A representation theory of arboreal categories

Tomáš Jakl [a] (j.w.w. Luca Reggio)

Czech Technical University

27 July 2026

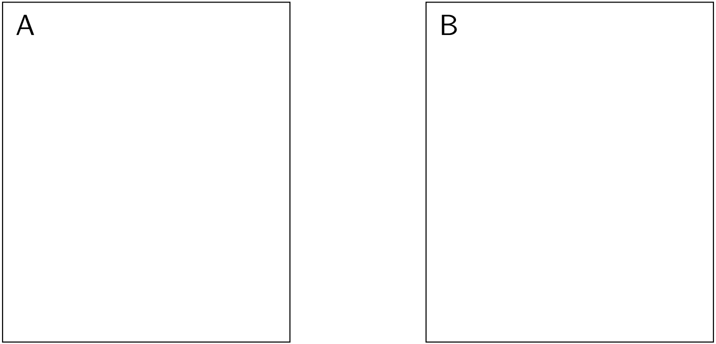
TACL 2026, Krakow

[a] This project has received funding from the EU's Horizon Europe research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101111373.

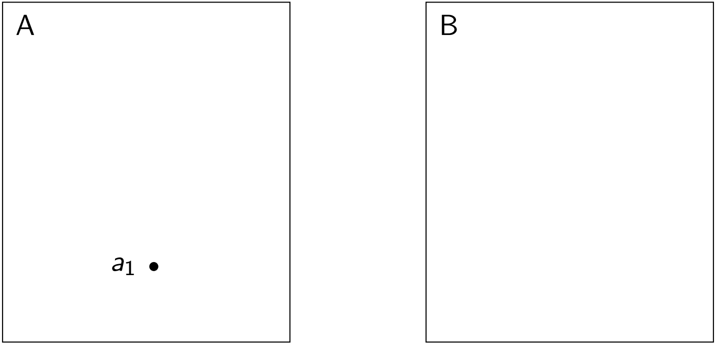
Funded by the European Union

Game Comonads

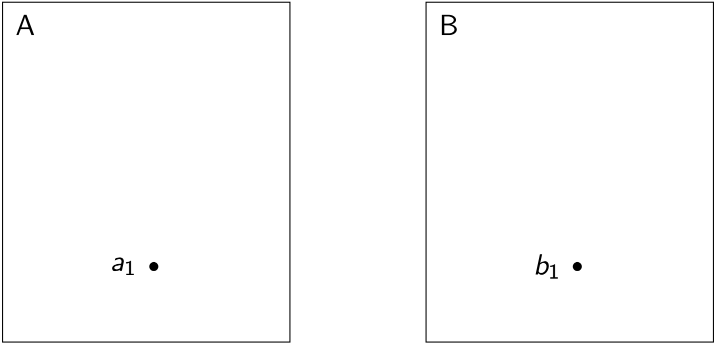
k-round Ehrenfeucht–Fraïssé game



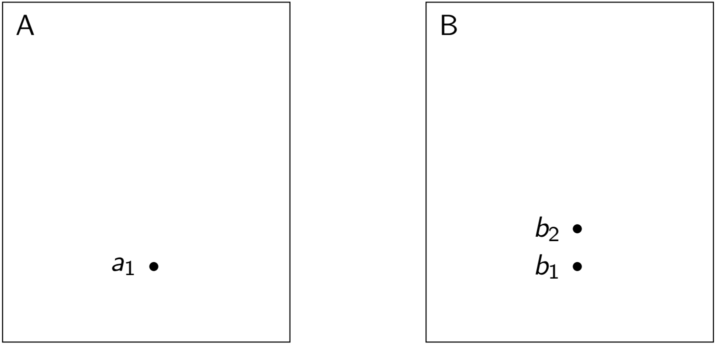
k-round Ehrenfeucht–Fraïssé game



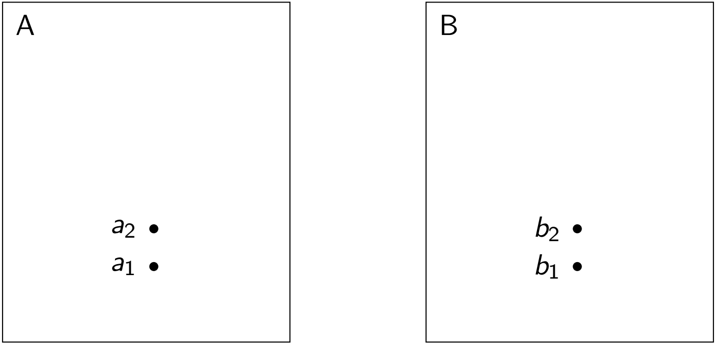
