ALMOST EVERY SIMPLY TYPED LAMBDA TERM HAS A LONG BETA-REDUCTION SEQUENCE

RYOMA SIN'YA
KAZUYUKI ASADA
NAOKI KOBAYASHI
TAKESHI TSUKADA
(THE UNIVERSITY OF TOKYO)

MOTIVATION

- A simply-typed term can have a very long **β**-reduction sequence.
  - *k*-EXP in the size of terms of order *k* [Beckmann 2001].

$$0\text{-EXP}(n) = n$$
$$(m+1)\text{-EXP}(n) = 2^{m\text{-EXP}(\mathrm{n})}$$

- How many terms have such long **β**-reduction sequences?

MOTIVATION

- A simply-typed term can have a very long **β**-reduction sequence.
  - *k*-EXP in the size of terms of order *k* [Beckmann 2001].

e.g. $$(Twice)^n \underbrace{Twice \cdots Twice}_{k-2 \text{ times}}(\lambda x.bxx)((\lambda x.x)c)$$

$$\text{where } Twice = \lambda f.\lambda x.f(f\,x)$$

- How many terms have such long **β**-reduction sequences?

SIDE REMARK

- The work has been motivated by quantitative analysis of the complexity of higher-order model checking (HOMC).

HIGHER-ORDER MODEL CHECKING [Ong 2006]

- Input : tree automaton $\mathcal{A}$ and **λ**Y-term *t*.
  Output : **YES** if $\mathcal{A}$ accepts the infinite tree represented by *t*, **NO** otherwise.
  Complexity: *k*-EXPTIME-complete for order-*k* **λ**Y-terms.

- We want to (*dis*)prove: HOMC can be efficiently solved for ***almost every input***.

RELATED WORK

- Quantitative analysis of untyped terms:
  - Almost every **λ**-term is strongly normalizing (**SN**), but almost every SK-combinatory term is not **SN** [David *et al.* 2009].
  - Almost every de Bruijn **λ**-term is not **SN** [Bendkowski *et al.* 2015].
  - Empirical results: almost every **λ**-term is not **β**-normal, untypable [Grygiel-Lescanne 2013].

RELATED WORK

- Quantitative analysis of untyped terms:
  - Almost every **λ**-term is strongly normalizing (**SN**), but almost every SK-combinatory term is not **SN** [David *et al.* 2009].
  - Almost every de Bruijn **λ**-term is not **SN** [Bendkowski *et al.* 2015].
  - Empirical results: almost every **λ**-term is not **β**-normal, untypable [Grygiel-Lescanne 2013].
- Quantitative analysis of typed terms: little is known.

OUTLINE

- Introduction
- Our result
- Proof of our result
- Conclusion

OUR RESULT

For $k, \iota, \xi \geq 2$ *and* $k \leq \iota$,

$$\lim_{n\to\infty} \frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k,\iota,\xi)\} \mid \beta(t) \geq (k-2)\text{-EXP}(n)\}}{\#\Lambda_n^\alpha(k,\iota,\xi)} = 1.$$

$\Lambda_n^\alpha(k,\iota,\xi)$: the set of **α**-equivalence classes of size-*n* terms such that:

(1) the *order* is at most k.

(2) the *number of arguments* (*internal arity*) is at most ι.

(3) the number of *distinct variables* is at most ξ.

OUR RESULT

the maximum length of **β**-reduction sequences of *t*.

For $k, \iota, \xi \geq 2$ *and* $k \leq \iota$,

$$\lim_{n\to\infty} \frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k,\iota,\xi)\} \mid \beta(t) \geq (k-2)\text{-EXP}(n)\}}{\#\Lambda_n^\alpha(k,\iota,\xi)} = 1.$$

$\Lambda_n^\alpha(k,\iota,\xi)$: the set of **α**-equivalence classes of size-*n* terms such that:

(1) the *order* is at most k.

(2) the *number of arguments* (*internal arity*) is at most ι.

(3) the number of *distinct variables* is at most ξ.

OUR RESULT

the maximum length of **β**-reduction sequences of *t*.

For $k, \iota, \xi \geq 2$ *and* $k \leq \iota$,

$$\lim_{n\to\infty} \frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k,\iota,\xi)\} \mid \beta(t) \geq (k-2)\text{-EXP}(n)\}}{\#\Lambda_n^\alpha(k,\iota,\xi)} = 1.$$

$\Lambda_n^\alpha(k,\iota,\xi)$: the set of **α**-equivalence classes of size-*n* terms such that:

(1) the *order* is at most k.

(2) the *number of arguments* (*internal arity*) is at most ι.

(3) the number of *distinct variables* is at most ξ.

OUR RESULT

the maximum length of **β**-reduction sequences of *t*.

For $k, \iota, \xi \geq 2$ *and* $k \leq \iota$,

$$\lim_{n\to\infty} \frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k,\iota,\xi)\} \mid \beta(t) \geq (k-2)\text{-EXP}(n)\}}{\#\Lambda_n^\alpha(k,\iota,\xi)} = 1.$$

Almost every term of size *n* and order at most *k* has a **β**-reduction sequence of length (*k*-2)-EXP(*n*).

$\Lambda_n^\alpha(k,\iota,\xi)$: the set of **α**-equivalence classes of size-*n* terms such that:

(1) the *order* is at most k.

(2) the *number of arguments* (*internal arity*) is at most ι.

(3) the number of *distinct variables* is at most ξ.

OUR RESULT

For $k, \iota, \xi \geq 2$ *and* $k \leq \iota$,

$$\lim_{n \to \infty} \frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k, \iota, \xi)\} \mid \beta(t) \geq (k-2)\text{-EXP}(n)\}}{\#\Lambda_n^\alpha(k, \iota, \xi)} = 1.$$

$\Lambda_n^\alpha(k, \iota, \xi)$: the set of **α**-equivalence classes of size-*n* terms such that:

(1) the *order* is at most k.

(2) the *number of arguments* (*internal arity*) is at most ι.

(3) the number of *distinct variables* is at most ξ.

NUMBER OF DISTINCT VARIABLES

- $\#\mathbf{V}(t)$: the # of variables in *t* excluding **unused variables**.
- For an **α**-equivalence class $[t]_\alpha$,

$$\#\mathbf{V}_\alpha([t]_\alpha) \triangleq \min\{\#\mathbf{V}(t') \mid t' \in [t]_\alpha\}$$

NUMBER OF DISTINCT VARIABLES

- $\#\mathbf{V}(t)$: the # of variables in *t* **excluding unused variables**.
- For an **α**-equivalence class $[t]_\alpha$,

$$\#\mathbf{V}_\alpha([t]_\alpha) \triangleq \min\{\#\mathbf{V}(t') \mid t' \in [t]_\alpha\}$$

Example

$$\#\mathbf{V}_\alpha([(\lambda z.z)\lambda y.x]_\alpha) = \mathbf{?}$$

NUMBER OF DISTINCT VARIABLES

- $\#\mathbf{V}(t)$: the # of variables in *t* excluding **unused variables**.
- For an **α**-equivalence class $[t]_\alpha$,

$$\#\mathbf{V}_\alpha([t]_\alpha) \triangleq \min\{\#\mathbf{V}(t') \mid t' \in [t]_\alpha\}$$

Example

$$\#\mathbf{V}_\alpha([(\lambda z.z)\lambda y.x]_\alpha) = \mathbf{?}$$

$$\#\mathbf{V}((\lambda z.z)\lambda y.x) = \#\{x, z\} = 2$$

NUMBER OF DISTINCT VARIABLES

- $\#\mathbf{V}(t)$: the # of variables in *t* **excluding unused variables**.
- For an **α**-equivalence class $[t]_\alpha$,

$$\#\mathbf{V}_\alpha([t]_\alpha) \triangleq \min\{\#\mathbf{V}(t') \mid t' \in [t]_\alpha\}$$

<u>Example</u>

$$\#\mathbf{V}_\alpha([(\lambda z.z)\lambda y.x]_\alpha) = \textbf{?}$$

$$\#\mathbf{V}((\underline{\lambda z.z})\lambda y.x) = \#\{x, z\} = 2$$

NUMBER OF DISTINCT VARIABLES

- $\#\mathbf{V}(t)$: the # of variables in *t* excluding **unused variables**.
- For an **α**-equivalence class $[t]_\alpha$,

$$\#\mathbf{V}_\alpha([t]_\alpha) \triangleq \min\{\#\mathbf{V}(t') \mid t' \in [t]_\alpha\}$$

<u>Example</u>

$$\#\mathbf{V}_\alpha([(\lambda z.z)\lambda y.x]_\alpha) = \textbf{?}$$

$$\#\mathbf{V}((\underline{\lambda z.z})\lambda y.x) = \#\{x, z\} = 2$$

$$\#\mathbf{V}((\underline{\lambda x.x})\lambda y.x) = \#\{x\} = 1$$

NUMBER OF DISTINCT VARIABLES

- $\#\mathbf{V}(t)$: the # of variables in *t* excluding **unused variables**.
- For an **α**-equivalence class $[t]_\alpha$,

$$\#\mathbf{V}_\alpha([t]_\alpha) \triangleq \min\{\#\mathbf{V}(t') \mid t' \in [t]_\alpha\}$$

<u>Example</u>

$$\#\mathbf{V}_\alpha([(\lambda z.z)\lambda y.x]_\alpha) = 1$$

$$\#\mathbf{V}((\underline{\lambda z.z})\lambda y.x) = \#\{x, z\} = 2$$

$$\#\mathbf{V}((\underline{\lambda x.x})\lambda y.x) = \#\{x\} = 1$$

OUR RESULT

For $k, \iota, \xi \geq 2$ *and* $k \leq \iota$,

$$\lim_{n \to \infty} \frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k, \iota, \xi)\} \mid \beta(t) \geq (k-2)\text{-EXP}(n)\}}{\#\Lambda_n^\alpha(k, \iota, \xi)} = 1.$$

$\Lambda_n^\alpha(k, \iota, \xi)$: the set of **α**-equivalence classes of size-*n* terms such that:

(1) the *order* is at most k.

