boldsymbol{a}]$ $\quad$ $\overline{\wp}$ wins if $\mathfrak{A} \not\models \varphi[\boldsymbol{a}]$

## the natural protocol for moves in $\mathrm{MC}^{\boldsymbol{k}}(\mathfrak{A})$

reflecting inductive definition of semantics

in position $(\boldsymbol{a}, \varphi, \wp)$:

| | |
|---|---|
| $\varphi = \varphi_1 \wedge \varphi_2$ | $\overline{\wp}$'s move:<br>$\overline{\wp}$ moves to $(\boldsymbol{a}, \varphi_1, \wp)$ or to $(\boldsymbol{a}, \varphi_2, \wp)$ |
| $\varphi = \varphi_1 \vee \varphi_2$ | $\wp$'s move:<br>$\wp$ moves to $(\boldsymbol{a}, \varphi_1, \wp)$ or to $(\boldsymbol{a}, \varphi_2, \wp)$ |
| $\varphi = \forall x_i \psi$ | $\overline{\wp}$'s move:<br>$\overline{\wp}$ moves to $(\boldsymbol{a}\frac{a}{i}, \psi, \wp)$ for some $a \in A$ |
| $\varphi = \exists x_i \psi$ | $\wp$'s move:<br>$\wp$ moves to $(\boldsymbol{a}\frac{a}{i}, \psi, \wp)$ for some $a \in A$ |
| $\varphi = \neg \psi$ | no-one's move:<br>game continues from $(\boldsymbol{a}, \psi, \overline{\wp})$ |

**Theorem:** $\wp$ **has winning strategy in** $(a, \varphi, \wp)$ **iff** $\mathfrak{A} \models \varphi[a]$

## model checking game and model checking complexity

consider *combined complexity* of deciding $\mathfrak{A} \models \varphi[\boldsymbol{a}]$
in terms of input size $||\mathfrak{A}, \boldsymbol{a}|| + ||\varphi||$

strategy search in (game graph associated with)
model checking game leads to

- **Ptime** algorithm for **model checking FO$^{\boldsymbol{k}}$**
  the problem is Ptime complete for fixed $k$
- **Pspace** algorithm for **model checking FO**
  the problem is Pspace complete

with many variations for other logics,
often yielding algorithms of optimal worst case complexity

## model theoretic comparison games: Ehrenfeucht–Fraïssé

recall general proviso: all vocabularies finite & relational

how similar are $\mathfrak{A}, \boldsymbol{a}$ and $\mathfrak{B}, \boldsymbol{b}$ ?

### the FO Ehrenfeucht–Fraïssé game $\mathrm{G}(\mathfrak{A}, \boldsymbol{a}; \mathfrak{B}, \boldsymbol{b})$

**players**: $\mathbf{I}/\mathbf{II}$ *challenger / defender* of similarity claim

**positions**: $(\mathfrak{A}, \boldsymbol{a}; \mathfrak{B}, \boldsymbol{b})$, $\boldsymbol{a}, \boldsymbol{b} \in \bigcup_n A^n \times B^n$

$$\left.\begin{array}{r} \boldsymbol{a} = (a_1, \ldots, a_n) \\ \boldsymbol{b} = (b_1, \ldots, b_n) \end{array}\right\} \text{ marked in } \mathfrak{A}/\mathfrak{B} \text{ with pebbles}$$

**single round**: $\mathbf{I}$ chooses to play in $\mathfrak{A}$ or $\mathfrak{B}$
and places next pebble in that structure
$\mathbf{II}$ must place pebble in opposite structure

net effect: $(\mathfrak{A}, \boldsymbol{a}; \mathfrak{B}, \boldsymbol{b}) \longmapsto (\mathfrak{A}, \boldsymbol{a}a; \mathfrak{B}, \boldsymbol{b}b)$

**win/lose**: $\mathbf{II}$ loses in $(\boldsymbol{a}; \boldsymbol{b})$ if
$p\colon \boldsymbol{a} \mapsto \boldsymbol{b}$ *not* a local isomorphism $p\colon \mathfrak{A}{\restriction}\boldsymbol{a} \simeq \mathfrak{B}{\restriction}\boldsymbol{b}$

## Ehrenfeucht–Fraïssé game and elementary equivalence

$G^m(\mathfrak{A}, \boldsymbol{a}; \mathfrak{B}, \boldsymbol{b})$: $m$-round game starting from $(\mathfrak{A}, \boldsymbol{a}; \mathfrak{B}, \boldsymbol{b})$
**II** wins if she survives $m$ rounds

$G^\infty(\mathfrak{A}, \boldsymbol{a}; \mathfrak{B}, \boldsymbol{b})$: unbounded game starting from $(\mathfrak{A}, \boldsymbol{a}; \mathfrak{B}, \boldsymbol{b})$
**II** wins if she can respond indefinitely

## degrees of similarity in terms of game:

$\mathfrak{A}, \boldsymbol{a} \simeq_m \mathfrak{B}, \boldsymbol{b}$ :⇔ **II** has winning strategy in $G^m(\mathfrak{A}, \boldsymbol{a}; \mathfrak{B}, \boldsymbol{b})$

$\mathfrak{A}, \boldsymbol{a} \simeq_\omega \mathfrak{B}, \boldsymbol{b}$ :⇔ **II** has winning strategy in all $G^m(\mathfrak{A}, \boldsymbol{a}; \mathfrak{B}, \boldsymbol{b})$

$\mathfrak{A}, \boldsymbol{a} \simeq_\infty \mathfrak{B}, \boldsymbol{b}$ :⇔ **II** has winning strategy in $G^\infty(\mathfrak{A}, \boldsymbol{a}; \mathfrak{B}, \boldsymbol{b})$

## degrees of elementary indistinguishability:

$\mathfrak{A}, \boldsymbol{a} \equiv_m \mathfrak{B}, \boldsymbol{b}$ : eq. in FO up to quantifier rank $m$

$\mathfrak{A}, \boldsymbol{a} \equiv \mathfrak{B}, \boldsymbol{b}$ : eq. in FO

$\mathfrak{A}, \boldsymbol{a} \equiv_\infty \mathfrak{B}, \boldsymbol{b}$ : eq. in infinitary first-order logic $\mathrm{FO}_\infty = L_{\infty\omega}$

## Ehrenfeucht–Fraïssé and Karp Theorems:

$$\mathfrak{A}, \boldsymbol{a} \simeq_m \mathfrak{B}, \boldsymbol{b} \quad \Leftrightarrow \quad \mathfrak{A}, \boldsymbol{a} \equiv_m \mathfrak{B}, \boldsymbol{b} \qquad (*)$$

$$\mathfrak{A}, \boldsymbol{a} \simeq_\omega \mathfrak{B}, \boldsymbol{b} \quad \Leftrightarrow \quad \mathfrak{A}, \boldsymbol{a} \equiv \mathfrak{B}, \boldsymbol{b}$$

$$\mathfrak{A}, \boldsymbol{a} \simeq_\infty \mathfrak{B}, \boldsymbol{b} \quad \Leftrightarrow \quad \mathfrak{A}, \boldsymbol{a} \equiv_\infty \mathfrak{B}, \boldsymbol{b}$$

moreover $\left\{ \begin{array}{c} \equiv \text{ and } \equiv_\infty \\ \\ \simeq_\omega \text{ and } \simeq_\infty \end{array} \right\}$ coincide in $\omega$-saturated structures

classical completeness test

## proof ingredients for $(*)$:

$(\Rightarrow)$ $\mathfrak{A}, \boldsymbol{a} \not\equiv_m \mathfrak{B}, \boldsymbol{b} \Rightarrow$ $\mathbf{I}$ has won, or can force $\mathfrak{A}, \boldsymbol{a}a \not\equiv_{m-1} \mathfrak{B}, \boldsymbol{b}b$ in one round

$(\Leftarrow)$ $\simeq_m$-class of $\mathfrak{A}, \boldsymbol{a}$ definable by qr $m$ formula $\chi(\boldsymbol{x}) = \chi^m_{\mathfrak{A},\boldsymbol{a}}$ describing back-and-forth conditions

s.t. $\quad \mathfrak{B} \models \chi[\boldsymbol{b}] \quad \Leftrightarrow \quad \mathfrak{B}, \boldsymbol{b} \simeq_m \mathfrak{A}, \boldsymbol{a}$

## formalising the back-and-forth conditions (inductively)

$$\chi^{m+1}_{\mathfrak{A},\boldsymbol{a}}(\boldsymbol{x}) =$$

$$\underbrace{\bigwedge_{a\in A} \exists y \;\; \chi^{m}_{\mathfrak{A},\boldsymbol{a}a}(\boldsymbol{x},y)}_{\text{forth: responses for challenges in } \mathfrak{A}} \quad \wedge \quad \underbrace{\forall y \; \bigvee_{a\in A} \;\; \chi^{m}_{\mathfrak{A},\boldsymbol{a}a}(\boldsymbol{x},y)}_{\text{back: responses for challenges in } \mathfrak{B}}$$

**forth**:
responses for
challenges in $\mathfrak{A}$

**back**:
responses for
challenges in $\mathfrak{B}$

NB: $\bigwedge$ and $\bigvee$ effectively finite even for infinite $A$!

$$\mathfrak{B} \models \chi^{m+1}_{\mathfrak{A},\boldsymbol{a}}[\boldsymbol{b}] \;\Leftrightarrow\; \mathfrak{B},\boldsymbol{b} \simeq_{m+1} \mathfrak{A},\boldsymbol{a}$$

## inexpressibility via games: example

the class of even length finite linear orderings is not FO-definable (among the class of finite linear orderings)

show that for all sufficiently large lengths $n, n'$:

$$\boxed{\mathfrak{A} = \big(\{1, \dots, n\}, <\big) \simeq_m \big(\{1, \dots, n'\}, <\big) = \mathfrak{B}}$$