k-round Ehrenfeucht–Fraïssé game



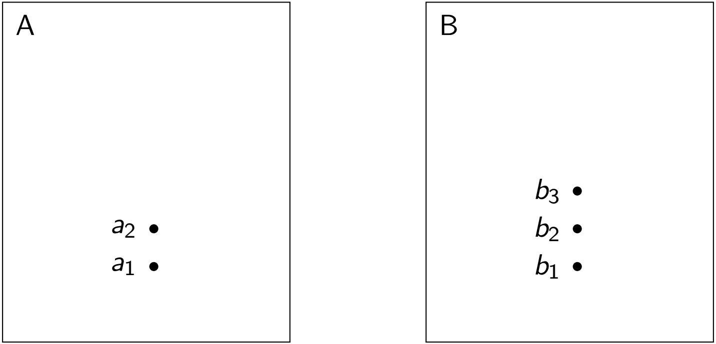
k-round Ehrenfeucht–Fraïssé game



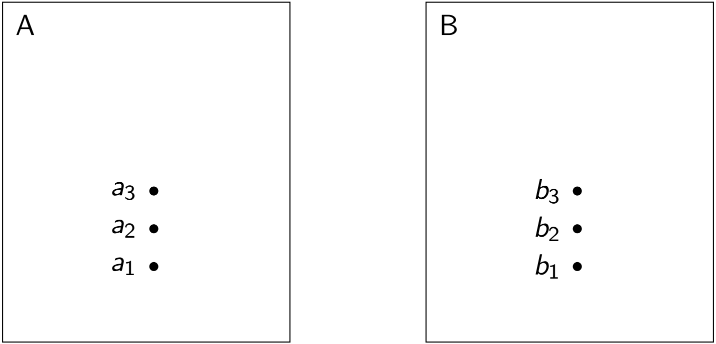
k-round Ehrenfeucht–Fraïssé game



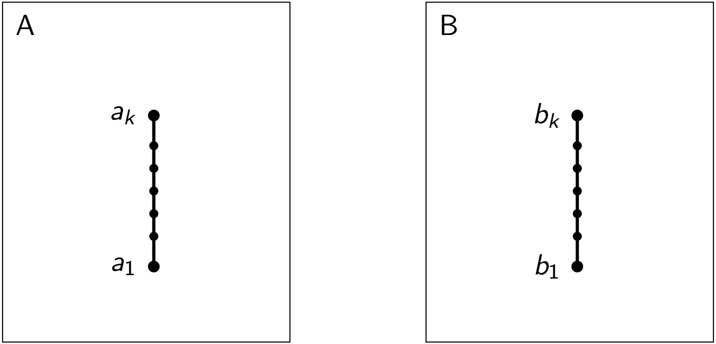
k-round Ehrenfeucht–Fraïssé game



k-round Ehrenfeucht–Fraïssé game

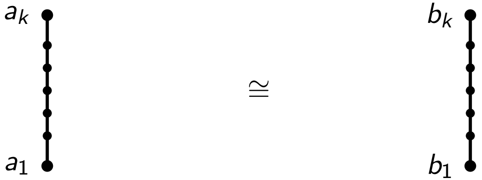


k-round Ehrenfeucht–Fraïssé game



k-round Ehrenfeucht–Fraïssé game

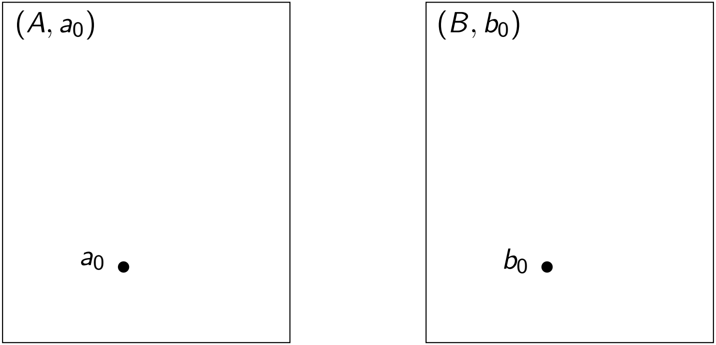
k-round Ehrenfeucht–Fraïssé game



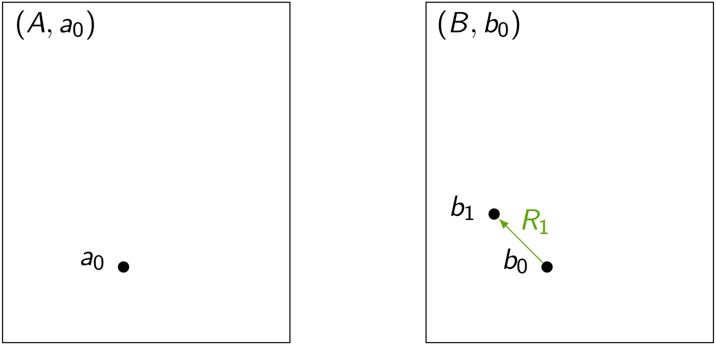
Theorem

$A \equiv_{\mathrm{FO}_k} B$ iff Duplicator wins in the k-round E–F game.