(2) the *number of arguments* (*internal arity*) is at most ι.

(3) the number of *distinct variables* is at most ξ.

$$\#\mathbf{V}_\alpha([t]_\alpha) \leq \xi \;\; \textit{for every} \;\; [t]_\alpha \in \Lambda_n^\alpha(k, \iota, \xi)$$

OUTLINE

- Introduction
- Our result
- Proof of our result
- Conclusion

OVERVIEW OF OUR PROOF

- Almost every term contains a certain “context” that has a very long **β**-reduction sequence.

OVERVIEW OF OUR PROOF

- Almost every term contains a certain “context” that has a very long **β**-reduction sequence.

OVERVIEW OF OUR PROOF

- Almost every term contains a certain "context" that has a very long **β**-reduction sequence.

- Inspired by **Infinite Monkey Theorem**: for any word *x*, almost every word contains *x* as a subword.

OUTLINE

- Introduction
- Our result
- Proof of our result
  - Idea
  - Infinite Monkey Theorem
  - Decomposition of terms
  - Sketch of the proof
- Conclusion

PROOF IDEA

1. Parameterizing Infinite Monkey Theorem.

2. Extending (1) to **λ**-terms.

3. Constructing “explosive context” that generates a long **β**-reduction sequence.

INFINITE MONKEY THEOREM

For any word x *over an alphabet A,*

$$\lim_{n\to\infty} \frac{\#\{w \in A^n \mid x \sqsubseteq w\}}{\#A^n} = 1.$$

$x \sqsubseteq w \iff w = uxv$ for some words $u, v \in A^*$.

INFINITE MONKEY THEOREM

For any word x *over an alphabet A,*

$$\lim_{n\to\infty} \frac{\#\{w \in A^n \mid x \sqsubseteq w\}}{\#A^n} = 1.$$

$x \sqsubseteq w \;\Leftrightarrow\; w = uxv$ for some words $u, v \in A^*$.

IDEA1: PARAMETERIZING INFINITE MONKEY THEOREM

For any family of words $(x_n)_n$ *over A such that* $|x_n| = \lceil \log^{(2)}(n) \rceil$,

$$\lim_{n \to \infty} \frac{\#\{w \in A^n \mid x_n \sqsubseteq w\}}{\#A^n} = 1.$$

$\log^{(2)}(n) = \log(\log(n))$

IDEA2: EXTENDING IDEA1 TO TERMS

For any family of contexts $(C_n)_n$ *such that*
$|C_n| = \lceil \log^{(2)}(n) \rceil$,

$$\lim_{n\to\infty} \frac{\#\{[t]_\alpha \in \Lambda^\alpha_n(k,\iota,\xi)\} \mid C_n \preceq t\}}{\#\Lambda^\alpha_n(k,\iota,\xi)} = 1.$$

if $k,\iota,\xi \geq 2$.

$C \preceq t \iff t = C'[C[t']]$ for some context C' and term t'.

IDEA3: CONSTRUCTING "EXPLOSIVE" CONTEXT

- For parameters *n* and *k*, we define the explosive context $^{k}_{n}$ of order-*k* as:

$$\lambda x.\Big((Twice)^n\,\underbrace{Twice\cdots Twice}_{k-2\ \text{times}}\,Dup(Id\,[\,])\Big)$$

$$\text{where}\quad Twice = \lambda f.\lambda x.f(f\,x)$$

$$Dup = \lambda x.(\lambda y.\lambda z.y)xx \quad\text{and}\quad Id = \lambda x.x$$

IDEA3: CONSTRUCTING "EXPLOSIVE" CONTEXT

- For parameters *n* and *k*, we define the explosive context 💣^k_n of order-*k* as:

$$\lambda x.\big((Twice)^n \underbrace{Twice \cdots Twice}_{k-2 \text{ times}} Dup(Id\,[\,])\big)$$

$$\text{where} \quad Twice = \lambda f.\lambda x.f(f\,x)$$

$$Dup = \lambda x.(\lambda y.\lambda z.y)xx \quad \text{and} \quad Id = \lambda x.x$$

- It has the following "explosive property":

$$\beta\left(💣^k_n\right) \geq k\text{-EXP(n)}$$

IDEA3: CONSTRUCTING "EXPLOSIVE" CONTEXT

- For parameters *n* and *k*, we define the explosive context 💣$^{k}_{n}$ of order-*k* as:

$$\lambda x.\big((Twice)^n \underbrace{Twice \cdots Twice}_{k-2 \text{ times}} Dup(Id\,[\,])\big)$$

$$\text{where} \quad Twice = \lambda f.\lambda x.f(f\,x)$$

$$Dup = \lambda x.(\lambda y.\lambda z.y)xx \quad \text{and} \quad Id = \lambda x.x$$

- It has the following "explosive property":

💣$^{k}_{n} \preceq t \Rightarrow k\text{-EXP}(n) \leq \beta(t).$

HARVEST

For $k, \iota, \xi \geq 2$ *and* $k \leq \iota$,

$$\lim_{n\to\infty} \frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k,\iota,\xi)\} \mid \text{💣}^{k}_{\lceil \log^{(2)}(n) \rceil} \preceq t\}}{\#\Lambda_n^\alpha(k,\iota,\xi)} = 1.$$

HARVEST

For $k, \iota, \xi \geq 2$ *and* $k \leq \iota$*,*

$$\lim_{n \to \infty} \frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k, \iota, \xi)\} \mid \text{💣}^{k}_{\lceil \log^{(2)}(n) \rceil} \preceq t\}}{\#\Lambda_n^\alpha(k, \iota, \xi)} = 1.$$

$\Downarrow$ A direct corollary of the explosive property:

$$\text{💣}^{k}_{\lceil \log^{(2)}(n) \rceil} \preceq t \Rightarrow (k-2)\text{-EXP}(n) \leq \beta(t).$$

Almost every term of size *n* and order at most *k* has a **β**-reduction sequence of length *(k-2)-EXP(n)*.

PROOF IDEA

1. Parameterizing Infinite Monkey Theorem.

2. Extending (1) to **λ**-terms.

3. Constructing “explosive context” that generates a long **β**-reduction sequence.

PROOF IDEA

1. Parameterizing Infinite Monkey Theorem.

2. Extending (1) to **λ**-terms.

Most technical part

3. Constructing "explosive context" that generates a long **β**-reduction sequence.

PROOF IDEA

1. Parameterizing Infinite Monkey Theorem.

We first give a proof of (1), because it **clarify** the overall structure of the proof of (2).

