Introduction to Collapsible Pushdown Automata and Higher-Order Recursion Schemes

Paweł Parys

University of Warsaw

Higher-order pushdown automata [Maslov 1974] - definition

A 1-stack is an ordinary stack. A 2-stack (resp. \((n+1)\)-stack) is a stack of 1-stacks (resp. \(n\)-stack).
Higher-order pushdown automata [Maslov 1974] - definition

A 1-stack is an ordinary stack. A 2-stack (resp. \((n+1)\)-stack) is a stack of 1-stacks (resp. \(n\)-stack).

**Operations on 2-stacks:** \(s_i\) are 1-stacks. Top of stack is on right.

\[
push_2 : \quad [s_1 \ldots s_{i-1} s_i] \quad \rightarrow \quad [s_1 \ldots s_{i-1} s_i s_i]
\]

\[
pop_2 : \quad [s_1 \ldots s_{i-1} s_i] \quad \rightarrow \quad [s_1 \ldots s_{i-1}]
\]
A 1-stack is an ordinary stack. A 2-stack (resp. \((n+1)\)-stack) is a stack of 1-stacks (resp. \(n\)-stack).

**Operations on 2-stacks:** \(s_i\) are 1-stacks. Top of stack is on right.

\[
\text{push}_2 : \ [s_1...s_{i-1}s_i] \rightarrow [s_1...s_{i-1}s_is_i] \\
\text{pop}_2 : \ [s_1...s_{i-1}s_i] \rightarrow [s_1...s_{i-1}] \\
\text{push}_1 x : \ [s_1...s_{i-1}[a_1...a_{j-1}a_j]] \rightarrow [s_1...s_{i-1}[a_1...a_{j-1}a_jx]] \\
\text{pop}_1 : \ [s_1...s_{i-1}[a_1...a_{j-1}a_j]] \rightarrow [s_1...s_{i-1}[a_1...a_{j-1}]]
\]
Higher-order pushdown automata [Maslov 1974] - definition

A 1-stack is an ordinary stack. A 2-stack (resp. (n+1)-stack) is a stack of 1-stacks (resp. n-stack).

Operations on 2-stacks: $s_i$ are 1-stacks. Top of stack is on right.

- **push**$_2$: \[s_1...s_{i-1}s_i\] → \[s_1...s_{i-1}s_is_i\]
- **pop**$_2$: \[s_1...s_{i-1}s_i\] → \[s_1...s_{i-1}\]
- **push**$_1$: \[s_1...s_{i-1}[a_1...a_{j-1}a]\] → \[s_1...s_{i-1}[a_1...a_{j-1}a, x]\]
- **pop**$_1$: \[s_1...s_{i-1}[a_1...a_{j-1}a]\] → \[s_1...s_{i-1}[a_1...a_{j-1}]\]

An order-$n$ PDA has an order-$n$ stack, and has $push_i$ and $pop_i$ for each $1 \leq i \leq n$.

The next operation depends on the topmost stack symbol, the state, and the next letter on the input.
Higher-order pushdown automata - example

Language: \( \{b^{2^k} : k \in \mathbb{N}\} \)
- order 2
- 3 stack symbols: \( \bot, x, \# \)

\((\_, q_1) \xrightarrow{\varepsilon} (q_1, \text{push}_1(x))\)

any stack symbol

\(q_1 \quad \bot\)
Higher-order pushdown automata - example

Language: \( \{ b^{2^k} : k \in \mathbb{N} \} \)
- order 2
- 3 stack symbols: \( \bot, x, \# \)

\( (\_, q_1) \xrightarrow{\varepsilon} (q_1, \text{push}_1(x)) \)

any stack symbol
Higher-order pushdown automata - example

Language: \( \{ b^{2^k} : k \in \mathbb{N} \} \)

- order 2
- 3 stack symbols: \( \perp, x, \# \)

\[ (\_, q_1) \xrightarrow{\epsilon} (q_1, \text{push}_1(x)) \]

any stack symbol
Higher-order pushdown automata - example

Language: $\{b^{2^k} : k \in \mathbb{N}\}$

- order 2
- 3 stack symbols: $\bot$, $x$, #

$(\_, q_1) \xrightarrow{\varepsilon} (q_1, \text{push}_1(x))$

any stack symbol
Higher-order pushdown automata - example

Language: \( \{ b^{2^k} : k \in \mathbb{N} \} \)

- order 2
- 3 stack symbols: \( \bot, x, \# \)

\[
\begin{array}{|c|c|}
\hline
\# & \# \\
\bot & \bot \\
x & x \\
x & x \\
x & x \\
\hline
\end{array}
\]

\( q_4 \)

- \( (_,q_1) \xrightarrow{\varepsilon} (q_1,\text{push}_1(x)) \)
- \( (_,q_1) \xrightarrow{\varepsilon} (q_2,\text{push}_1(\#)) \)
- \( (\#,q_2) \xrightarrow{\varepsilon} (q_3,\text{push}_2) \)
- \( (\#,q_3) \xrightarrow{\varepsilon} (q_4,\text{pop}_1) \)

Order-2 stack
Higher-order pushdown automata - example

Language: \( \{ b^{2^k} : k \in \mathbb{N} \} \)

- order 2
- 3 stack symbols: \( \bot, x, \# \)

\[
\begin{array}{c|c|c|c}
\# & \# & x & x \\
\hline
x & x & x & x \\
\hline
\bot & \bot & \bot & \bot \\
\end{array}
\]

\[
\begin{align*}
(\_,q_1) & \xrightarrow{\varepsilon} (q_1, \text{push}_1(x)) \\
(\_,q_1) & \xrightarrow{\varepsilon} (q_2, \text{push}_1(\#)) \\
(\#,q_2) & \xrightarrow{\varepsilon} (q_3, \text{push}_2) \\
(\#,q_3) & \xrightarrow{\varepsilon} (q_4, \text{pop}_1) \\
(x,q_4) & \xrightarrow{\varepsilon} (q_5, \text{pop}_1) \\
(\_,q_5) & \xrightarrow{\varepsilon} (q_4, \text{push}_2)
\end{align*}
\]
Higher-order pushdown automata - example

Language: \( \{ b^{2^k} : k \in \mathbb{N} \} \)

- order 2
- 3 stack symbols: \( \perp, x, \# \)

\[
\begin{array}{c}
\# \\
\times \\
\times \\
\times \\
\times \\
\times \\
\times \\
\times \\
\times \\
\times
\end{array}
\]

\( q_4 \)

\[
(\_ , q_1) \xrightarrow{\varepsilon} (q_1 , \text{push}_1(x))
\]

\[
(\_ , q_1) \xrightarrow{\varepsilon} (q_2 , \text{push}_1(\#))
\]

\[
(\# , q_2) \xrightarrow{\varepsilon} (q_3 , \text{push}_2)
\]

\[
(\# , q_3) \xrightarrow{\varepsilon} (q_4 , \text{pop}_1)
\]

\[
(x , q_4) \xrightarrow{\varepsilon} (q_5 , \text{pop}_1)
\]

\[
(\_ , q_5) \xrightarrow{\varepsilon} (q_4 , \text{push}_2)
\]
Higher-order pushdown automata - example

Language: \( \{ b^{2^k} : k \in \mathbb{N} \} \)

- order 2
- 3 stack symbols: \( \bot, x, \# \)

\[
\begin{align*}
(q_4, x) &\xrightarrow{\varepsilon} (q_5, \text{pop}_1) \\
(q_4, \#) &\xrightarrow{\varepsilon} (q_4, \text{push}_2) \\
(q_1, \#) &\xrightarrow{\varepsilon} (q_2, \text{push}_1(\#)) \\
(q_1, x) &\xrightarrow{\varepsilon} (q_2, \text{push}_1(x)) \\
(q_2, x) &\xrightarrow{\varepsilon} (q_3, \text{push}_2) \\
(q_3, x) &\xrightarrow{\varepsilon} (q_4, \text{push}_1) \\
(q_5, x) &\xrightarrow{\varepsilon} (q_6, \text{pop}_1) \\
(q_5, \#) &\xrightarrow{\varepsilon} (q_6, \text{push}_2)
\end{align*}
\]
Higher-order pushdown automata - example

Language: \( \{b^{2^k} : k \in \mathbb{N}\} \)
- order 2
- 3 stack symbols: \( \perp, x, \# \)

\[
\begin{array}{cccc}
\# & x & x & x \\
\# & x & x & x \\
\# & x & x & x \\
\# & x & x & x \\
\hline
q_4 & \perp & \perp & \perp & \perp
\end{array}
\]

\[
(\_,q_1) \xrightarrow{\varepsilon} (q_1,\text{push}_1(x))
\]
\[
(\_,q_1) \xrightarrow{\varepsilon} (q_2,\text{push}_1(\#))
\]
\[
(\#,q_2) \xrightarrow{\varepsilon} (q_3,\text{push}_2)
\]
\[
(\#,q_3) \xrightarrow{\varepsilon} (q_4,\text{pop}_1)
\]
\[
(x,q_4) \xrightarrow{\varepsilon} (q_5,\text{pop}_1)
\]
\[
(\_,q_5) \xrightarrow{\varepsilon} (q_4,\text{push}_2)
\]
\[
(\_,q_4) \xrightarrow{b} (q_4,\text{pop}_2)
\]

Input: \( b \)
Higher-order pushdown automata - example

Language: \( \{ b^{2^k} : k \in \mathbb{N} \} \)

- order 2
- 3 stack symbols: \( \bot, x, \# \)

\[
\begin{array}{c}
\text{Input: } b \ b \\
\end{array}
\]
Higher-order pushdown automata - example

Language: \( \{ b^{2^k} : k \in \mathbb{N} \} \)

- order 2
- 3 stack symbols: \( \bot, x, # \)

Transition rules:

\[
\begin{align*}
(_, q_1) & \rightarrow (q_1, \text{push}_1(x)) \\
(_, q_1) & \rightarrow (q_2, \text{push}_1(#)) \\
(#, q_2) & \rightarrow (q_3, \text{push}_2) \\
(#, q_3) & \rightarrow (q_4, \text{pop}_1) \\
(x, q_4) & \rightarrow (q_5, \text{pop}_1) \\
(_, q_5) & \rightarrow (q_4, \text{push}_2) \\
(\bot, q_4) & \rightarrow (q_4, \text{pop}_2) \\
(#, q_4) & \rightarrow (q_{\text{acc}}, \text{id})
\end{align*}
\]

Input: \( b \; b \; b \; b \; b \; b \; b \; b \)

Diagram:

```

q_4

\[ \begin{array}{c}
\bot \\
\text{x} \\
\text{x} \\
\text{x} \\
\# \\
\end{array} \]
```
Higher-order pushdown automata

“Traditional” view:
- a nondeterministic HOPDA recognizing a language of words, as on previous slides

“Modern” view:
- a deterministic HOPDA generating a single tree (node-labeled, ranked, ordered, usually infinite)

One can also consider configuration graphs of HOPDA – not in this talk.
Higher-order pushdown automata - example