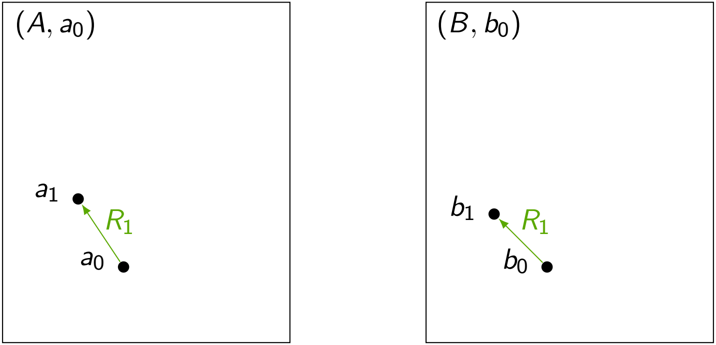
k-round bisimulation game



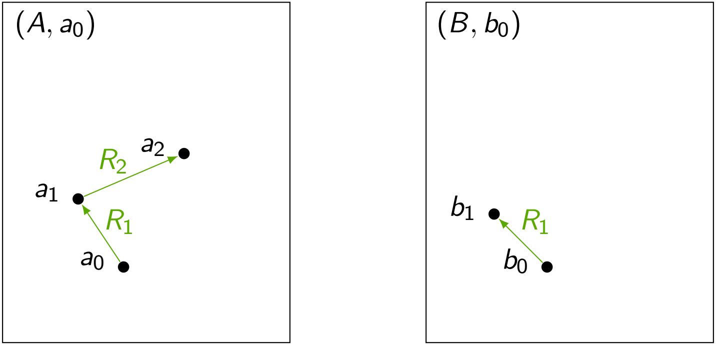
k-round bisimulation game



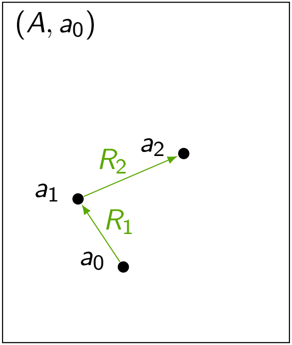
k-round bisimulation game

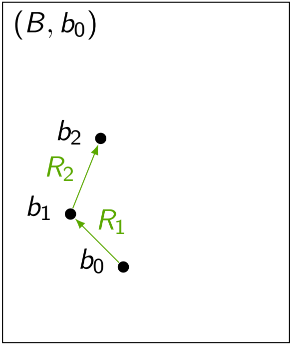


k-round bisimulation game



k-round bisimulation game





k-round bisimulation game

Check if $\quad a_i \in P^A \iff b_i \in P^B \qquad (\forall P, i)$

k-round bisimulation game

Check if $\quad a_i \in P^A \iff b_i \in P^B \qquad (\forall P, i)$

Theorem

$(A, a_0) \equiv_{\mathrm{ML}_k} (B, b_0)$ *iff Duplicator wins the k-round bisim. game.*