2. Extending (1) to **λ**-terms.

Most technical part

3. Constructing "explosive context" that generates a long **β**-reduction sequence.

OUTLINE

- Introduction
- Our result
- Proof of our result
  - Idea
  - Infinite Monkey Theorem
  - Decomposition of terms
  - Sketch of the proof
- Conclusion

http://en.wikipedia.org/wiki/Infinite_monkey_theorem

PROOF OF MONKEY THEOREM FOR WORDS

For any word x *over an alphabet A,*

$$\lim_{n\to\infty} \frac{\#\{w \in A^n \mid x \sqsubseteq w\}}{\#A^n} = 1.$$

$\because$

It suffice to show that:

$$\frac{\#\{w \in A^n \mid x \not\sqsubseteq w\}}{\#A^n} \to 0 \ (n \to \infty)$$

PROOF OF MONKEY THEOREM FOR WORDS

$\because$

$$\frac{\#\{w \in A^n \mid x \not\sqsubseteq w\}}{\#A^n} \overset{?}{\to} 0 \ (n \to \infty)$$

PROOF OF MONKEY THEOREM FOR WORDS

$\because$ Let $\ell = |x|, w \in A^n$.

$$\frac{\#\{w \in A^n \mid x \not\sqsubseteq w\}}{\#A^n} \overset{?}{\to} 0 \ (n \to \infty)$$

PROOF OF MONKEY THEOREM FOR WORDS

$\because$ Let $\ell = |x|, w \in A^n$.

$$w = w_1 \, w_2 \, \cdots \, w_{\lfloor n/\ell \rfloor} \, w'$$

$$\frac{\#\{w \in A^n \mid x \not\sqsubseteq w\}}{\#A^n} \overset{?}{\to} 0 \; (n \to \infty)$$

PROOF OF MONKEY THEOREM FOR WORDS

$\because$ Let $\ell = |x|, w \in A^n$.

$$w = \boxed{w_1} w_2 \cdots w_{\lfloor n/\ell \rfloor} w'$$

$$\underbrace{}_{\ell}$$

$$\frac{\#\{w \in A^n \mid x \not\sqsubseteq w\}}{\#A^n} \overset{?}{\to} 0 \ (n \to \infty)$$

PROOF OF MONKEY THEOREM FOR WORDS

$\because$ Let $\ell = |x|, w \in A^n$.

$$w = \underbrace{w_1}_{\ell}\underbrace{w_2}_{\ell} \cdots w_{\lfloor n/\ell \rfloor} w'$$

$$\frac{\#\{w \in A^n \mid x \not\sqsubseteq w\}}{\#A^n} \overset{?}{\to} 0 \ (n \to \infty)$$

PROOF OF MONKEY THEOREM FOR WORDS

$\because$ Let $\ell = |x|, w \in A^n$.

$$w = \underbrace{w_1}_{\ell}\underbrace{w_2}_{\ell} \cdots \underbrace{w_{\lfloor n/\ell \rfloor}}_{\ell} w'$$

$$\frac{\#\{w \in A^n \mid x \not\sqsubseteq w\}}{\#A^n} \overset{?}{\to} 0 \ (n \to \infty)$$

PROOF OF MONKEY THEOREM FOR WORDS

$\because$ Let $\ell = |x|, w \in A^n$.

$$w = \underbrace{w_1}_{\ell}\underbrace{w_2}_{\ell} \cdots \underbrace{w_{\lfloor n/\ell \rfloor}}_{\ell}\overbrace{w'}^{(n \bmod \ell)<\ell}$$

$$\frac{\#\{w \in A^n \mid x \not\sqsubseteq w\}}{\#A^n} \overset{?}{\to} 0 \ (n \to \infty)$$

PROOF OF MONKEY THEOREM FOR WORDS

$\because$ Let $\ell = |x|, w \in A^n$.

$$w = \underbrace{w_1}_{\ell}\underbrace{w_2}_{\ell} \cdots \underbrace{w_{\lfloor n/\ell \rfloor}}_{\ell}\overbrace{w'}^{(n \bmod \ell) < \ell}$$

$$\frac{\#\{w \in A^n \mid x \not\sqsubseteq w\}}{\#A^n}$$

$$\leq \frac{\#\{w \in A^n \mid w_i \neq x \text{ for all } i \leq \lfloor n/\ell \rfloor\}}{\#A^n}$$

PROOF OF MONKEY THEOREM FOR WORDS

$\because$ Let $\ell = |x|, w \in A^n$.

$$w = \underbrace{w_1}_{\ell}\underbrace{w_2}_{\ell} \cdots \underbrace{w_{\lfloor n/\ell \rfloor}}_{\ell} \overbrace{w'}^{(n \bmod \ell) < \ell}$$

$$\frac{\#\{w \in A^n \mid x \not\sqsubseteq w\}}{\#A^n}$$

$$\leq \frac{\#\{w \in A^n \mid w_i \neq x \text{ for all } i \leq \lfloor n/\ell \rfloor\}}{\#A^n}$$

$$= \left(1 - \frac{1}{\#A^\ell}\right)^{\lfloor n/\ell \rfloor}$$

PROOF OF MONKEY THEOREM FOR WORDS

$\because$ Let $\ell = |x|, w \in A^n$.

$$w = \underbrace{w_1}_{\ell}\underbrace{w_2}_{\ell} \cdots \underbrace{w_{\lfloor n/\ell \rfloor}}_{\ell} \overbrace{w'}^{(n \bmod \ell) < \ell}$$

$$\frac{\#\{w \in A^n \mid x \not\sqsubseteq w\}}{\#A^n}$$

$$\leq \frac{\#\{w \in A^n \mid w_i \neq x \text{ for all } i \leq \lfloor n/\ell \rfloor\}}{\#A^n}$$

$$= \left(1 - \frac{1}{\#A^\ell}\right)^{\lfloor n/\ell \rfloor} \to 0 \ (n \to \infty)$$

$\therefore$

DECOMPOSITION OF WORDS

cf.

$$A^n \ni w = \underbrace{\boxed{w_1}}_{\ell}\underbrace{\boxed{w_2}}_{\ell} \cdots \underbrace{\boxed{w_{\lfloor n/\ell \rfloor}}}_{\ell} \overbrace{\boxed{w'}}^{(n \bmod \ell) < \ell}$$

- Previous proof is based on a **"good"** decomposition of words.
  - This good decomposition is induced by the following coproduct-product form:

$$A^n \cong \coprod_{w' \in A^{(n \bmod \ell)}} \prod_{i \leq \lfloor n/\ell \rfloor} A^\ell$$

  - This point of view forms the basis of the later extensions.

DECOMPOSITION OF WORDS

cf.

$$A^n \ni w = \underbrace{w_1}_{\ell}\underbrace{w_2}_{\ell}\cdots\underbrace{w_{\lfloor n/\ell \rfloor}}_{\ell}\overbrace{w'}^{(n \bmod \ell)<\ell}$$

$$\begin{array}{ccccc} w & & w' & \& & (w_1, w_2, \cdots, w_{\lfloor n/\ell \rfloor}) \\ \cap & & \cap & & \cap \\ A^n & \cong & \coprod_{w' \in A^{(n \bmod \ell)}} & & \prod_{i \leq \lfloor n/\ell \rfloor} A^\ell \end{array}$$

DECOMPOSITION OF WORDS

cf.

$$A^n \ni w = \underbrace{w_1}_{\ell}\underbrace{w_2}_{\ell}\cdots\underbrace{w_{\lfloor n/\ell \rfloor}}_{\ell}\overbrace{w'}^{(n \bmod \ell)<\ell}$$

$$\begin{array}{ccccc} w & & w' & \& & (w_1, w_2, \cdots, w_{\lfloor n/\ell \rfloor}) \\ \cap & & \cap & & \cap \\ A^n & \cong & \coprod_{w' \in A^{(n \bmod \ell)}} & & \prod_{i \leq \lfloor n/\ell \rfloor} A^\ell \end{array}$$

Residual part (coproduct part)

DECOMPOSITION OF WORDS

cf.

$$A^n \ni w = \underbrace{w_1}_{\ell}\underbrace{w_2}_{\ell}\cdots\underbrace{w_{\lfloor n/\ell \rfloor}}_{\ell}\overbrace{w'}^{(n \bmod \ell)<\ell}$$

$$\begin{array}{ccccc} w & & w' & \& & (w_1, w_2, \cdots, w_{\lfloor n/\ell \rfloor}) \\ \Cap & & \Cap & & \Cap \\ A^n & \cong & \coprod_{w' \in A^{(n \bmod \ell)}} & & \prod_{i \leq \lfloor n/\ell \rfloor} A^{\ell} \end{array}$$

Residual part (coproduct part)

Decomposed parts (product parts)

PROOF OF INFINITE MONKEY THEOREM (REVISED)

$\because$ Let $\ell = |x|$.

$$\frac{\#\{w \in A^n \mid x \not\sqsubseteq w\}}{\#A^n}$$

cf. $A^n \cong \coprod_{w' \in A^{(n \bmod \ell)}} \prod_{i \leq \lfloor n/\ell \rfloor} A^\ell$

(residual part) **(decomposed parts)**

PROOF OF INFINITE MONKEY THEOREM (REVISED)

$\because$ Let $\ell = |x|$.

$$\frac{\#\{w \in A^n \mid x \not\sqsubseteq w\}}{\#A^n} = \frac{\#\coprod_{w' \in A^{(n \bmod \ell)}} \prod_{i \leq \lfloor n/\ell \rfloor} A^\ell}{\#A^n}$$

cf. $A^n \cong \coprod_{w' \in A^{(n \bmod \ell)}} \prod_{i \leq \lfloor n/\ell \rfloor} A^\ell$

(residual part) (decomposed parts)

PROOF OF INFINITE MONKEY THEOREM (REVISED)

$\because$ Let $\ell = |x|$.

