

# Applying automata and games to simply typed lambda calculus

Colin Stirling  
cps@inf.ed.ac.uk

LFCS  
School of Informatics  
University of Edinburgh

EXPRESS/SOS, Newcastle, 3rd September 2012

# Methods for verifying finite and infinite state systems

- ▶ Notable success in Computer Science
- ▶ **model checking + equivalence checking**
- ▶ System = finite/infinite state transition graph

# Methods for verifying finite and infinite state systems

- ▶ Notable success in Computer Science
- ▶ model checking + equivalence checking
- ▶ System = finite/infinite state transition graph
- ▶ Model checking: does state  $s$  have property  $\Phi$  ?
- ▶ apply automata/game theoretic techniques to solve it: mostly computing monadic fixed points, reachability sets by traversing graph (possibly repeatedly)

# Methods for verifying finite and infinite state systems

- ▶ Notable success in Computer Science
- ▶ **model checking + equivalence checking**
- ▶ System = finite/infinite state transition graph
- ▶ Model checking: **does state  $s$  have property  $\Phi$  ?**
- ▶ apply automata/game theoretic techniques to solve it: mostly computing monadic fixed points, reachability sets by traversing graph (possibly repeatedly)
- ▶ Equivalence checking: **is state  $s$  equivalent to  $t$  ?**
- ▶ Mostly computing dyadic fixed points e.g. bisimulations to solve it. May need algebraic/combinatorial properties of reachability sets/generators of graph

# Active research goal: transfer these techniques to

finite/infinite state systems with **binding**

1. Deciding observational equivalence for fragments of idealized Algol and ML (w.r.t. finite value sets)  
[Ghica, McCusker 2000; Ong 2002; Hopkins, Murawski, Ong 2012; ...]

# Active research goal: transfer these techniques to

finite/infinite state systems with **binding**

1. Deciding observational equivalence for fragments of idealized Algol and ML (w.r.t. finite value sets)  
[Ghica, McCusker 2000; Ong 2002; Hopkins, Murawski, Ong 2012; ...]
2. Model checking higher-order trees/schemes  
[Knapik, Niwinski, Urzyczyn 2002; Caucal 2002; Ong 2006; Hague, Murawski, Ong, Serre 2008, Kobayashi 2009; Kobayashi, Ong 2009; Kobayashi 2011; ...]

# Active research goal: transfer these techniques to

finite/infinite state systems with **binding**

1. Deciding observational equivalence for fragments of idealized Algol and ML (w.r.t. finite value sets)  
[Ghica, McCusker 2000; Ong 2002; Hopkins, Murawski, Ong 2012; ...]
2. Model checking higher-order trees/schemes  
[Knapik, Niwinski, Urzyczyn 2002; Caucal 2002; Ong 2006; Hague, Murawski, Ong, Serre 2008, Kobayashi 2009; Kobayashi, Ong 2009; Kobayashi 2011; ...]
3. : : :
4. Application of tree automata and games to deciding higher-order matching  
[Comon, Jurski 1997; Stirling 2005-2012, work described here]

# Simply typed $\lambda$ -calculus

- ▶ Fundamental exemplar of binding



# Simply typed $\lambda$ -calculus

- ▶ Fundamental exemplar of binding
- ▶ Types  $A ::= \mathbf{0} \mid A \rightarrow A$ 
  - ▶  $\mathbf{0}$  single base type (for simplicity)
  - ▶  $A \rightarrow B$  type of functions from  $A$  to  $B$

# Simply typed $\lambda$ -calculus

- ▶ Fundamental exemplar of binding
- ▶ Types  $A ::= \mathbf{0} \mid A \rightarrow A$ 
  - ▶  $\mathbf{0}$  single base type (for simplicity)
  - ▶  $A \rightarrow B$  type of functions from  $A$  to  $B$
- ▶ If  $A \neq \mathbf{0}$  then  $A$  has form  $A_1 \rightarrow \dots \rightarrow A_n \rightarrow \mathbf{0}$   
abbreviated to  $(A_1, \dots, A_n, \mathbf{0})$

# Simply typed $\lambda$ -calculus

- ▶ Fundamental exemplar of binding
- ▶ Types  $A ::= \mathbf{0} \mid A \rightarrow A$ 
  - ▶  $\mathbf{0}$  single base type (for simplicity)
  - ▶  $A \rightarrow B$  type of functions from  $A$  to  $B$
- ▶ If  $A \neq \mathbf{0}$  then  $A$  has form  $A_1 \rightarrow \dots \rightarrow A_n \rightarrow \mathbf{0}$   
abbreviated to  $(A_1, \dots, A_n, \mathbf{0})$
- ▶ Order of  $\mathbf{0}$  is 1;
- ▶ Order of  $(A_1, \dots, A_n, \mathbf{0})$  is  $k + 1$  where  $k$  is maximum of orders of  $A_i$ s

# Terms of simply type $\lambda$ -calculus

Variables and constants each have a unique type (Church style)

The smallest set  $T$  of simply typed terms is:

# Terms of simply type $\lambda$ -calculus

Variables and constants each have a unique type (Church style)

The smallest set  $T$  of simply typed terms is:

1. if  $x$  ( $f$ ) has type  $A$  then  $x : A \in T$  ( $f : A \in T$ )
2. if  $t : B \in T$  and  $x : A \in T$  then  $\lambda x.t : A \rightarrow B \in T$
3. if  $t : A \rightarrow B \in T$  and  $u : A \in T$  then  $(tu) : B \in T$

# Terms of simply type $\lambda$ -calculus

Variables and constants each have a unique type (Church style)

The smallest set  $T$  of simply typed terms is:

1. if  $x$  ( $f$ ) has type  $A$  then  $x : A \in T$  ( $f : A \in T$ )
2. if  $t : B \in T$  and  $x : A \in T$  then  $\lambda x.t : A \rightarrow B \in T$
3. if  $t : A \rightarrow B \in T$  and  $u : A \in T$  then  $(tu) : B \in T$

- ▶ order of  $t : A = \text{order } A$
- ▶ closed  $t : A$  no free variables
- ▶  $t, t' : A$  are  $\alpha$ -equivalent renamings of each other

## Dynamics: reduction

$$\begin{array}{l} (\beta) \quad (\lambda x.t)v \rightarrow_{\beta} t\{v/x\} \quad \{\cdot/\cdot\} \text{ Substitution} \\ (\eta) \quad \lambda x.(tx) \rightarrow_{\eta} t \quad x \text{ not free in } t \end{array}$$

## Dynamics: reduction

$$\begin{array}{l} (\beta) \quad (\lambda x.t)v \rightarrow_{\beta} t\{v/x\} \quad \{\cdot/\cdot\} \text{ Substitution} \\ (\eta) \quad \lambda x.(tx) \rightarrow_{\eta} t \quad x \text{ not free in } t \end{array}$$

### ► Facts

Strong normalisation and confluence (of  $\rightarrow_{\beta}$ ,  $\rightarrow_{\eta}$ ,  $\rightarrow_{\beta\eta}$ )



# Dynamics: reduction

$$\begin{array}{l} (\beta) \quad (\lambda x.t)v \rightarrow_{\beta} t\{v/x\} \quad \{\cdot/\cdot\} \text{ Substitution} \\ (\eta) \quad \lambda x.(tx) \rightarrow_{\eta} t \quad x \text{ not free in } t \end{array}$$

## ► Facts

Strong normalisation and confluence (of  $\rightarrow_{\beta}$ ,  $\rightarrow_{\eta}$ ,  $\rightarrow_{\beta\eta}$ )

## ► Equivalence: $t =_{\beta\eta} t'$ if there are $s, s'$

$$\begin{array}{l} t \rightarrow_{\beta\eta}^* s \\ t' \rightarrow_{\beta\eta}^* s' \\ \parallel_{\alpha} \end{array}$$

# Long Normal Forms

- ▶  $t$  in  $\beta$ -normal form if no  $t'$  such that  $t \rightarrow_{\beta} t'$

# Long Normal Forms

- ▶  $t$  in  $\beta$ -normal form if no  $t'$  such that  $t \rightarrow_{\beta} t'$
- ▶  $t$  in  $\eta$ -long  $\beta$ -normal form if  $t$  is in  $\beta$ -normal form and there is no  $t'$  such that  $t' \rightarrow_{\eta} t$

# Long Normal Forms

- ▶  $t$  in  $\beta$ -normal form if no  $t'$  such that  $t \rightarrow_{\beta} t'$
- ▶  $t$  in  $\eta$ -long  $\beta$ -normal form if  $t$  is in  $\beta$ -normal form and there is no  $t'$  such that  $t' \rightarrow_{\eta} t$
- ▶ Inf =  $\eta$ -long  $\beta$ -normal form
- ▶ If  $t, v$  are in Inf and  $tv \rightarrow_{\beta}^* s$  in  $\beta$ -normal form then  $s$  is in Inf  
No  $\eta$ -reductions when terms in Inf

# Long Normal Forms

- ▶  $t$  in  $\beta$ -normal form if no  $t'$  such that  $t \rightarrow_{\beta} t'$
- ▶  $t$  in  $\eta$ -long  $\beta$ -normal form if  $t$  is in  $\beta$ -normal form and there is no  $t'$  such that  $t' \rightarrow_{\eta} t$
- ▶ Inf =  $\eta$ -long  $\beta$ -normal form
- ▶ If  $t, v$  are in Inf and  $tv \rightarrow_{\beta}^* s$  in  $\beta$ -normal form then  $s$  is in Inf  
No  $\eta$ -reductions when terms in Inf
- ▶  $T_A(C)$  is the set of closed Inf terms of type  $A$  whose constants belong to  $C$ .

## Example

- ▶  $x : (\mathbf{0}, \mathbf{0}) \lambda x.x : ((\mathbf{0}, \mathbf{0}), \mathbf{0}, \mathbf{0})$  in  $\beta$ -normal form not Inf

# Example

- ▶  $x : (\mathbf{0}, \mathbf{0}) \quad \lambda x.x : ((\mathbf{0}, \mathbf{0}), \mathbf{0}, \mathbf{0})$  in  $\beta$ -normal form not Inf
- ▶  $z : \mathbf{0}$  then  $\lambda z.xz : (\mathbf{0}, \mathbf{0})$  and  $\lambda xz.xz : ((\mathbf{0}, \mathbf{0}), \mathbf{0}, \mathbf{0})$  are in Inf

## Example

- ▶  $x : (\mathbf{0}, \mathbf{0}) \lambda x.x : ((\mathbf{0}, \mathbf{0}), \mathbf{0}, \mathbf{0})$  in  $\beta$ -normal form not Inf
- ▶  $z : \mathbf{0}$  then  $\lambda z.xz : (\mathbf{0}, \mathbf{0})$  and  $\lambda xz.xz : ((\mathbf{0}, \mathbf{0}), \mathbf{0}, \mathbf{0})$  are in Inf
- ▶ Notation:  $\lambda x_1 \dots x_k.t$  abbreviates  $\lambda x_1 \dots \lambda x_k.t$
- ▶ Monster type  $M = (((\mathbf{0}, \mathbf{0}), \mathbf{0}), \mathbf{0}, \mathbf{0}, \mathbf{0})$ , has order 5
- ▶  $\lambda xy.x(\lambda z_1.x(\lambda z_2.z_1y)) \in T_M(\emptyset)$   
when  $x : (((\mathbf{0}, \mathbf{0}), \mathbf{0}), \mathbf{0})$ ,  $z_i : (\mathbf{0}, \mathbf{0})$  and  $y : \mathbf{0}$ .