• order 2
• 3 stack symbols: ⊥, x, #

\[
\begin{array}{c}
\text{x} \\
\text{x} \\
\text{x} \\
\text{⊥}
\end{array}
\quad q_1
\]

**nondeterminism** – what to do next?

\[
(\_, q_1) \xrightarrow{\varepsilon} (q_1, \text{push}_1(x))
\]
\[
(\_, q_1) \xrightarrow{\varepsilon} (q_2, \text{push}_1(\#))
\]
\[
(\#, q_2) \xrightarrow{\varepsilon} (q_3, \text{push}_2)
\]
\[
(\#, q_3) \xrightarrow{\varepsilon} (q_4, \text{pop}_1)
\]
\[
(x, q_4) \xrightarrow{\varepsilon} (q_5, \text{pop}_1)
\]
\[
(\_, q_5) \xrightarrow{\varepsilon} (q_4, \text{push}_2)
\]
\[
(\_, q_4) \xrightarrow{b} (q_4, \text{pop}_2)
\]
\[
(\#, q_4) \xrightarrow{\varepsilon} (q_{\text{acc}}, \text{id})
\]
Higher-order pushdown automata - example

- order 2
- 3 stack symbols: ⊥, x, #

\[
\begin{array}{c|ccc}
    & x & x & x \\
q_1 & \_
\end{array}
\]

\[
\begin{align*}
(q_1, \text{push}_1(x)) \\
(q_2, \text{push}_1(\#)) \\
(\#, q_2) & \xrightarrow{\varepsilon} (q_3, \text{push}_2) \\
(\#, q_3) & \xrightarrow{\varepsilon} (q_4, \text{pop}_1) \\
(x, q_4) & \xrightarrow{\varepsilon} (q_5, \text{pop}_1) \\
(\_, q_5) & \xrightarrow{\varepsilon} (q_4, \text{push}_2) \\
(\_, q_4) & \xrightarrow{b} (q_4, \text{pop}_2) \\
(\#, q_4) & \xrightarrow{\varepsilon} (q_{\text{acc}}, \text{id})
\end{align*}
\]
Higher-order pushdown automata - example

- order 2
- 3 stack symbols: $\perp$, $x$, $#$

$\begin{array}{c}
  \text{x} \\
  \text{x} \\
  \text{x} \\
  \perp \\
\end{array}$

$q_1$

$(\perp, q_1) \xrightarrow{a} (q_1, \text{push}_1(x))$

$(\#, q_2) \xrightarrow{\varepsilon} (q_3, \text{push}_2)$

$(\#, q_3) \xrightarrow{\varepsilon} (q_4, \text{pop}_1)$

$(x, q_4) \xrightarrow{\varepsilon} (q_5, \text{pop}_1)$

$(\perp, q_5) \xrightarrow{\varepsilon} (q_4, \text{push}_2)$

$(\#, q_4) \xrightarrow{b} (q_4, \text{pop}_2)$

$(\#, q_4) \xrightarrow{c} (q_4, \text{push}_2)$

Letter $c$ of rank 0, instead of an accepting state
Higher-order pushdown automata - example

- order 2
- 3 stack symbols: ⊥, x, #

**Tree-generating HOPDA - definition**
From every pair of stack symbol & state there is either:
- one \( \varepsilon \)-transition
- one transition reading a letter of rank \( k \), resulting in \( k \) (ordered) pairs of state & operation.
Higher-order pushdown automata - example

- order 2
- 3 stack symbols: \( \perp, x, \# \)

Generated tree:

\[
\begin{align*}
(a, q_1) & \xrightarrow{a} (q_2, \text{push}_1 (\#)) \\
(\#, q_2) & \xrightarrow{\varepsilon} (q_3, \text{push}_2) \\
(\#, q_3) & \xrightarrow{\varepsilon} (q_4, \text{pop}_1) \\
(x, q_4) & \xrightarrow{\varepsilon} (q_5, \text{pop}_1) \\
(\_, q_5) & \xrightarrow{\varepsilon} (q_4, \text{push}_2) \\
(\_, q_4) & \xrightarrow{b} (q_4, \text{pop}_2) \\
(\#, q_4) & \xrightarrow{c}
\end{align*}
\]
Higher-order recursion schemes

pushdown automata $\xrightarrow{\text{generalization}}$ higher-order pushdown automata

custom-free grammars $\xrightarrow{\text{generalization}}$ higher-order recursion schemes
Higher-order recursion schemes - definition

Nonterminals may take arguments, that can be then used on the right side of productions.
Nonterminals may take arguments, that can be then used on the right side of productions.

Every nonterminal (every argument) has assigned some type.

Types:

\[ \alpha ::= o \mid \alpha \to \beta \]

- \( o \) – type of a tree
- \( o \to o \) – type of a function that takes a tree, and produces a tree
- \( o \to (o \to o) \to o \) – type of a function that takes a tree and a function of type \( o \to o \), and produces a tree

abbreviation of \( o \to ((o \to o) \to o) \)
Higher-order recursion schemes - definition

Nonterminals may take arguments, that can be then used on the right side of productions.

Every nonterminal (every argument) has assigned some type.

Types:

\[ \alpha ::= o \mid \alpha \rightarrow \beta \]

Order:

\[ \text{ord}(o) = 0 \]
\[ \text{ord}(\alpha_1 \rightarrow ... \rightarrow \alpha_k \rightarrow o) = 1 + \max(\text{ord}(\alpha_1), \ldots, \text{ord}(\alpha_k)) \]

- \( \text{ord}(o) = 0 \),
- \( \text{ord}(o \rightarrow o) = \text{ord}(o \rightarrow o \rightarrow o) = 1 \),
- \( \text{ord}(o \rightarrow (o \rightarrow o) \rightarrow o) = 2 \)
Higher-order recursion schemes – example

Ranked alphabet:
\[ a^{0\rightarrow o\rightarrow o} \text{ of rank 2}, \ b^{0\rightarrow o} \text{ of rank 1}, \ c^{o} \text{ of rank 0} \]

Nonterminals:
\[ S^{o} \text{ (starting), } A^{(o \rightarrow o) \rightarrow o}, \ D^{(o \rightarrow o) \rightarrow o \rightarrow o} \]
Higher-order recursion schemes – example (of order 2)

Ranked alphabet:
\[ a^o \rightarrow o \rightarrow o \text{ of rank 2}, \ b^o \rightarrow o \text{ of rank 1}, \ c^o \text{ of rank 0} \]

Nonterminals:
\[ S^o \text{ (starting), } A^{(o \rightarrow o) \rightarrow o}, \ D^{(o \rightarrow o) \rightarrow o \rightarrow o} \]

Order of a HORS = maximal order of (a type of) its nonterminal
Higher-order recursion schemes – example (of order 2)

Ranked alphabet:
\[ a^{0 \rightarrow 0 \rightarrow 0} \text{ of rank 2}, \quad b^{0 \rightarrow o} \text{ of rank 1}, \quad c^o \text{ of rank 0} \]

Nonterminals:
\[ S^o (\text{starting}), \quad A^{(o \rightarrow o) \rightarrow o}, \quad D^{(o \rightarrow o) \rightarrow o \rightarrow o} \]

Rules:
\[
S \rightarrow A \ b \\
A \ f \rightarrow a \ (A \ (D \ f)) \ (f \ c) \\
D \ f \ x \rightarrow f \ (f \ x)
\]

It is required that:
1) types are respected
   e.g. \( D \) of type \((o \rightarrow o) \rightarrow o \rightarrow o\) is applied to \( f \) of type \(o \rightarrow o\),
   \( A \) of type \((o \rightarrow o) \rightarrow o\) is applied to \( D \ f \) of type \(o \rightarrow o\), etc.
2) right side of every rule is of type \( o \)
Higher-order recursion schemes – example (of order 2)

Ranked alphabet:
- $a^0 \rightarrow o \rightarrow o$ of rank 2, $b^0 \rightarrow o$ of rank 1, $c^0$ of rank 0

Nonterminals:
- $S^0$ (starting), $A^{(o \rightarrow o) \rightarrow o}$, $D^{(o \rightarrow o) \rightarrow o \rightarrow o}$

Rules:
- $S \rightarrow A \ b$
- $A \ f \rightarrow a \ (A \ (D \ f)) \ (f \ c)$
- $D \ f \ x \rightarrow f \ (f \ x)$

$S \rightarrow A \ b \rightarrow a \ (A \ (D \ b)) \ (b \ c)$
Higher-order recursion schemes – example (of order 2)

Ranked alphabet:
- $a^o \rightarrow o \rightarrow o$ of rank 2,
- $b^o \rightarrow o$ of rank 1,
- $c^o$ of rank 0

Nonterminals:
- $S^o$ (starting), $A^{(o \rightarrow o) \rightarrow o}$, $D^{(o \rightarrow o) \rightarrow o \rightarrow o}$

Rules:
- $S \rightarrow A \ b$
- $A \ f \rightarrow a \ (A \ (D \ f)) \ (f \ c)$
- $D \ f \ x \rightarrow f \ (f \ x)$

$S \rightarrow A \ b \rightarrow a \ (A \ (D \ b)) \ (b \ c)$
Higher-order recursion schemes – example (of order 2)

Ranked alphabet:
- $a^0→o→o$ of rank 2,
- $b^0→o$ of rank 1,
- $c^o$ of rank 0

Nonterminals:
- $S^o$ (starting), $A^{(o→o)→o}$, $D^{(o→o)→o→o}$

Rules:
- $S → A b$
- $A f → a (A (D f)) (f c)$
- $D f x → f (f x)$

\[
S → A b → a (A (D b)) (b c)
\]
Higher-order recursion schemes – example (of order 2)

Ranked alphabet:
- \(a^{o \rightarrow o \rightarrow o}\) of rank 2,
- \(b^{o \rightarrow o}\) of rank 1,
- \(c^{o}\) of rank 0

Nonterminals:
- \(S^{o}\) (starting),
- \(A^{(o \rightarrow o) \rightarrow o}\),
- \(D^{(o \rightarrow o) \rightarrow o \rightarrow o}\)

Rules:
\[
\begin{align*}
S & \rightarrow A \ b \\
A \ f & \rightarrow a \ (A \ (D \ f)) \ (f \ c) \\
D \ f \ x & \rightarrow f \ (f \ x)
\end{align*}
\]

\[
\begin{align*}
S & \rightarrow A \ b \rightarrow a \ (A \ (D \ b)) \ (b \ c) \\
A \ (D \ b) & \rightarrow a \ (A \ (D \ (D \ b))) \ (D \ b \ c)
\end{align*}
\]
Higher-order recursion schemes – example (of order 2)

Ranked alphabet:
\[ a^{0 \rightarrow 0 \rightarrow 0} \text{ of rank 2}, \ b^{0 \rightarrow 0} \text{ of rank 1}, \ c^0 \text{ of rank 0} \]

Nonterminals:
\[ S^0 \text{ (starting)}, \ A^{(0 \rightarrow 0) \rightarrow 0}, \ D^{(0 \rightarrow 0) \rightarrow 0 \rightarrow 0} \]