$$\frac{\#\{w \in A^n \mid x \not\sqsubseteq w\}}{\#A^n} = \frac{\# \coprod_{w' \in A^{(n \bmod \ell)}} \prod_{i \leq \lfloor n/\ell \rfloor} A^\ell}{\#A^n}$$

$$\leq \frac{\# \coprod_{w' \in A^{(n \bmod \ell)}} \prod_{i \leq \lfloor n/\ell \rfloor} \left(A^\ell \setminus \{x\}\right)}{\#A^n}$$

cf. $A^n \cong \coprod_{w' \in A^{(n \bmod \ell)}} \prod_{i \leq \lfloor n/\ell \rfloor} A^\ell$

(residual part) (decomposed parts)

PROOF OF INFINITE MONKEY THEOREM (REVISED)

$\because$ Let $\ell = |x|$.

$$\frac{\#\{w \in A^n \mid x \not\trianglelefteq w\}}{\#A^n} = \frac{\#\coprod_{w' \in A^{(n \bmod \ell)}} \prod_{i \le \lfloor n/\ell \rfloor} A^\ell}{\#A^n}$$

$$\le \frac{\#\coprod_{w' \in A^{(n \bmod \ell)}} \prod_{i \le \lfloor n/\ell \rfloor} \left(A^\ell \setminus \{x\}\right)}{\#A^n}$$

PROOF OF INFINITE MONKEY THEOREM (REVISED)

$\because$ Let $\ell = |x|$.

$$\frac{\#\{w \in A^n \mid x \not\sqsubseteq w\}}{\#A^n} = \frac{\#\coprod_{w' \in A^{(n \bmod \ell)}} \prod_{i \leq \lfloor n/\ell \rfloor} A^\ell}{\#A^n}$$

$$\leq \frac{\#\coprod_{w' \in A^{(n \bmod \ell)}} \prod_{i \leq \lfloor n/\ell \rfloor} (A^\ell \setminus \{x\})}{\#A^n}$$

$$= \left(1 - \frac{1}{\#A^\ell}\right)^{\lfloor n/\ell \rfloor}$$

PROOF OF INFINITE MONKEY THEOREM (REVISED)

$\because$ Let $\ell = |x|$.

$$\frac{\#\{w \in A^n \mid x \not\sqsubseteq w\}}{\#A^n} = \frac{\#\coprod_{w' \in A^{(n \bmod \ell)}} \prod_{i \le \lfloor n/\ell \rfloor} A^\ell}{\#A^n}$$

$$\le \frac{\#\coprod_{w' \in A^{(n \bmod \ell)}} \prod_{i \le \lfloor n/\ell \rfloor} (A^\ell \setminus \{x\})}{\#A^n}$$

$$= \left(1 - \frac{1}{\#A^\ell}\right)^{\lfloor n/\ell \rfloor} \to 0 \ (n \to \infty) \qquad \therefore$$

PROOF OF PARAMETERIZED INFINITE MONKEY THEOREM FOR WORDS

For any family of words $(x_n)_n$ *over A such that* $|x_n| = \lceil \log^{(2)}(n) \rceil$,

$$\lim_{n\to\infty} \frac{\#\{w \in A^n \mid x_n \sqsubseteq w\}}{\#A^n} = 1.$$

$$\because \frac{\#\{w \in A^n \mid x_n \not\sqsubseteq w\}}{\#A^n}$$

$$\leq \frac{\#\{w \in A^n \mid \text{every decomposed part} \neq x\}}{\#A^n}$$

cf. $$A^n \cong \coprod_{w' \in A^{(n \bmod \lceil \log^{(2)}(n) \rceil)}} \prod_{i \leq \lfloor n / \lceil \log^{(2)}(n) \rceil \rfloor} A^{\lceil \log^{(2)}(n) \rceil}$$

(residual part) **(decomposed parts)**

PROOF OF PARAMETERIZED INFINITE MONKEY THEOREM FOR WORDS

For any family of words $(x_n)_n$ *over A such that*
$|x_n| = \lceil \log^{(2)}(n) \rceil$,

$$\lim_{n\to\infty} \frac{\#\{w \in A^n \mid x_n \sqsubseteq w\}}{\#A^n} = 1.$$

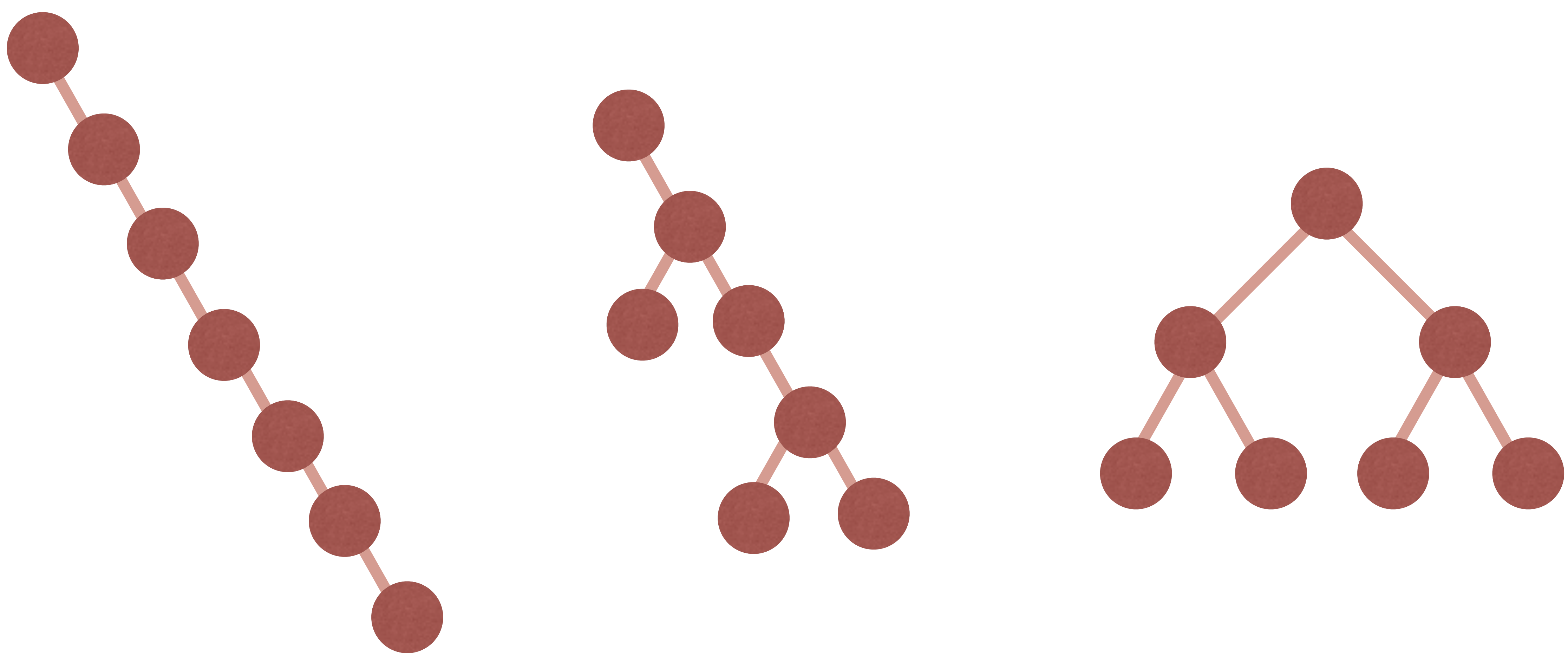
$$\because \frac{\#\{w \in A^n \mid x_n \not\sqsubseteq w\}}{\#A^n}$$

$$\leq \frac{\#\{w \in A^n \mid \text{every decomposed part} \neq x\}}{\#A^n}$$

$$= \left(1 - \frac{1}{A^{\lceil \log^{(2)}(n) \rceil}}\right)^{\lfloor n / \lceil \log^{(2)}(n) \rceil \rfloor} \to 0 \ (n \to \infty) \qquad \therefore$$

CHALLENGE IN PROVING PARAMETERISED MONKEY THEOREM FOR TERMS

- How to obtain such a "good" decomposition for the set of **λ**-terms $\Lambda_n^{\alpha}(k, \iota, \xi)$?
  - **Non-trivial** since terms have various shapes:



OUTLINE

- Introduction
- Our result
- Proof of our result
  - Idea
  - Infinite Monkey Theorem
  - Decomposition of terms
  - Sketch of the proof
- Conclusion