# Decision questions

- ▶ Higher-order unification
- ▶  $v = u$  contains free variables  $x_1, \dots, x_n$
- ▶ Solution  $\theta = \{t_1/x_1, \dots, t_n/x_n\}$  such that  $v\theta =_{\beta\eta} u\theta$   
Simultaneous substitution

# Decision questions

- ▶ Higher-order unification
- ▶  $v = u$  contains free variables  $x_1, \dots, x_n$
- ▶ Solution  $\theta = \{t_1/x_1, \dots, t_n/x_n\}$  such that  $v\theta =_{\beta\eta} u\theta$   
Simultaneous substitution
- ▶ Decision question: given  $v = u$ , does it have a solution ?
- ▶ Order is max order of the  $x_i$ s

# Decision questions

- ▶ Higher-order unification
- ▶  $v = u$  contains free variables  $x_1, \dots, x_n$
- ▶ Solution  $\theta = \{t_1/x_1, \dots, t_n/x_n\}$  such that  $v\theta =_{\beta\eta} u\theta$   
Simultaneous substitution
- ▶ Decision question: given  $v = u$ , does it have a solution ?
- ▶ Order is max order of the  $x_i$ s
- ▶ Undecidable (even at order 2) [Huet 1972; Goldfarb 1981]

## Decision questions (cont)

- ▶ Higher-order matching
- ▶  $v = u$  contains free variables  $x_1, \dots, x_n$  BUT  $u$  closed
- ▶ Solution  $\theta = \{t_1/x_1, \dots, t_n/x_n\}$  such that  $v\theta =_{\beta\eta} u$   
W.l.o.g assume  $v, u : \mathbf{0}$

## Decision questions (cont)

- ▶ Higher-order matching
- ▶  $v = u$  contains free variables  $x_1, \dots, x_n$  BUT  $u$  closed
- ▶ Solution  $\theta = \{t_1/x_1, \dots, t_n/x_n\}$  such that  $v\theta =_{\beta\eta} u$   
W.l.o.g assume  $v, u : \mathbf{0}$
- ▶ Decision question: given  $v = u$ , does it have a solution ?
- ▶ Order is max order of the  $x_i$ s

## Decision questions (cont)

- ▶ Higher-order matching
- ▶  $v = u$  contains free variables  $x_1, \dots, x_n$  BUT  $u$  closed
- ▶ Solution  $\theta = \{t_1/x_1, \dots, t_n/x_n\}$  such that  $v\theta =_{\beta\eta} u$   
W.l.o.g assume  $v, u : \mathbf{0}$
- ▶ Decision question: given  $v = u$ , does it have a solution ?
- ▶ Order is max order of the  $x_i$ s
- ▶ Huet conjecture decidable [Huet 1976]
- ▶ Up to order 4 decidable + special cases [Huet 1976, Dowek 1993, Padovani 2000, ...]

## Decision questions (cont)

- ▶ Higher-order matching
- ▶  $v = u$  contains free variables  $x_1, \dots, x_n$  BUT  $u$  closed
- ▶ Solution  $\theta = \{t_1/x_1, \dots, t_n/x_n\}$  such that  $v\theta =_{\beta\eta} u$   
W.l.o.g assume  $v, u : \mathbf{0}$
- ▶ Decision question: given  $v = u$ , does it have a solution ?
- ▶ Order is max order of the  $x_i$ s
- ▶ Huet conjecture decidable [Huet 1976]
- ▶ Up to order 4 decidable + special cases [Huet 1976, Dowek 1993, Padovani 2000, ...]
- ▶ Undecidable for  $=_{\beta}$  [Loader 2003]

## Decision questions (cont)

- ▶ Higher-order matching
- ▶  $v = u$  contains free variables  $x_1, \dots, x_n$  BUT  $u$  closed
- ▶ Solution  $\theta = \{t_1/x_1, \dots, t_n/x_n\}$  such that  $v\theta =_{\beta\eta} u$   
W.l.o.g assume  $v, u : \mathbf{0}$
- ▶ Decision question: given  $v = u$ , does it have a solution ?
- ▶ Order is max order of the  $x_i$ s
- ▶ Huet conjecture decidable [Huet 1976]
- ▶ Up to order 4 decidable + special cases [Huet 1976, Dowek 1993, Padovani 2000, ...]
- ▶ Undecidable for  $=_{\beta}$  [Loader 2003]
- ▶ Decidable for all orders [Stirling 2006; 2009; 2012]



## Matching is essentially monadic

- ▶ Given  $v = u$  with  $x_1 : A_1, \dots, x_n : A_n$  in  $v$
- ▶ there is the matching problem  $x(\lambda x_1 \dots x_n.v) = u$   
where  $x : ((A_1, \dots, A_n, \mathbf{0}), \mathbf{0})$

# Matching is essentially monadic

- ▶ Given  $v = u$  with  $x_1 : A_1, \dots, x_n : A_n$  in  $v$
- ▶ there is the matching problem  $x(\lambda x_1 \dots x_n.v) = u$   
where  $x : ((A_1, \dots, A_n, \mathbf{0}), \mathbf{0})$
- ▶ Conceptually simpler problem: just one free variable
- ▶ Called "interpolation":  $x w_1 \dots w_k = u$   
where  $x : (B_1, \dots, B_k, \mathbf{0}), w_i : B_i, u : \mathbf{0}$  in Inf.
- ▶ Solution  $t$  in Inf such that  $tw_1 \dots w_k \rightarrow_{\beta}^* u$

# Matching is essentially monadic

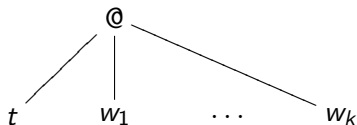
- ▶ Given  $v = u$  with  $x_1 : A_1, \dots, x_n : A_n$  in  $v$
- ▶ there is the matching problem  $x(\lambda x_1 \dots x_n. v) = u$   
where  $x : ((A_1, \dots, A_n, \mathbf{0}), \mathbf{0})$
  
- ▶ Canonical solution  $\lambda z. z t_1 \dots t_n$  where  $t_i$ s are closed  
Consequently  $v\{t_1/x_1, \dots, t_n/x_n\} \rightarrow_{\beta}^* u$   
So,  $\theta = \{t_1/x_1, \dots, t_n/x_n\}$  solves  $v = u$   
Reduces matching to interpolation

# Matching is essentially monadic

- ▶ Given  $v = u$  with  $x_1 : A_1, \dots, x_n : A_n$  in  $v$
- ▶ there is the matching problem  $x(\lambda x_1 \dots x_n.v) = u$   
where  $x : ((A_1, \dots, A_n, \mathbf{0}), \mathbf{0})$
  
- ▶ Canonical solution  $\lambda z.z t_1 \dots t_n$  where  $t_i$ s are closed  
Consequently  $v\{t_1/x_1, \dots, t_n/x_n\} \rightarrow_{\beta}^* u$   
So,  $\theta = \{t_1/x_1, \dots, t_n/x_n\}$  solves  $v = u$   
Reduces matching to interpolation
- ▶ Restrict constants in solution terms to be those in  $u$  plus fresh  $b : \mathbf{0}$   
Restricts to finitely many constants

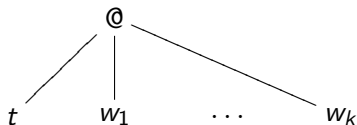
# Interpolation trees

- ▶  $t$  (of the right type and in Inf) a potential solution to  $xw_1 \dots w_k = u$  (Assume  $u$  does not have bound variables)
- ▶ Interpolation tree



# Interpolation trees

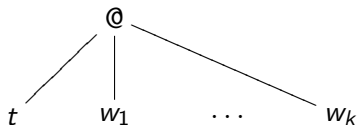
- ▶  $t$  (of the right type and in Inf) a potential solution to  $xw_1 \dots w_k = u$  (Assume  $u$  does not have bound variables)
- ▶ Interpolation tree



- ▶ Goal: understand the reduction of  $tw_1 \dots w_k$  to normal form by only examining the interpolation tree

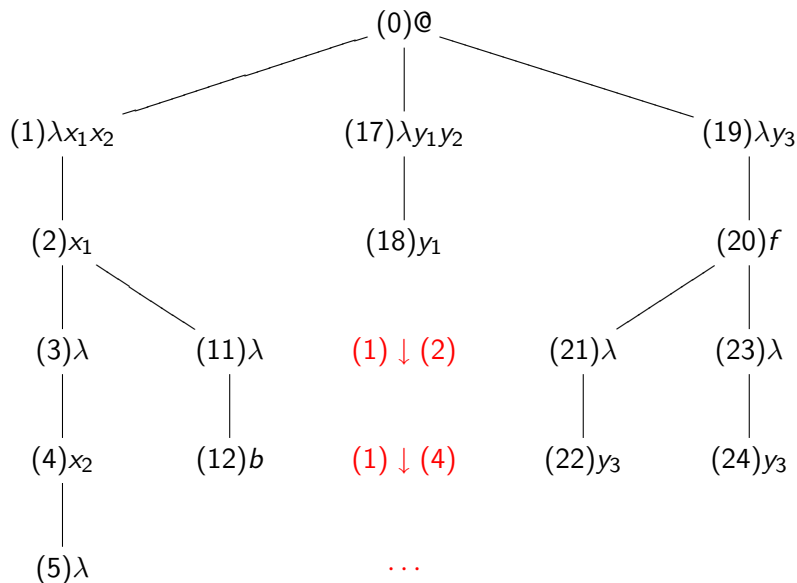
# Interpolation trees

- ▶  $t$  (of the right type and in Inf) a potential solution to  $xw_1 \dots w_k = u$  (Assume  $u$  does not have bound variables)
- ▶ Interpolation tree



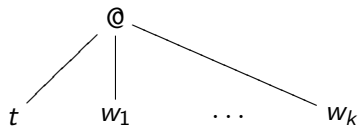
- ▶ Goal: understand the reduction of  $tw_1 \dots w_k$  to normal form by only examining the interpolation tree
- ▶  $t, w_1, \dots, w_k$  binding trees with dummy lambdas and binding relation  $\downarrow$  between nodes
- ▶  $n \downarrow m$  if  $n$  labelled  $\lambda \bar{y}$  binds  $y_j$ , label at  $m$

Example:  $x(\lambda y_1 y_2 . y_1)(\lambda y_3 . f y_3 y_3) = faa$



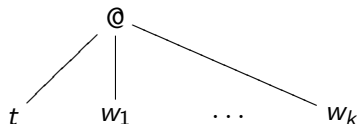


# Tree automata



- ▶ Are solutions recognisable by an automaton ?

# Tree automata



► Are solutions recognisable by an automaton ?