Game comonads in a nutshell

For a (well-behaved) model comparison game for logic $\mathscr{L}$

- unravelling of one structure following the rules of the game
  $\Rightarrow$ construction $\mathbb{C}\colon \mathcal{R}(\sigma) \to \mathcal{R}(\sigma)$

Game comonads in a nutshell

For a (well-behaved) model comparison game for logic $\mathscr{L}$

- unravelling of one structure following the rules of the game
  $\Rightarrow$ construction $\mathbb{C} \colon \mathcal{R}(\sigma) \to \mathcal{R}(\sigma)$

- $\mathbb{C}$ is a comonad $\Rightarrow$ adjunction

$$\begin{tikzcd} \mathrm{CoAlg}(\mathbb{C}) \arrow[d, bend right, "U^{\mathbb{C}}"'] \arrow[d, phantom, "\dashv"] \\ \mathcal{R}(\sigma) \arrow[u, bend right, "F^{\mathbb{C}}"'] \end{tikzcd}$$

Game comonads in a nutshell

For a (well-behaved) model comparison game for logic $\mathscr{L}$

- unravelling of one structure following the rules of the game
  $\Rightarrow$ construction $\mathbb{C}\colon \mathcal{R}(\sigma) \to \mathcal{R}(\sigma)$

- $\mathbb{C}$ is a comonad $\Rightarrow$ adjunction
$$\begin{tikzcd} \mathsf{CoAlg}(\mathbb{C}) \arrow[d, bend right, "U^{\mathbb{C}}"'] \arrow[d, phantom, "\dashv"] \\ \mathcal{R}(\sigma) \arrow[u, bend right, "F^{\mathbb{C}}"'] \end{tikzcd}$$

- $\mathsf{CoAlg}(\mathbb{C}) \approx$ tree-ordered structures wrt $\sqsubseteq$
  (not part of signature σ)

Game comonads in a nutshell

For a (well-behaved) model comparison game for logic $\mathscr{L}$

- unravelling of one structure following the rules of the game
  $\Rightarrow$ construction $\mathbb{C} \colon \mathcal{R}(\sigma) \to \mathcal{R}(\sigma)$

- $\mathbb{C}$ is a comonad $\Rightarrow$ adjunction

$$\begin{tikzcd} \mathrm{CoAlg}(\mathbb{C}) \arrow[d, bend right, "U^{\mathbb{C}}"'] \arrow[d, phantom, "\dashv"] \\ \mathcal{R}(\sigma) \arrow[u, bend right, "F^{\mathbb{C}}"'] \end{tikzcd}$$

- $\mathrm{CoAlg}(\mathbb{C}) \approx$ tree-ordered structures wrt $\sqsubseteq$
  (not part of signature σ)

- a $A \equiv_{\mathscr{L}} B$ iff bisimulation $F^{\mathbb{C}}(A) \sim F^{\mathbb{C}}(B)$

Example: comonads for FO_k and ML_k

For FO_k:

- $\mathcal{R}(\sigma) \xrightarrow{F^{\mathbb{E}_k}} \mathsf{CoAlg}(\mathbb{E}_k)$

  $F^{\mathbb{E}_k}(A) =$ sequences $\overline{a} = [a_1, \ldots, a_n]$ with $a_i \in A$ and $n \leq k$

- $\overline{a} \sqsubseteq \overline{b}$ iff $\overline{a}$ is a prefix of $\overline{b}$