DECOMPOSITION OF TERMS

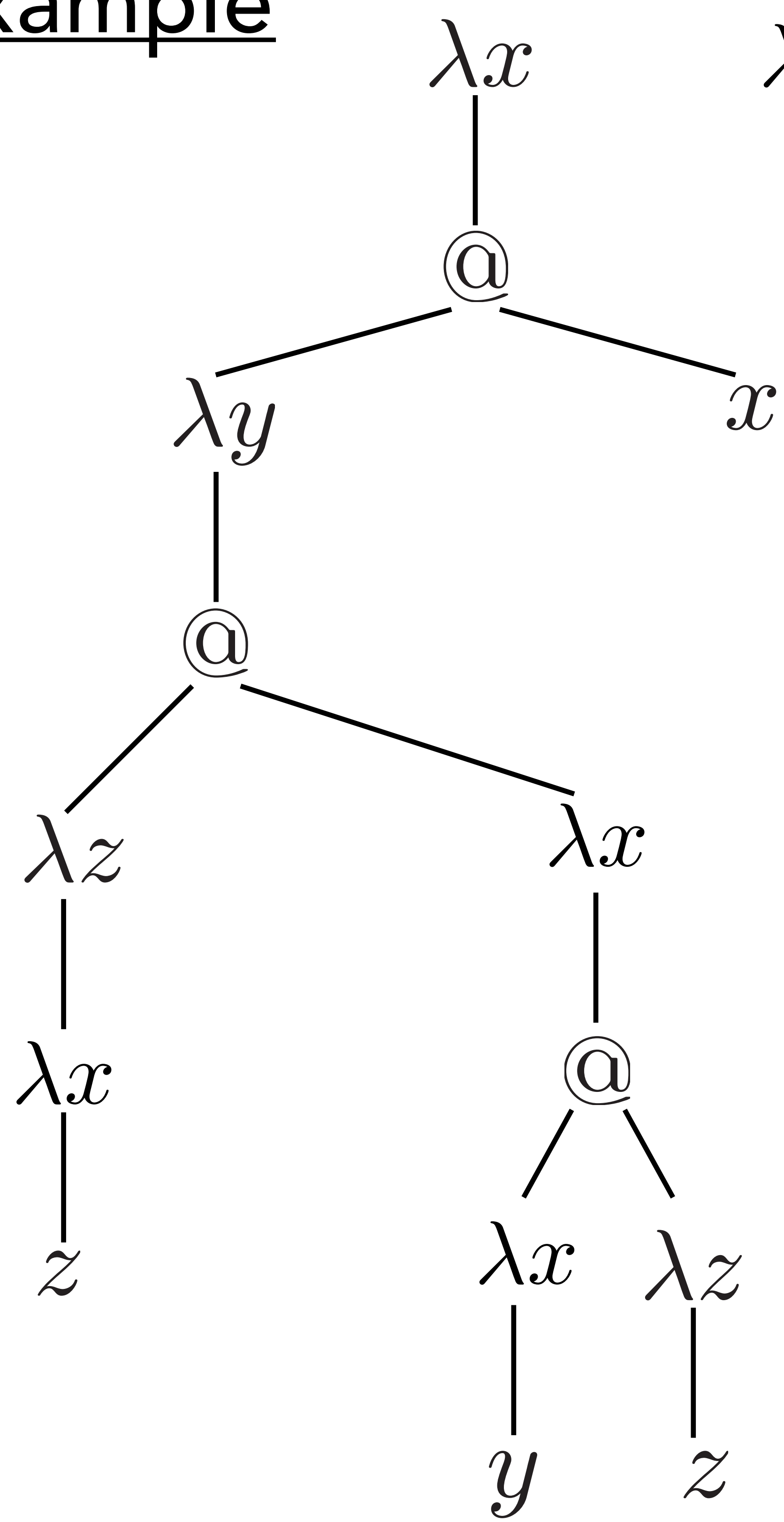
Example

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

DECOMPOSITION OF TERMS

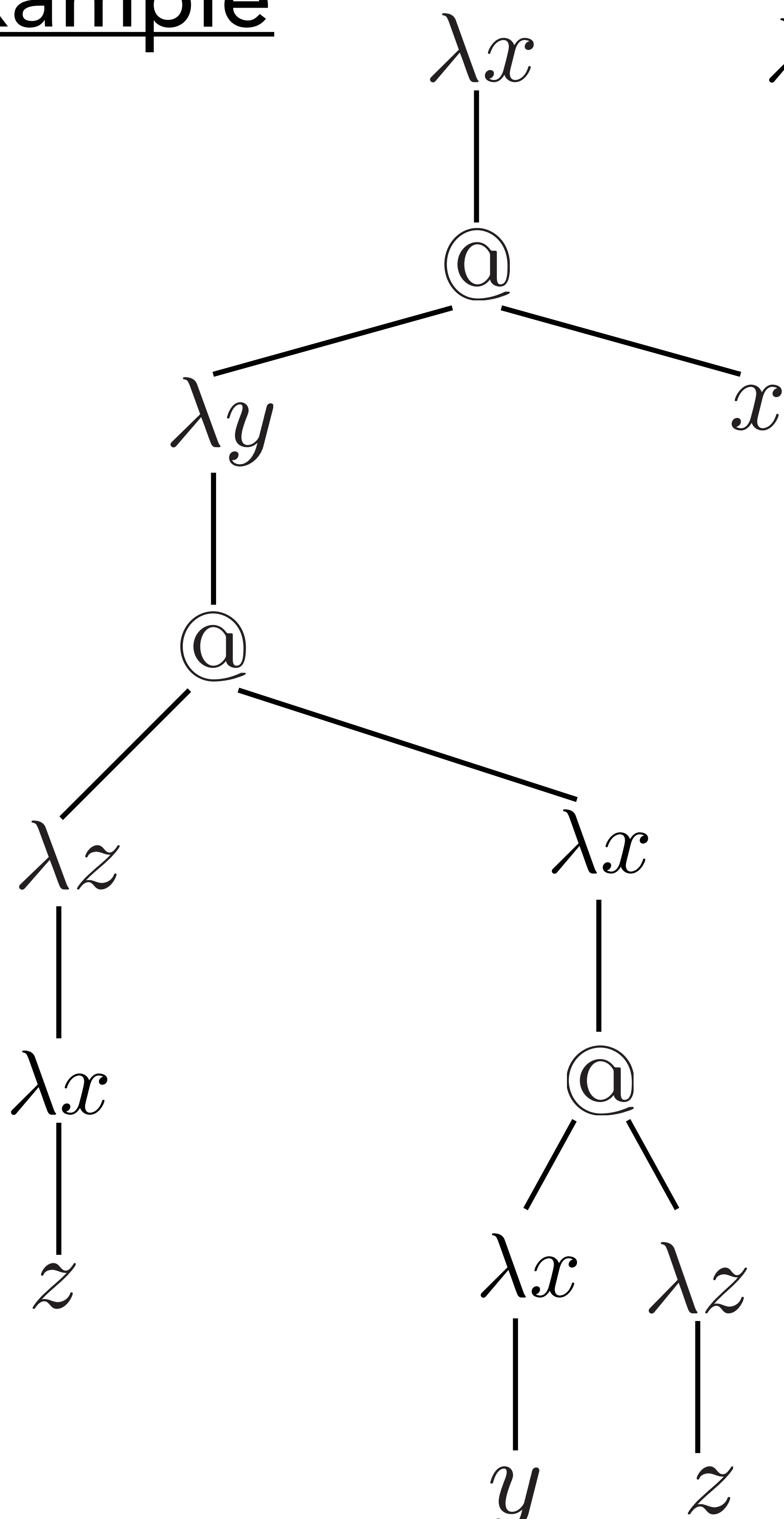
<u>Example</u>

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



DECOMPOSITION OF TERMS

Example



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

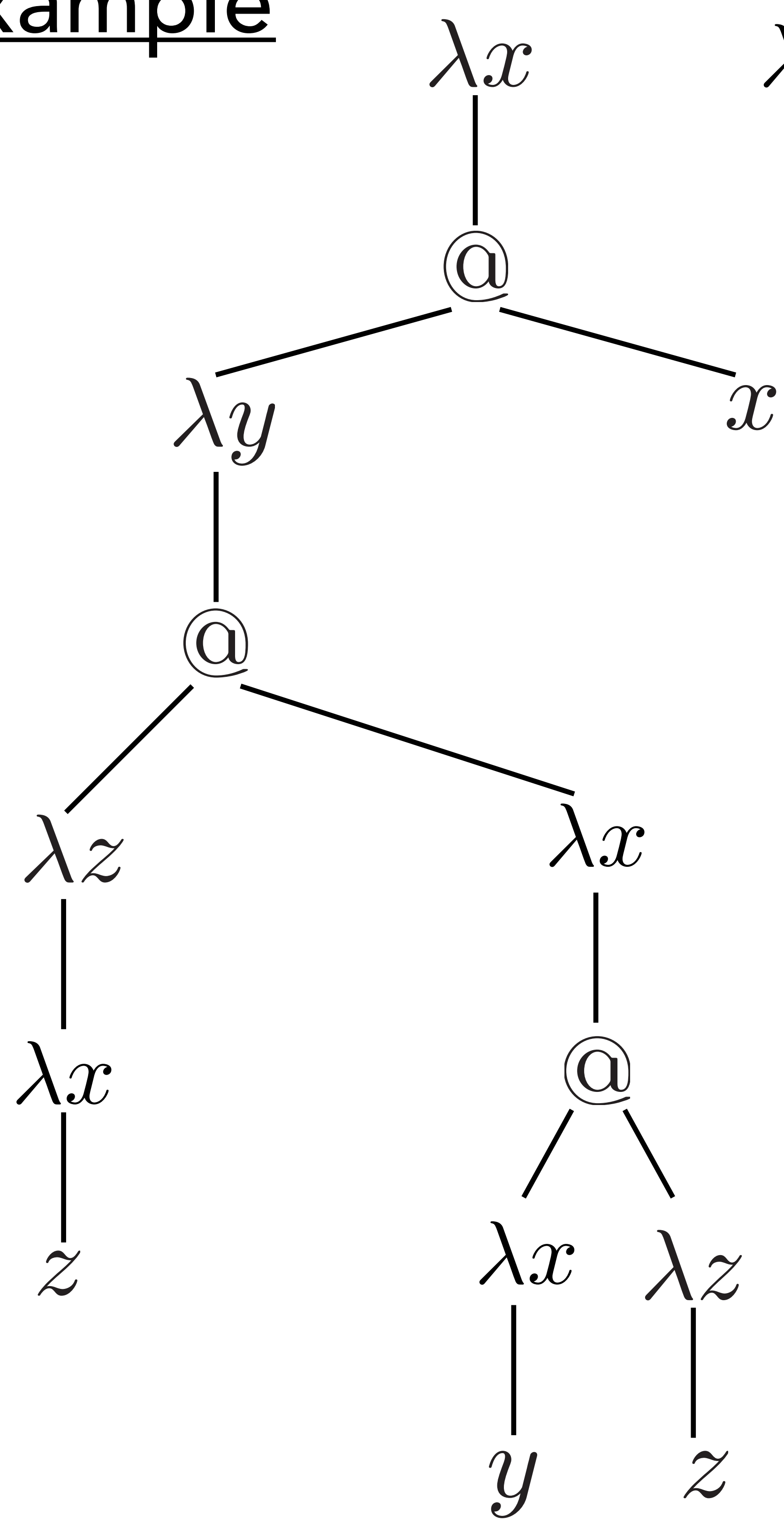
decomposition size $m = 3$

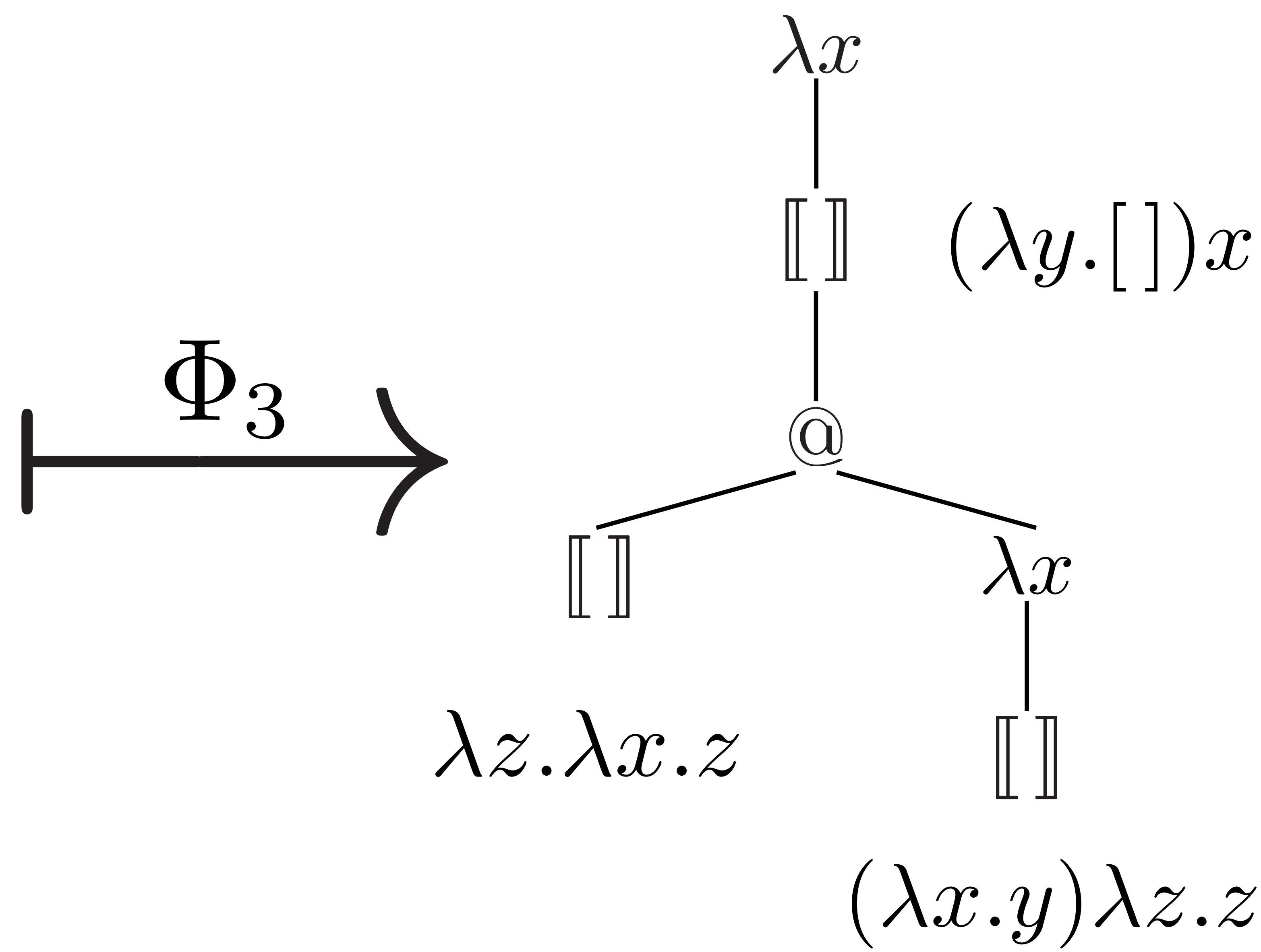