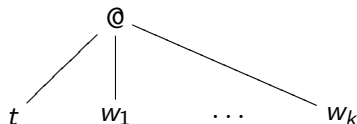
► Tree automaton

Finite sets: states  $Q$ , alphabet  $\Sigma$ , final states  $F \subseteq Q$ ,

transitions  $\Delta$  of form  $sq_1 \dots q_k \Rightarrow q$ ,

$k \geq 0$ ,  $s \in \Sigma$ ,  $\text{arity}(s) = k$  and  $q, q_i \in Q$

# Tree automata



- ▶ Are solutions recognisable by an automaton ?

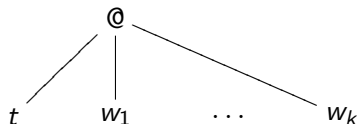
- ▶ Tree automaton

Finite sets: states  $Q$ , alphabet  $\Sigma$ , final states  $F \subseteq Q$ ,  
transitions  $\Delta$  of form  $sq_1 \dots q_k \Rightarrow q$ ,  
 $k \geq 0$ ,  $s \in \Sigma$ ,  $\text{arity}(s) = k$  and  $q, q_i \in Q$

- ▶ Using transitions label tree bottom-up with states

Automaton accepts tree iff root can be labelled with a final state

# Tree automata



- ▶ Are solutions recognisable by an automaton ?
- ▶ Tree automaton
  - Finite sets: states  $Q$ , alphabet  $\Sigma$ , final states  $F \subseteq Q$ ,
  - transitions  $\Delta$  of form  $sq_1 \dots q_k \Rightarrow q$ ,
  - $k \geq 0$ ,  $s \in \Sigma$ ,  $\text{arity}(s) = k$  and  $q, q_i \in Q$
- ▶ Using transitions label tree bottom-up with states
  - Automaton accepts tree iff root can be labelled with a final state
- ▶ Non-emptiness decidable; sets recognised are regular

## Tree automata for 4th-order [Comon, Jurski 97]

- ▶ 4th-order  $E$ ,  $xw_1 \dots w_k = u$  and  $C_E$  constants in  $u$  plus  $b$
- ▶ For finite alphabet  $\Sigma$  consider a potential solution term  
 $t = \lambda x_1 \dots x_k . s$

$$s ::= z_j^n \mid x_i v_1 \dots v_m \mid f s_1 \dots s_m \quad f \in C_E, m \geq 0$$
$$v ::= s \mid \lambda z_1^n \dots z_m^n . s \quad m \geq 0 \quad \text{NOTE: } z_j^n : \mathbf{0}$$

## Tree automata for 4th-order [Comon, Jurski 97]

- ▶ 4th-order  $E$ ,  $xw_1 \dots w_k = u$  and  $C_E$  constants in  $u$  plus  $b$
- ▶ For finite alphabet  $\Sigma$  consider a potential solution term  
 $t = \lambda x_1 \dots x_k . s$

$$s ::= z_j^n \mid x_i v_1 \dots v_m \mid f s_1 \dots s_m \quad f \in C_E, m \geq 0$$
$$v ::= s \mid \lambda z_1^n \dots z_m^n . s \quad m \geq 0 \quad \text{NOTE: } z_j^n : \mathbf{0}$$

- ▶ node of  $t$  labelled with a variable/constant is inaccessible if it does not contribute to normal form of  $tw_1 \dots w_k$   
(Therefore, can replace subterm  $t'$  rooted there with  $b : \mathbf{0}$ ; preserve normal form)

## Tree automata for 4th-order [Comon, Jurski 97]

- ▶ 4th-order  $E$ ,  $xw_1 \dots w_k = u$  and  $C_E$  constants in  $u$  plus  $b$
- ▶ For finite alphabet  $\Sigma$  consider a potential solution term  
 $t = \lambda x_1 \dots x_k . s$

$$s ::= z_j^n \mid x_i v_1 \dots v_m \mid f s_1 \dots s_m \quad f \in C_E, m \geq 0$$
$$v ::= s \mid \lambda z_1^n \dots z_m^n . s \quad m \geq 0 \quad \text{NOTE: } z_i^n : \mathbf{0}$$

- ▶ node of  $t$  labelled with a variable/constant is inaccessible if it does not contribute to normal form of  $tw_1 \dots w_k$  (Therefore, can replace subterm  $t'$  rooted there with  $b : \mathbf{0}$ ; preserve normal form)
- ▶ node of  $t$  is accessible; look at evaluation of  $t'$  which is normal form of  $t' \{w_1/x_1, \dots, w_k/x_k\}$  this is a subterm of  $u$  when subtrees can be replaced by leaves labelled with a variable  $z_j^i$

## Tree automata for 4th-order [Comon, Jurski 97]

- ▶ 4th-order  $E$ ,  $xw_1 \dots w_k = u$  and  $C_E$  constants in  $u$  plus  $b$
- ▶ For finite alphabet  $\Sigma$  consider a potential solution term  
 $t = \lambda x_1 \dots x_k . s$

$$s ::= z_j^n \mid x_i v_1 \dots v_m \mid f s_1 \dots s_m \quad f \in C_E, m \geq 0$$
$$v ::= s \mid \lambda z_1^n \dots z_m^n . s \quad m \geq 0 \quad \text{NOTE: } z_i^n : \mathbf{0}$$

- ▶ node of  $t$  labelled with a variable/constant is inaccessible if it does not contribute to normal form of  $tw_1 \dots w_k$  (Therefore, can replace subterm  $t'$  rooted there with  $b : \mathbf{0}$ ; preserve normal form)
- ▶ node of  $t$  is accessible; look at evaluation of  $t'$  which is normal form of  $t' \{w_1/x_1, \dots, w_k/x_k\}$  this is a subterm of  $u$  when subtrees can be replaced by leaves labelled with a variable  $z_j^i$
- ▶ Guarantees finite alphabet for accessible nodes: reuse variables



## Tree automata for 4th-order [Comon, Jurski 97] II

- ▶ **Finite states:**
  - inaccessible nodes:** – for “dont care”
  - accessible nodes:**  $U$  is subterms of  $u$  closed under replacement of subtrees with leaves labelled with finitely many different variables  $z_j^i : \mathbf{0}$

$$\{[e], [\lambda.e], [\lambda z_1^n \dots z_m^n.e], [\lambda x_1 \dots x_k.u] \mid e \in U \cup \{-}\}$$

## Tree automata for 4th-order [Comon, Jurski 97] II

- ▶ **Finite states:**
  - inaccessible nodes:** – for “dont care”
  - accessible nodes:**  $U$  is subterms of  $u$  closed under replacement of subtrees with leaves labelled with finitely many different variables  $z_j^i$  :  $\mathbf{0}$

$$\{[e], [\lambda.e], [\lambda z_1^n \dots z_m^n.e], [\lambda x_1 \dots x_k.u] \mid e \in U \cup \{-}\}$$

- ▶ **Final states**  $\{[\lambda x_1 \dots x_k.u]\}$

## Tree automata for 4th-order [Comon, Jurski 97] II

- ▶ **Finite states:**  
**inaccessible nodes:** – for “dont care”  
**accessible nodes:**  $U$  is subterms of  $u$  closed under replacement of subtrees with leaves labelled with finitely many different variables  $z_j^i$  :  $\mathbf{0}$

$$\{[e], [\lambda.e], [\lambda z_1^n \dots z_m^n.e], [\lambda x_1 \dots x_k.u] \mid e \in U \cup \{-}\}$$

- ▶ **Final states**  $\{[\lambda x_1 \dots x_k.u]\}$
- ▶ **Finite transitions:** such as

$$\begin{array}{llll} z \Rightarrow [z] & z \Rightarrow [-] & a \Rightarrow [a] & c \Rightarrow [-] \\ \lambda x_1[ga] \Rightarrow [\lambda x_1.ga] & x_1[\lambda.-][\lambda.gz] \Rightarrow [gz] & & \\ \dots & & & \end{array}$$

# Tree automata for 4th-order [Comon, Jurski 97] II

- ▶ **Finite states:**  
**inaccessible nodes:** – for “dont care”  
**accessible nodes:**  $U$  is subterms of  $u$  closed under replacement of subtrees with leaves labelled with finitely many different variables  $z_j^i : \mathbf{0}$

$$\{[e], [\lambda.e], [\lambda z_1^n \dots z_m^n.e], [\lambda x_1 \dots x_k.u] \mid e \in U \cup \{-}\}$$

- ▶ **Final states**  $\{[\lambda x_1 \dots x_k.u]\}$
- ▶ **Finite transitions:** such as

$$\begin{array}{l} z \Rightarrow [z] \quad z \Rightarrow [-] \quad a \Rightarrow [a] \quad c \Rightarrow [-] \\ \lambda x_1[ga] \Rightarrow [\lambda x_1.ga] \quad x_1[\lambda.-][\lambda.gz] \Rightarrow [gz] \\ \dots \end{array}$$

- ▶ **Theorem** The tree automaton associated with a 4th-order  $xw_1 \dots w_k = u$  accepts  $t$  iff  $t$  solves  $xw_1 \dots w_k = u$

# Tree automata for 5th-order: PROBLEMS

- ▶ **Finite alphabet ?** can stack 2nd-order variables

$\lambda xy.x(\lambda z_1.x(\lambda z_2.\dots(\lambda z_n.z_n(z_{n-1}(\dots z_1(y))\dots))\dots))$

Need an infinite alphabet

# Tree automata for 5th-order: PROBLEMS

- ▶ **Finite alphabet ?** can stack 2nd-order variables

$\lambda xy.x(\lambda z_1.x(\lambda z_2 \dots (\lambda z_n.z_n(z_{n-1}(\dots z_1(y)) \dots))) \dots))$

Need an infinite alphabet

- ▶ **Finite number of states ?** stack same 2nd-order variable

$z_n(z_n(\dots (z_n a) \dots))$

Number of evaluations can be infinite

# Tree automata for 5th-order: PROBLEMS

- ▶ **Finite alphabet ?** can stack 2nd-order variables

$\lambda xy.x(\lambda z_1.x(\lambda z_2 \dots (\lambda z_n.z_n(z_{n-1}(\dots z_1(y)) \dots)) \dots))$

Need an infinite alphabet

- ▶ **Finite number of states ?** stack same 2nd-order variable

$z_n(z_n(\dots (z_n a) \dots))$

Number of evaluations can be infinite

- ▶ **Overcome the first problem: employ binding trees**

$\lambda xy.x(\lambda z.x(\lambda z \dots x(\lambda z.z(\lambda.z(\dots \lambda.z(\lambda.y)) \dots)) \dots))$

↓ from the first node labelled  $\lambda z$  to the last node labelled  $z$ ,  
and so on

**Fact** For all  $A$  and finite  $C$ , there is a finite  $\Sigma$  such that every  $t \in T_A(C)$  up to  $\alpha$ -equivalence is a binding  $\Sigma$ -tree.