Rules:
\[ S \rightarrow A \ b \]
\[ A \ f \rightarrow a \ (A \ (D \ f)) \ (f \ c) \]
\[ D \ f \ x \rightarrow f \ (f \ x) \]
\[ S \rightarrow A \ b \rightarrow a \ (A \ (D \ b)) \ (b \ c) \]
\[ A \ (D \ b) \rightarrow a \ (A \ (D \ (D \ b))) \ (D \ b \ c) \]
Higher-order recursion schemes – example (of order 2)

Ranked alphabet:
\[ a^{0 \to o \to o} \text{ of rank 2, } b^{o \to o} \text{ of rank 1, } c^o \text{ of rank 0} \]

Nonterminals:
- \( S^o \) (starting), \( A^{(o \to o) \to o} \), \( D^{(o \to o) \to o \to o} \)

Rules:
- \( S \to A \ b \)
- \( A \ f \to a \ (A \ (D \ f)) \ (f \ c) \)
- \( D \ f \ x \to f \ (f \ x) \)

\begin{align*}
S & \to A \ b \to a \ (A \ (D \ b)) \ (b \ c) \\
A \ (D \ b) & \to a \ (A \ (D \ (D \ b))) \ (D \ b \ c) \\
D \ b \ c & \to b \ (b \ c)
\end{align*}
Higher-order recursion schemes – example (of order 2)

Ranked alphabet:
- $a^{o \rightarrow o \rightarrow o}$ of rank 2,
- $b^{o \rightarrow o}$ of rank 1,
- $c^{o}$ of rank 0

Nonterminals:
- $S^{o}$ (starting), $A^{(o \rightarrow o) \rightarrow o}$, $D^{(o \rightarrow o) \rightarrow o \rightarrow o}$

Rules:
- $S \rightarrow A \ b$
- $A \ f \rightarrow a \ (A \ (D \ f)) \ (f \ c)$
- $D \ f \ x \rightarrow f \ (f \ x)$

$S \rightarrow A \ b \rightarrow a \ (A \ (D \ b)) \ (b \ c)$
- $A \ (D \ b) \rightarrow a \ (A \ (D \ (D \ b))) \ (D \ b \ c)$
- $D \ b \ c \rightarrow b \ (b \ c)$
- $A \ (D \ (D \ b)) \rightarrow a \ (A \ (D \ (D \ (D \ b)))) \ (D \ (D \ b) \ c)$
- $D \ (D \ b) \ c \rightarrow D \ b \ (D \ b \ c) \rightarrow b \ (b \ (D \ b \ c))$
Higher-order recursion schemes – example (of order 2)

Ranked alphabet:

\hspace{0.5cm} a^{o \rightarrow o \rightarrow o} \text{ of rank } 2, \hspace{0.2cm} b^{o \rightarrow o} \text{ of rank } 1, \hspace{0.2cm} c^{o} \text{ of rank } 0

Nonterminals:

\hspace{0.5cm} S^{o} \text{ (starting)}, \hspace{0.2cm} A^{(o \rightarrow o) \rightarrow o}, \hspace{0.2cm} D^{(o \rightarrow o) \rightarrow o \rightarrow o}

Rules:

\begin{align*}
S & \rightarrow A \ b \\
A \ f & \rightarrow a \ (A \ (D \ f)) \ (f \ c) \\
D \ f \ x & \rightarrow f \ (f \ x)
\end{align*}

\begin{align*}
S & \rightarrow A \ b \rightarrow a \ (A \ (D \ b)) \ (b \ c) \\
A \ (D \ b) & \rightarrow a \ (A \ (D \ (D \ b))) \ (D \ b \ c) \\
D \ b \ c & \rightarrow b \ (b \ c) \\
A \ (D \ (D \ b)) & \rightarrow a \ (A \ (D \ (D \ (D \ b)))) \ (D \ (D \ b) \ c) \\
D \ (D \ b) \ c & \rightarrow D \ b \ (D \ b \ c) \rightarrow b \ (b \ (D \ b \ c))
\end{align*}
Higher-order recursion schemes

- Previous slides: a deterministic HORS generating a single tree.
Higher-order recursion schemes

- Previous slides: a deterministic HORS generating a single tree.
- One can also consider a nondeterministic HORS, recognizing a language of finite trees.
Higher-order recursion schemes

- Previous slides: a deterministic HORS generating a single tree.
- One can also consider a nondeterministic HORS, recognizing a language of finite trees.
- If every letter is of rank 1, except a single letter of rank 0, then these trees, consisting of a single branch, can be seen as words → the HORS recognizes a set of words.
Higher-order recursion schemes

- Previous slides: a deterministic HORS generating a single tree.
- One can also consider a nondeterministic HORS, recognizing a language of finite trees.
- If every letter is of rank 1, except a single letter of rank 0, then these trees, consisting of a single branch, can be seen as words → the HORS recognizes a set of words.

Example:

Alphabet: \( a \) of rank 2, \( b \) of rank 1, \( c \) of rank 0

Nonterminals: \( S^o \) (starting), \( A^{(o \rightarrow o) \rightarrow o} \), \( D^{(o \rightarrow o) \rightarrow o \rightarrow o} \)

Rules:

\[
S \rightarrow A \ b \\
A f \rightarrow a \ (A \ (D \ f)) \ (f \ c) \\
D \ f \ x \rightarrow f \ (f \ x)
\]

Recognized language: \( \{b^{2^k} : k \in \mathbb{N}\} \)
HOPDA vs HORS

Are these two formalisms equivalent?

higher-order pushdown automata $\iff$ higher-order recursion schemes
HOPDA vs HORS

Are these two formalisms equivalent?

higher-order pushdown automata \(\not\leftrightarrow\) higher-order recursion schemes

Not exactly!
HOPDA vs HORS

Are these two formalisms equivalent?

higher-order pushdown automata ⇔ safe higher-order recursion schemes

collapsible pushdown automata ⇔ higher-order recursion schemes

Theorem [Knapik, Niwiński, Urzyczyn 2002 & earlier results]
For every $n$, HOPDA of order $n$ and safe HORSes of order $n$ generate the same trees (recognize the same word languages);
[Caucal 2002] these are trees from the Caucal hierarchy, defined by iterating MSO interpretations and unfolding of graphs into trees.

Theorem [Hague, Murawski, Ong, Serre 2008]
For every $n$, collapsible HOPDA of order $n$ and HORSes of order $n$ generate the same trees (recognize the same word languages).
What is safety?

Restriction on terms appearing on right sides of rules:

- unrestricted terms:
  \[ M ::= a \mid x \mid A \mid MN \]

- safe terms:
  \[ M ::= a \mid x \mid A \mid MN_1 \ldots N_k \]
  only if \( \text{ord}(MN_1 \ldots N_k) \leq \text{ord}(N_i) \) for all \( i \)

In other words: if we apply an argument of some order \( k \),
then we have to apply also all arguments of order \( \geq k \)
What is safety?

Restriction on terms appearing on right sides of rules:

- unrestricted terms:
  \[ M ::= a \mid x \mid A \mid MN \]

- safe terms:
  \[ M ::= a \mid x \mid A \mid MN_1 \ldots N_k \]
  only if \( \text{ord}(MN_1 \ldots N_k) \leq \text{ord}(N_i) \) for all \( i \)

In other words: if we apply an argument of some order \( k \), then we have to apply also all arguments of order \( \geq k \)

Let's check for our example HORS:

\[
\begin{align*}
S & \rightarrow Ab \\
Af & \rightarrow a(A(Df))(fc) \\
Dfx & \rightarrow f(fx)
\end{align*}
\]
What is safety?

Restriction on terms appearing on right sides of rules:

• unrestricted terms:
  \[ M ::= a | x | A | M N \]

• safe terms:
  \[ M ::= a | x | A | M N_1 \ldots N_k \]
  only if \( \text{ord}(M N_1 \ldots N_k) \leq \text{ord}(N_i) \) for all \( i \)

In other words: if we apply an argument of some order \( k \), then we have to apply also all arguments of order \( \geq k \)

Let's check for our example HORS:

\[
\begin{align*}
S & \rightarrow A \ b \\
A \ f & \rightarrow a \ (A \ (D \ f)) \ (f \ c) \\
D \ f \ x & \rightarrow f \ (f \ x)
\end{align*}
\]

\( \text{ord}(D \ f) = 1 \leq 1 = \text{ord}(f) \rightarrow \text{OK} \)
What is safety?

Restriction on terms appearing on right sides of rules:

• unrestricted terms:
  \[ M ::= a \mid x \mid A \mid M \, N \]

• safe terms:
  \[ M ::= a \mid x \mid A \mid M \, N_1 \ldots \, N_k \]
  only if \( \text{ord}(M \, N_1 \ldots \, N_k) \leq \text{ord}(N_i) \) for all \( i \)

In other words: if we apply an argument of some order \( k \),
then we have to apply also all arguments of order \( \geq k \)

Let's check for our example HORS:

\[ \begin{align*}
S & \rightarrow A \, b \\
A \, f & \rightarrow a \, (A \, (D \, f)) \, (f \, c) \quad \checkmark \, \text{safe} \\
D \, f \, x & \rightarrow f \, (f \, x)
\end{align*} \]

\( \text{ord}(D \, f) = 1 \leq 1 = \text{ord}(f) \rightarrow \text{OK} \)
All other subterms are of order 0 \rightarrow \text{OK}
What is safety?

Restriction on terms appearing on right sides of rules:

• unrestricted terms:
  \[ M ::= a \mid x \mid A \mid M \, N \]

• safe terms:
  \[ M ::= a \mid x \mid A \mid M \, N_1 \ldots \, N_k \]
  only if \( \text{ord}(M \, N_1 \ldots \, N_k) \leq \text{ord}(N_i) \) for all \( i \)

In other words: if we apply an argument of some order \( k \), then we have to apply also all arguments of order \( \geq k \)

Example: Unsafe HORS (generating ”Urzyczyn's tree” \( U \)):
Types: \( a^{0\to o\to o}, b^{o\to o}, c^{o\to o}, d^{o}, e^{o}, S^{o}, F^{(o\to o)\to o\to o\to o} \)
Rules: \( S \to F \, b \, d \, e \)
\[ F \, f \, x \, y \to a \, (F \, (F \, f \, x) \, y \, (c \, y)) \, (a \, (f \, y) \, x) \]  
\( \times \) unsafe

(and not equivalent to any safe HORS)

\( \text{ord}(F \, f \, x) = 1 > 0 = \text{ord}(x) \)

\( F \) expects two order-0 arguments; we have applied one \( (x) \), but not the other)
Why safety helps?

**Theorem** [Knapik, Niwiński, Urzyczyn 2002; Blum, Ong 2007]
Substitution (hence $\beta$-reduction) in safe $\lambda$-calculus can be implemented without renaming bound variables.