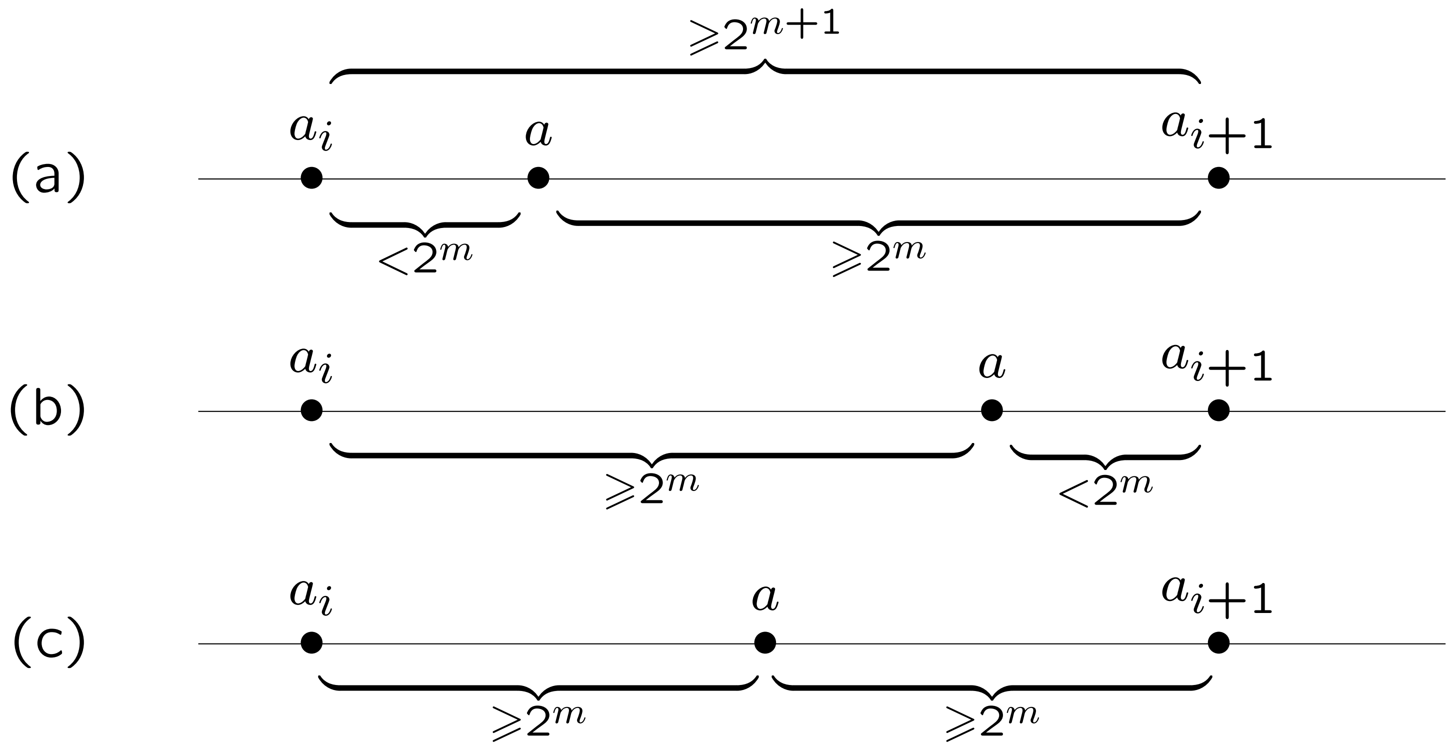
**II** can survive $m$ rounds from any position $(\mathfrak{A}, \boldsymbol{a}; \mathfrak{B}, \boldsymbol{b})$ such that

$$0 < a_1 < a_2 < \cdots < a_s < n+1$$

$$0 < b_1 < b_2 < \cdots < b_s < n'+1$$

with corresponding intervals of same length, or lengths $\geqslant 2^m$

**how to respond to challenge $a \in (a_i, a_{i+1})$ with $m$ further rounds to play**




in each case, **II** finds adequate response in $(b_i, b_{i+1})$ if similarly $b_{i+1} - b_i \geqslant 2^{m+1}$

## parity of finite linear orders not FO-definable:

$$\big(\{1, \ldots, 2^m\}, <\big) \ \simeq_m \ \big(\{1, \ldots, 2^m + 1\}, <\big)$$

## corollaries, via simple interpretations

also not definable in FO, e.g.:

- 2-colourability (of finite graphs)
- connectivity (of finite graphs)

cf. classical arguments (via compactness)
which only show non-definability over all graphs

## locality and modularity of games

sufficient conditions for $\simeq_q$ in suitable positions

### Gaifman graph and distance

with relational $\mathfrak{A} = (A, R^{\mathfrak{A}}, \dots)$ associate undirected graph $G(\mathfrak{A})$ on $A$ with edge $\{a, a'\}$ if $a \neq a'$ and $a, a' \in \boldsymbol{a}$ for some $\boldsymbol{a} \in R^{\mathfrak{A}}$

- $d(a, a')$: graph distance in $G(\mathfrak{A})$
- $N^\ell(a) := \{a' \in A \colon d(a, a') \leqslant \ell\}$ the *$\ell$-neighbourhood* of $a$;
  $N^\ell(\boldsymbol{a}) := \bigcup_i N^\ell(a_i)$
- $a_1, \dots, a_m$ *$\ell$-scattered* if $d(a_i, a_j) > 2\ell$ for $i \neq j$

the theorems of Hanf and Gaifman establish $\simeq_q$
on the basis of suitable degrees of local similarity

**modularity of E-F game w.r.t. Gaifman locality**

## theorems of Hanf and Gaifman

modularity of game in terms of local views:

**Hanf**: same numbers of realisations for each local isomorphism type *FMT only*

**Gaifman**: indistinguishability w.r.t. local behaviour near distinguished parameters and of scattered tuples up to some radius/size/quantifier rank

## Hanf's theorem

finite relational $\mathfrak{A}$ and $\mathfrak{B}$ are $\ell$-Hanf-equivalent, $\mathfrak{A} \approx^{\ell}_{\text{Hanf}} \mathfrak{B}$,
if for all isomorphism types $\iota$:

$$|\{a \in A \colon \mathfrak{A}\restriction N^{\ell}(a) \simeq \iota\}| = |\{b \in B \colon \mathfrak{B}\restriction N^{\ell}(b) \simeq \iota\}|$$

let $\ell_0 := 0$ and $\ell_{k+1} = 3\ell_k + 1$ for $k \leqslant q$, $\mathfrak{A} \approx^{\ell_q}_{\text{Hanf}} \mathfrak{B}$,

then $\mathbf{II}$ can survive for $k$ rounds from positions
$(\mathfrak{A}, \boldsymbol{a}; \mathfrak{B}, \boldsymbol{b})$ such that $\mathfrak{A}\restriction N^{\ell_k}(\boldsymbol{a}), \boldsymbol{a} \simeq \mathfrak{B}\restriction N^{\ell_k}(\boldsymbol{b}), \boldsymbol{b}$

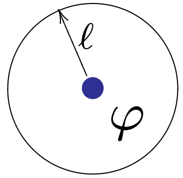
in particular: $\boxed{\mathfrak{A} \approx^{\ell_q}_{\text{Hanf}} \mathfrak{B} \quad \Rightarrow \quad \mathfrak{A} \simeq_q \mathfrak{B}}$

**example:**

connectivity of finite graphs not definable in existential MSO

## levels of local equivalence: Gaifman-equivalence

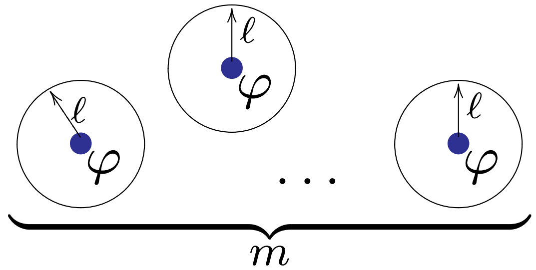
**(L) local FO formulae**: $\varphi^\ell(\boldsymbol{x}) := [\varphi(\boldsymbol{x})]^{N^\ell(\boldsymbol{x})}$




relativisation to $N^\ell(\boldsymbol{x})$
asserting local properties about $\boldsymbol{x}$

**(S) basic local FO sentences**:

asserting existence of $\ell$-scattered $m$-tuple within some $\varphi^\ell[\mathfrak{A}]$




$\mathfrak{A}, \boldsymbol{a} \equiv^\ell_{q,m} \mathfrak{B}, \boldsymbol{b}$: **(L)/(S)** agreement to $\begin{cases} \text{radius} & \ell \\ \text{qfr rank} & q \\ \text{scatter size} & m \end{cases}$

**finite index approximation to $\equiv$**
based on local properties / scattered tuples view

## Gaifman's theorem

- **every FO-formula $\varphi(x)$ equivalent to boolean comb. of local formulae (L) and basic local sentences (S)**
- **every FO-formula $\varphi(x)$ is preserved under $\equiv^{\boldsymbol{\ell}}_{\boldsymbol{q,m}}$ for sufficiently large parameters $\boldsymbol{\ell, q, m}$**

use the $\equiv^{\ell}_{q,m}$ as locality-sensitive finite index approximations to $\equiv$

## proof: modularity of strategies

in $\mathfrak{A} \equiv^{L}_{Q,m} \mathfrak{B}$ [**(S)**-conditions]

**II** has choices to lead game in one round from

$$\mathfrak{A}\restriction N^{\ell_{k+1}}(\boldsymbol{a}),\boldsymbol{a} \;\equiv_{q_{k+1}}\; \mathfrak{B}\restriction N^{\ell_{k+1}}(\boldsymbol{b}),\boldsymbol{b}$$

to $$\mathfrak{A}\restriction N^{\ell_k}(\boldsymbol{a}a),\boldsymbol{a}a \;\equiv_{q_k}\; \mathfrak{B}\restriction N^{\ell_k}(\boldsymbol{b}b),\boldsymbol{b}b$$ [**(L)**-conditions]

where $|\boldsymbol{a}| = |\boldsymbol{b}| < m$; and w.r.t. suitable sequence $(\ell_k, q_k)$

# I B: Variations and some Fragments of FO

**FO too weak**: connectivity, simple properties of strings, ...

**FO too strong**: $\equiv$ coincides with $\simeq$ in finite structures
SAT(FO) and FINSAT(FO) undecidable

**FO ill-adapted**: no smooth model theory
nor good algorithmic behaviour
over important non-elementary classes

look to alternative logics/levels of expressiveness

and to well-behaved fragments and their extensions
over well-behaved classes of models

## some classical fragments of FO

$\exists^*$**FO**: existential FO — classically associated with extension preservation

$\forall^*$**FO**: universal FO — substructure preservation

$\exists^*$**FO**$^+$: existential positive — homomorphism preservation

## less classical fragments of FO

**FO**$^{\boldsymbol{k}}$: $k$-variable FO
quantitative access restriction

algorithmically relevant
prominent in FMT
non-trivial $\equiv^k$

**ML**: modal logic as a fragment of FO
qualitative access restriction
restricted, relativised quantification

bisimulation preservation
algorithmically tame
smooth FMT

## classical extensions of FO

**MSO**, monadic second-order

interesting level of expressiveness
tractable over important classes

**fixed-point extensions**

adding relational recursion
rather an "extension scheme"
$\longrightarrow$ more in part II

here now look at **$\mathrm{FO}^k$, MSO, ML and their games**

## $\mathrm{FO}^k$ and the $k$-pebble game

**positions**: $(\mathfrak{A}, \boldsymbol{a}; \mathfrak{B}, \boldsymbol{b})$ with $\boldsymbol{a} \in A^k$, $\boldsymbol{b} \in B^k$
$k$ pebbles in each structure