# Tree automata for 5th-order: PROBLEMS

- ▶ **Finite alphabet ?** can stack 2nd-order variables

$\lambda xy.x(\lambda z_1.x(\lambda z_2 \dots (\lambda z_n.z_n(z_{n-1}(\dots z_1(y)) \dots))) \dots))$

Need an infinite alphabet

- ▶ **Finite number of states ?** stack same 2nd-order variable

$z_n(z_n(\dots (z_n a) \dots))$

Number of evaluations can be infinite

- ▶ **Overcome the first problem: employ binding trees**

$\lambda xy.x(\lambda z.x(\lambda z \dots x(\lambda z.z(\lambda.z(\dots \lambda.z(\lambda.y)) \dots))) \dots))$

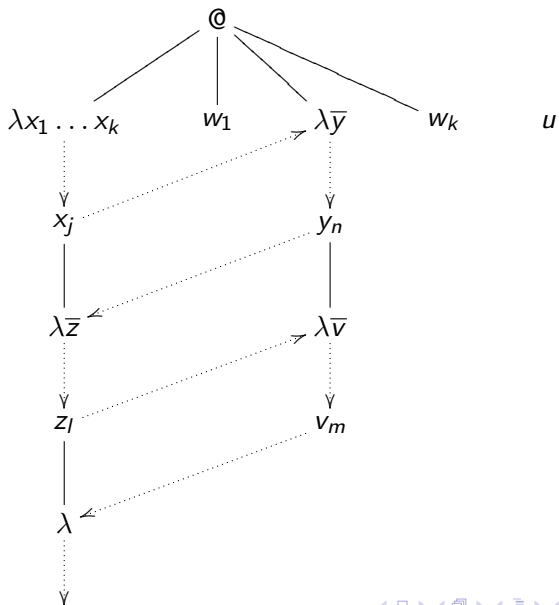
↓ from the first node labelled  $\lambda z$  to the last node labelled  $z$ ,  
and so on

**Fact** For all  $A$  and finite  $C$ , there is a finite  $\Sigma$  such that every  $t \in T_A(C)$  up to  $\alpha$ -equivalence is a binding  $\Sigma$ -tree.

- ▶ **Second problem ?** need finer analysis than evaluation of an accessible node



# Games: automaton moving around interpolation tree



## Game/automaton $G(t, E)$

- ▶ Play: sequence  $n_1q_1\theta_1, \dots, n_lq_l\theta_l$ 
  - ▶  $n_i$  is a node of the interpolation tree,
  - ▶  $q_i$  is a state  $[u']$  where  $u'$  subterm of  $u$  or final
  - ▶  $\theta_i$  is look-up table: tells where to jump when at a variable

# Game/automaton $G(t, E)$

- ▶ Play: sequence  $n_1 q_1 \theta_1, \dots, n_l q_l \theta_l$ 
  - ▶  $n_i$  is a node of the interpolation tree,
  - ▶  $q_i$  is a state  $[u']$  where  $u'$  subterm of  $u$  or final
  - ▶  $\theta_i$  is look-up table: tells where to jump when at a variable
- ▶ **Initial position** at  $\textcircled{u}$ ,  $q_1 = [u]$  and  $\theta_1$  is empty
- ▶  **$\forall$  loses the play if the final state  $q_l = [\exists]$**

# Game/automaton $G(t, E)$

- ▶ Play: sequence  $n_1q_1\theta_1, \dots, n_lq_l\theta_l$ 
  - ▶  $n_i$  is a node of the interpolation tree,
  - ▶  $q_i$  is a state  $[u']$  where  $u'$  subterm of  $u$  or final
  - ▶  $\theta_i$  is look-up table: tells where to jump when at a variable
- ▶ **Initial position** at  $@$ ,  $q_1 = [u]$  and  $\theta_1$  is empty
- ▶  **$\forall$  loses the play if the final state  $q_l = [\exists]$**
- ▶ Current position is  $n[r]\theta$ ; **next position**
  - ▶  $@$  then  $n1[r]\theta'$  where  $\theta' = \theta\{((n2, \dots, n(k+1)), \theta)/n1\}$
  - ▶  $\lambda\bar{y}$  then  $n1[r]\theta$

# Game/automaton $G(t, E)$

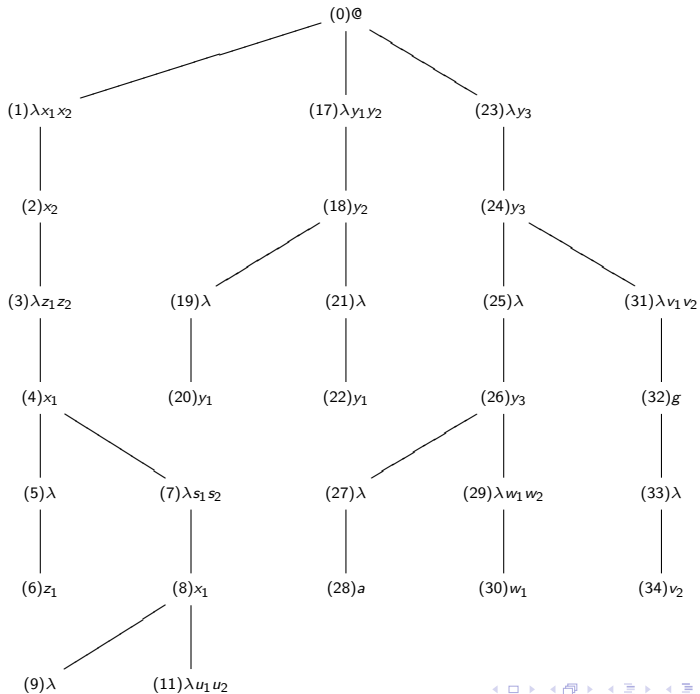
- ▶ Play: sequence  $n_1q_1\theta_1, \dots, n_lq_l\theta_l$ 
  - ▶  $n_i$  is a node of the interpolation tree,
  - ▶  $q_i$  is a state  $[u']$  where  $u'$  subterm of  $u$  or final
  - ▶  $\theta_i$  is look-up table: tells where to jump when at a variable
- ▶ **Initial position** at  $\textcircled{0}$ ,  $q_1 = [u]$  and  $\theta_1$  is empty
- ▶  $\forall$  loses the play if the final state  $q_l = [\exists]$
- ▶ Current position is  $n[r]\theta$ ; **next position**
  - ▶  $\textcircled{0}$  then  $n1[r]\theta'$  where  $\theta' = \theta\{((n2, \dots, n(k+1)), \theta)/n1\}$
  - ▶  $\lambda\bar{y}$  then  $n1[r]\theta$
  - ▶  $a : \mathbf{0}$  if  $r = a$  then  $n[\exists]\theta$  else  $n[\forall]\theta$
  - ▶  $f : (B_1, \dots, B_p, \mathbf{0})$  if  $r = fr_1 \dots r_p$  then  $\forall$  chooses  $j \in \{1, \dots, p\}$  and  $n_j[r_j]\theta$  else  $n[\forall]\theta$

# Game/automaton $G(t, E)$

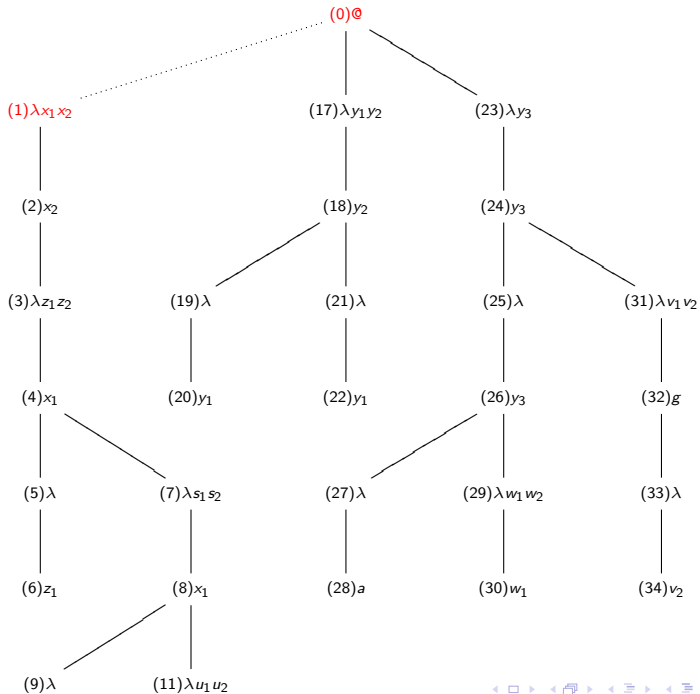
- ▶ Play: sequence  $n_1q_1\theta_1, \dots, n_lq_l\theta_l$ 
  - ▶  $n_i$  is a node of the interpolation tree,
  - ▶  $q_i$  is a state  $[u']$  where  $u'$  subterm of  $u$  or final
  - ▶  $\theta_i$  is look-up table: tells where to jump when at a variable
- ▶ **Initial position** at  $\textcircled{0}$ ,  $q_1 = [u]$  and  $\theta_1$  is empty
- ▶  $\forall$  loses the play if the final state  $q_l = [\exists]$
- ▶ Current position is  $n[r]\theta$ ; **next position**
  - ▶  $\textcircled{0}$  then  $n1[r]\theta'$  where  $\theta' = \theta\{((n2, \dots, n(k+1)), \theta)/n1\}$
  - ▶  $\lambda\bar{y}$  then  $n1[r]\theta$
  - ▶  $a : \mathbf{0}$  if  $r = a$  then  $n[\exists]\theta$  else  $n[\forall]\theta$
  - ▶  $f : (B_1, \dots, B_p, \mathbf{0})$  if  $r = fr_1 \dots r_p$  then  $\forall$  chooses  $j \in \{1, \dots, p\}$  and  $n_j[r_j]\theta$  else  $n[\forall]\theta$
  - ▶  $y_j : \mathbf{0}$  if  $m \downarrow n$  and  $\theta(m) = ((m_1, \dots, m_l), \theta')$  then  $m_j[r]\theta'$
  - ▶  $y_j : (B_1, \dots, B_p, \mathbf{0})$  if  $m \downarrow n$  and  $\theta(m) = ((m_1, \dots, m_l), \theta')$  then  $m_j[r]\theta''$  where  $\theta'' = \theta'\{((n1, \dots, np), \theta)/m_j\}$