Bad example: when you substitute $(\lambda x.y\ x) [a\ x\ x/y]$, it is necessary to change the first two $x$ to some other variable name.
Collapsible pushdown automata

- Every stack symbol has an identifier.
- $push_1x$ pushes symbol $x$ with a fresh identifier.
- $push_k$ for $k \geq 2$ copy symbols with their identifiers.
Collapsible pushdown automata

- Every stack symbol has an identifier.
- \( \text{push}_1 x \) pushes symbol \( x \) with a fresh identifier.
- \( \text{push}_k \) for \( k \geq 2 \) copy symbols with their identifiers.
- New operation \( \text{collapse}_k \):
  - remove from the topmost order-\( k \) stack all order-\( (k-1) \) stacks containing a copy of the topmost stack symbol.
Collapsible pushdown automata

- Every stack symbol has an identifier.
- $push_1 x$ pushes symbol $x$ with a fresh identifier.
- $push_k$ for $k \geq 2$ copy symbols with their identifiers.
- New operation $collapse_k$:
  remove from the topmost order-$k$ stack all order-$(k-1)$ stacks containing a copy of the topmost stack symbol.
Collapsible pushdown automata

- Every stack symbol has an identifier.
- $push_1 x$ pushes symbol $x$ with a fresh identifier.
- $push_k$ for $k \geq 2$ copy symbols with their identifiers.
- New operation $\text{collapse}_k$:
  remove from the topmost order-$k$ stack all order-$(k-1)$ stacks containing a copy of the topmost stack symbol.

$push_1 x$
Collapsible pushdown automata

• Every stack symbol has an identifier.
• $\text{push}_1 x$ pushes symbol $x$ with a fresh identifier.
• $\text{push}_k$ for $k \geq 2$ copy symbols with their identifiers.
• New operation $\text{collapse}_k$:
  remove from the topmost order-$k$ stack all order-$(k-1)$ stacks containing a copy of the topmost stack symbol.

\[
\text{push}_1 x \\
\text{push}_2
\]
Collapsible pushdown automata

- Every stack symbol has an identifier.
- $push_1 x$ pushes symbol $x$ with a fresh identifier.
- $push_k$ for $k \geq 2$ copy symbols with their identifiers.
- New operation $\text{collapse}_k$:
  remove from the topmost order-$k$ stack all order-$(k-1)$ stacks containing a copy of the topmost stack symbol.

\[
\begin{array}{c}
push_1 x \\
push_2 \\
push_2
\end{array}
\]

\[
\begin{array}{ccc}
x,1 & x,1 & x,1 \\
\bot & \bot & \bot
\end{array}
\]
Collapsible pushdown automata

- Every stack symbol has an identifier.
- $push_1x$ pushes symbol $x$ with a fresh identifier.
- $push_k$ for $k \geq 2$ copy symbols with their identifiers.
- New operation $collapse_k$:
  remove from the topmost order-$k$ stack all order-$(k-1)$ stacks containing a copy of the topmost stack symbol.

```
push_1x
push_2
push_2
pop_1
```

```
x,1
x,1
pop
pop
```
Collapsible pushdown automata

- Every stack symbol has an identifier.
- $\text{push}_1x$ pushes symbol $x$ with a fresh identifier.
- $\text{push}_k$ for $k \geq 2$ copy symbols with their identifiers.
- New operation $\text{collapse}_k$: 
  remove from the topmost order-$k$ stack all order-$(k-1)$ stacks containing a copy of the topmost stack symbol.

```
push_1x
push_2
push_2
pop_1
push_1x
```

```
x,1   x,1   x,2
   ⊥     ⊥     ⊥
```
Collapsible pushdown automata

- Every stack symbol has an identifier.
- $\text{push}_1x$ pushes symbol $x$ with a fresh identifier.
- $\text{push}_k$ for $k \geq 2$ copy symbols with their identifiers.
- New operation $\text{collapse}_k$:
  
  remove from the topmost order-$k$ stack all order-$(k-1)$ stacks containing a copy of the topmost stack symbol.

<table>
<thead>
<tr>
<th>$\text{x,1}$</th>
<th>$\text{x,1}$</th>
<th>$\text{x,2}$</th>
<th>$\text{x,2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bot$</td>
<td>$\bot$</td>
<td>$\bot$</td>
<td>$\bot$</td>
</tr>
</tbody>
</table>

$\text{push}_1x$

$\text{push}_2$

$\text{push}_2$

$\text{pop}_1$

$\text{push}_1x$

$\text{push}_2$
Collapsible pushdown automata

- Every stack symbol has an identifier.
- \( \text{push}_1 x \) pushes symbol \( x \) with a fresh identifier.
- \( \text{push}_k \) for \( k \geq 2 \) copy symbols with their identifiers.
- New operation \( \text{collapse}_k \):
  remove from the topmost order-\( k \) stack all order-(\( k-1 \)) stacks containing a copy of the topmost stack symbol.
Collapsible pushdown automata

- Every stack symbol has an identifier.
- $push_1x$ pushes symbol $x$ with a fresh identifier.
- $push_k$ for $k \geq 2$ copy symbols with their identifiers.
- New operation $collapse_k$:
  remove from the topmost order-$k$ stack all order-$(k-1)$ stacks containing a copy of the topmost stack symbol.

\[
\begin{array}{cccc}
\text{x,1} & \text{x,1} & \text{x,2} & \text{z,3} \\
\bot & \bot & \bot & \bot \\
\end{array}
\]

\[
\begin{array}{cccc}
\text{z,3} & \text{z,3} & \text{x,2} & \bot \\
\bot & \bot & \bot & \bot \\
\end{array}
\]

\[
\begin{array}{cccc}
push_1x \\
push_2 \\
push_2 \\
pop_1 \\
push_1x \\
push_2 \\
push_1z \\
push_2 \\
\end{array}
\]
Collapsible pushdown automata

- Every stack symbol has an identifier.
- $push_1 x$ pushes symbol $x$ with a fresh identifier.
- $push_k$ for $k \geq 2$ copy symbols with their identifiers.
- New operation $\text{collapse}_k$:
  remove from the topmost order-$k$ stack all order-$(k-1)$ stacks containing a copy of the topmost stack symbol.

```
push_1 x
push_2
push_2
z,3
pop_1
push_1 x
push_2
push_2
push_1 z
push_2
pop_1
```
Collapsible pushdown automata

• Every stack symbol has an identifier.
• $push_1x$ pushes symbol $x$ with a fresh identifier.
• $push_k$ for $k \geq 2$ copy symbols with their identifiers.
• New operation $\text{collapse}_k$:
  remove from the topmost order-$k$ stack all order-$(k-1)$ stacks containing a copy of the topmost stack symbol.

```
push_1x
push_2
push_2
pop_1
push_1x
push_2
push_1z
push_2
pop_1
collapse_2
```
Collapsible pushdown automata

How collapse can be useful? – Urzyczyn's language U
(≈ branches in the Urzyczyn's tree)

alphabet: [, ], *

U contains words of the form:

```
[ [ ] [ [ ] [ [ ] ] ] ] ****
```

- segment A forms a prefix of a well-bracketed word that ends in [ not matched in the entire word
- segment B forms a well-bracketed word
- the number of stars in C equals the number of brackets in A
Collapsible pushdown automata

How collapse can be useful? – Urzyczyn's language U
(≈ branches in the Urzyczyn's tree)

Words in U:
A) a prefix of a well-bracketed word
B) a well-bracketed word
C) as many stars as brackets in part A

→ one stack symbol
→ first-order stack counts the number of currently open brackets
→ a copy is done after each bracket
Collapsible pushdown automata

How collapse can be useful? – Urzyczyn's language $U$

($\approx$ branches in the Urzyczyn's tree)

Words in $U$:
A) a prefix of a well-bracketed word
B) a well-bracketed word
C) as many stars as brackets in part A

→ one stack symbol
→ first-order stack counts the number of currently open brackets
→ a copy is done after each bracket

order-2 stack

$\left[ \left[ \right] \left[ \right] \left[ \right] \left[ \right] \left[ \right] \left[ \right] \right] \ast \ast \ast \ast \ast$
Collapsible pushdown automata

How collapse can be useful? – Urzyczyn's language U

\( \approx \) branches in the Urzyczyn's tree

Words in U:
A) a prefix of a well-bracketed word
B) a well-bracketed word
C) as many stars as brackets in part A

- one stack symbol
- first-order stack counts the number of currently open brackets
- a copy is done after each bracket

\[ \quad \text{order-2 stack} \quad \]

\[ [ [ ] [ [ ] [ [ ] [ [ ] ] ] ] ] \ast \ast \ast \ast \ast \]
Collapsible pushdown automata

How collapse can be useful? – Urzyczyn's language U
(≈ branches in the Urzyczyn's tree)

Words in U:
A) a prefix of a well-bracketed word
B) a well-bracketed word
C) as many stars as brackets in part A

→ one stack symbol
→ first-order stack counts the number of currently open brackets
→ a copy is done after each bracket
Collapsible pushdown automata

How collapse can be useful? – Urzyczyn's language $U$

$(\approx$ branches in the Urzyczyn's tree$)$

Words in $U$:
A) a prefix of a well-bracketed word
B) a well-bracketed word
C) as many stars as brackets in part A

→ one stack symbol
→ first-order stack counts the number of currently open brackets
→ a copy is done after each bracket
→ on the first star make the collapse
→ count the number of stacks

Collapse = remove all stacks on which the topmost symbol is present
Collapsible pushdown automata

How collapse can be useful? – Urzyczyn's language U
(≈ branches in the Urzyczyn's tree)

Words in U:
A) a prefix of a well-bracketed word
B) a well-bracketed word
C) as many stars as brackets in part A

→ one stack symbol
→ first-order stack counts the number of currently open brackets
→ a copy is done after each bracket
→ on the first star make the collapse
→ count the number of stacks

Collapse = remove all stacks on which the topmost symbol is present
Collapsible pushdown automata

How collapse can be useful? – Urzyczyn's language U  
(≈ branches in the Urzyczyn's tree)

Words in U: 
A) a prefix of a well-bracketed word
B) a well-bracketed word
C) as many stars as brackets in part A

- one stack symbol
- first-order stack counts the number of currently open brackets
- a copy is done after each bracket
- on the first star make the collapse
- count the number of stacks

Remark:  
A nondeterministic order-2 PDA without collapse can recognize U,  
as it can guess when is the beginning of the “B” part.  
But not a deterministic HOPDA without collapse, of any order!  
(This means that the Urzyczyn's tree cannot be generated by a HOPDA)
Expressivity questions

Tree\((n)\) = trees generated by HORSes (CPDA) of order \(n\)
SafeTree\((n)\) = trees generated by safe HORSes (HOPDA) of order \(n\)

\[
\begin{align*}
\text{SafeTree}(0) & \subseteq \text{SafeTree}(1) \subseteq \text{SafeTree}(2) \subseteq \text{SafeTree}(3) \subseteq \ldots \\
\text{Tree}(0) & \subseteq \text{Tree}(1) \subseteq \text{Tree}(2) \subseteq \text{Tree}(3) \subseteq \ldots
\end{align*}
\]

Lang\((n)\) = word languages recogn. by HORSes (CPDA) of order \(n\)
SafeLang\((n)\) = word lang. rec. by safe HORSes (HOPDA) of order \(n\)

\[
\begin{align*}
\text{SafeLang}(0) & \subseteq \text{SafeLang}(1) \subseteq \text{SafeLang}(2) \subseteq \text{SafeLang}(3) \subseteq \ldots \\
\text{Lang}(0) & \subseteq \text{Lang}(1) \subseteq \text{Lang}(2) \subseteq \text{Lang}(3) \subseteq \ldots
\end{align*}
\]
Expressivity questions