**single round**:

**I** selects one pebble in one structure to move
**II** moves corresponding pebble in opposite structure

net effect: $(\mathfrak{A}, \boldsymbol{a}; \mathfrak{B}, \boldsymbol{b}) \longmapsto (\mathfrak{A}, \boldsymbol{a}\frac{a}{i}; \mathfrak{B}, \boldsymbol{b}\frac{b}{i})$ for round played with pebble $i$

winning conditons as before

$\mathfrak{A}, \boldsymbol{a} \simeq_m^k \mathfrak{B}, \boldsymbol{b}$ $:\Leftrightarrow$ **II** has winning strategy for $m$-round game from position $(\mathfrak{A}, \boldsymbol{a}; \mathfrak{B}, \boldsymbol{b})$

## characteristic formulae for $k$-pebble game

$\chi^m_{\mathfrak{A},\boldsymbol{a}}(\boldsymbol{x}) \in \mathrm{FO}^k$ s.t. $\quad \mathfrak{B}, \boldsymbol{b} \simeq^k_m \mathfrak{A}, \boldsymbol{a} \quad \Leftrightarrow \quad \mathfrak{B} \models \chi^m_{\mathfrak{A},\boldsymbol{a}}[\boldsymbol{b}]$

inductively put

$$\chi^{m+1}_{\mathfrak{A},\boldsymbol{a}}(\boldsymbol{x}) = \chi^m_{\mathfrak{A},\boldsymbol{a}}(\boldsymbol{x}) \wedge$$

$$\bigwedge_{1 \leqslant i \leqslant k} \Big( \underbrace{\bigwedge_{a \in A} \exists x_i \, \chi^m_{\mathfrak{A},\boldsymbol{a}\frac{a}{i}}(\boldsymbol{x})}_{\textbf{forth}\text{: challenges in } \mathfrak{A}} \quad \wedge \quad \underbrace{\forall x_i \bigvee_{a \in A} \chi^m_{\mathfrak{A},\boldsymbol{a}\frac{a}{i}}(\boldsymbol{x})}_{\textbf{back}\text{: challenges in } \mathfrak{B}} \Big)$$

## $\mathrm{FO}^k$ Ehrenfeucht–Fraïssé theorem

$\mathfrak{A}, \boldsymbol{a} \simeq^k_m \mathfrak{B}, \boldsymbol{b}$ iff $\mathfrak{A}, \boldsymbol{a} \equiv^k_m \mathfrak{B}, \boldsymbol{b}$ & variants for $\simeq^k_\omega$ and $\simeq^k_\infty$

remark: over finite $\mathfrak{A}, \mathfrak{B}$: $\mathfrak{A}, \boldsymbol{a} \simeq^k_n \mathfrak{B}, \boldsymbol{b} \Rightarrow \mathfrak{A}, \boldsymbol{a} \simeq^k_\infty \mathfrak{B}, \boldsymbol{b}$ for $n > \max(|A|^k, |B|^k)$

## $\mathrm{FO}^k$ Ehrenfeucht–Fraïssé theorem

$\mathfrak{A}, \boldsymbol{a} \simeq^k_m \mathfrak{B}, \boldsymbol{b}$ iff $\mathfrak{A}, \boldsymbol{a} \equiv^k_m \mathfrak{B}, \boldsymbol{b}$

## examples:

- linear order of length $n$ characterised up to $\simeq$ by $\mathrm{FO}^2$-sentence of qr $n+1$ (Poizat)
- the class of all finite linear orderings is closed under $\simeq^2_\omega$, but not definable in $\mathrm{FO}^2_\infty$ (even among finite structures); transitivity *really* requires 3 variables.

## MSO and its Ehrenfeucht–Fraïssé game

positions $(\mathfrak{A}, \boldsymbol{P}, \boldsymbol{a}; \mathfrak{B}, \boldsymbol{Q}, \boldsymbol{b})$
with marked subsets $\boldsymbol{P}/\boldsymbol{Q}$ (colours) and elements $\boldsymbol{a}/\boldsymbol{b}$ (pebbles)

two kinds of moves: element moves/set moves (**I**'s choice)

everything else entirely analogous,
considering $\equiv_m^{\mathsf{MSO}}$ w.r.t. (mixed) quantifier rank $m$
in relation to $\simeq_m^{\mathsf{MSO}}$ (**II** has strategy for $m$ rounds)

## MSO Ehrenfeucht–Fraïssé theorem

$\mathfrak{A}, \boldsymbol{P}, \boldsymbol{a} \simeq_m^{\mathsf{MSO}} \mathfrak{B}, \boldsymbol{Q}, \boldsymbol{b}$ iff $\mathfrak{A}, \boldsymbol{P}, \boldsymbol{a} \equiv_m^{\mathsf{MSO}} \mathfrak{B}, \boldsymbol{Q}, \boldsymbol{b}$

## example: expressiveness of MSO: Büchi's theorem

| | | |
|---|---|---|
| words over alphabet $\Sigma$ | — | finite linear orderings with monadic colours (for letters) |
| $\Sigma$-languages | — | classes of such word structures |
| run of finite automaton with states $q \in Q$ | — | colouring of word structure with $(P_q)_{q \in Q}$ |

## Büchi's theorem

**regular languages/recognisability by automata**
**=** MSO-**definability over finite linear orderings**

i.e., MSO admits model checking by finite automata
and captures algorithmic power of finite automata

this extends to $\omega$-word-structures and to trees

**MSO:** **modularity of strategies**
**model theoretic (de)composition arguments**

---

here: in the context of word structures

**concatenation/ordered sums:**
for word structures $\mathfrak{A} = (A, <^{\mathfrak{A}}, \boldsymbol{P}^{\mathfrak{A}})$; $\mathfrak{B} = (A, <^{\mathfrak{B}}, \boldsymbol{P}^{\mathfrak{B}})$:

$\mathfrak{A} \oplus \mathfrak{B}$: disjoint union of universes $A$ and $B$
$<^{\mathfrak{A}}$ followed by $<^{\mathfrak{B}}$
disjoint union of $\boldsymbol{P}$

$$\xrightarrow{\mathfrak{A},<^{\mathfrak{A}}} \quad \oplus \quad \xrightarrow{\mathfrak{B},<^{\mathfrak{B}}}$$

**strategy composition**:

$$\mathfrak{A} \equiv_m^{\mathsf{MSO}} \mathfrak{A}' \text{ and } \mathfrak{B} \equiv_m^{\mathsf{MSO}} \mathfrak{B}' \quad \Rightarrow \quad \mathfrak{A} \oplus \mathfrak{B} \equiv_m^{\mathsf{MSO}} \mathfrak{A}' \oplus \mathfrak{B}'$$

$\Rightarrow$ $\equiv_m^{\mathsf{MSO}}$ induces finite index congruence
on the word monoid $(\Sigma^*, \cdot, \epsilon)$

## MSO: consequences of modularity (over word structures)

- $\equiv_m^{\mathsf{MSO}}$ induces finite index congruence on the word monoid $(\Sigma^*, \cdots, \epsilon)$
- MSO model checking by automata
- MSO-definable languages are regular
- pumping arguments for MSO/FO-definable languages
- SAT(MSO) in word models decidable

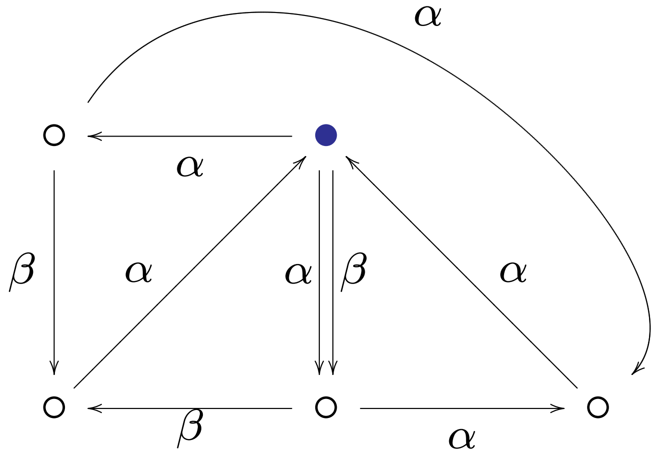
with analogous results for $\omega$-word-models and trees

## ML and the bisimulation game

**the structures**: **edge- and vertex-coloured directed graphs**
transition systems/Kripke structures

$$\mathfrak{A} = \big(\boldsymbol{A}, (\boldsymbol{E_\alpha}), (\boldsymbol{P_i})\big)$$

| | | |
|---|---|---|
| $a \in A$ | nodes | states/possible worlds |
| $E_\alpha^{\mathfrak{A}} \subseteq A^2$ | edge relations | transition/accessibility relations |
| $P_i^{\mathfrak{A}} \subseteq A$ | unary predicates | basic state properties/propositions |




in particular: game graphs

## the bisimulation game (mother of back&forth)

**positions**: $(\mathfrak{A}, a; \mathfrak{B}, b)$ *one* node marked in each structure

**single round**: **I** chooses $\alpha$,
moves pebble along $E_\alpha$-edge in $\mathfrak{A}$ or in $\mathfrak{B}$
**II** has to respond in opposite structure

**win/lose**: lose when stuck
**II** loses in $(\mathfrak{A}, a; \mathfrak{B}, b)$ with $\boldsymbol{P}$-inequivalent $a$, $b$

## bisimulation game and equivalences

$\mathfrak{A}, a \sim^\ell \mathfrak{B}, b$ :⇔ **II** has winning strategy in $G^\ell(\mathfrak{A}, a; \mathfrak{B}, b)$
**$\ell$-bisimilarity**

$\mathfrak{A}, a \sim^\omega \mathfrak{B}, b$ :⇔ **II** has winning strategy in all $G^\ell(\mathfrak{A}, a; \mathfrak{B}, b)$

$\mathfrak{A}, a \sim \mathfrak{B}, b$ :⇔ **II** has winning strategy in $G^\infty(\mathfrak{A}, a; \mathfrak{B}, b)$
**bisimilarity**