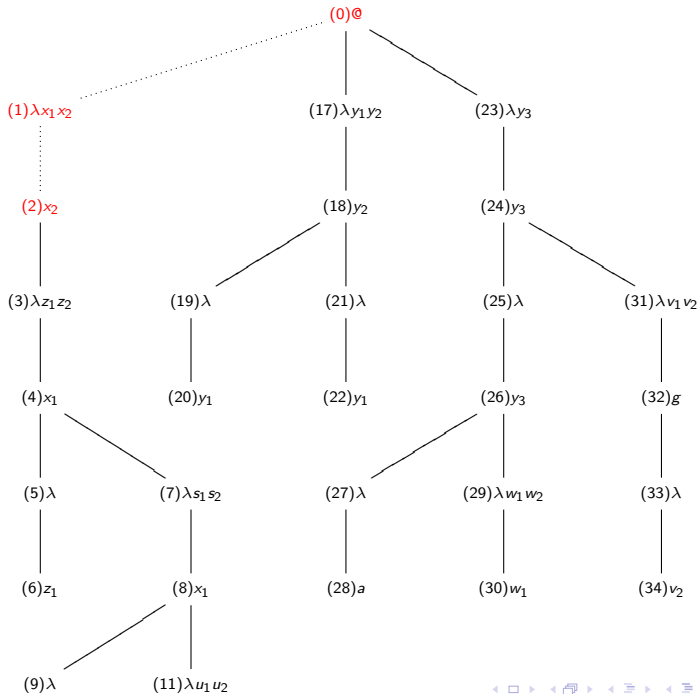
# Game/automaton $G(t, E)$

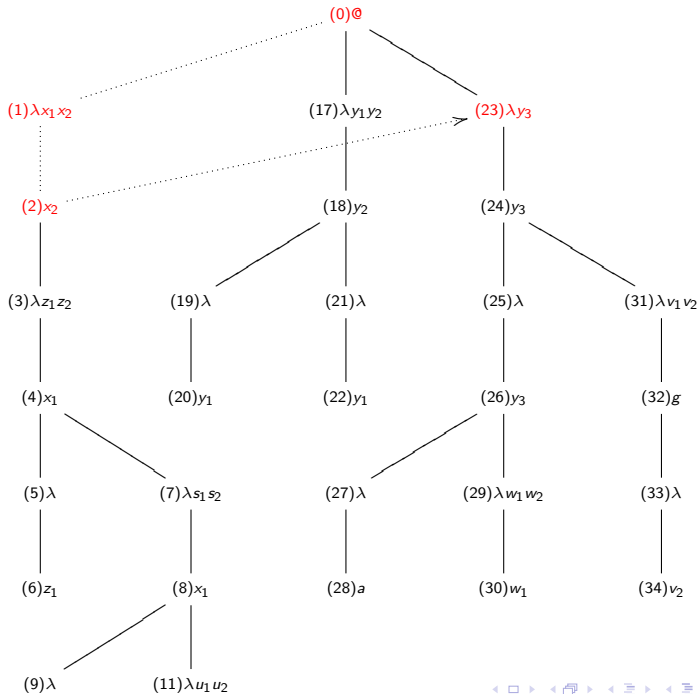
- ▶ Play: sequence  $n_1q_1\theta_1, \dots, n_lq_l\theta_l$ 
  - ▶  $n_i$  is a node of the interpolation tree,
  - ▶  $q_i$  is a state  $[u']$  where  $u'$  subterm of  $u$  or final
  - ▶  $\theta_i$  is look-up table: tells where to jump when at a variable
- ▶ **Initial position** at  $\textcircled{0}$ ,  $q_1 = [u]$  and  $\theta_1$  is empty
- ▶  $\forall$  loses the play if the final state  $q_l = [\exists]$
- ▶ Current position is  $n[r]\theta$ ; **next position**
  - ▶  $\textcircled{0}$  then  $n1[r]\theta'$  where  $\theta' = \theta\{((n2, \dots, n(k+1)), \theta)/n1\}$
  - ▶  $\lambda\bar{y}$  then  $n1[r]\theta$
  - ▶  $a : \mathbf{0}$  if  $r = a$  then  $n[\exists]\theta$  else  $n[\forall]\theta$
  - ▶  $f : (B_1, \dots, B_p, \mathbf{0})$  if  $r = fr_1 \dots r_p$  then  $\forall$  chooses  $j \in \{1, \dots, p\}$  and  $n_j[r_j]\theta$  else  $n[\forall]\theta$
  - ▶  $y_j : \mathbf{0}$  if  $m \downarrow n$  and  $\theta(m) = ((m_1, \dots, m_l), \theta')$  then  $m_j[r]\theta'$
  - ▶  $y_j : (B_1, \dots, B_p, \mathbf{0})$  if  $m \downarrow n$  and  $\theta(m) = ((m_1, \dots, m_l), \theta')$  then  $m_j[r]\theta''$  where  $\theta'' = \theta'\{((n1, \dots, np), \theta)/m_j\}$
- ▶ Player  $\forall$  loses every play in  $G(t, E)$  iff  $t$  solves  $E$