- $A \equiv_{\mathrm{FO}_k} B$ iff $F^{\mathbb{E}_k}(A) \sim F^{\mathbb{E}_k}(B)$

For ML_k:

- $\mathcal{R}(\sigma) \xrightarrow{F^{\mathbb{M}_k}} \mathsf{CoAlg}(\mathbb{M}_k)$

  $F^{\mathbb{M}_k}(A) =$ unravel A to depth k

- $A \equiv_{\mathrm{ML}_k} B$ iff $F^{\mathbb{M}_k}(A) \sim F^{\mathbb{M}_k}(B)$

Example: comonads for FO_k and ML_k

For FO_k:

- $\mathcal{R}(\sigma) \xrightarrow{F^{\mathbb{E}_k}} \mathsf{CoAlg}(\mathbb{E}_k)$

  $F^{\mathbb{E}_k}(A) =$ sequences $\overline{a} = [a_1, \ldots, a_n]$ with $a_i \in A$ and $n \leq k$

- $\overline{a} \sqsubseteq \overline{b}$ iff $\overline{a}$ is a prefix of $\overline{b}$

- $A \equiv_{\mathrm{FO}_k} B$ iff $F^{\mathbb{E}_k}(A) \sim F^{\mathbb{E}_k}(B)$

For ML_k:

- $\mathcal{R}(\sigma) \xrightarrow{F^{\mathbb{M}_k}} \mathsf{CoAlg}(\mathbb{M}_k)$

  $F^{\mathbb{M}_k}(A) =$ unravel A to depth k

- $A \equiv_{\mathrm{ML}_k} B$ iff $F^{\mathbb{M}_k}(A) \sim F^{\mathbb{M}_k}(B)$

Similar pattern for:

bounded variable count, monadic second order, hybrid logic, guarded fragments, generalised quantifiers, description logic, restricted conjunction, ...

Arboreal categories

Expressing bisimilarity, categorical game

We need a category $\mathscr{A}$ (e.g. $\mathscr{A} = \mathsf{CoAlg}(\mathbb{C})$) with a class

- $\rightarrowtail \subseteq \mathscr{A}$ of embeddings
- $\mathscr{A}_p \subseteq \mathscr{A}$ of paths

  i.e. objects with a finite linear chain of $\rightarrowtail$-subobjects

Expressing bisimilarity, categorical game

We need a category $\mathscr{A}$ (e.g. $\mathscr{A} = \mathsf{CoAlg}(\mathbb{C})$) with a class

- $\rightarrowtail \subseteq \mathscr{A}$ of embeddings
- $\mathscr{A}_p \subseteq \mathscr{A}$ of paths

  i.e. objects with a finite linear chain of $\rightarrowtail$-subobjects

Then $X \sim Y$ in $\mathscr{A}$ if (for $P \in \mathscr{A}_p$)

$$X \longleftarrowtail P \rightarrowtail\longrightarrow Y$$

Expressing bisimilarity, categorical game

We need a category $\mathscr{A}$ (e.g. $\mathscr{A} = \mathsf{CoAlg}(\mathbb{C})$) with a class

- $\rightarrowtail \subseteq \mathscr{A}$ of embeddings
- $\mathscr{A}_p \subseteq \mathscr{A}$ of paths

  i.e. objects with a finite linear chain of $\rightarrowtail$-subobjects

Then $X \sim Y$ in $\mathscr{A}$ if (for $P, Q \in \mathscr{A}_p$)

$$\begin{tikzcd} X & P \arrow[l, tail] \arrow[r, tail] \arrow[d, tail, dashed] & Y \\ & Q \arrow[ul, tail, dashed] & \end{tikzcd}$$

Expressing bisimilarity, categorical game

We need a category $\mathscr{A}$ (e.g. $\mathscr{A} = \mathrm{CoAlg}(\mathbb{C})$) with a class

- $\rightarrowtail \subseteq \mathscr{A}$ of embeddings
- $\mathscr{A}_p \subseteq \mathscr{A}$ of paths

  i.e. objects with a finite linear chain of $\rightarrowtail$-subobjects

Then $X \sim Y$ in $\mathscr{A}$ if (for $P, Q \in \mathscr{A}_p$)

$$\begin{tikzcd}
X & P \arrow[l, tail] \arrow[r, tail] \arrow[d, tail] & Y \\
& Q \arrow[ul, tail] \arrow[ur, tail, dashed] &
\end{tikzcd}$$

Expressing bisimilarity, following Joyal–Nielsen–Winskel

Then $X \leftrightarrow Y$ in $\mathscr{A}$ if exists a span

$$X \longleftarrow Z \longrightarrow Y$$

of open morphisms.