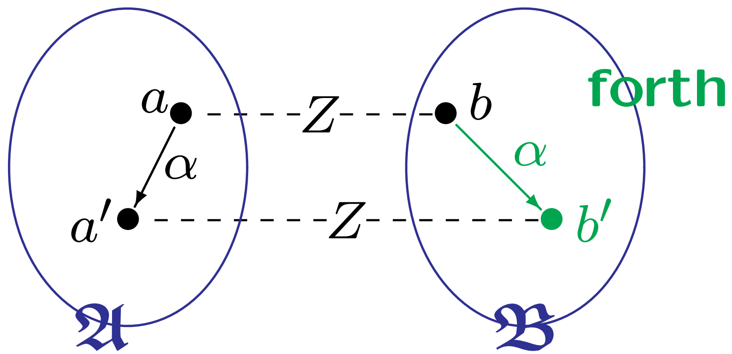
## back&forth in bisimulation

$\mathfrak{A}, a \sim \mathfrak{B}, b$ iff

- $a \simeq b$ (same colours w.r.t. $\boldsymbol{P}^{\mathfrak{A}}/\boldsymbol{P}^{\mathfrak{B}}$)
- for all $a \xrightarrow{\alpha} a'$ in $\mathfrak{A}$ there is $b \xrightarrow{\alpha} b'$ in $\mathfrak{B}$: $\mathfrak{A}, a' \sim \mathfrak{B}, b'$
- for all $b \xrightarrow{\alpha} b'$ in $\mathfrak{B}$ there is $a \xrightarrow{\alpha} a'$ in $\mathfrak{A}$: $\mathfrak{A}, a' \sim \mathfrak{B}, b'$

**back & forth system $\boldsymbol{Z \subseteq A \times B}$:**

non-det. winning strategy for **II**
witnessing bisimulation equivalence

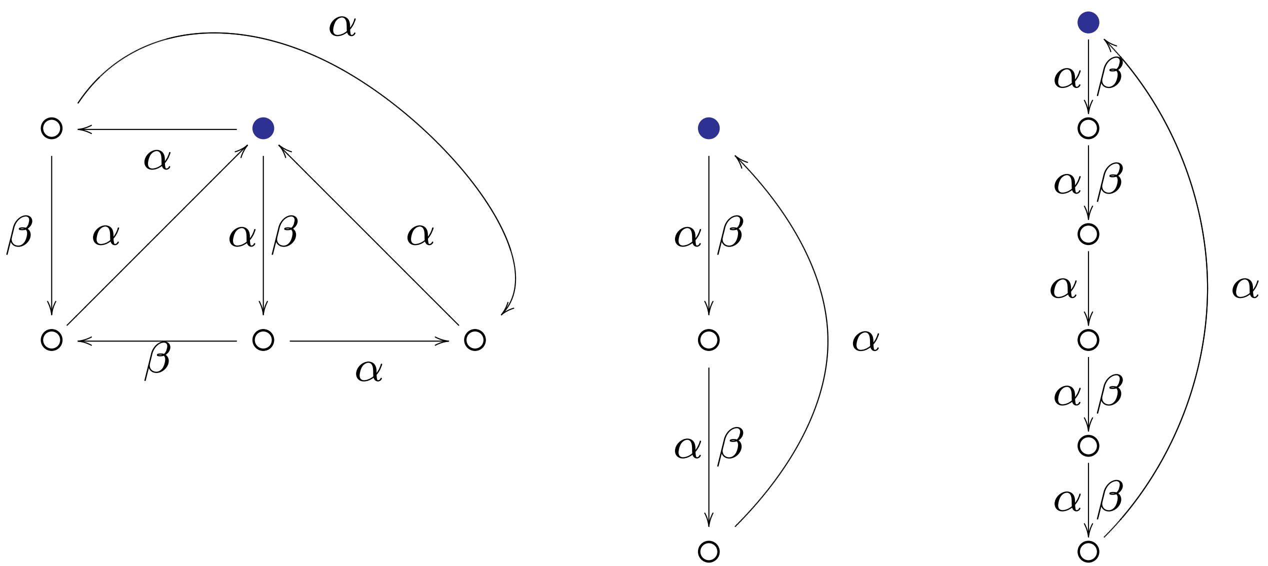



**largest bisimulation**

greatest fixed point $Z^\infty$ w.r.t. the back&forth conditions

$\mathfrak{A}, a \sim \mathfrak{B}, b$ iff $(a, b) \in Z^\infty$

## example of bisimulation equivalence




different traditions: bisimulation: Hennessy/Milner/Park
zig-zag equivalence: van Benthem
Ehrenfeucht–Fraïssé back&forth

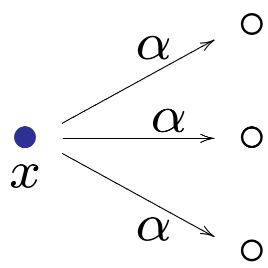
**which logic?**

## basic modal logic ML

atomic formulae: $\top$, $\bot$, $p_i$ (vertex colours $P_i$)
boolean connectives: $\vee$, $\wedge$, $\neg$, $\rightarrow$, ...
relativised quantification: $\langle\boldsymbol{\alpha}\rangle$, $[\boldsymbol{\alpha}]$

$$\langle\boldsymbol{\alpha}\rangle\boldsymbol{\psi}(x): \quad \exists y\Big((x \xrightarrow{\alpha} y) \wedge \psi(y)\Big) \equiv \exists y\Big(E_\alpha xy \wedge \psi(y)\Big)$$

$$[\boldsymbol{\alpha}]\boldsymbol{\psi}(x): \quad \forall y\Big((x \xrightarrow{\alpha} y) \rightarrow \psi(y)\Big) \equiv \forall y\Big(E_\alpha xy \rightarrow \psi(y)\Big)$$




$+$ variations (modalities w.r.t. derived edge relations)

NB: $\mathbf{ML} \subseteq \mathbf{FO^2}$ via standard translation

## modal Ehrenfeucht–Fraïssé and Karp theorems

$$\mathfrak{A}, a \sim^{\ell} \mathfrak{B}, a \quad \Leftrightarrow \quad \mathfrak{A}, \boldsymbol{a} \equiv_{\ell}^{\mathsf{ML}} \mathfrak{B}, b \qquad (*)$$

$$\mathfrak{A}, a \sim^{\omega} \mathfrak{B}, b \quad \Leftrightarrow \quad \mathfrak{A}, a \equiv^{\mathsf{ML}} \mathfrak{B}, b$$

$$\mathfrak{A}, a \sim \mathfrak{B}, b \quad \Leftrightarrow \quad \mathfrak{A}, a \equiv^{\mathsf{ML}_\infty} \mathfrak{B}, b$$

moreover, $\sim^{\omega}$ and $\sim$ coincide in
- $\omega$-saturated structures
- ML saturated structures
- finitely branching structures

$(*)$ key: formulae $\chi^{\ell}_{\mathfrak{A},a} \in \mathsf{ML}_{\ell}$ characterising $\sim^{\ell}$ class of $\mathfrak{A}, a$

## the modal back&forth conditions

inductively put

$$\chi^{\ell+1}_{\mathfrak{A},a} \;=\; \chi^{\ell}_{\mathfrak{A},a} \;\wedge\; \bigwedge_{\alpha} \Big( \underbrace{\bigwedge_{a' \in R_\alpha[a]} \langle \alpha \rangle \; \chi^{\ell}_{\mathfrak{A},a'}}_{\textbf{forth}\text{: challenges in } \mathfrak{A}} \;\wedge\; \underbrace{[\alpha] \bigvee_{a' \in R_\alpha[a]} \chi^{\ell}_{\mathfrak{A},a'}}_{\textbf{back}\text{: challenges in } \mathfrak{B}} \Big)$$

- $\chi^{\ell+1}_{\mathfrak{A},a} \in \mathsf{ML}_{\ell+1}$
- $\chi^{\ell+1}_{\mathfrak{A},a}$ such that $\mathfrak{B}, b \models \chi^{\ell+1}_{\mathfrak{A},a} \;\Leftrightarrow\; \mathfrak{B}, b \sim^{\ell+1} \mathfrak{A}, a$

## view other games through modal glasses

with back&forth game setting associate game graphs $\mathfrak{G}(\mathfrak{A})$ such that

$$\mathfrak{G}(\mathfrak{A}), \boldsymbol{a} \sim \mathfrak{G}(\mathfrak{B}), \boldsymbol{b} \quad \Leftrightarrow \quad \textbf{II} \text{ has winning strategy in } G^{\infty}(\mathfrak{A}, \boldsymbol{a}; \mathfrak{B}, \boldsymbol{b})$$

e.g., for $k$-pebble game: $\mathfrak{G}(\mathfrak{A}) = \Big(A^k, (R_i)_{1 \leqslant i \leqslant k}, (P_\rho)_{\rho \in \mathsf{atp}}\Big)$

**view $\sim$ (and its approximations $\sim^\ell$)**
**as back&forth equivalence of games**

in this sense, e.g., view correspondence:

$\simeq \;=\; \simeq_\omega$ over $\omega$-saturated structures

$\sim \;=\; \sim^\omega$ (Hennessy–Milner property)
for associated game graphs

## II A: Preservation and Expressive Completeness

recall **Q2: What can be expressed in $L$ ?**
definability, expressive power, measured against
- other logics
- **semantic criteria**
- complexity criteria

classical example: **Łos–Tarski theorem**

$\varphi(x) \in$ **FO preserved under extensions** $\Leftrightarrow$ $\varphi \equiv \tilde{\varphi} \in \exists^*$-**FO**

$\Leftarrow$ : obvious

$\Rightarrow$ : expressive completeness of $\exists^*$-FO
for extension-robust properties

classical proof: compactness/elementary extns

## expressive completeness issues: **classical and elsewhere**

characterisation theorems (like Łos–Tarski)
— *not* robust w.r.t. underlying class of structures
— not even w.r.t. restriction to $\mathcal{C}_0 \subseteq \mathcal{C}$

**preservation is robust, expressive completeness is not**

$$\varphi \ast\text{-invariant within } \mathcal{C}_0 \quad \not\Rightarrow \quad \varphi \ast\text{-invariant within } \mathcal{C}$$

$$\varphi \equiv \tilde{\varphi} \text{ within } \mathcal{C}_0 \quad \not\Rightarrow \quad \varphi \equiv \tilde{\varphi} \text{ within } \mathcal{C}$$

e.g., **Łos–Tarski thm fails in FMT** (Tait, Gurevich)

exhibit FO-definable class of structures,
whose finite members are robust under extension,
but not existentially FO-definable (among finite structures)
with infinitely many minimal finite models

## further examples

- **$FO^2$** and invariance under 2-pebble game equivalence $\simeq^2$

$FO/{\simeq^2} \;\equiv\; FO^2$ classically

the usual compactness argument, $\omega$-saturated extensions

but **not in FMT**

finite linear orderings

- **ML** and invariance under bisimulation $\sim$

$FO/{\sim} \;\equiv\;$ ML classically

**van Benthem 83**

the usual compactness argument, $\omega$-saturated extensions

and also $FO/{\sim} \;\equiv\;$ ML **(FIN)**

**Rosen 97**

game based model constructions
new proof below

with many variations

still $\sim$-invariance in finite $\not\Rightarrow$ $\sim$-invariance throughout

**$\mathbf{FO}/{\sim} \equiv \mathbf{ML}$** classically as well as in FMT

for FO definable properties:

**bisimulation invariance = definability in ML**

i.e., for $\varphi(x) \in$ FO:

$$\begin{aligned} & \varphi \sim \text{invariant} \\ \Leftrightarrow\ & \varphi \text{ equivalent to some } \tilde{\varphi} \in \text{ML} \\ \Leftrightarrow\ & \varphi \sim^{\ell} \text{invariant for some } \ell \text{ (!)} \end{aligned}$$

characterising ML $\subseteq$ FO and effective syntax for FO$/{\sim}$

**ML is the first-order logic of games/process behaviour**

## $\mathbf{FO}/{\sim} \equiv \mathbf{ML}$

**preservation: $\mathbf{ML} \subseteq \mathbf{FO}/{\sim}$**

---

$\varphi \in \mathbf{ML}_\ell$ **invariant under** $\sim^\ell$ Ehrenfeucht-Fraïssé

**expressive completeness: $\mathbf{FO}/{\sim} \subseteq \mathbf{ML}$**

---

**proof methods**

classical: compactness

**constructive: explicit model constructions**
**Ehrenfeucht-Fraïssé: FO vs ML**

## infinite vs. finite game equivalence — as in $\sim / \sim^\ell$

| full equivalence | approximants | of finite index |
| --- | --- | --- |
| $\leftrightarrows$ | $\leftrightarrows^\ell$ | |
| **infinite games** | **finite, $\ell$-round games** | |

## Ehrenfeucht-Fraïssé analysis of $\leftrightarrows^\ell$

$\longrightarrow$ approximants to full characterisation thm

$$\mathrm{FO}/{\leftrightarrows^\ell} \equiv \mathcal{L}_\ell \quad \text{as in} \quad \mathrm{FO}/{\sim^\ell} \equiv \mathrm{ML}_\ell$$

full characterisation thm equivalent to **compactness property**

**$\leftrightarrows$ invariance $\Rightarrow$ $\leftrightarrows^\ell$ invariance for some $\ell$**

## classical proofs: compactness of FO

**based on convergence $\leftrightharpoons^{\ell} \longrightarrow \leftrightharpoons$**

in $*$-models (e.g., $\omega$ saturated) where $\leftrightharpoons^{\omega} := \bigcap_{\ell} \leftrightharpoons^{\ell}$ is $\leftrightharpoons$

for $\leftrightharpoons$ invariant $\varphi$:

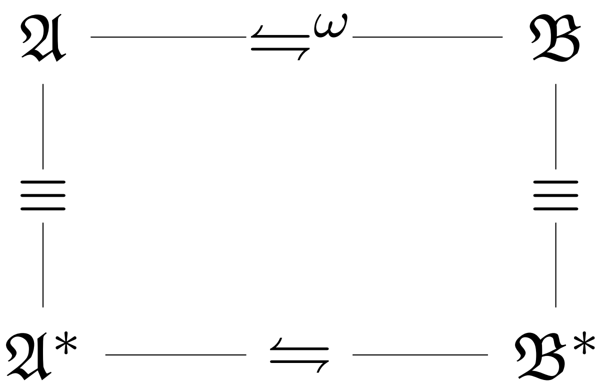
$$
\begin{array}{cccl}
\mathfrak{A}_{\ell} & \leftrightharpoons^{\ell} & \mathfrak{B}_{\ell} & \text{(one } \ell \text{ at a time)} \\
\mathfrak{A} & \leftrightharpoons^{\omega} & \mathfrak{B} & \text{(all } \ell \text{ simultaneously)} \\
\mathfrak{A}^{*} & \leftrightharpoons & \mathfrak{B}^{*} & \\
\models & & \not\models & \# \\
\varphi & & \varphi &
\end{array}
$$

non-constructive (indirect)
does not go through in fmt

## orthogonal approach to expressive completeness proofs

instead of

**via full $\equiv$ to full $\leftrightharpoons$**

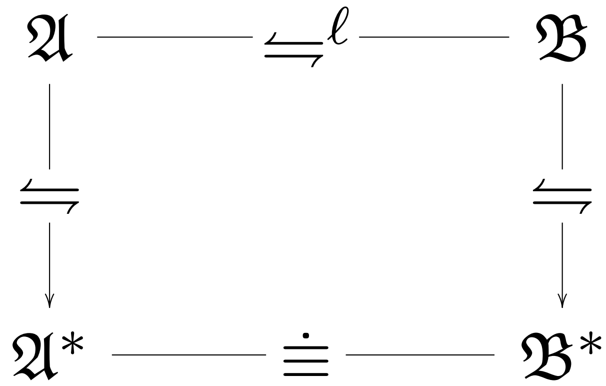



prep: $(\leftrightharpoons^{\ell})_{\ell\in\omega} \longrightarrow \leftrightharpoons^{\omega}$

upgrading via $\omega$-saturation

try

**via full $\leftrightharpoons$ to approximate $\equiv$**




direct upgrading

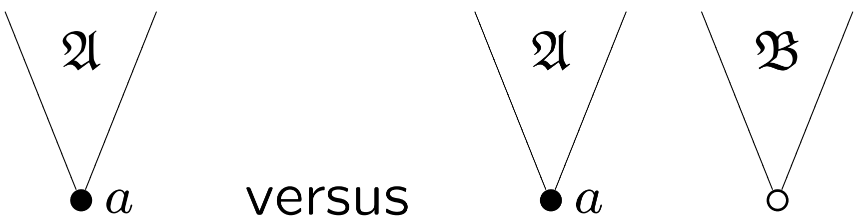
## aside: new stand-alone proof for van Benthem-Rosen

reduces input from classical model theory to Ehrenfeucht-Fraïssé
$\longrightarrow$ valid classically as well as in fmt

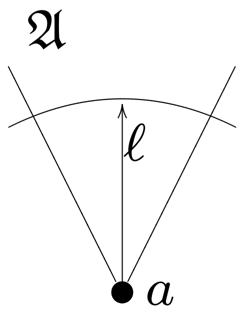
**(0)** $\varphi$ **$\sim$ invariant $\Rightarrow$ $\varphi$ invariant under disjoint unions**

**(1)** $\varphi$ **$\sim$ invariant $\Rightarrow$ $\varphi$ $\ell$-local for $\ell \leqslant 2^{\mathrm{qr}(\varphi)}$** **(E-F)**

**(2)** $\varphi$ **$\sim$ invariant & $\ell$-local $\Rightarrow$ $\varphi \sim^{\ell}$ invariant**



invariance under disjoint union



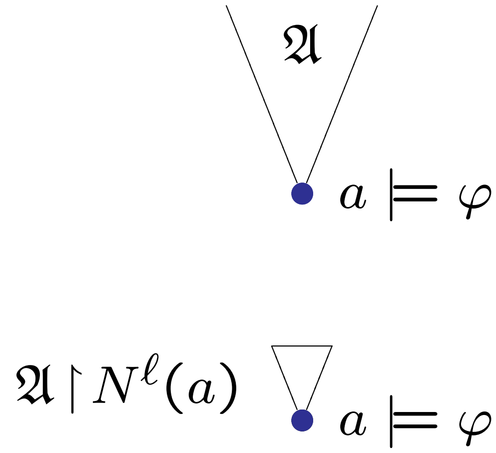
$\ell$-locality

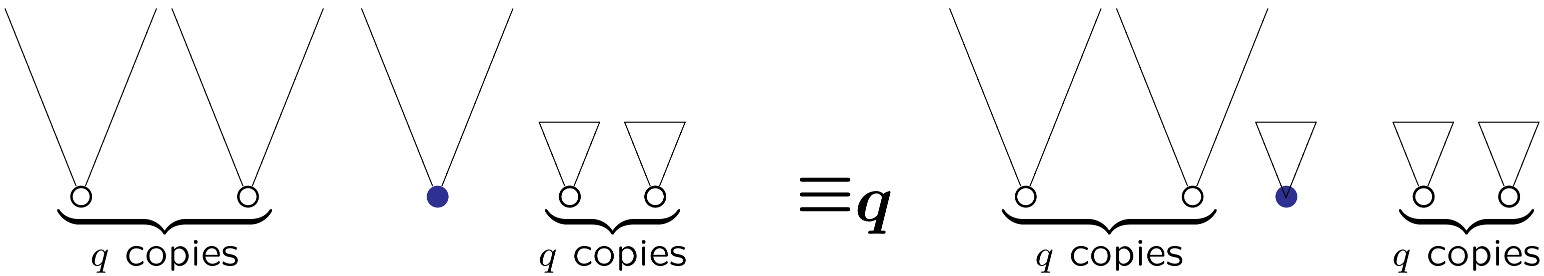
## the Ehrenfeucht-Fraïssé argument

**(1)** $\varphi \sim$ **invariant** $\Rightarrow$ $\varphi$ $\ell$**-local for** $\ell = 2^{\mathrm{qr}(\varphi)}$

show

$$\mathfrak{A}, a \models \varphi \quad \text{iff} \quad \mathfrak{A} \restriction N^{\ell}(a), a \models \varphi$$




play $q$ rounds respecting critical distance $d_m = 2^{q-m}$ in round $m$

## (2) $\sim$ invariant & $\ell$-local $\Rightarrow$ $\sim^{\ell}$ invariant

here an almost trivial case of upgrading $\sim^{\ell}$ to $\ell$-local isomorphism

$$\begin{array}{ccc} \mathfrak{A}, a & \overset{\sim^{\ell}}{\text{———}} & \mathfrak{B}, b \\ \Big| \sim & & \Big| \sim \\ \tilde{\mathfrak{A}}, \tilde{a} & \overset{\simeq^{(\ell)}}{\text{———}} & \tilde{\mathfrak{B}}, \tilde{b} \end{array}$$

$\downarrow$ partial unravelling to depth $\ell$
& adjustment of multiplicities

**challenge**: uniform locality for finer, global variants of $\sim$
upgrade to appropriate levels of $\equiv$ rather than $\simeq$
$\rightarrow$ locality and levels of Gaifman equivalence $\equiv^{\ell}_{q,m}$

## generic idea: upgrading $\leftrightharpoons^{\ell}$ to $\equiv^{\ell'}_{q,m}$

$$\mathfrak{A}, a \quad \overset{\leftrightharpoons^{\ell}}{\text{———}} \quad \mathfrak{B}, b$$

$$\leftrightharpoons \qquad\qquad \leftrightharpoons$$

$$\tilde{\mathfrak{A}}, \tilde{a} \quad \overset{\equiv^{\ell'}_{q,m}}{\text{———}} \quad \tilde{\mathfrak{B}}, \tilde{b}$$

**local control over FO**
up to quantifier rank $q$

$\varphi$ **preserved under** $\equiv^{\ell'}_{q,m}$ **and** $\leftrightharpoons$ **invariant**