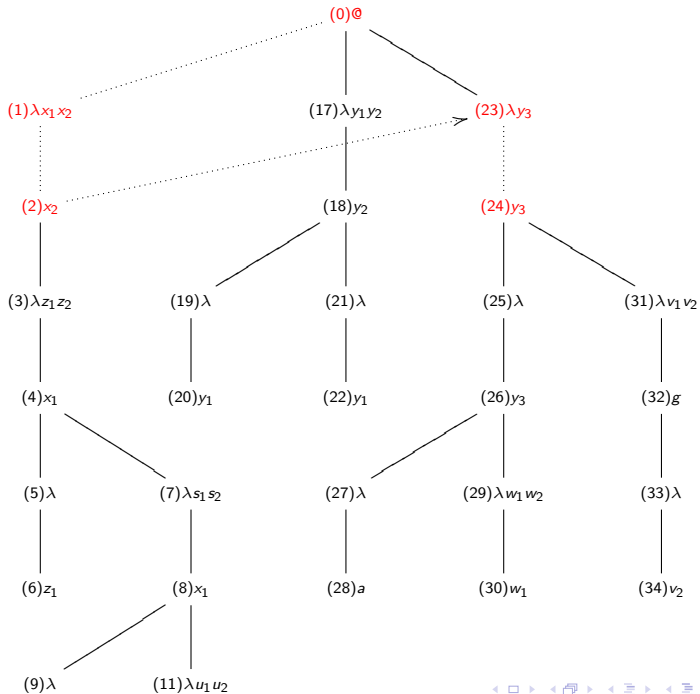


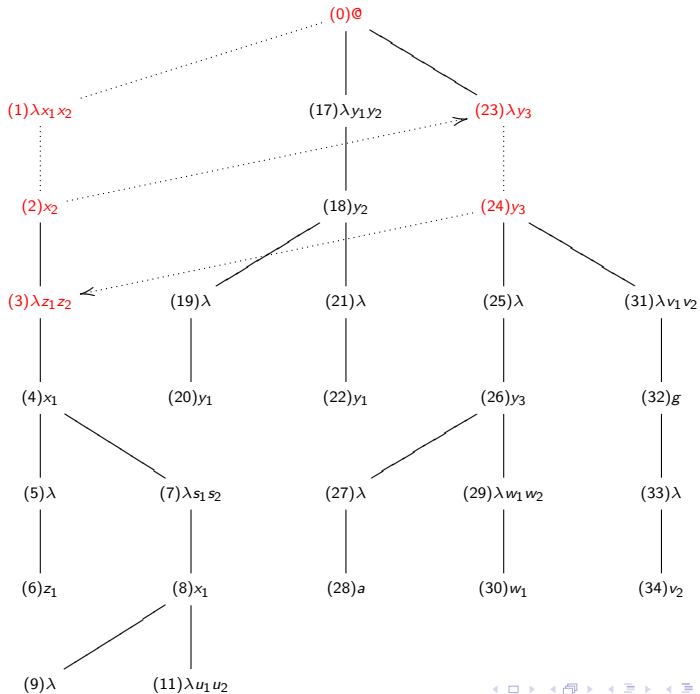


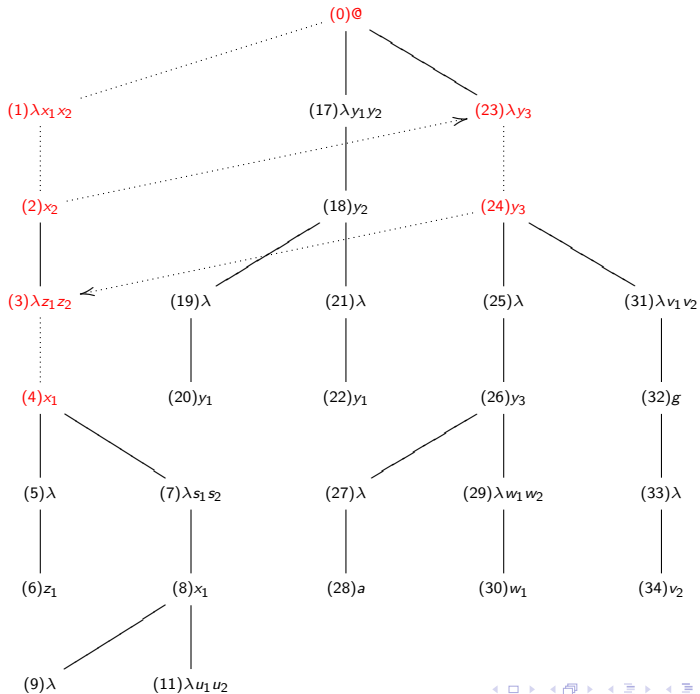


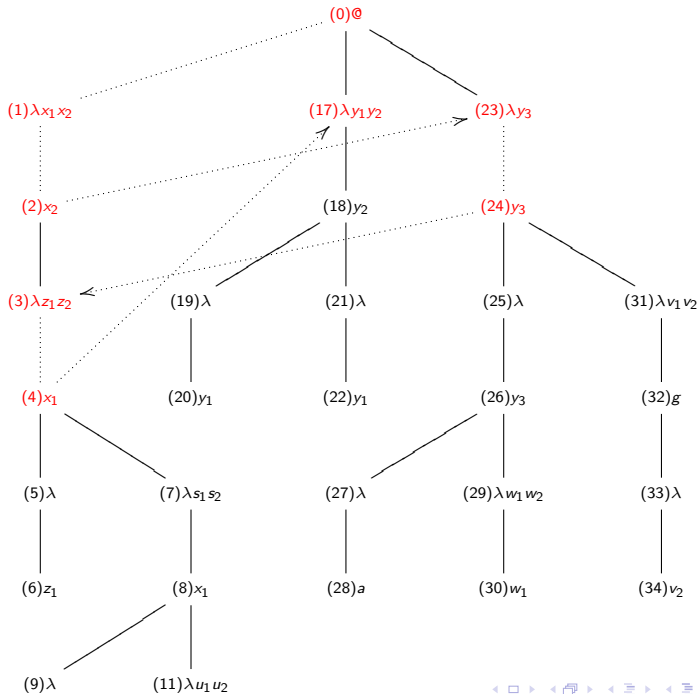


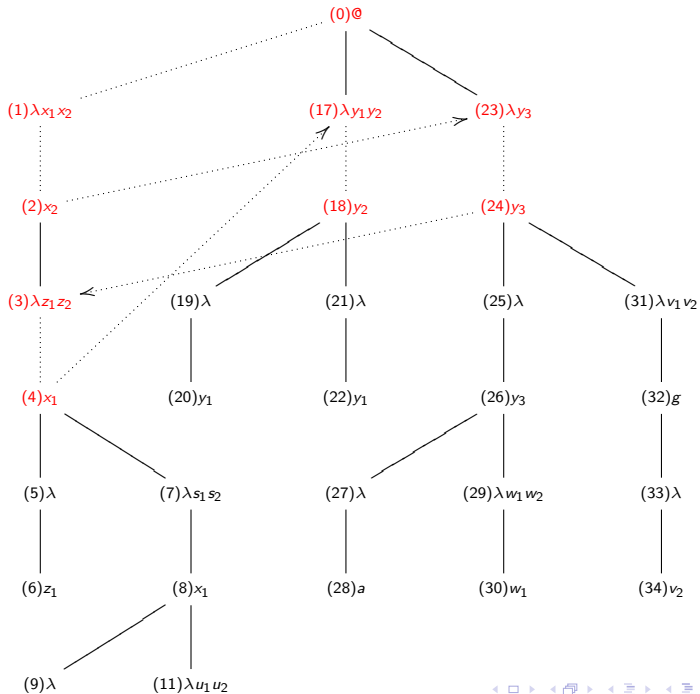




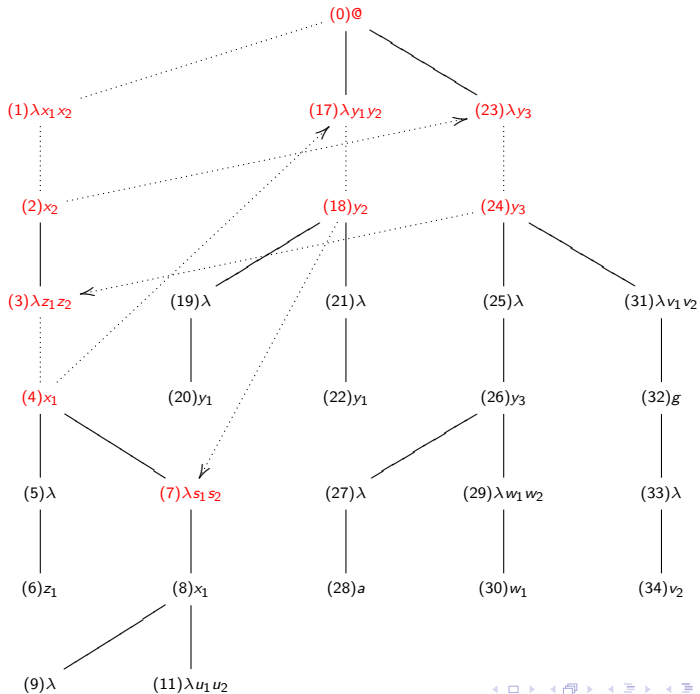


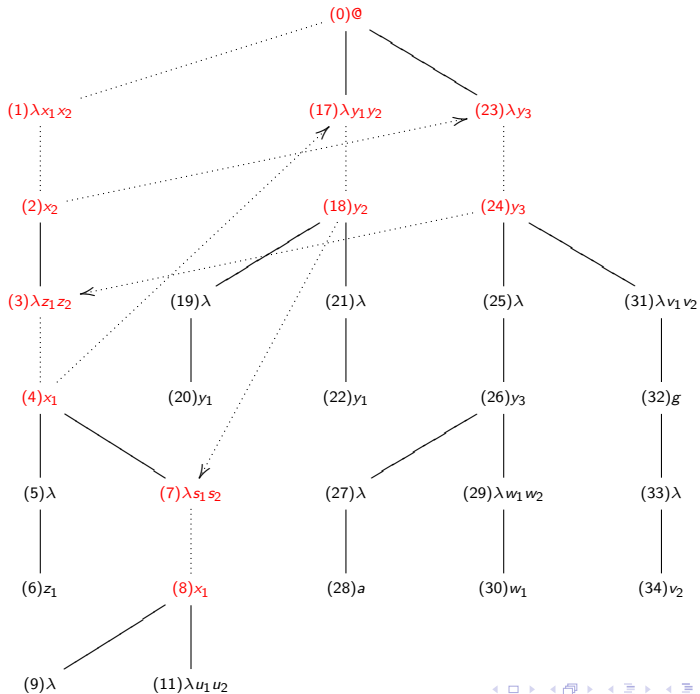


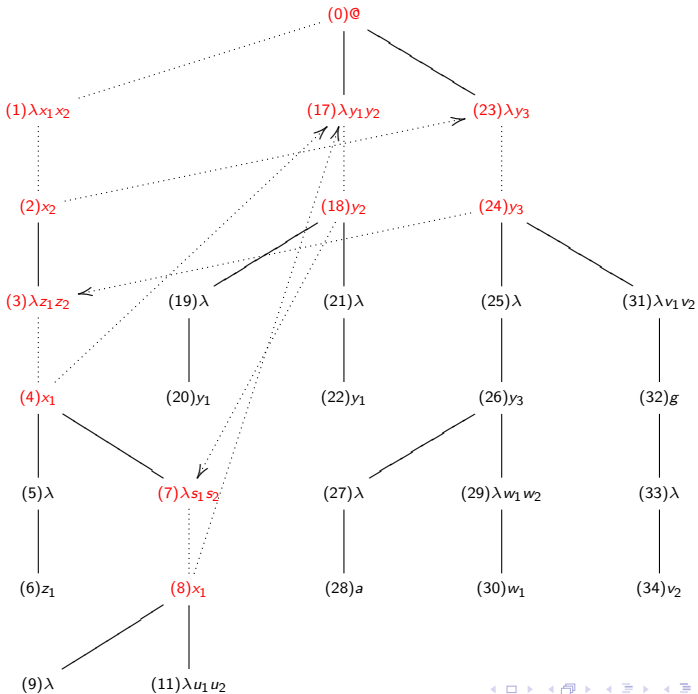


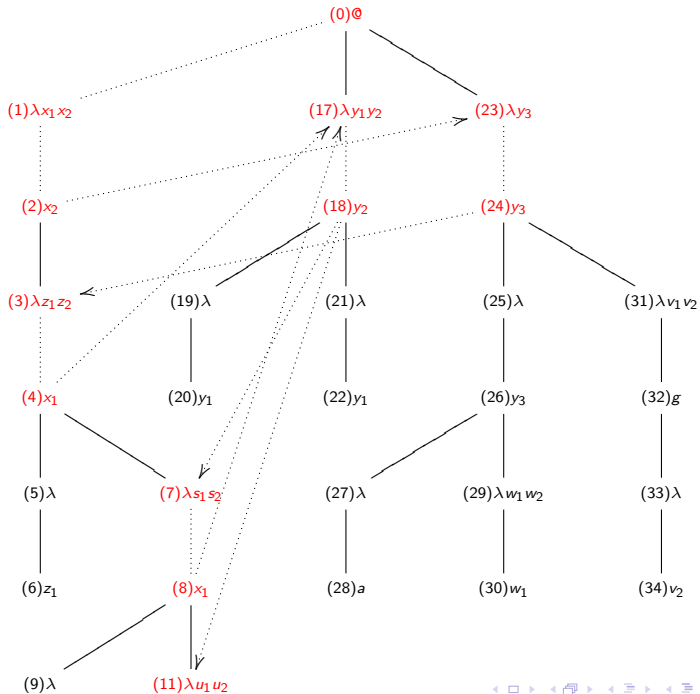












# Application of games to define tree automata

- ▶ Two problems extending Comon+ Jurski's result
  1. Ensuring finitely many states in automaton
  2. Ensuring finite alphabet in automaton

# Application of games to define tree automata

- ▶ Two problems extending Comon+ Jurski's result
  1. Ensuring finitely many states in automaton
  2. Ensuring finite alphabet in automaton
- ▶ Instead of using evaluation of a node for states use (version of) variable profiles from [Ong 2006]
- ▶ Variable profile: abstraction from sequences of positions from a node  $\{(x_2, ga, \{(y_3, ga, \{(z_2, ga, \{(v_2, a, \emptyset)\})\})\})\}$

# Application of games to define tree automata

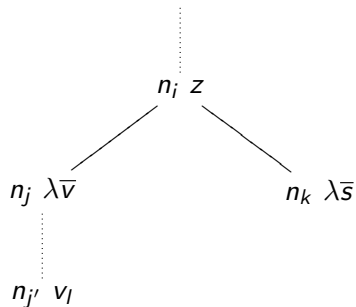
- ▶ Two problems extending Comon+ Jurski's result
  1. Ensuring finitely many states in automaton
  2. Ensuring finite alphabet in automaton
- ▶ Instead of using evaluation of a node for states use (version of) **variable profiles** from [Ong 2006]
- ▶ **Variable profile: abstraction from sequences of positions from a node**  $\{(x_2, ga, \{(y_3, ga, \{(z_2, ga, \{(v_2, a, \emptyset)\})\})\})\}$
- ▶ **Theorem** The set of solutions of  $E$  **built out of a fixed finite alphabet of variables** is recognised by a tree automaton [Stirling 2007]

# Application of games to define tree automata

- ▶ Two problems extending Comon+ Jurski's result
  1. Ensuring finitely many states in automaton
  2. Ensuring finite alphabet in automaton
- ▶ Instead of using evaluation of a node for states use (version of) **variable profiles** from [Ong 2006]
- ▶ **Variable profile: abstraction from sequences of positions from a node**  $\{(x_2, ga, \{(y_3, ga, \{(z_2, ga, \{(v_2, a, \emptyset)\})\})\})\}$
- ▶ **Theorem** The set of solutions of  $E$  **built out of a fixed finite alphabet of variables** is recognised by a tree automaton [Stirling 2007]
- ▶ **Theorem** The set of solutions of  $E$  is recognised by a special alternating binding tree automaton **BUT non-emptiness of such automata undecidable** [Stirling 2009; Ong, Tzevelekos 2009]

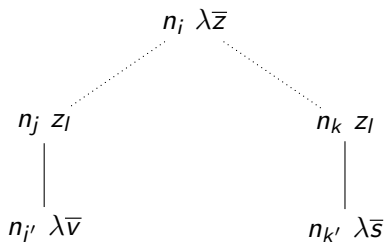


# Uniformity properties of game playing I



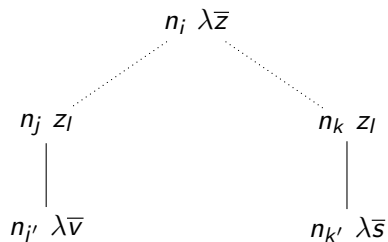
- ▶ If play is at  $n_i$  then at  $n_j$  and then later at  $n_k$  then inbetween there must have been a position at an  $n_{j'}$  bound by  $n_j$

## Uniformity properties of game playing II



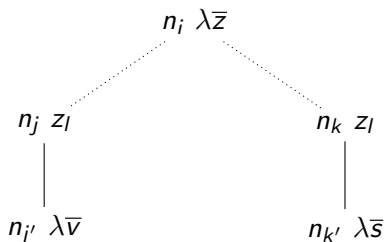
- ▶ If play is at  $n_i$  then at  $n_j$  and then later at  $n_{j'}$ ; then at  $n_k$  then there must be a corresponding sequence to that between  $n_j$  and  $n_{j'}$  between  $n_k$  and  $n_{k'}$  unless there is a different  $\forall$  choice

## Uniformity properties of game playing II



- ▶ If play is at  $n_i$  then at  $n_j$  and then later at  $n_{j'}$ ; then at  $n_k$  then there must be a corresponding sequence to that between  $n_j$  and  $n_{j'}$  between  $n_k$  and  $n_{k'}$  unless there is a different  $\forall$  choice
- ▶ Especially relevant if  $n_k$  is somewhere below  $n_j$  (“embedded”)