Where $f\colon Z \to Y$ is open if (for $P, Q \in \mathscr{A}_p$)

$$\begin{tikzcd} Z \arrow[r, "f"] & Y \\ P \arrow[u] \arrow[r] & Q \arrow[u] \arrow[ul, dashed] \end{tikzcd}$$

Arboreal categories

Theorem (Abramsky–Reggio, 2021)

In arboreal categories with binary products, $X \sim Y$ iff $X \leftrightarrow Y$.

A category $\mathscr{A}$ with a stable[1] proper factorisation system $(\twoheadrightarrow, \rightarrowtail)$ is arboreal if

- paths are dense in $\mathscr{A}$
- it has $\mathbf{0}$, tree-colimits and paths are tree-connected
- 2-out-of-3 for paths $\quad$ ($P \to Q \to R$ quotient $\Rightarrow$ $P \to Q$ quotient)

[1] Stability of quotients along embeddings.

Arboreal categories

Theorem (Abramsky–Reggio, 2021)
In arboreal categories with binary products, $X \sim Y$ iff $X \leftrightarrow Y$.

A category $\mathscr{A}$ with a stable[1] proper factorisation system $(\twoheadrightarrow, \rightarrowtail)$ is arboreal if

- paths are dense in $\mathscr{A}$
- it has $\mathbf{0}$, tree-colimits and paths are tree-connected
- 2-out-of-3 for paths ($P \to Q \to R$ quotient $\Rightarrow P \to Q$ quotient)

Arboreal categories useful framework for results in logic, e.g.

- Feferman–Vaught–Mostowski thm, equi-rank Homomorphism Preservation Theorem, Hodges' word construction

[1]Stability of quotients along embeddings.

The usual situation

unravelling functor

$$\mathscr{X} \xrightarrow{\mathsf{F}} \mathscr{A}$$

The usual situation

unravelling functor

$$\mathscr{X} \xrightarrow{\quad \mathsf{F} \quad} \mathscr{A}$$

typically σ-structures

arboreal category
($\sim$ expressing logic)

Representing arboreal categories

Since paths are dense

$$\mathscr{A} \xhookrightarrow{\textit{fully faithful}} \widehat{\mathscr{A}_p} \qquad A \mapsto \hom(I(-), A)$$

for $I \colon \mathscr{A}_p \hookrightarrow \mathscr{A}$

Since paths are dense

$$\mathscr{A} \xhookrightarrow{\textit{fully faithful}} \widehat{\mathscr{A}_p} \qquad\qquad A \mapsto \hom(I(-), A)$$

$$\text{for } I \colon \mathscr{A}_p \hookrightarrow \mathscr{A}$$

but not for the restriction $\mathscr{A}_p^e \;:=\; \mathscr{A}_p \cap (\rightarrowtail)$

$$\mathscr{A} \longrightarrow \widehat{\mathscr{A}_p^e} \qquad\qquad A \mapsto \hom(I^e(-), A)$$

$$\text{for } I^e \colon \mathscr{A}_p^e \hookrightarrow \mathscr{A}$$

Key observations, 1

- Subobject of a path is a path $\implies$ fact. sys. restricts to $\mathscr{A}_p$

$$P \longrightarrow Q \qquad = \qquad P \twoheadrightarrow S \rightarrowtail Q$$

$\Rightarrow$ Information about quotients missing in $\mathscr{A} \to \widehat{\mathscr{A}_p^e}$

Key observations, 1

- Subobject of a path is a path $\implies$ fact. sys. restricts to $\mathscr{A}_p$

$$P \longrightarrow Q \quad = \quad P \longrightarrow\!\!\!\!\!\rightarrow S \rightarrowtail Q$$