Tree($n$) = trees generated by HORSes (CPDA) of order $n$
SafeTree($n$) = trees generated by safe HORSes (HOPDA) of order $n$

\[
\begin{align*}
\text{SafeTree}(0) & \subseteq \text{SafeTree}(1) \subseteq \text{SafeTree}(2) \subseteq \text{SafeTree}(3) \subseteq \ldots \\
\text{Tree}(0) & \subseteq \text{Tree}(1) \subseteq \text{Tree}(2) \subseteq \text{Tree}(3) \subseteq \ldots
\end{align*}
\]

regular trees

Lang($n$) = word languages recogn. by HORSes (CPDA) of order $n$
SafeLang($n$) = word lang. rec. by safe HORSes (HOPDA) of order $n$

\[
\begin{align*}
\text{SafeLang}(0) & \subseteq \text{SafeLang}(1) \subseteq \text{SafeLang}(2) \subseteq \text{SafeLang}(3) \subseteq \ldots \\
\text{Lang}(0) & \subseteq \text{Lang}(1) \subseteq \text{Lang}(2) \subseteq \text{Lang}(3) \subseteq \ldots
\end{align*}
\]

regular languages
Expressivity questions

Tree($n$) = trees generated by HORSes (CPDA) of order $n$
SafeTree($n$) = trees generated by safe HORSes (HOPDA) of order $n$

SafeTree(0) $\subseteq$ SafeTree(1) $\subseteq$ SafeTree(2) $\subseteq$ SafeTree(3) $\subseteq$ ...
Tree(0) $\subseteq$ Tree(1) $\subseteq$ Tree(2) $\subseteq$ Tree(3) $\subseteq$ ...
regular trees $|$ context-free trees

Lang($n$) = word languages recogn. by HORSes (CPDA) of order $n$
SafeLang($n$) = word lang. rec. by safe HORSes (HOPDA) of order $n$

SafeLang(0) $\subseteq$ SafeLang(1) $\subseteq$ SafeLang(2) $\subseteq$ SafeLang(3) $\subseteq$ ...
Lang(0) $\subseteq$ Lang(1) $\subseteq$ Lang(2) $\subseteq$ Lang(3) $\subseteq$ ...
regular languages $|$ context-free languages

regular languages $=$ context-free languages

23
Expressivity questions

Tree($n$) = trees generated by HORSes (CPDA) of order $n$
SafeTree($n$) = trees generated by safe HORSes (HOPDA) of order $n$

SafeTree(0) $\subseteq$ SafeTree(1) $\subseteq$ SafeTree(2) $\subseteq$ SafeTree(3) $\subseteq$ ...  
Tree(0) $\subseteq$ Tree(1) $\subseteq$ Tree(2) $\subseteq$ Tree(3) $\subseteq$ ... 
regular trees $\subseteq$ context-free trees

Lang($n$) = word languages recogn. by HORSes (CPDA) of order $n$
SafeLang($n$) = word lang. rec. by safe HORSes (HOPDA) of order $n$

SafeLang(0) $\subseteq$ SafeLang(1) $\subseteq$ SafeLang(2) $\subseteq$ SafeLang(3) $\subseteq$ ...  
Lang(0) $\subseteq$ Lang(1) $\subseteq$ Lang(2) $\subseteq$ Lang(3) $\subseteq$ ... 
regular languages $\subseteq$ context-free languages $\subseteq$ indexed languages
Expressivity questions

Tree(n) = trees generated by HORSes (CPDA) of order $n$
SafeTree(n) = trees generated by safe HORSes (HOPDA) of order $n$

SafeTree(0) $\subseteq$ SafeTree(1) $\subseteq$ SafeTree(2) $\subseteq$ SafeTree(3) $\subseteq$ ...
Tree(0) $\subseteq$ Tree(1) $\subseteq$ Tree(2) $\subseteq$ Tree(3) $\subseteq$ ...
regular $\quad$ context-free
trees

Lang(n) = word languages recogn. by HORSes (CPDA) of order $n$
SafeLang(n) = word lang. rec. by safe HORSes (HOPDA) of order $n$

SafeLang(0) $\subseteq$ SafeLang(1) $\subseteq$ SafeLang(2) $\subseteq$ SafeLang(3) $\subseteq$ ...
Lang(0) $\subseteq$ Lang(1) $\subseteq$ Lang(2) $\subseteq$ Lang(3) $\subseteq$ ...
regular $\quad$ context-free $\quad$ indexed
languages $\quad$ languages $\quad$ languages

Are these hierarchies strict?
Expressivity questions

\[
\begin{align*}
\text{SafeTree}(0) \subseteq & \text{SafeTree}(1) \subseteq \text{SafeTree}(2) \subseteq \text{SafeTree}(3) \subseteq \ldots \\
\text{Tree}(0) \subseteq & \text{Tree}(1) \subseteq \text{Tree}(2) \subseteq \text{Tree}(3) \subseteq \ldots \\
\text{SafeLang}(0) \subseteq & \text{SafeLang}(1) \subseteq \text{SafeLang}(2) \subseteq \text{SafeLang}(3) \subseteq \ldots \\
\text{Lang}(0) \subseteq & \text{Lang}(1) \subseteq \text{Lang}(2) \subseteq \text{Lang}(3) \subseteq \ldots 
\end{align*}
\]

Are these hierarchies strict?

**Theorem** [Engelfriet 1991]
For every \( n \), SafeLang\( (n) \neq \text{SafeLang}(n+1) \), and thus also SafeTree\( (n) \neq \text{SafeTree}(n+1) \).

Separating language: correct sequences of operations of order-(\( n+1 \)) HOPDA (including the topmost stack symbol after every step).

Proof: “Simple trick” using the fact that reachability for order-\( n \) HOPDA is in (\( n-1 \))-EXPTIME, while reachability for order-(\( n+1 \)) HOPDA is \( n \)-EXPTIME-hard.
Expressivity questions

\[
\text{SafeTree}(0) \subseteq \text{SafeTree}(1) \subseteq \text{SafeTree}(2) \subseteq \text{SafeTree}(3) \subseteq \ldots \\
\text{Tree}(0) \subseteq \text{Tree}(1) \subseteq \text{Tree}(2) \subseteq \text{Tree}(3) \subseteq \ldots \\
\text{SafeLang}(0) \subseteq \text{SafeLang}(1) \subseteq \text{SafeLang}(2) \subseteq \text{SafeLang}(3) \subseteq \ldots \\
\text{Lang}(0) \subseteq \text{Lang}(1) \subseteq \text{Lang}(2) \subseteq \text{Lang}(3) \subseteq \ldots 
\]

Are these hierarchies strict?

**Theorem** [Engelfriet 1991]
For every \( n \), \( \text{SafeLang}(n) \neq \text{SafeLang}(n+1) \), and thus also \( \text{SafeTree}(n) \neq \text{SafeTree}(n+1) \).

Separating language: correct sequences of operations of order-\((n+1)\) HOPDA (including the topmost stack symbol after every step).

Proof: “Simple trick” using the fact that reachability for order-\( n \) HOPDA is in \((n-1)\)-EXPTIME, while reachability for order-\((n+1)\) HOPDA is \( n \)-EXPTIME-hard.

The same proof works for CPDA.
Thus \( \text{Tree}(n) \neq \text{Tree}(n+1) \) & \( \text{Lang}(n) \neq \text{Lang}(n+1) \).
Expressivity questions

SafeTree(0) \subsetneq \text{SafeTree}(1) \subsetneq \text{SafeTree}(2) \subsetneq \text{SafeTree}(3) \subsetneq \ldots
\quad \text{Tree}(0) \subsetneq \text{Tree}(1) \subsetneq \text{Tree}(2) \subsetneq \text{Tree}(3) \subsetneq \ldots
\quad \text{SafeLang}(0) \subsetneq \text{SafeLang}(1) \subsetneq \text{SafeLang}(2) \subsetneq \text{SafeLang}(3) \subsetneq \ldots
\quad \text{Lang}(0) \subsetneq \text{Lang}(1) \subsetneq \text{Lang}(2) \subsetneq \text{Lang}(3) \subsetneq \ldots

Are these hierarchies strict?

Another separator:
\[ T_n = \text{tree with branches } a^{k b^{\exp_n(k)} c}, \text{ where } \exp_n(k) = 2^{2^{n^{2^k}}} \]

We have SafeTree(n+1) \ni T_n \not\in \text{Tree}(n).
Expressivity questions

SafeTree(0) ⊆ SafeTree(1) ⊆ SafeTree(2) ⊆ SafeTree(3) ⊆ ...
Tree(0) ⊆ Tree(1) ⊆ Tree(2) ⊆ Tree(3) ⊆ ...

SafeLang(0) ⊆ SafeLang(1) ⊆ SafeLang(2) ⊆ SafeLang(3) ⊆ ...
Lang(0) ⊆ Lang(1) ⊆ Lang(2) ⊆ Lang(3) ⊆ ...

Are these hierarchies strict?

Another separator:
\[ T_n = \text{tree with branches } a^k b^{\exp_n(k)} c, \text{ where } \exp_n(k) = 2^{2^{k-1}} \]
We have SafeTree(n+1) \ni T_n \notin Tree(n).

For languages we do not know:
SafeLang(n+1) \ni \{ b^{\exp_n(k)} : k \in \mathbb{N} \} \notin Lang(n).

Open problem: a pumping lemma for nondeterministic HORSes.
Expressivity questions

SafeTree(0) $\subsetneq$ SafeTree(1) $\subsetneq$ SafeTree(2) $\subsetneq$ SafeTree(3) $\subsetneq$ ...
\[\text{Tree}(0) \subsetneq \text{Tree}(1) \subsetneq \text{Tree}(2) \subsetneq \text{Tree}(3) \subsetneq \ldots\]

SafeLang(0) $\subsetneq$ SafeLang(1) $\subsetneq$ SafeLang(2) $\subsetneq$ SafeLang(3) $\subsetneq$ ...
\[\text{Lang}(0) \subsetneq \text{Lang}(1) \subsetneq \text{Lang}(2) \subsetneq \text{Lang}(3) \subsetneq \ldots\]

Is safety really a restriction?
Expressivity questions

SafeTree(0) ⊈ SafeTree(1) ⊈ SafeTree(2) ⊈ SafeTree(3) ⊈ ... 
Tree(0) ⊈ Tree(1) ⊈ Tree(2) ⊈ Tree(3) ⊈ ...

SafeLang(0) ⊈ SafeLang(1) ⊈ SafeLang(2) ⊈ SafeLang(3) ⊈ ... 
Lang(0) ⊈ Lang(1) ⊈ Lang(2) ⊈ Lang(3) ⊈ ...

Is safety really a restriction?