$\Rightarrow$ $\varphi$ $\leftrightharpoons^{\ell}$ **invariant**

## works for . . .

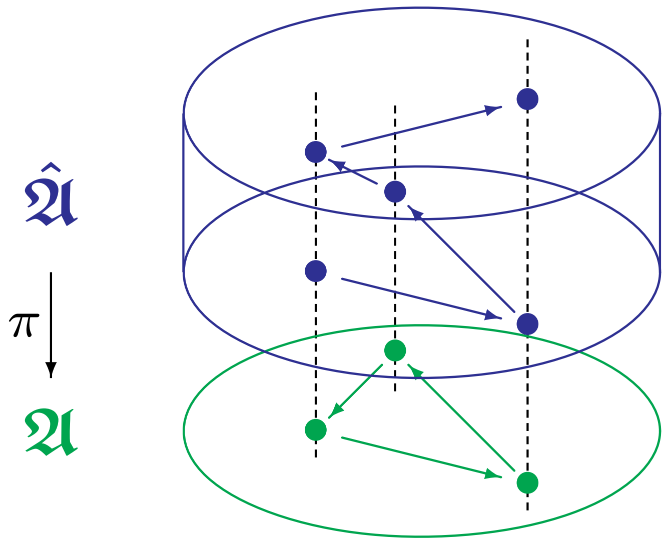
| | | classical and in FMT |
|---|---|---|
| $\sim_\forall$ | global (forward) bisimulation | $\mathbf{FO}/{\sim_\forall} \equiv \mathbf{ML}[\forall]$ |
| $\approx$ | global two-way bisimulation | $\mathbf{FO}/{\approx} \equiv \mathbf{ML}[-,\forall]$ |
| $\sim\, = \,\sim_\forall$ | | $\mathbf{FO}/{\sim} \equiv \mathbf{ML}[\forall]$ **over rooted frames** |

## from $\leftrightarrows^{\ell}$ to local control over FO

### locally acyclic covers

instead of (infinite) tree unravellings

homomorphism $\pi\colon \hat{\mathfrak{A}} \longrightarrow \mathfrak{A}$
whose graph induces a
two-way global bisimulation



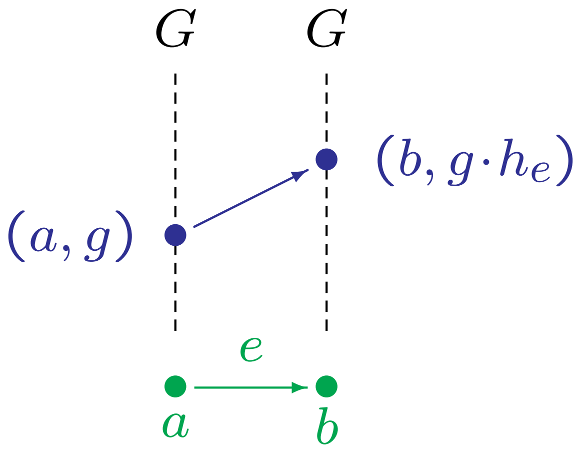

**NB: two-way unravellings are (infinite) acyclic covers**

## theorem

**any [finite] transition system admits a cover by a [finite] $\ell$-locally acyclic transition system.**

proof: "fibre bundle" over base system
using group whose Cayley graph
has no short cycles

[polynomial blow-up for fixed $\ell$]



## further variations

non-trivial locality to no apparent locality

- **classical frame properties**: symmetry, reflexivity, transitivity

  **equivalence frames (S5)**
  (modified locality arguments) Dawar, O_ LICS 05

  **transitive (and tree-like) frames**
  (decomposition arguments) Dawar, O_; recently right

- **challenge**: beyond transition systems

  **guarded logics and hypergraph bisimulations**
  (major open problems of a combinatorial nature)

## **example:** decomposition based techniques

e.g.: **upgrading $\sim^{\ell}$ to $\equiv_q$ in $\prec$-trees or $\preccurlyeq$-trees**

finite irreflexive/reflexive transitive $\mathfrak{A}, a$ unravel to
finite $\prec/\preccurlyeq$-trees $s(\mathfrak{A}, a)$ with boosted multiplicities

$$\mathfrak{A}, a \quad \overset{\sim^{\ell}}{\text{———}} \quad \mathfrak{B}, b$$

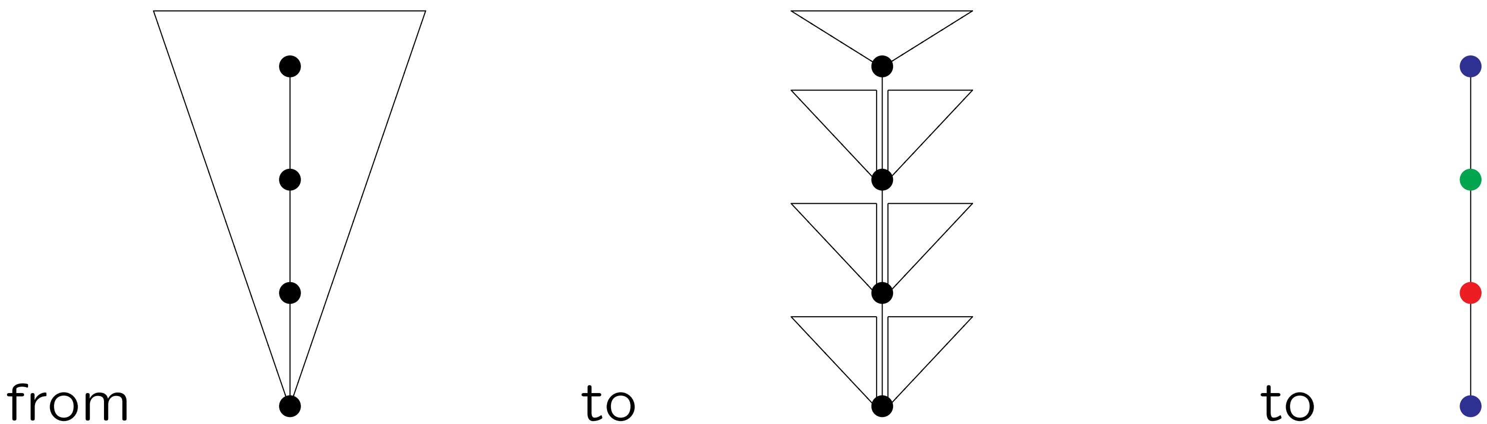
$$\sim \Big| \qquad\qquad\qquad \Big| \sim$$

$$s(\mathfrak{A}, a) \quad \overset{\equiv_q}{\text{———}} \quad s(\mathfrak{B}, b)$$

$\longleftarrow$ saturated $\prec/\preccurlyeq$-trees

in suitably saturated finite (!) $\prec/\preccurlyeq$-trees $s(\mathfrak{A}, a), s(\mathfrak{B}, b)$:
**establish $\equiv_q$ via games and path decompositions
instead of plain locality argument**

## decomposition & game argument — from trees to words




**pumping lemma** (Ehrenfeucht-Fraïssé):

bound on length of relevant words realised in $s(\mathfrak{A}, a)$
**finiteness property**

$\longrightarrow$ inductive bound on $\ell$
for which $\sim^{\ell}$ governs $\equiv_q$

## if reflexivity is not prescribed: something new happens

the *interesting* mistake in DO LICS 05

$$\varphi(x) = \exists y(Exy \wedge Eyy)$$

- $\sim$ invariant over finite (!) transitive frames
- *not* $\sim^\ell$ invariant for any $\ell$

while FO/$\sim$ $\equiv$ ML over the class of all transitive frames,
FO/$\sim$ $\not\equiv$ ML over the class of finite transitive frames

instead, a new modality emerges:

$$\Diamond^*\varphi \equiv \exists y\big(Exy \wedge Eyy \wedge \varphi(y)\big)$$

with associated $\sim_*$ / $\sim_*^\ell$

$\mathfrak{A}, a \sim \mathfrak{B}, b \;\Rightarrow\; \mathfrak{A}, a \sim_* \mathfrak{B}, b$ for finite (!) transitive frames

but $\mathfrak{A}, a \sim^\ell \mathfrak{B}, b \;\not\Rightarrow\; \mathfrak{A}, a \sim_*^1 \mathfrak{B}, b$ for any $\ell$

## with the new modality $\Diamond^*$

$\sim_*^{\ell'}$ can be upgraded to $\sim^{\ell}$ in expansions with reflexivity predicate

and to $\equiv_q$ in these

## new

Dawar, O_ 07

$\mathbf{FO}/\!\sim \; \equiv \mathbf{ML}[\Diamond^*]$ over $\left\{ \begin{array}{l} \textbf{finite transitive frames} \\ \textbf{finite transitive tree-like frames} \end{array} \right.$

versus (classically)

$\mathbf{FO}/\!\sim \; \equiv \mathbf{ML}$ over **all transitive frames**

**excursion:**

---

## locality criteria and explicit model constructions
**from FMT to the study of well-behaved classes**

examples of classical thereoms regained

### Łos–Tarski extension preservation

| |
|---|
| $\varphi(x) \in$ **FO preserved under extensions** $\quad \Leftrightarrow \quad \varphi \equiv \tilde{\varphi} \in \exists^*$-**FO** |

valid over special classes of finite structures (Atserias, Dawar, Grohe 05)

### Lyndon–Tarski homomorphism preservation

| |
|---|
| $\varphi(x) \in$ **FO preserved under homomorphisms** $\quad \Leftrightarrow \quad \varphi \equiv \tilde{\varphi} \in \exists^*$-**FO**$^+$ |

valid over special classes of finite structures (Atserias, Dawar, Kolaitis 04)

valid in FMT (Rossman 05)

## **extension preservation in special classes**

$\mathcal{C}$ a $\subseteq$-closed class of finite structures
$\varphi \in$ FO preserved under extensions in $\mathcal{C}$

need: finitely many $\subseteq$-minimal elements in $\varphi[\mathcal{C}]$
then $\varphi$ equivalent to disjunction over
$\exists$-closure of algebraic diagrams

## **homomorphism preservation in special classes**

need: finitely many $\subseteq_{\mathrm{w}}$-minimal elements in $\varphi[\mathcal{C}]$
then $\varphi$ equivalent to disjunction over
$\exists$-closure of positive algebraic diagrams

**expressive completeness**:
**bounds on size of minimal models**
**through locality based criteria**

## notions of wideness

Atserias, Dawar, Grohe, Kolaitis 04/05 — Ajtai, Gurevich 89

**$\mathfrak{A}$ $(\ell, m)$-wide**: $\mathfrak{A}$ contains $\ell$-scattered subset of size $m$
a property of the Gaifman the graph

**$\mathcal{C}$ wide**: for all $\ell, m$ exists $N$:
$\mathfrak{A} \in \mathcal{C}$, $|\mathfrak{A}| \geqslant N \;\Rightarrow\; \mathfrak{A}$ $(\ell, m)$-wide

relax to
**$\mathcal{C}$ almost wide**: wide up to constant
number of elements
e.g., trees