## Uniformity properties of game playing II



- ▶ If play is at  $n_i$  then at  $n_j$  and then later at  $n_{j'}$ ; then at  $n_k$  then there must be a corresponding sequence to that between  $n_j$  and  $n_{j'}$  between  $n_k$  and  $n_{k'}$  unless there is a different  $\forall$  choice
- ▶ Especially relevant if  $n_k$  is somewhere below  $n_j$  (“embedded”)
- ▶ Property not enforced in a tree automaton

# Application of games to decidability of matching

- ▶ Combinatorial argument based on uniformities of play

# Application of games to decidability of matching

- ▶ Combinatorial argument based on uniformities of play
- ▶ Two steps in proof assuming an arbitrary solution term

# Application of games to decidability of matching

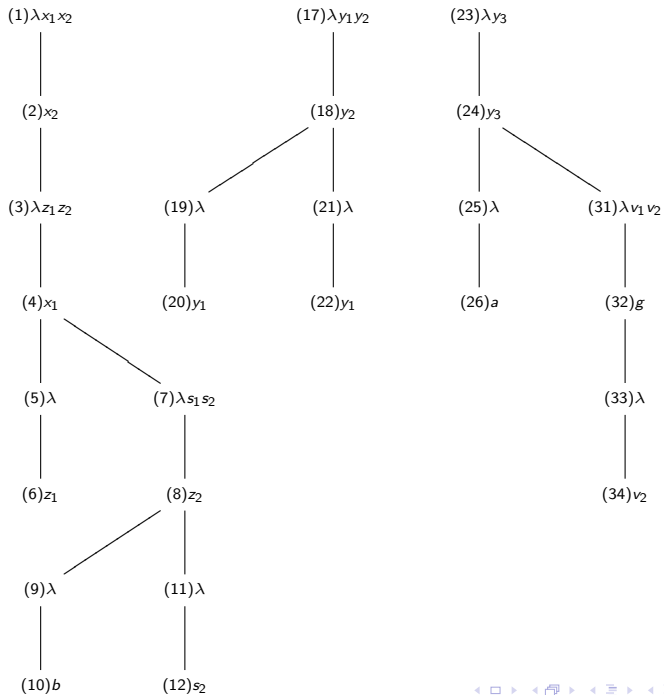
- ▶ Combinatorial argument based on uniformities of play
- ▶ Two steps in proof assuming an arbitrary solution term
  1. partition each play

# Application of games to decidability of matching

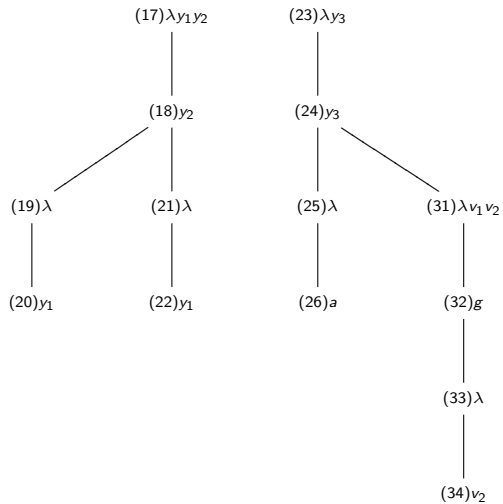
- ▶ Combinatorial argument based on uniformities of play
- ▶ Two steps in proof assuming an arbitrary solution term
  1. partition each play
- ▶ 5th-order example

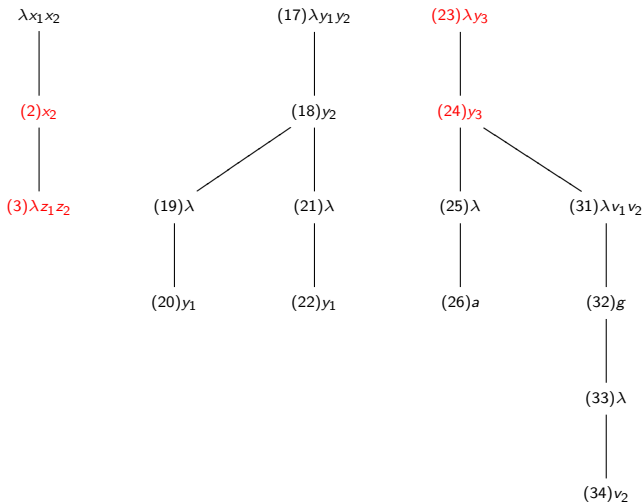
$$x(\lambda y_1 y_2 . y_2 y_1 y_1)(\lambda y_3 . y_3 a(\lambda v_1 v_2 . g v_2)) = ga$$

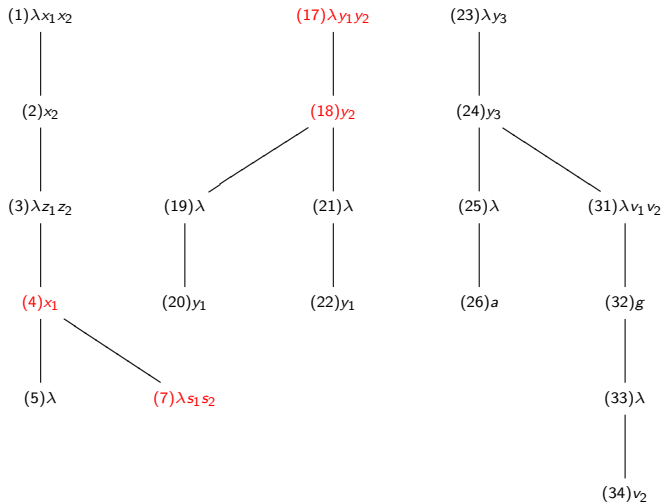


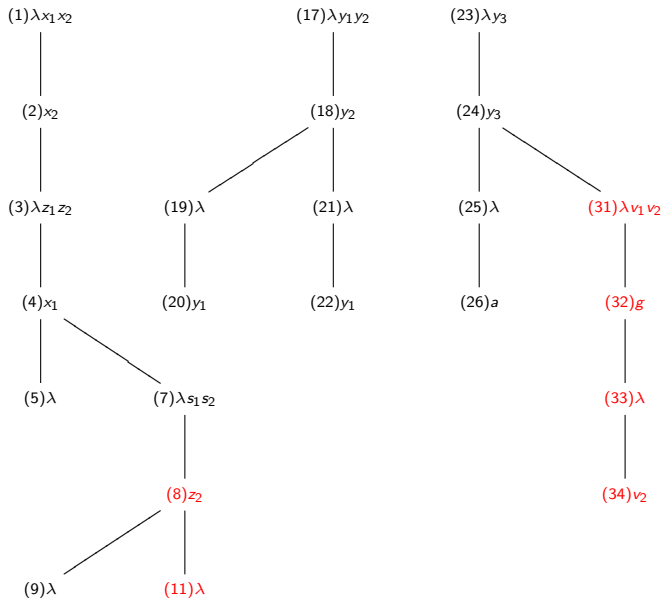


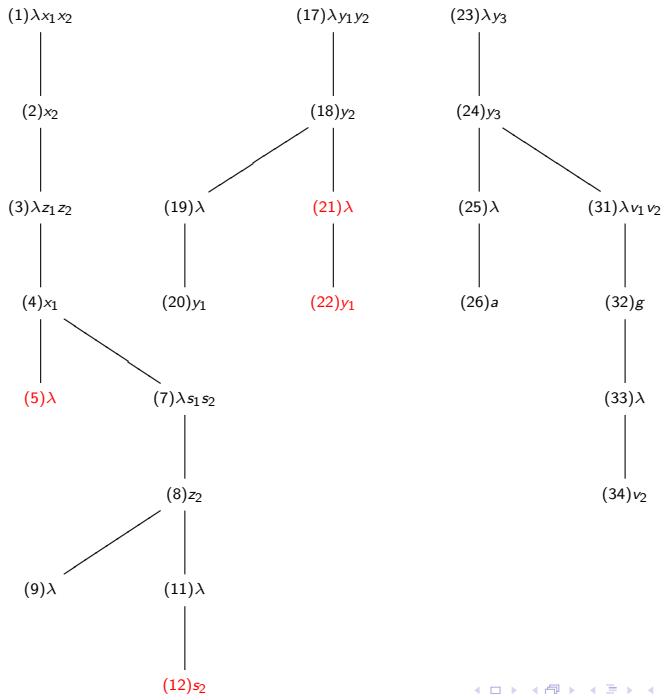
(1)  $\lambda_{x_1 x_2}$

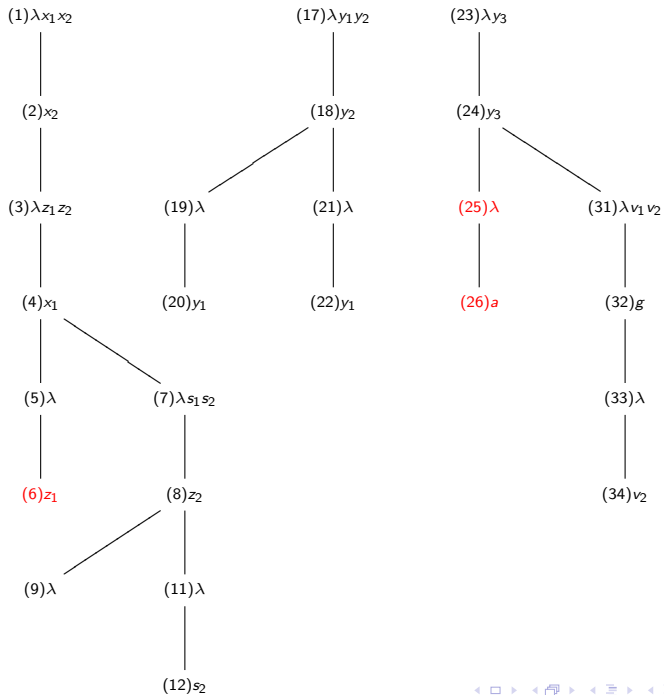












# Application of games to decidability of matching

- ▶ Combinatorial argument based on uniformities of play.
- ▶ Two steps in proof assuming an arbitrary solution term
  1. partition each play
  