DECOMPOSITION OF TERMS

<u>Example</u>

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

decomposition size $m = 3$



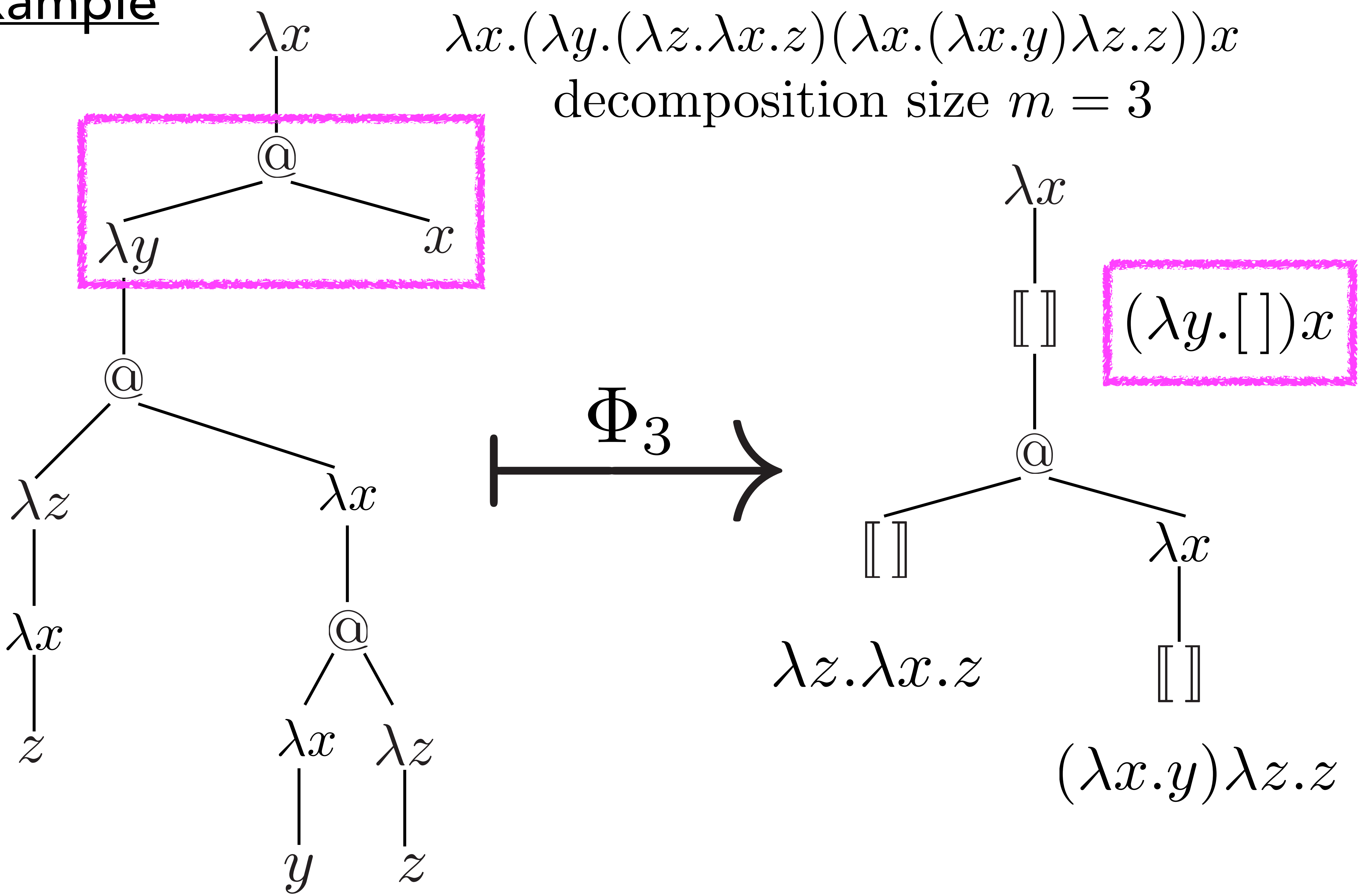


DECOMPOSITION OF TERMS

Example

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

decomposition size $m = 3$

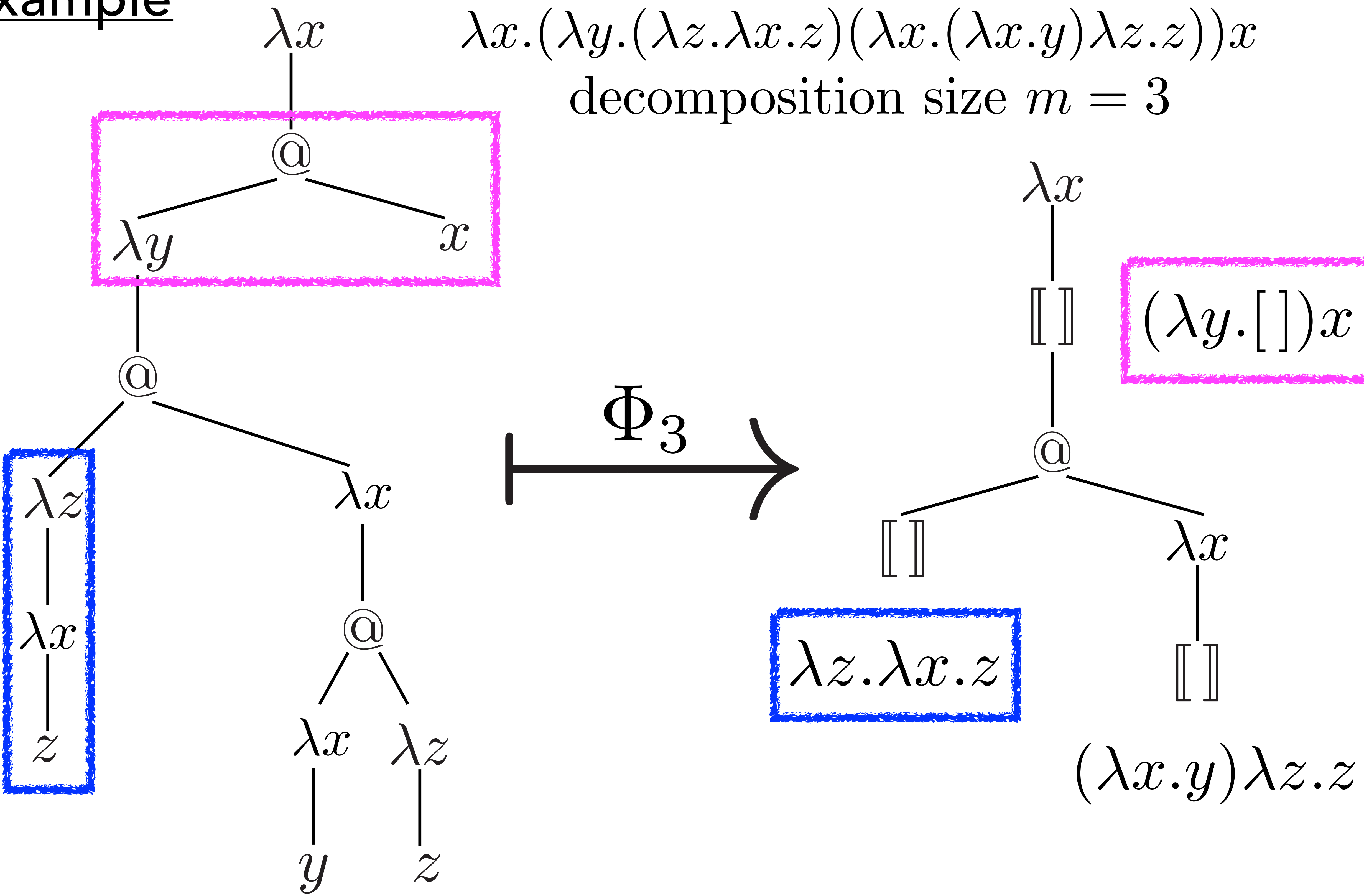


DECOMPOSITION OF TERMS

Example

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

decomposition size $m = 3$

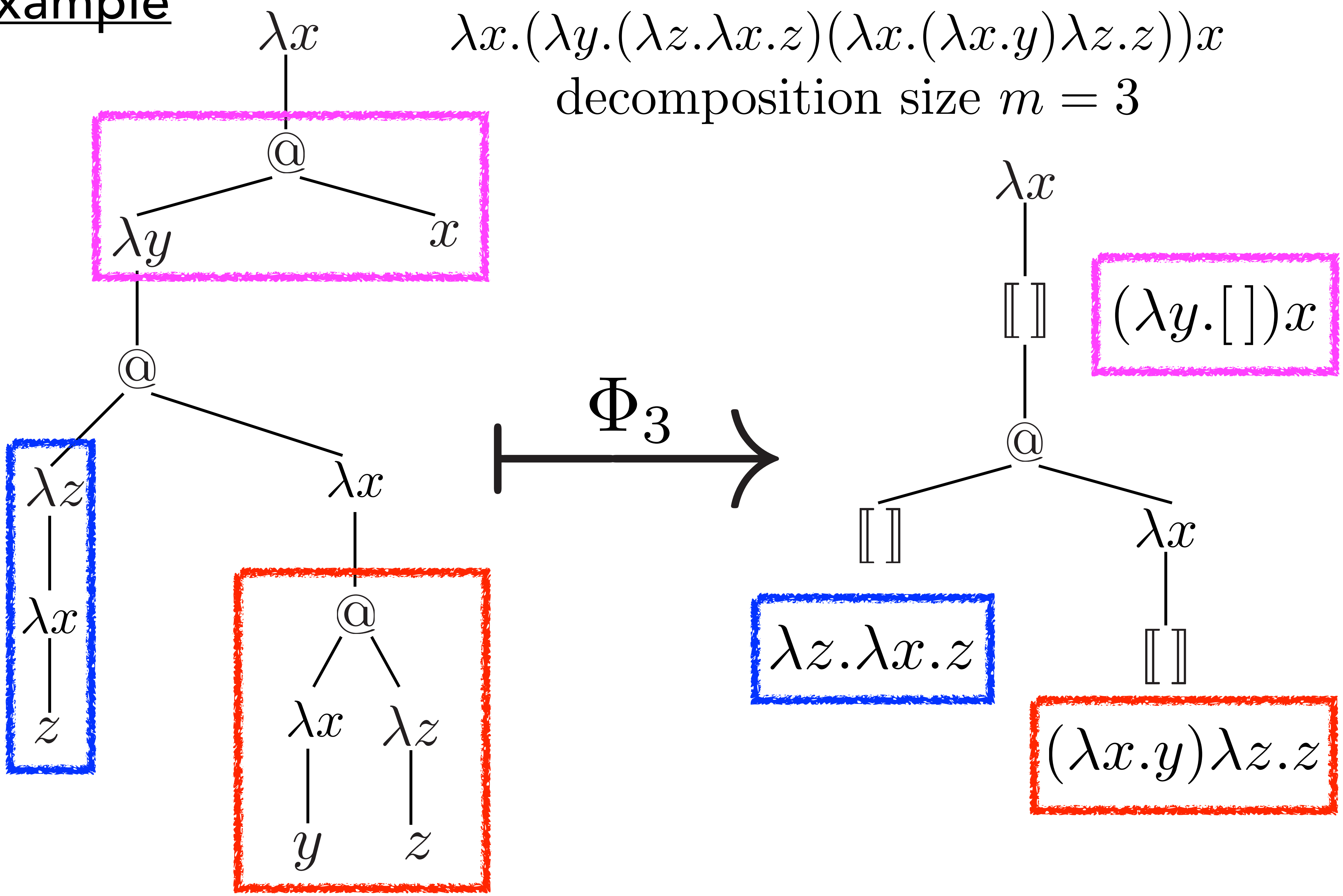