For trees – yes.
Example: Urzyczyn's tree $U$

$\text{Tree}(2) \ni U \notin \text{SafeTree}(n)$ for every $n$  [P. 2012]

For word languages – open problem (e.g. SafeLang(3) $\not\cong$ Lang(3))
Expressivity questions

\[
\text{SafeLang}(0) \not\subseteq \text{SafeLang}(1) \not\subseteq \text{SafeLang}(2) \not\subseteq \text{SafeLang}(3) \not\subseteq \cdots
\]

\[
\text{Lang}(0) \not\subseteq \text{Lang}(1) \not\subseteq \text{Lang}(2) \not\subseteq \text{Lang}(3) \not\subseteq \cdots
\]

Are these languages context-sensitive?
Expressivity questions

SafeLang(0) ⊈ SafeLang(1) ⊈ SafeLang(2) ⊈ SafeLang(3) ⊈ ... ⊆ CSens

Lang(0) ⊈ Lang(1) ⊈ Lang(2) ⊈ Lang(3) ⊈ ... ⊆ CSens

Are these languages context-sensitive?

CSens = context-sensitive languages (type-1 in the Chomsky hierarchy)

SafeLang(n) ⊆ CSens, for every n  [Inaba, Maneth 2008]
Expressivity questions

\[
\begin{align*}
\text{SafeLang}(0) & \not\subseteq \text{SafeLang}(1) \not\subseteq \text{SafeLang}(2) \not\subseteq \text{SafeLang}(3) \not\subseteq \cdots \subseteq \text{CSens} \\
\text{Lang}(0) & \not\subseteq \text{Lang}(1) \not\subseteq \text{Lang}(2) \not\subseteq \text{Lang}(3) \not\subseteq \cdots \subseteq \text{CSens}
\end{align*}
\]

Are these languages context-sensitive?

CSens = context-sensitive languages (type-1 in the Chomsky hierarchy)

SafeLang\((n) \subseteq \text{CSens}, \text{ for every } n\) \ [Inaba, Maneth 2008]

Lang\((3) \subseteq \text{CSens}\) \ [Kobayashi, Inaba, Tsukada 2014]
Expressivity questions

SafeLang(0) ⊄ SafeLang(1) ⊄ SafeLang(2) ⊄ SafeLang(3) ⊄ ... ⊆ CSens

Lang(0) ⊄ Lang(1) ⊄ Lang(2) ⊄ Lang(3) ⊄ ... ⊆ CSens

Are these languages context-sensitive?

CSens = context-sensitive languages (type-1 in the Chomsky hierarchy)

SafeLang(n) ⊆ CSens, for every n  [Inaba, Maneth 2008]

Lang(3) ⊆ CSens  [Kobayashi, Inaba, Tsukada 2014]

Lang(n) ⊄ CSens for \( n \geq 4 \)  – open problem
Expressivity questions

Are these languages context-sensitive?

Lang(n) ⊆ CSens for n ≥ 4 – open problem

This inclusion is “almost obvious”:
• Recall that CSens = languages recognized by a nondeterministic Turing machine in linear space.
Expressivity questions

\[
\text{Safe\text{Lang}(0)} \not\subseteq \text{Safe\text{Lang}(1)} \not\subseteq \text{Safe\text{Lang}(2)} \not\subseteq \text{Safe\text{Lang}(3)} \not\subseteq \ldots \subseteq \text{CSens}
\]

\[
\text{Lang(0)} \not\subseteq \text{Lang(1)} \not\subseteq \text{Lang(2)} \not\subseteq \text{Lang(3)} \not\subseteq \ldots \subseteq \text{CSens}
\]

Are these languages context-sensitive?

\[
\text{Lang}(n) \not\subseteq \text{CSens} \text{ for } n \geq 4 \quad \text{– open problem}
\]

This inclusion is “almost obvious”:

- Recall that CSens = languages recognized by a nondeterministic Turing machine in linear space.
- Consider the following algorithm: starting from the initial nonterminal, follow nondeterministically rules of the HORS, trying to derive the input word.
Expressivity questions

Are these languages context-sensitive?

\[
\text{SafeLang}(0) \not\subseteq \text{SafeLang}(1) \not\subseteq \text{SafeLang}(2) \not\subseteq \text{SafeLang}(3) \not\subseteq \ldots \subseteq \text{CSens}
\]

\[
\text{Lang}(0) \not\subseteq \text{Lang}(1) \not\subseteq \text{Lang}(2) \not\subseteq \text{Lang}(3) \not\subseteq \ldots \subseteq \text{CSens}
\]

Are these languages context-sensitive?

\[\text{Lang}(n) \not\subseteq \text{CSens} \text{ for } n \geq 4 \text{ – open problem}\]

This inclusion is “almost obvious”:

- Recall that \(\text{CSens} = \) languages recognized by a nondeterministic Turing machine in linear space.
- Consider the following algorithm: starting from the initial nonterminal, follow nondeterministically rules of the HORS, trying to derive the input word.
- It works well if all intermediate terms are smaller than the derived word (= input word).
- The “only difficulty”: describe/eliminate nonterminals that are “not productive”, i.e., that do not increase the size of the derived word.
<table>
<thead>
<tr>
<th>Algorithmic questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem: MSO model-checking</td>
</tr>
<tr>
<td>Input: MSO formula $\phi$, HORS $S$</td>
</tr>
<tr>
<td>Output: does $\phi$ hold in the tree generated by $S$?</td>
</tr>
</tbody>
</table>
**Algorithmic questions**

<table>
<thead>
<tr>
<th>Problem: MSO model-checking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input: MSO formula $\phi$, HORS $S$</td>
</tr>
<tr>
<td>Output: does $\phi$ hold in the tree generated by $S$?</td>
</tr>
</tbody>
</table>

**Theorem:** MSO model-checking is decidable.

- [Knapik, Niwiński, Urzyczyn 2002] – safe schemes only
- [Knapik, Niwiński, Urzyczyn, Walukiewicz 2005] – order-2 only
- [Ong 2006] – via game semantics
- [Hague, Murawski, Ong, Serre 2008] – via collapsible pushdown automata
- [Broadbent, Ong 2009] – global model-checking
- [Kobayashi, Ong 2009] – via a type system
- [Broadbent, Carayol, Ong, Serre 2010] – MSO reflection
- [Salvati, Walukiewicz 2011] – via Krivine machine
- [Carayol, Serre 2012] – MSO selection
- [Salvati, Walukiewicz 2015] – model for $\lambda Y$-calculus

...
Algorithmic questions

Problem: MSO model-checking
Input: MSO formula $\phi$, HORS $S$
Output: does $\phi$ hold in the tree generated by $S$?

Theorem: MSO model-checking is decidable.

Complexity:
- nonelementary when $\phi \in \text{MSO}$
Algorithmic questions

Problem: MSO model-checking
Input: MSO formula $\phi$, HORS $S$
Output: does $\phi$ hold in the tree generated by $S$?

**Theorem:** MSO model-checking is decidable.

**Complexity:**
- nonelementary when $\phi \in$ MSO
- $n$-EXPTIME-complete when $\phi$ is given as a $\mu$-calculus formula or a parity automaton, and the scheme is of order $n$
Algorithmic questions

Problem: MSO model-checking
Input: MSO formula $\phi$, HORS $S$
Output: does $\phi$ hold in the tree generated by $S$?

Theorem: MSO model-checking is decidable.

Complexity:
- nonelementary when $\phi \in \text{MSO}$
- $n$-EXPTIME-complete when $\phi$ is given as a $\mu$-calculus formula or a parity automaton, and the scheme is of order $n$
- $(n-1)$-EXPTIME-complete for reachability properties (is “a” present in the tree)
Algorithmic questions

Problem: MSO model-checking
Input: MSO formula $\phi$, HORS $S$
Output: does $\phi$ hold in the tree generated by $S$?

Theorem: MSO model-checking is decidable.

Complexity:
- nonelementary when $\phi \in$ MSO
- $n$-EXPTIME-complete when $\phi$ is given as a $\mu$-calculus formula or a parity automaton, and the scheme is of order $n$
- $(n-1)$-EXPTIME-complete for reachability properties (is "a" present in the tree)
- polynomial when $n$, $\phi$, and maximal arity of a nonterminal fixed
Algorithmic questions

Problem: MSO model-checking
Input: MSO formula $\phi$, HORS $S$
Output: does $\phi$ hold in the tree generated by $S$?

Theorem: MSO model-checking is decidable.

Complexity:
- nonelementary when $\phi \in$ MSO
- $n$-EXPTIME-complete when $\phi$ is given as a $\mu$-calculus formula or a parity automaton, and the scheme is of order $n$
- $(n-1)$-EXPTIME-complete for reachability properties (is “a” present in the tree)
- polynomial when $n$, $\phi$, and maximal arity of a nonterminal fixed
- despite high complexity, solvable in practice (see next talk)
Algorithmic questions

Theorem: MSO model-checking is decidable.

Idea of a proof

Input: alternating parity automaton $A$, HORS $S$
Question: does $A$ accept the tree generated by $S$?

We refine simple types into intersection types of the form:

- $o$ is refined to $q \in Q$ (a state)
- $\alpha \rightarrow \beta$ is refined to $\{(\tau_1, m_1), \ldots, (\tau_k, m_k)\} \rightarrow \tau$ where $\tau_i$ refines $\alpha$, $\tau$ refines $\beta$, $m_i$ is a priority

Intuition: a function with type $\{(q_1, m_1), (q_2, m_2)\} \rightarrow q$ (refining $o \rightarrow o$):

the smallest priority on these paths is $m_1$
the smallest priority on this path is $m_2$
Algorithmic questions

**Theorem:** MSO model-checking is decidable.

[Knapik, Niwiński, Urzyczyn 2002] – safe schemes only
[Knapik, Niwiński, Urzyczyn, Walukiewicz 2005] – order-2 only
[Ong 2006] – via game semantics
[Hague, Murawski, Ong, Serre 2008] – via collapsible pushdown automata
[Broadbent, Ong 2009] – global model-checking
[Kobayashi, Ong 2009] – via a type system
[Broadbent, Carayol, Ong, Serre 2010] – MSO reflection
[Salvati, Walukiewicz 2011] – via Krivine machine
[Carayol, Serre 2012] – MSO selection
[Salvati, Walukiewicz 2015] – model for \( \lambda Y \)-calculus

---

**MSO reflection**

Input: MSO formula \( \phi(x) \), HORS \( S \)
Output: HORS \( S' \) generating the same tree as \( S \), where the nodes \( x \) in which \( \phi(x) \) holds are marked.
Algorithmic questions

Theorem: MSO model-checking is decidable.

[Knapik, Niwiński, Urzyczyn 2002] – safe schemes only
[Knapik, Niwiński, Urzyczyn, Walukiewicz 2005] – order-2 only
[Ong 2006] – via game semantics
[Hague, Murawski, Ong, Serre 2008] – via collapsible pushdown automata
[Broadbent, Ong 2009] – global model-checking
[Kobayashi, Ong 2009] – via a type system
[Broadbent, Carayol, Ong, Serre 2010] – MSO reflection
[Salvati, Walukiewicz 2011] – via Krivine machine
[Carayol, Serre 2012] – MSO selection
[Salvati, Walukiewicz 2015] – model for λY-calculus

MSO reflection
Input: MSO formula $\phi(x)$, HORS $S$
Output: HORS $S'$ generating the same tree as $S$, where the nodes $x$ in which $\phi(x)$ holds are marked.