## theorem

Atserias, Dawar, Kolaitis 04

**any class of graphs with excluded minor is almost wide**

## homomorphism preservation

Atserias, Dawar, Kolaitis 04
Rossman 05

**theorem** Ajtai, Gurevich

$\mathcal{C}$ closed under substructures and disjoint unions
$\varphi \in$ FO preserved under homomorphisms on $\mathcal{C}$

$\Rightarrow$

minimal models of $\varphi$ cannot be $(\ell, m)$-wide (suitable $\ell, m$)
similarly, even up to removal of any fixed number of elements

**corollary**

over almost wide $\mathcal{C}$: $\longrightarrow$ bound on size of minimal models
$\longrightarrow$ finitely many minimal models
$\longrightarrow$ positive $\exists^*$ definability

**homomorphism preservation thm in restriction to $\mathcal{C}$**

## extension preservation

Atserias, Dawar, Grohe 05

**can bound size of minimal models over:**

- classes of structures with acyclic Gaifman graphs
- **all wide $\mathcal{C}$,   e.g., bounded degree graphs**
- **$\mathcal{C}_k$ (treewidth $k$)**

size bounds on minimal models via Gaifman:

in large $\mathfrak{A} \models \varphi$ find

$$\mathfrak{A}_0 \subsetneq \mathfrak{A} \subseteq \widehat{\mathfrak{A}}$$

$$\mathfrak{A}_0 \equiv^{\ell}_{q,m} \widehat{\mathfrak{A}} \qquad \Rightarrow \mathfrak{A}_0 \models \varphi$$

finite chain construction!

remark: Łos–Tarski fails over planar finite graphs

## homomorphism preservation: new classical proof and FMT

**homomomorphism preservation** Rossman 05

for any $\varphi \in \mathrm{FO}$:

**classically, with extra value:**

| | | |
|---|---|---|
| $\varphi$ **preserved under homomorphisms** | $\Leftrightarrow$ | $\varphi \equiv \tilde{\varphi} \in \exists^*$-$\mathbf{FO}^+$ <br> $\mathrm{qr}(\varphi') = \mathrm{qr}(\varphi)$ (!) |

**in FMT:**

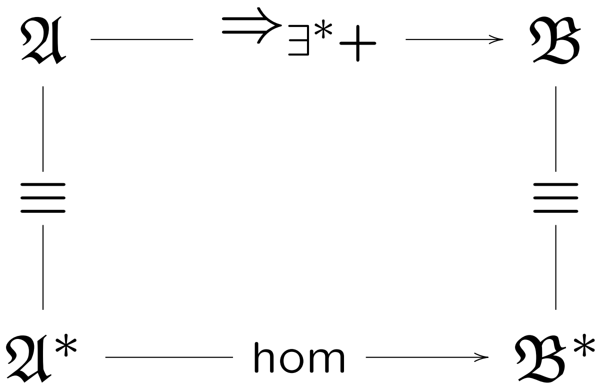
| | | |
|---|---|---|
| $\varphi$ **preserved under homomorphisms** | $\Leftrightarrow$ | $\varphi \equiv \tilde{\varphi} \in \exists^*$-$\mathbf{FO}^+$ <br> with non-elementary gap in qr |

**method**: **existential positive types & saturation (chain)**

> compactness property in finite structures:
> large finite degree of saturation suffices

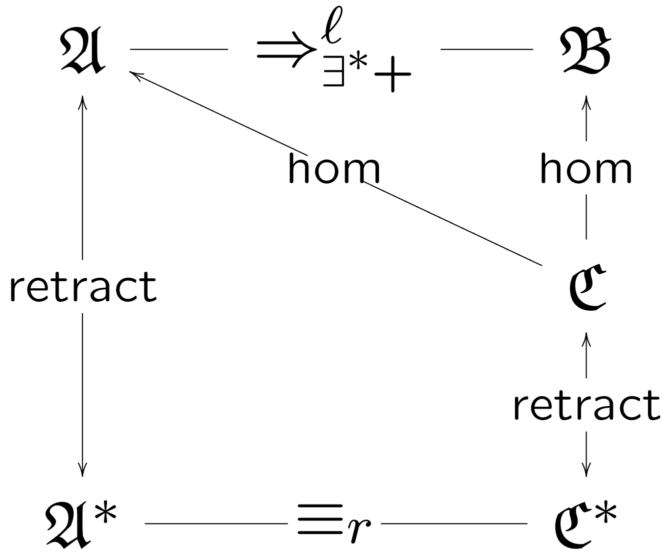
instead of

**via full $\equiv$ to hom**




upgrading via $\omega$-saturation

**via hom to approximate $\equiv$**




finite $\mathfrak{A}^*$: $\ell(r)$ non-elementary

infinite $\mathfrak{A}^*$: $\ell = r$

## II B: Relational Recursion

recall

**Q2: What can be expressed in $L$ ?**

definability, expressive power, measured against

- other logics
- semantic criteria
- **complexity criteria**

FO too weak to express
algorithmically very basic properties
like reachability, connectivity

**FO static and local**

$\longrightarrow$ add recursion mechanisms
especially fixed points of monotone operators
like $\varphi(X, x) = Px \vee \exists y(Exy \wedge Xy)$

## least fixed points of monotone operators

with $\varphi(X, \boldsymbol{x})$, $X$ and $\boldsymbol{x}$ of arity $r$, associate operator over $\mathfrak{A}$

$$\begin{aligned} \varphi^{\mathfrak{A}} \colon \mathcal{P}(A^r) &\longrightarrow \mathcal{P}(A^r) \\ P &\longmapsto \varphi^{\mathfrak{A}}[P] := \big\{\boldsymbol{a} \in A^r : \mathfrak{A} \models \varphi[P, \boldsymbol{a}]\big\} \end{aligned}$$

$\varphi$ is positive in $X$
$\Rightarrow$ $\varphi^{\mathfrak{A}}$ is monotone $\quad (P \subseteq P' \Rightarrow \varphi^{\mathfrak{A}}[P] \subseteq \varphi^{\mathfrak{A}}[P'])$
$\Rightarrow$ $\varphi^{\mathfrak{A}}$ possesses unique least and greatest fixed points

**least fixpoint** $\boxed{(\mu_X \varphi)[\mathfrak{A}] = \bigcap\{P \subseteq A^r : \varphi^{\mathfrak{A}}[P] = P\}}$

also as limit of
**inductive stages**: $(\mu_X \varphi)[\mathfrak{A}] = \bigcup_\alpha X^\alpha[\mathfrak{A}]$ where

$$\begin{aligned} X^0[\mathfrak{A}] &= \emptyset \\ X^{\alpha+1}[\mathfrak{A}] &= \varphi^{\mathfrak{A}}[X^\alpha[\mathfrak{A}]] \\ X^\lambda[\mathfrak{A}] &= \bigcup_{\alpha<\lambda} X^\alpha[\mathfrak{A}] \end{aligned}$$

## background on fixed point logics

key examples

## least fixed point logic LFP:

extension of FO by $\mu/\nu$ for $X$-positive operators

e.g.: $\mu_X(Exy \vee \exists z(Xxz \wedge Xzy))$ defines TC($E$)

as expressive as (more general) IFP extension
for inductive definitions (Gurevich–Shelah/Kreutzer)

## modal $\mu$-calculus $\mathrm{L}_\mu$:

extension of ML by $\mu/\nu$ for (monadic) $X$-positive operators

e.g.: $\mu_X(\Box X)$ defines well-founded support for $R^{-1}$

*the* unifying framework for the
most important process/game/temporal logics
— also a fragment of MSO

## Immerman–Vardi theorem

for properties of finite, *linearly ordered* structures:

**Ptime properties $\equiv$ LFP definable properties**

**Ptime model checking** fixed points reached within polynomially many steps

**expressive completeness** simulation of polynomially bounded TM computations in fixed point recursion *over ordered domains*

**Janin–Walukiewicz theorem**

$\mathbf{MSO}/{\sim} \;\equiv\; \mathbf{L}_\mu$

compare $\mathrm{FO}/{\sim} \equiv \mathrm{ML}$
at first-order level

**expressive completeness**: tree automata for MSO and $\mathrm{L}_\mu$

**descriptive complexity in the modal world:**

$\mathbf{Ptime}/{\sim} \;\equiv\; \mathbf{L}_\mu{}^{\omega}$

higher-arity variant of $\mathrm{L}_\mu$
for $\sim$-invariant Ptime

**expressive completeness**: definable ordering of $\sim$ quotients
and reduction to Immerman–Vardi

## boundedness of fixed point recursions

$\varphi(X, \boldsymbol{x})$ positive in $X$; fixed point process with stages $X^\alpha$

**closure ordinal:** $\boldsymbol{\gamma[\varphi, \mathfrak{A}] = \min_\alpha\big(X^{\alpha+1}[\mathfrak{A}] = X^\alpha[\mathfrak{A}]\big)}$

$\boldsymbol{\varphi(X, x)}$ **bounded:** $\boldsymbol{\exists n \in \mathbb{N}}$ **s.t.** $\boldsymbol{\gamma[\varphi, \mathfrak{A}] < n}$ **for all** $\boldsymbol{\mathfrak{A}}$

$\varphi(X, x) \in$ FO bounded $\Rightarrow$ recursion spurious

$\Rightarrow$ $\mu_X\varphi \equiv \varphi^n$ uniformly FO

## boundedness and definability

### Barwise–Moschovakis theorem

for any $X$-positive FO formula $\varphi(X, x)$
the following are equivalent:

(i) $\mu_X\varphi$ **bounded**

(ii) $\mu_X\varphi$ **uniformly FO definable**

(iii) $\mu_X\varphi[\mathfrak{A}]$ **FO definable in each** $\mathfrak{A}$

relativises to natural fragments: $\forall^*$-FO, $\exists^*$-FO, FO$^k$, ML, . . .

relativises to elementary classes: acyclic, $\mathcal{C}_k$ (treewidth $k$), . . .

proof: compactness argument
$\gamma[\varphi, \mathfrak{A}] \leqslant \omega$ in $\omega$-saturated $\mathfrak{A}$

## boundedness as a decision problem

for a class $\mathcal{F}$ of FO formulae:

$\mathrm{BDD}(\mathcal{F})$

| |
|---|
| **given** $\varphi(X, x) \in \mathcal{F}$ |
| **decide if** $\mu_X \varphi$ **is bounded** |

- **SAT reducible to BDD** for natural fragments $\mathcal{F}$
- **BDD a generalised SAT problem**: $(\varphi^{n+1} \wedge \neg \varphi^{n})$ for all $n \in \mathbb{N}$
- **few decidable cases**, even for monadic recursion