$\Rightarrow$ Information about quotients missing in $\mathscr{A} \to \widehat{\mathscr{A}_p^e}$

- If $P \twoheadrightarrow Q$ then P and Q have the same height
(= length of their chain of subobjects)

Key observations, 1

- Subobject of a path is a path $\implies$ fact. sys. restricts to $\mathscr{A}_p$

$$P \longrightarrow Q \quad = \quad P \twoheadrightarrow S \rightarrowtail Q$$

$\Rightarrow$ Information about quotients missing in $\mathscr{A} \to \widehat{\mathscr{A}_p^e}$

- If $P \twoheadrightarrow Q$ then P and Q have the same height
(= length of their chain of subobjects)

- $\mathscr{A}_p^e$ can be seen as a forest

$$\begin{tikzcd}
 & P_1 & & P_3 \\
P_0 \arrow[ur, tail] \arrow[r, tail] \arrow[dr, tail] & P'_1 \arrow[r, tail] \arrow[dr, tail] & P_2 \arrow[ur, tail] \arrow[r, tail] \arrow[dr, tail] & P'_3 \\
 & P''_1 & P'_2 & P''_3
\end{tikzcd}$$

Key observations, 1

- Subobject of a path is a path $\implies$ fact. sys. restricts to $\mathscr{A}_p$

$$P \longrightarrow Q \quad = \quad P \twoheadrightarrow S \rightarrowtail Q$$

  $\Rightarrow$ Information about quotients missing in $\mathscr{A} \to \widehat{\mathscr{A}_p^e}$

- If $P \twoheadrightarrow Q$ then P and Q have the same height (= length of their chain of subobjects)

- $\mathscr{A}_p$ can be seen as a forest with quotients between paths of the same height

$$\begin{tikzcd}
 & P_1 & & P_3 \\
P_0 \arrow[ur, tail] \arrow[r, tail] \arrow[dr, tail] & P_1' \arrow[u, two heads] \arrow[r, tail] \arrow[dr, tail] & P_2 \arrow[ur, tail] \arrow[r, tail] \arrow[dr, tail] & P_3' \\
 & P_1'' \arrow[u, two heads] & P_2' & P_3'' \arrow[u, two heads] \arrow[uu, bend right, two heads]
\end{tikzcd}$$

Key observations, 2

- Forest $(F, \leq) \approx$ functors $\Phi \colon \mathbb{N}^{\mathrm{op}} \to \mathbf{Set}$

$$F \mapsto \Phi_F(n) = \{x \in F \mid \text{ height of } x \text{ is } n\}$$

- Given (suitable) arboreal cat. $\mathscr{A}$, define $\Phi_{\mathscr{A}} \colon \mathbb{N}^{\mathrm{op}} \to \mathbf{Set}$
  - $\Phi_{\mathscr{A}}(n) = \{$ paths of $\mathscr{A}$ of height n $\}$
  - $\Phi_{\mathscr{A}}(n \leq m) \colon P \mapsto$ the prefix of P of height n

- Together with the quotient order:

$$\Phi_{\mathscr{A}} \colon \mathbb{N}^{\mathrm{op}} \to \mathbf{Pos}$$

where $P \leq Q$ in $\Phi_{\mathscr{A}}(n)$ $\iff$ exists $P \twoheadrightarrow Q$ in $\mathscr{A}$

Arboreal categories from functors

Given any

$$\Phi\colon \mathbb{N}^{\mathrm{op}} \to \mathbf{Pos}$$

where $\Phi(0)$ has a bottom.