MSO selection
Input: MSO formula $\phi(X)$, HORS $S$
Output: HORS $S'$ generating the same tree as $S$ with some nodes marked so that $\phi(X)$ holds for the set $X$ of marked nodes.
Algorithmic questions

Theorem: MSO model-checking is decidable.

[Knapik, Niwiński, Urzyczyn 2002] – safe schemes only
[Knapik, Niwiński, Urzyczyn, Walukiewicz 2005] – order-2 only
[Ong 2006] – via game semantics
[Hague, Murawski, Ong, Serre 2008] – via collapsible pushdown automata
[Broadbent, Ong 2009] – global model-checking
[Kobayashi, Ong 2009] – via a type system
[Broadbent, Carayol, Ong, Serre 2010] – MSO reflection
[Salvati, Walukiewicz 2011] – via Krivine machine
[Carayol, Serre 2012] – MSO selection
[Salvati, Walukiewicz 2015] – model for λY-calculus

...
What is a model?

Input: MSO formula $\phi$, HORS $S$

Output: a finite set $D_\alpha$ for every sort $\alpha$
What is a model?

Input: MSO formula $\phi$, HORS $S$

Output: a finite set $D_\alpha$ for every sort $\alpha$,
a value $[M] \in D_\alpha$ for every term $M$ sort $\alpha$
What is a model?

Input: MSO formula $\phi$, HORS $S$

Output: a finite set $D_\alpha$ for every sort $\alpha$,

a value $[M] \in D_\alpha$ for every term $M$ sort $\alpha$

has to be compositional: $[MN]$ determined by $[M]$ and $[N]$.

$[M]$ depends on valuation of free variables of $M$
What is a model?

Input: MSO formula \( \phi \), HORS \( S \)

Output: a finite set \( D_\alpha \) for every sort \( \alpha \),
a value \([M] \in D_\alpha\) for every term \( M \) sort \( \alpha \)
has to be compositional: \([MN]\) determined by \([M]\) and \([N]\).

([M] depends on valuation of free variables of \( M \))

(model \( \Rightarrow \) reflection)

(we enrich the scheme so that every term “knows” its value in the model)
**What is a model?**

Input: MSO formula $\phi$, HORS $S$

Output: a finite set $D_\alpha$ for every sort $\alpha$, a value $[M] \in D_\alpha$ for every term $M$ sort $\alpha$

has to be compositional: $[MN]$ determined by $[M]$ and $[N]$.

$[M]$ depends on valuation of free variables of $M$

---

(model $\iff$ reflection $\iff$ transfer theorem)

(we enrich the scheme so that every term “knows” its value in the model)
What is a model?

Input: MSO formula $\phi$, HORS $S$
Output: a finite set $D_\alpha$ for every sort $\alpha$,

a value $[M] \in D_\alpha$ for every term $M$ sort $\alpha$

has to be compositional: $[MN]$ determined by $[M]$ and $[N]$.

$[M]$ depends on valuation of free variables of $M$

model $\Rightarrow$ reflection

$\Leftrightarrow$

transfer theorem

Basing on $\phi$ one can construct $\phi'$ such that $\phi'$ holds in a closed term $M$ of sort $o$ iff $\phi$ holds in the tree generated from $M$.

(term $M$)  \rightarrow \leftarrow (tree generated from $M$)

$\phi' \leftarrow \phi$

(special case: $M =$ starting nonterminal)
Beyond MSO?

Problem: WMSO+U model-checking
Input: WMSO+U formula $\phi$, HORS $S$
Output: does $\phi$ hold in the tree generated by $S$?

*Ongoing work:* WMSO+U model-checking is decidable.

MSO+U = Weak MSO (set quantifiers range over finite sets only)
+ new quantifier U

where: $UX.\phi$ means that $\phi$ holds for some arbitrarily large finite sets $X$
**Downward closure**

Let $L$ be a set of words. Its downward closure $L\downarrow$ contains all words that can be obtained from words in $L$ by removing some letters.

E.g. $L = \{abc\}$, $L\downarrow = \{e,a,b,c,ab,bc,ac,abc\}$
Downward closure
Let $L$ be a set of words. Its downward closure $L\downarrow$ contains all words that can be obtained from words in $L$ by removing some letters.

E.g. $L=\{abc\}$, $L\downarrow=\{e,a,b,c,ab,bc,ac,abc\}$

Higman's lemma: the downward closure of any set $L$ is a regular language.
**Downward closure**

Let $L$ be a set of words. Its downward closure $L\downarrow$ contains all words that can be obtained from words in $L$ by removing some letters.

E.g. $L=\{abc\}$, $L\downarrow=\{e,a,b,c,ab,bc,ac,abc\}$

**Higman's lemma:** the downward closure of any set $L$ is a regular language.

Quest: Given a scheme $S$ recognizing $L$, compute $L\downarrow$. 
**Downward closure**

Let $L$ be a set of words. Its downward closure $L\downarrow$ contains all words that can be obtained from words in $L$ by removing some letters.

E.g. $L=\{abc\}$, $L\downarrow=\{e,a,b,c,ab,bc,ac,abc\}$

**Higman's lemma:** the downward closure of any set $L$ is a regular language.

**Quest:** Given a scheme $S$ recognizing $L$, compute $L\downarrow$.

- Trivial but useless: Compute a scheme $S'$ recognizing $L\downarrow$. 
**Downward closure**

Let $L$ be a set of words. Its downward closure $L\downarrow$ contains all words that can be obtained from words in $L$ by removing some letters.

E.g. $L=\{abc\}$, $L\downarrow=\{e,a,b,c,ab,bc,ac,abc\}$

**Higman's lemma:** the downward closure of any set $L$ is a regular language.

Quest: Given a scheme $S$ recognizing $L$, compute $L\downarrow$.
- Trivial but useless: Compute a scheme $S'$ recognizing $L\downarrow$.
- Real quest: Compute an NFA $A$ recognizing $L\downarrow$. 
**Downward closure**

Let $L$ be a set of words. Its downward closure $L\downarrow$ contains all words that can be obtained from words in $L$ by removing some letters.

E.g. $L=\{abc\}$, $L\downarrow=\{e, a, b, c, ab, bc, ac, abc\}$

**Higman's lemma:** the downward closure of any set $L$ is a regular language.

**Quest:** Given a scheme $S$ recognizing $L$, compute $L\downarrow$.
- Trivial but useless: Compute a scheme $S'$ recognizing $L\downarrow$.
- Real quest: Compute an NFA $A$ recognizing $L\downarrow$.

**Theorem** [Zetzsche 2015, Hague, Kochems, Ong 2016, Clemente, P., Salvati, Walukiewicz 2016]

Given a scheme $S$ recognizing $L$, one can compute an NFA $A$ recognizing $L\downarrow$. 
**Downward closure**

**Theorem**
Given a scheme $S$ recognizing $L$, one can compute an NFA $A$ recognizing $L^\downarrow$.

Some ideas:
- For every regular language $K$ we check whether $L^\downarrow = K$. 
**Downward closure**

**Theorem**
Given a scheme $S$ recognizing $L$, one can compute an NFA $A$ recognizing $L\downarrow$.

Some ideas:
- For every regular language $K$ we check whether $L\downarrow = K$.
- Easy to test whether $L\downarrow \subseteq K$, i.e. $L\downarrow \cap \overline{K} = \emptyset$. 
**Downward closure**

**Theorem**

Given a scheme \( S \) recognizing \( L \), one can compute an NFA \( A \) recognizing \( L\downarrow \).

Some ideas:

- For every regular language \( K \) we check whether \( L\downarrow = K \).
- Easy to test whether \( L\downarrow \subseteq K \), i.e. \( L\downarrow \cap \overline{K} = \emptyset \).
- \( L\downarrow \) (so \( K \) as well) is necessarily a finite union of languages of the form \( S_i = A_0^*a_1^*A_1^*a_2^*...A_{k-1}^*a_k^*A_k^* \). It remains to check whether \( S_i \subseteq L\downarrow \) for all \( i \).
**Downward closure**

**Theorem**
Given a scheme $S$ recognizing $L$, one can compute an NFA $A$ recognizing $L\downarrow$.

Some ideas:
- For every regular language $K$ we check whether $L\downarrow=K$.
- Easy to test whether $L\downarrow\subseteq K$, i.e. $L\downarrow\cap \overline{K}=\emptyset$.
- $L\downarrow$ (so $K$ as well) is necessarily a finite union of languages of the form $S_i=A_0^*a_1^*A_1^*a_2^*\ldots A_{k-1}^*a_k^*A_k^*$. It remains to check whether $S_i\subseteq L\downarrow$ for all $i$.
- By transforming the scheme, this reduces to the **diagonal problem**:

Input: a scheme $S$ recognizing $L\subseteq a_1^*a_2^*\ldots a_k^*$ (with different letters)

Question: does $L\downarrow=a_1^*a_2^*\ldots a_k^*$?

(in other words: is it the case that for every $n$ we have in $L$ words with more than $n$ appearances of every letter?)

*This is the actual problem to be solved.*
The diagonal problem

Input: a scheme $S$ recognizing $L \subseteq a_1^*a_2^*...a_k^*$ (with different letters)

Question: does $L \downarrow = a_1^*a_2^*...a_k^*$?

How to solve it?

a scheme $S$ of order $n$ with \underline{step 1} \rightarrow a scheme $S$ of order $n-1$ with

a word written on a branch this word written in leaves
The diagonal problem

Input: a scheme $S$ recognizing $L \subseteq a_1^*a_2^*...a_k^*$ (with different letters)

Question: does $L \updownarrow = a_1^*a_2^*...a_k^*$?

How to solve it?

- a scheme $S$ of order $n$ with a word written on a branch

  step 1

- a scheme $S$ of order $n-1$ with this word written in leaves

Example:

$S \rightarrow A \ e$

$A \ x \rightarrow a \ (A \ (b \ x))$

$A \ x \rightarrow x$

(rank 1: $a, b$; rank 0: $e$

$S \rightarrow \land \ A \ e$

$A \rightarrow \land \ a \ (\land \ A \ b)$

$A \rightarrow \bullet$

(rank 2: $\land$; rank 0: $a, b, e, \bullet$)
The diagonal problem

Input: a scheme $S$ recognizing $L \subseteq a_1^* a_2^* \cdots a_k^*$ (with different letters)

Question: does $L \downarrow = a_1^* a_2^* \cdots a_k^*$?

How to solve it?

a scheme $S$ of order $n$ with a word written on a branch → a scheme $S$ of order $n-1$ with this word written in leaves

Example:

- $S \rightarrow A \ e$
- $A \ x \rightarrow a \ (A \ (b \ x))$
- $A \ x \rightarrow x$

(rank 1: $a, b$; rank 0: $e$)

- $S \rightarrow \land \ A \ e$
- $A \rightarrow \land \ a \ (\land \ A \ b))$
- $A \rightarrow \bullet$

(rank 2: $\land$; rank 0: $a, b, e, \bullet$)

Idea: 1) Observe that an argument of type $o$ can be used at most once.
The diagonal problem

Input: a scheme $S$ recognizing $L \subseteq a_1^*a_2^*...a_k^*$ (with different letters)
Question: does $L \downarrow = a_1^*a_2^*...a_k^*$?