## decidability vs. undecidability for monadic BDD

| undecidable | decidable |
|---|---|
| $\exists^*$-**FO**<br>**existential, positive with inequality**<br>Gaifman, Mairson, Sagiv, Vardi 87 | $\exists^*$-**FO**$^+$<br>**pure existential positive**<br>Cosmadakis, Gaifman, Kanellakis, Vardi 95 |
| $\mathrm{FO}^2$<br>**two variables**<br>Kolaitis, O_ 98 | ML<br>**modal**<br>O_ 98, improved 06 |
| $\forall^*$-**FO**<br>**universal, mixed polarities or with equality**<br>O_ 06 | $\forall^*$-**FO**$^-$<br>**universal, single polarities without equality**<br>O_ 06 |

can encode tilings

decidable via tree codings

## locality and boundedness in tree-like structures

NB: monadic fixed points are MSO definable

## local MSO = local FO

in acyclic relational structures (trees):

$\varphi(x) \in$ **MSO local** $\Rightarrow$ $\varphi(x) \equiv \tilde{\varphi}(x) \in$ **FO** game argument

in particular, for $\varphi(X) \in$ ML:

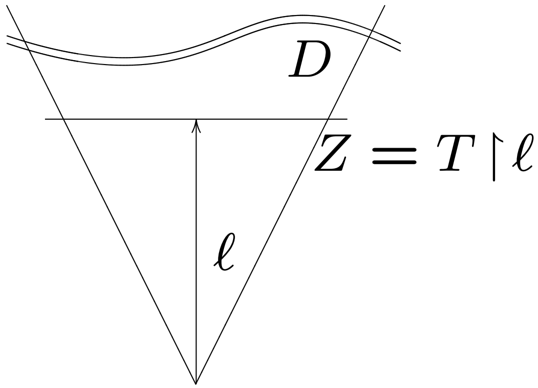
$$
\begin{aligned}
& \varphi \text{ bounded} \\
\Rightarrow\ & \mu_X\varphi\ \ell\text{-local for some } \ell \\
\Rightarrow\ & \mu_X\varphi \text{ FO-definable} \\
\Rightarrow\ & \mu_X\varphi \text{ ML-definable} \\
\Rightarrow\ & \varphi \text{ bounded}
\end{aligned}
$$

all equivalent

## tree-locality of $\psi \in$ MSO

$\exists \ell \in \mathbb{N}$ such that for all trees $T$,
and all initial $D \subseteq T$ with $D \supseteq T{\upharpoonright}\ell$:

$$T \models \psi \text{ iff } T{\upharpoonright}D \models \psi$$




## towards a reduction to the MSO-theory of $T_\omega$

$$\left.\begin{array}{l} Z \text{ initial and for all } \boldsymbol{I} \text{ and all initial } D: \\ Z \subseteq D \longrightarrow \big(\psi[\boldsymbol{I}] \leftrightarrow \psi[\boldsymbol{I}{\upharpoonright}D]\big) \end{array}\right\} \ \eta(Z) \in \text{MSO}$$

$$\psi \text{ **tree-local** iff } T_\omega \models \exists Z \left( \boxed{Z \text{ bounded}} \wedge \eta(Z) \right)$$

not MSO

## König's lemma for regular expansions of $T_\omega$

**for regular $(\boldsymbol{T_\omega, Z})$** (finite number of subtrees up to $\simeq$) with initial $Z \subseteq T_\omega$ t.f.a.e.:

(i) **$Z$ path-finite** (no infinte path within $Z$)

(ii) **$Z$ bounded** ($Z \subseteq T{\restriction}\ell$ for some $\ell \in \mathbb{N}$)

**tree-locality criterion in MSO-Th$(\boldsymbol{T_\omega})$:**

$$T_\omega \models \exists Z(\varphi_{\text{path-fin}}(Z) \wedge \eta(Z))$$

$\Leftrightarrow$ $(T_\omega, Z) \models \varphi_{\text{path-fin}}(Z) \wedge \eta(Z)$ for some $Z \subseteq T_\omega$

$\Leftrightarrow$ $(T_\omega, Z) \models \varphi_{\text{path-fin}}(Z) \wedge \eta(Z)$ for some regular $(T_\omega, Z)$

$\Leftrightarrow$ $T_\omega \models \exists Z \Big( Z \text{ bounded } \wedge \ \eta(Z) \Big)$

$\longrightarrow$ decidability of BDD(ML) via locality and MSO-Th$(T_\omega)$

## deciding monadic $\mathrm{BDD}(\mathbf{FO})$ over acyclic structures

Kreutzer, O_, Schweikardt ICALP 07

### decidable BDD

$\mathcal{C}$ (any FO-definable sublass of) the class of all acyclic structures

for $X$-positive $\varphi(X, x) \in$ FO,

decide whether $\begin{cases} \varphi(X,x) & \text{is bounded over } \mathcal{C} \\ \mu_X \varphi(X,x) & \text{is FO over } \mathcal{C} \end{cases}$

**methods**:

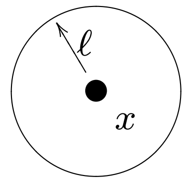
locality analysis of $\varphi$ (Gaifman$^{+}$)
locality testing for phases of purely local iteration (MSO-based)
Barwise-Moschovakis (FO-based)

**open**: treewidth $k$ // trees // finite acyclic // . . .

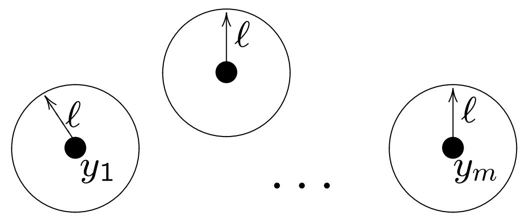
## Gaifman's theorem

$\varphi(X, x) \in$ FO equivalent to boolean combination of FO-formulae of two types

**(L)** $\chi^{(\ell)}(X, x)$ asserting properties of $N^{\ell}(x)$




**(S)** assertions about existence of $\ell$-scattered tuples $y_1, \ldots, y_m$ within some $\chi^{(\ell)}[\mathfrak{A}, X]$




**respecting positivity in $X$?**

example: $\varphi(X, x) = \exists y (y \neq x \wedge Xy)$

## respecting positivity in $X$?

- $X$-positive $\varphi(X,x) \not\equiv X$-positive b.c. of (L)/(S)

  $X$-positive type (L) may not suffice

- $\varphi(X)$ $X$-positive $\equiv$ $X$-positive b.c. of (S)

  Dawar/Grohe/Kreutzer/Schweikardt LICS 06

- for $X$-positive $\varphi(X,x)$:
  **unrestricted (L)-parts + only $X$-pos. (S)-parts**

example:

$$\exists y(y \neq x \wedge Xy) \equiv \begin{cases} \quad Xx & \wedge \quad \exists y_1 y_2 (d(y_1,y_2) > 0 \wedge Xy_1 \wedge Xy_2) \\ \vee & \\ \neg Xx & \wedge \quad \exists y_1 X y_1 \end{cases}$$

## leading to generic format:

$$\boldsymbol{\varphi(X, x) = \bigvee_i \big( \underbrace{\varphi_i^{(\ell)}(X, x)}_{\textsf{(L)}} \wedge \psi_i(X) \big)}$$

$\varphi_i^{(\ell)}(X, x)$: local about $x$, but not necesssarily $X$-positive

$\psi_i(X)$: $X$-positive guards for local components

**idea**: decompose iteration on $\varphi$ into phases
of purely local iterations driven by
$\varphi_i^{(\ell)}$ switched on by $\psi_i(X)$

## phase analysis (indication in generic example)

$$\varphi(X, x) = \big(\ \varphi_1^{(\ell)}(X, x)\ \wedge\ \psi_1(X)\ \big) \vee \big(\ \varphi_2^{(\ell)}(X, x)\ \wedge\ \psi_2(X)\ \big)$$

| **detecting unboundedness** over $\mathfrak{A}$ such that | | **through** |
|---|---|---|
| (0) $\mathfrak{A} \models \neg\psi_1[\emptyset] \wedge \neg\psi_2[\emptyset]$ | — | |
| (1) $\mathfrak{A} \models \psi_1[\emptyset] \wedge \psi_2[\emptyset]$ | driven by $\varphi_1^{(\ell)} \vee \varphi_2^{(\ell)}$ | **LT** |
| (2) $\mathfrak{A} \models \psi_1[\emptyset] \wedge \neg\psi_2[\varphi^\infty]$ | driven by $\varphi_1^{(\ell)}$ | **LT** |
| (3) $\mathfrak{A} \models \psi_1[\emptyset] \wedge \psi_2[\varphi^\infty]$ | two phases (!) | |
| (a) $\varphi_1^{(\ell)} \vee \psi_2$ unbdd | subsumed in (2) | **LT** |
| (b) $\varphi_1^{(\ell)} \vee \psi_2$ bdd | subsumed in (1) up to initialisation | **LT** |

**LT**: locality testing

## why not any better yet?

treading on thin ice:

- Barwise–Moschovakis fails for $\begin{cases}\text{trees (finite or infinite)}\\ \text{finite acyclic structures}\end{cases}$
- "locality implies FO" fails for treewidth 3 graphs

on the other hand, decidability of BDD in bounded treewidth would have great explanatory power . . .

**model theoretic games** and **model constructions**

| |
|---|
| work in all sorts of interesting classes ignored by classical model theory |

| |
|---|
| for many issues, there are interesting classes other than just elementary |

**locality** and its role in mediating game analysis curiously under-exposed in classical model theory

**explicit model constructions** can replace classical arguments in surprising manners

## selected references

**Finite Model Theory**, Ebbinghaus, Flum, (2nd ed) Springer 1999

**Finite Model Theory**
Lecture Notes at `www.mathematik.tu-darmstadt.de/~otto`, 2005/06

**Elementary proof of the van Benthem–Rosen characterisation theorem**, O_, TUD online preprint 2342, 2004

**Model theory of modal logic**, Goranko&O_, in:
**Handbook of Modal Logic**, Elsevier 2006

**Modal and guarded characterisation theorems over finite transition systems**, O_, APAL 2004

**Bisimulation invariance and finite structures**
O_, LNL 27, Logic Colloquium 2002, ASL 2006

**Modal characterisation theorems over special classes of frames**
Dawar&O_, LICS 2005, $\longrightarrow$ preprint of journal version (in prep) 2007

## selected references ctd.

**On presevation under homomorphisms and unions of conjunctive queries**
Atserias,Dawar&Kolaitis, PODS 2004 and JACM 53, 2006.

**Preservation under extensions on well-behaved finite structures**
Atserias,Dawar&Grohe, ICALP 2005, LNCS 3580

**Existential positive types and preservation under homomorphisms**
Rossman, LICS 2005

**Capturing bisimulation-invariant Ptime**, O_, TCS 1999

**Boundedness of monadic FO over acyclic structures**
Kreutzer,O_&Schweikardt, ICALP 2007.

**The Boundedness Problem for Monadic Universal FO**, O_, LICS 2006