- ▶ 5th-order example

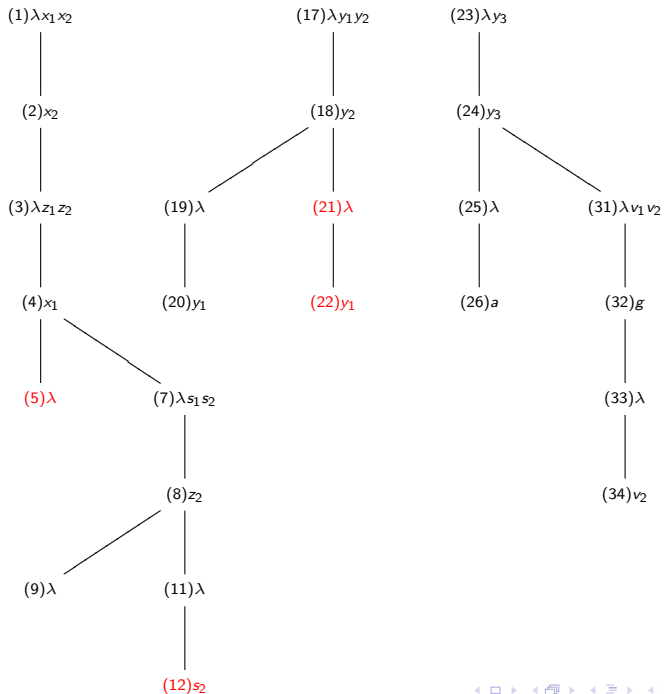
$$x(\lambda y_1 y_2 . y_2 y_1 y_1)(\lambda y_3 . y_3 a(\lambda v_1 v_2 . g v_2)) = ga$$

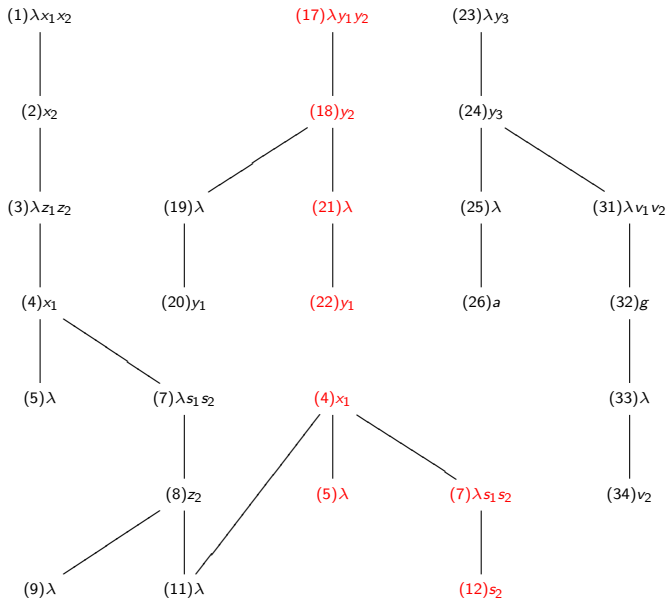


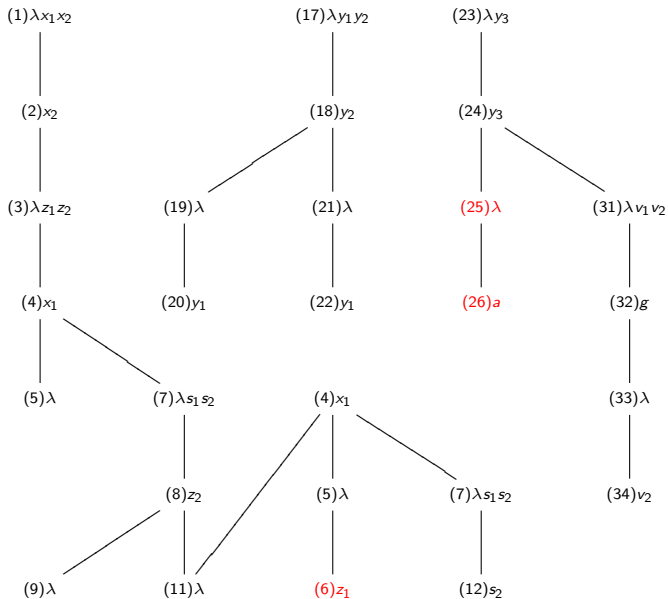
# Application of games to decidability of matching

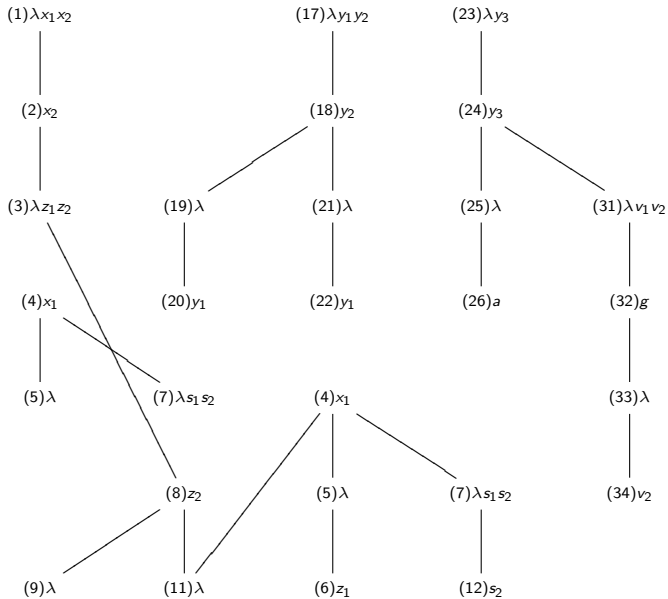
- ▶ Combinatorial argument based on uniformities of play.
- ▶ Two steps in proof assuming an arbitrary solution term
  1. partition each play
  2. unfold into a small solution (by adding prefixes) and then removing redundant parts of term
- ▶ 5th-order example

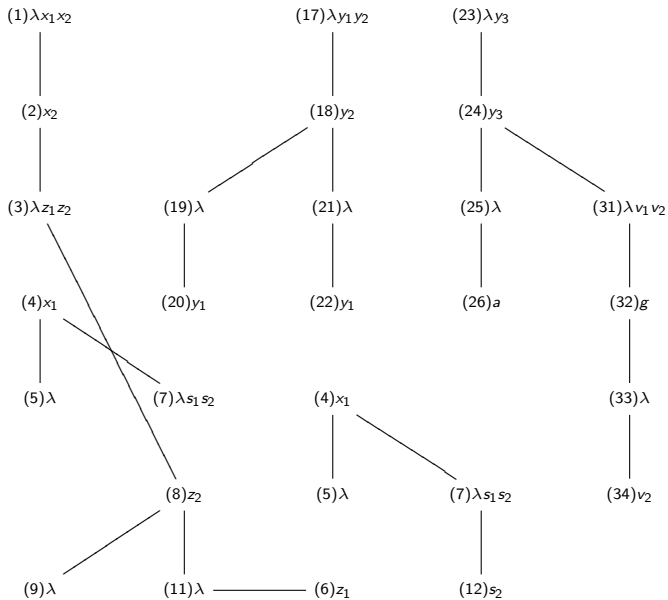
$$x(\lambda y_1 y_2 . y_2 y_1 y_1)(\lambda y_3 . y_3 a(\lambda v_1 v_2 . g v_2)) = ga$$

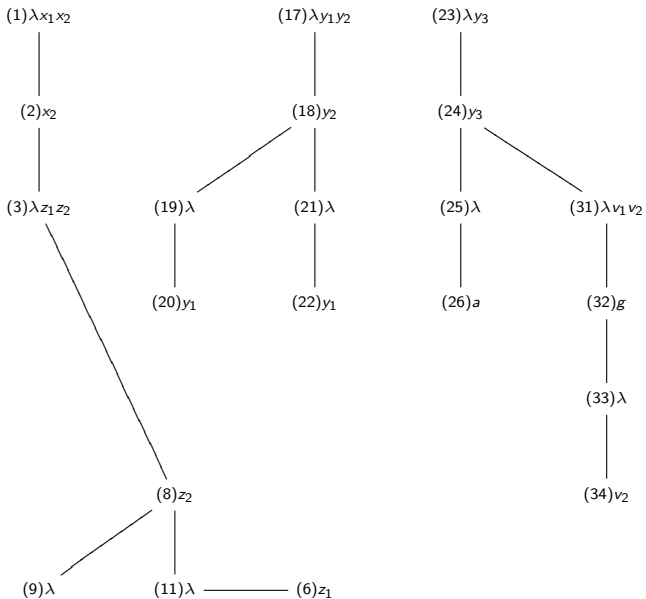












# Application of games to decidability of matching

- ▶ Combinatorial argument based on uniformities of play.
- ▶ Two steps in proof assuming an arbitrary solution term
  1. partition each play
  2. unfold into a small solution (by adding prefixes) and then remove redundant parts of term



# Application of games to decidability of matching

- ▶ Combinatorial argument based on uniformities of play.
- ▶ Two steps in proof assuming an arbitrary solution term
  1. partition each play
  2. unfold into a small solution (by adding prefixes) and then remove redundant parts of term
- ▶ Small term property

# Application of games to decidability of matching

- ▶ Combinatorial argument based on uniformities of play.
- ▶ **Two steps in proof assuming an arbitrary solution term**
  1. partition each play
  2. unfold into a small solution (by adding prefixes) and then remove redundant parts of term
- ▶ **Small term property**
- ▶ **Theorem** If  $x : A$  and  $A$  has order  $2n + 2$  or  $2n + 3$  and arity  $m$  then  $xw_1 \dots w_k = u$  has a (canonical) solution iff it has a (canonical) solution of depth at most  $4n^2 m^{2n-2} |u|$

# The Retract Problem

- ▶ Type  $A$  is a **retract** of  $B$ ,  $A \trianglelefteq B$ , if there are  $t : A \rightarrow B$  and  $s : B \rightarrow A$  such that  $s(tx) =_{\beta\eta} x$

# The Retract Problem

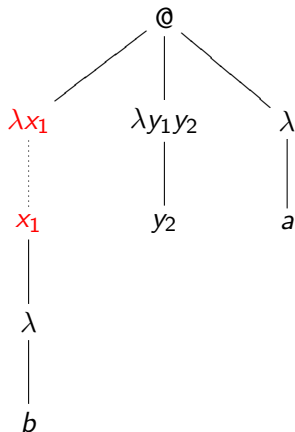
- ▶ Type  $A$  is a **retract** of  $B$ ,  $A \trianglelefteq B$ , if there are  $t : A \rightarrow B$  and  $s : B \rightarrow A$  such that  $s(tx) =_{\beta\eta} x$
- ▶ **Problem:** Is there a proof system that characterises  $A \trianglelefteq B$  ?
- ▶ Solved for affine (linear) case [Liguro, Piperno, Statman 1992; Regnier Urzyczyn 2002]

# The Retract Problem

- ▶ Type  $A$  is a **retract** of  $B$ ,  $A \trianglelefteq B$ , if there are  $t : A \rightarrow B$  and  $s : B \rightarrow A$  such that  $s(tx) =_{\beta\eta} x$
- ▶ **Problem:** Is there a proof system that characterises  $A \trianglelefteq B$  ?
- ▶ Solved for affine (linear) case [Liguro, Piperno, Statman 1992; Regnier Urzyczyn 2002]
- ▶ **Work in progress:** using games, there is a proof system for the general case

## Games only work for long normal forms

- ▶  $x : ((\mathbf{0}, \mathbf{0}, \mathbf{0}), \mathbf{0}, \mathbf{0})$   $\beta$ -interpolation problem  $x(\lambda y_1 y_2 . y_2) a = a$



- ▶  $\lambda x_1 . x_1 b (\lambda y_1 y_2 . y_2) a \rightarrow_{\beta}^* a$
- ▶ No natural game?