How to solve it?

\[ \text{a scheme } S \text{ of order } n \text{ with } \quad \text{step 1} \quad \text{a scheme } S \text{ of order } n-1 \text{ with this word written in leaves} \]

Example:

\[
\begin{align*}
S & \rightarrow A \ e \\
A \ x & \rightarrow a \ (A \ (b \ x)) \\
A \ x & \rightarrow x
\end{align*}
\]

(rank 1: $a, b$; rank 0: $e$)

\[
\begin{align*}
S & \rightarrow \land \ A \ e \\
A & \rightarrow \land \ a \ (\land \ A \ b)) \\
A & \rightarrow \bullet
\end{align*}
\]

(rank 2: $\land$; rank 0: $a, b, e, \bullet$)

Idea: 1) Observe that an argument of type $o$ can be used at most once.
2) All arguments of type $o$ are dropped ($\Rightarrow$ order decreases).
3) Every subterm $MN$ with $N$ of type $o$ can be replaced
   a) either by $\land MN$ (when the argument is used in $M$),
   b) or by $M$ (when the argument is ignored in $M$).
The diagonal problem

Input: a scheme $S$ recognizing $L \subseteq a_1^*a_2^*...a_k^*$ (with different letters)

Question: does $L \downarrow = a_1^*a_2^*...a_k^*$?

How to solve it?

1. a scheme $S$ of order $n$ with a word written on a branch
2. step 1
3. a scheme $S$ of order $n-1$ with this word written in leaves

Example:

\[ S \rightarrow A e \]
\[ A x \rightarrow a (A (b x)) \]
\[ A x \rightarrow x \]

(rank 1: $a, b$; rank 0: $e$)

\[ S \rightarrow \land A e \]
\[ A \rightarrow \land a (\land A b) \]
\[ A \rightarrow \bullet \]

(rank 2: $\land$; rank 0: $a, b, e, \bullet$)

Idea: 1) Observe that an argument of type $o$ can be used at most once.
2) All arguments of type $o$ are dropped ($\Rightarrow$ order decreases).
3) Every subterm $M N$ with $N$ of type $o$ can be replaced
   a) either by $\land M N$ (when the argument is used in $M$),
   b) or by $M$ (when the argument is ignored in $M$).
4) Additional work is required to choose correctly a) or b).
The diagonal problem
Input: a scheme $S$ recognizing $L \subseteq a_1^*a_2^*...a_k^*$ (with different letters)
Question: does $L \downarrow = a_1^*a_2^*...a_k^*$?

How to solve it?

1. a scheme $S$ of order $n$ with a word written on a branch
2. a scheme $S$ of order $n-1$ with this word written in leaves

a scheme $S$ of order $n-1$ with a similar word written on a branch
The diagonal problem

Input: a scheme $S$ recognizing $L \subseteq a_1^* a_2^* ... a_k^*$ (with different letters)

Question: does $L \downarrow = a_1^* a_2^* ... a_k^*$?

How to solve it?

a scheme $S$ of order $n$ with a word written on a branch

step 1

a scheme $S$ of order $n-1$ with this word written in leaves

a scheme $S$ of order $n-1$ with a similar word written on a branch

step 2

Example:

Idea:
1) Choose (nondeterministically) only one branch.
2) For every removed subtree with $a$, write a new $a$ just above.
The diagonal problem

Input: a scheme $S$ recognizing $L \subseteq a_1^*a_2^*...a_k^*$ (with different letters)

Question: does $L\downarrow = a_1^*a_2^*...a_k^*$?

How to solve it?

1. a scheme $S$ of order $n$ with a word written on a branch → step 1 → a scheme $S$ of order $n-1$ with this word written in leaves
2. a scheme $S$ of order $n-1$ with a similar word written on a branch

Example:

Idea:
1) Choose (nondeterministically) only one branch.
2) For every removed subtree with $a$, write a new $a$ just above.
3) The number of $a$'s decreases at most logarithmically, if the branch is chosen correctly (always go to the subtree with more $a$'s).
The diagonal problem

Input: a scheme $S$ recognizing $L \subseteq a_1^* a_2^* ... a_k^*$ (with different letters)

Question: does $L \downarrow = a_1^* a_2^* ... a_k^*$?

How to solve it?

1. a scheme $S$ of order $n$ with a word written on a branch
2. a scheme $S$ of order $n-1$ with this word written in leaves
3. a scheme $S$ of order $n-1$ with a similar “word” written on $|\Sigma|$ branches

Example:

Idea:
1) Choose (nondeterministically) only one branch.
2) For every removed subtree with $a$, write a new $a$ just above.
3) The number of $a$'s decreases at most logarithmically, if the branch is chosen correctly (always go to the subtree with more $a$'s).

We have to this for every letter $\Rightarrow |\Sigma|$ branches
The diagonal problem

Input: a scheme $S$ recognizing $L \subseteq a_1^* a_2^* ... a_k^*$ (with different letters)

Question: does $L \downarrow = a_1^* a_2^* ... a_k^*$?

How to solve it?

a scheme $S$ of order $n$ with a “word” written on $|\Sigma|$ branches

step 1

a scheme $S$ of order $n-1$ with this “word” written in leaves

step 2

a scheme $S$ of order $n-1$ with a similar “word” written on $|\Sigma|$ branches

Example:

Idea:
1) Choose (nondeterministically) only one branch.
2) For every removed subtree with $a$, write a new $a$ just above.
3) The number of $a$’s decreases at most logarithmically, if the branch is chosen correctly (always go to the subtree with more $a$'s).
We have to this for every letter $\Rightarrow |\Sigma|$ branches
The diagonal problem
Input: a scheme \( S \) recognizing \( L \subseteq a_1^*a_2^*...a_k^* \) (with different letters)
Question: does \( L \downarrow = a_1^*a_2^*...a_k^* \)?

How to solve it?
a scheme \( S \) of order \( n \) with
a “word” written on \(|\Sigma|\) branches

step 1

a scheme \( S \) of order \( n-1 \) with
this “word” written in leaves

step 2

a scheme \( S \) of order \( n-1 \) with
a similar “word” written on \(|\Sigma|\) branches

Repeat these steps until the order drops down to 0,
and solve the diagonal problem for a regular language.
**Downward closure**

**Theorem**
Given a scheme $S$ recognizing $L$, one can compute an NFA $A$ recognizing $L\downarrow$.

**Motivation?**
0. It gives a simple abstraction of the language recognized by a scheme.
**Downward closure**

**Theorem**
Given a scheme $S$ recognizing $L$, one can compute an NFA $A$ recognizing $L\downarrow$.

**Motivation?**
0. It gives a simple abstraction of the language recognized by a scheme.
1. It is undecidable whether $L=A^*$, $L_1=L_2$, etc.
   But we can check this approximately, by checking whether $L\downarrow=A^*$, $L_1\downarrow=L_2\downarrow$, etc.
**Downward closure**

**Theorem**
Given a scheme $S$ recognizing $L$, one can compute an NFA $A$ recognizing $L\downarrow$.

**Motivation?**
0. It gives a simple abstraction of the language recognized by a scheme.
1. It is undecidable whether $L=A^*$, $L_1=L_2$, etc.
   But we can check this approximately, by checking whether $L\downarrow=A^*$, $L_1\downarrow=L_2\downarrow$, etc.
2. The problem “is there a piecewise testable language (i.e., boolean combination of downward closed languages) containing $L_1$ and not intersecting with $L_2$” reduces to the diagonal problem [Czerwiński, Martens, van Rooijen, Zeitoun 2015]. This gives a more refined approximation for disjointness of $L_1$ and $L_2$ than the test $L_1\downarrow\cap L_2\downarrow=\emptyset$. 
**Theorem**

Given a scheme $S$ recognizing $L$, one can compute an NFA $A$ recognizing $L\downarrow$.

**Motivation?**

0. It gives a simple abstraction of the language recognized by a scheme.

1. It is undecidable whether $L=\overline{A}$, $L_1=\overline{L_2}$, etc.
   
   But we can check this approximately, by checking whether $L\downarrow=\overline{A}$, $L_1\downarrow=\overline{L_2}$, etc.

2. The problem “is there a piecewise testable language (i.e., boolean combination of downward closed languages) containing $L_1$ and not intersecting with $L_2$” reduces to the diagonal problem [Czerwiński, Martens, van Rooijen, Zeitoun 2015]. This gives a more refined approximation for disjointness of $L_1$ and $L_2$ than the test $L_1\downarrow\cap L_2\downarrow=\emptyset$.

3. Consider a system with one leader and some (unspecified) number of contributors, that communicate via common register (read or write, without any locks). The reachability problem in such system reduces to computation of the downward closure [La Torre, Muscholl, Walukiewicz 2015]. *(Yesterday’s talk – downward closure no longer needed)*

---

**Downward closure**
**Downward closure**

**Theorem**
Given a scheme $S$ recognizing $L$, one can compute an NFA $A$ recognizing $L\downarrow$.

**Complexity?**
**Downward closure**

**Theorem**
Given a scheme $S$ recognizing $L$, one can compute an NFA $A$ recognizing $L_{\downarrow}$.

**Complexity?**
- The diagonal problem $\Rightarrow$ probably $(n-1)$-EXPTIME for schemes of order $n$ (ongoing work)
**Downward closure**

**Theorem**
Given a scheme $S$ recognizing $L$, one can compute an NFA $A$ recognizing $L\downarrow$.

**Complexity?**
- The diagonal problem $\Rightarrow$ probably $(n-1)$-EXPTIME for schemes of order $n$ (ongoing work)
- Computation of downward closure $\Rightarrow$ open problem
  - We need to bound the maximal size of the downward closure (a pumping lemma is needed).
- Lower bound: checking whether $L_1\downarrow = L_2\downarrow$ or $L_1\downarrow \subseteq L_2\downarrow$ is co-$n$-NEXPTIME-hard [Zetzsche 2016]
**Downward closure**

**Theorem**
Given a scheme $S$ recognizing $L$, one can compute an NFA $A$ recognizing $L \downarrow$.

**Complexity?**
- The diagonal problem $\Rightarrow$ probably $(n-1)$-EXPTIME for schemes of order $n$ (ongoing work)
- Computation of downward closure $\Rightarrow$ open problem
  We need to bound the maximal size of the downward closure (a pumping lemma is needed).
- Lower bound: checking whether $L_1 \downarrow = L_2 \downarrow$ or $L_1 \downarrow \subseteq L_2 \downarrow$ is co-$n$-NEXPTIME-hard [Zetzsche 2016]

**Another open problem**: computation of downward closure for schemes recognizing languages of trees.
(By Kruskal's tree theorem the downward closure of any language of trees is a regular language.)
Thank you!