Define category $\mathrm{Ar}(\Phi)$ with objects the Φ-labelled trees, ie.

$$(T, t\colon T \to \Phi)$$

(viewing T as a functor $\mathbb{N}^{\mathrm{op}} \to \mathbf{Set} \hookrightarrow \mathbf{Pos}$)

Morphisms in $\mathrm{Ar}(\Phi)$ are

$$\begin{tikzcd} T \arrow[rr, "f"] \arrow[dr, "t"'] & \arrow[d, phantom, "\leq"] & S \arrow[dl, "s"] \\ & \Phi & \end{tikzcd}$$

Representation theorems

Theorem

For an arboreal category $\mathscr{A}$ *with a set of paths and* $|\operatorname{hom}(P, Q)| \leq 1$

$$\mathscr{A} \cong \mathsf{Ar}(\Phi_{\mathscr{A}}).$$

For $\Phi \colon \mathbb{N}^{op} \to \mathbf{Pos}$ *where* $\Phi(0)$ *has a bottom*

$$\Phi \cong \Phi_{\mathsf{Ar}(\Phi)}.$$

Consequences

New examples? (Future work)

(1) Given a hyperdoctrine $H \colon \mathscr{C} \to \mathbf{DLat}$ with an enumeration of variables $\mathbb{N} \hookrightarrow \mathscr{C}$, we have

$$\Phi = \mathbb{N}^{\mathsf{op}} \hookrightarrow \mathbb{C}^{\mathsf{op}} \to \mathbf{DLat}^{\mathsf{op}} \to \mathbf{Prie} \to \mathbf{Pos}$$

Then $\mathsf{Ar}(\Phi) \;\approx\;$ theories of H

(2) Proof trees from

$$\Phi(n) = \{[\varphi_1, \ldots, \varphi_n] \mid T, \varphi_1, \ldots, \varphi_{n-1} \vdash \varphi_n \text{ using one rule}\}$$

New examples? (Future work)

(3) Unravelling of coalgebras, given $\alpha\colon A \to F(A)$

$$\Phi(n) = \{[x_1, \ldots, x_n] \mid x_{i+1} \in \tau(\alpha(x_i))\}$$

where $\tau\colon F(A) \to \mathcal{P}(A)$ is the support of F.

(...) ...

What is your favourite functor $\mathbb{N}^{\mathsf{op}} \to \mathbf{Pos}$?

Consequences of the representation theorem, 1

Consider a pushout

$$\begin{array}{ccc} A & \xrightarrow{g} & B \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle \overline{f}} \\ C & \xrightarrow{\overline{g}} & D \end{array}$$

If f, g open then so are $\overline{f}, \overline{g}$

$\Longrightarrow$ Define quotients $A/_\sim$ by the largest bisimulation.

$\Longrightarrow$ Coinduction on types as $F^{\mathbb{C}}(X)/_\sim \cong$ types

If f, g embeddings so are $\overline{f}, \overline{g}$

$\Longrightarrow$ Free amalgamations.

$\Longrightarrow$ Fraïssé limits (if $\mathscr{A}$ cocomplete).

Consequences of the representation theorem, 2

- If $\Phi(0) = \{\star\}$ then

$$\mathbf{Trees}/|\Phi| \;\subseteq\; \mathrm{Ar}(\Phi)$$

  I.e. such arboreal category contains a topos as a wide subcategory, on its pathwise-embeddings.

- $\mathrm{Ar}(\Phi)$ is order-enriched: $f \leq g : (T, t) \to (S, s)$ if $sf \leq sg$

- $\mathrm{Ar}(\Phi)$ is fibred over **Trees**, via the forgetful

$$\mathrm{Ar}(\Phi) \to \mathbf{Trees}, \quad (T, t) \mapsto T.$$

- ...

Conclusion

- Representation theorem
  - allows to easily construct counterexamples
  - validates our intuition e.g. description of (open) pathwise-embeddings, embeddings, ...
  - many direct consequences
- Future work:
  - Develop examples
  - Dropping $\quad |\mathrm{hom}(P, Q)| \leq 1 \quad \Longrightarrow \quad \Phi \colon \mathbb{N}^{\mathrm{op}} \to \mathbf{Cat}$
- Connections to locally finitely presentable categories:

Categorical Model Theory	Categorical Finite MT
lfp/accessible categories	arboreal categories
compact objects directed colimits	paths tree-colimits
stability, algebraic theories, Gabriel–Ulmer, ...	game comonads, bisimulation, pres. thms., FVM, ...

Thank you!