DECOMPOSITION OF TERMS

Example

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

decomposition size $m = 3$

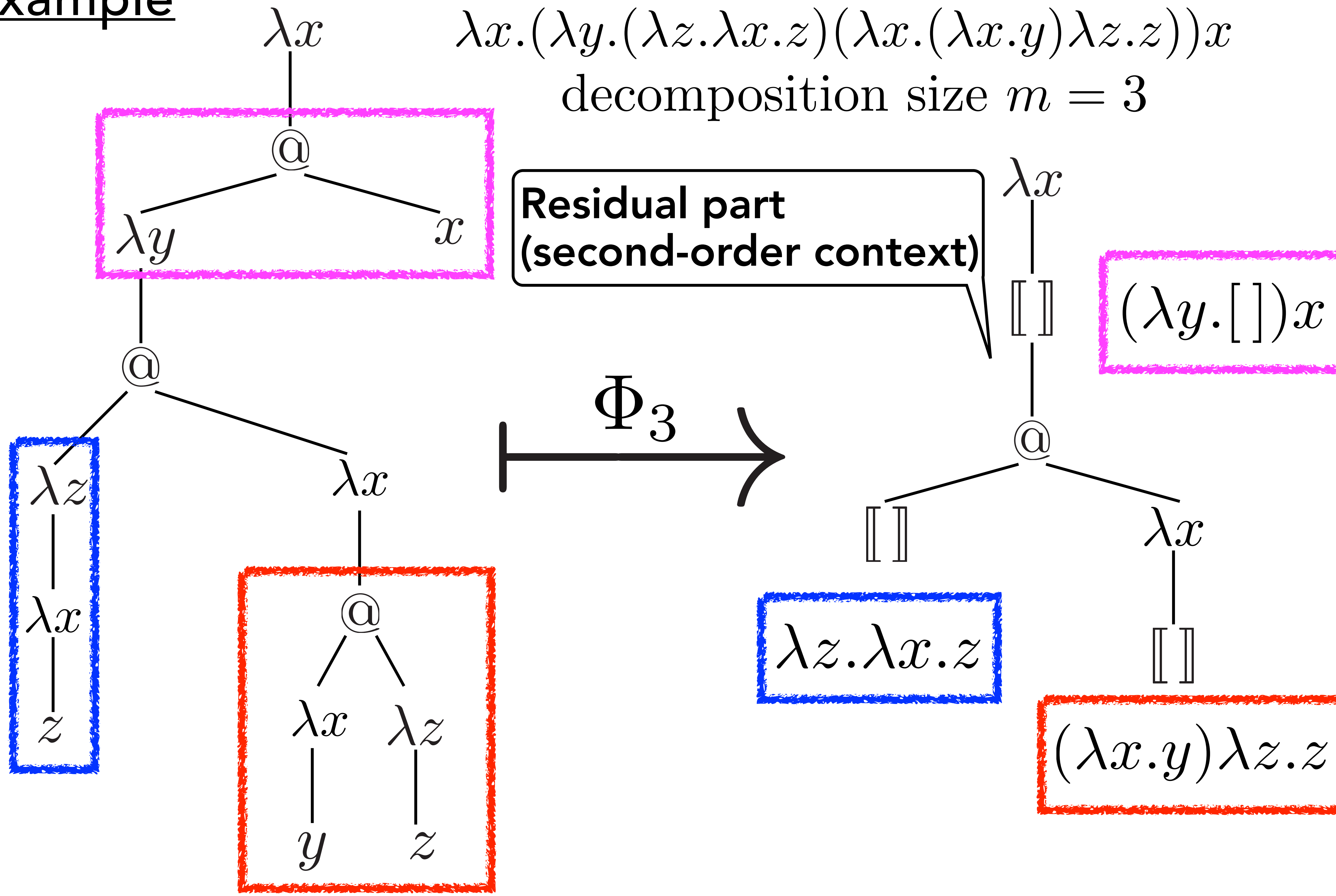


DECOMPOSITION OF TERMS

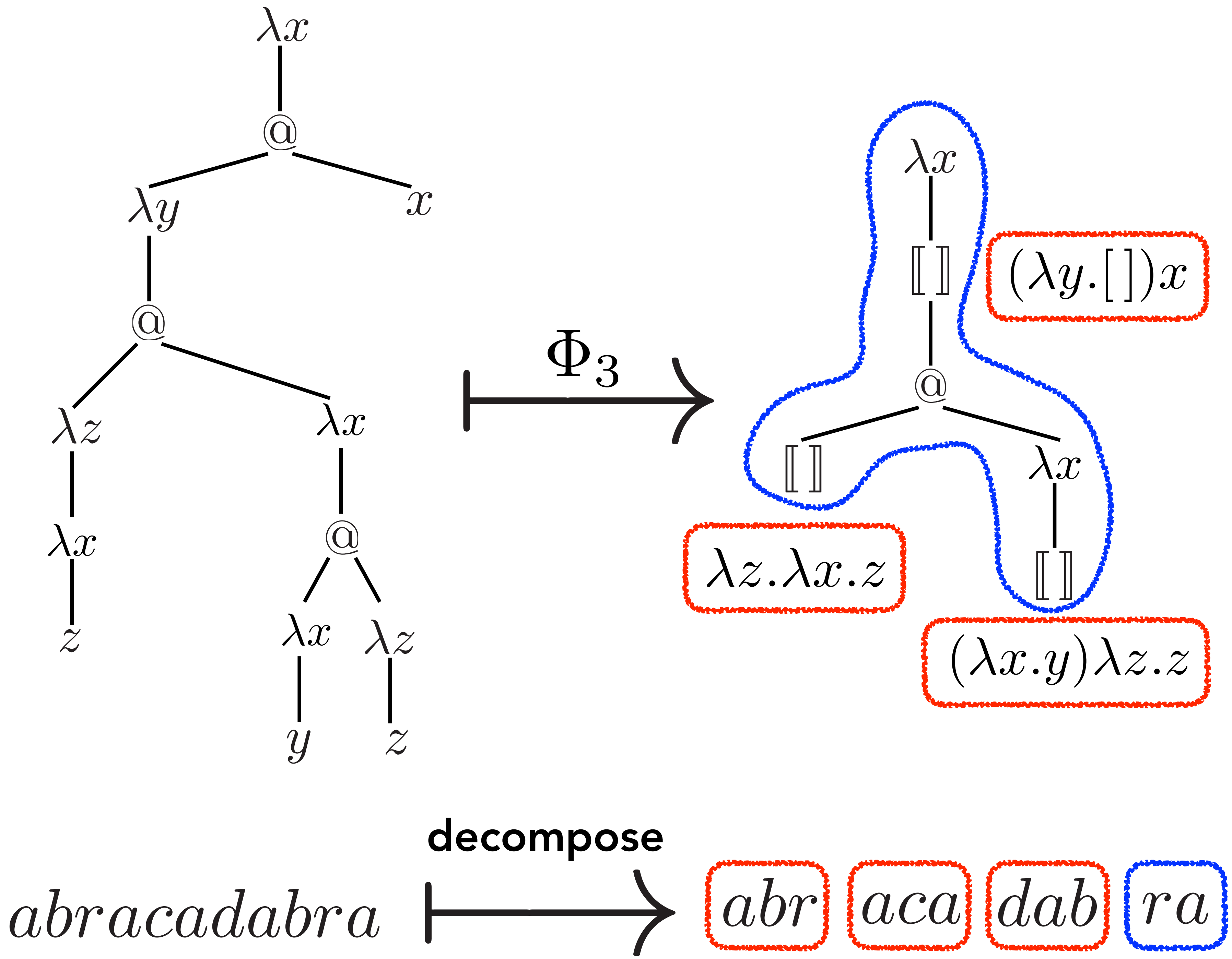
Example

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

decomposition size $m = 3$



ANALOGY BETWEEN THE DECOMPOSITION OF TERMS AND WORDS



* Decomposed part

* Residual part

DECOMPOSITION LEMMA

For $k, \iota, \xi \geq 0$ *and* $n \geq m \geq 2$,

$$\Lambda^{\alpha}_{n}(k, \iota, \xi) \cong \coprod_{E \in \mathcal{B}^{n}_{m}} \prod_{i \leq \mathrm{shn}(E)} U^{m}_{E.i}$$

cf. $$A^{n} \cong \coprod_{w \in A^{(n \bmod m)}} \prod_{i \leq \lfloor \frac{n}{m} \rfloor} A^{m}$$

DECOMPOSITION LEMMA

For $k, \iota, \xi \geq 0$ *and* $n \geq m \geq 2$,

$$\Lambda^\alpha_n(k, \iota, \xi) \cong \coprod_{E \in \mathcal{B}^n_m} \prod_{i \leq \mathrm{shn}(E)} U^m_{E.i}$$

some set of second-order contexts

cf. $$A^n \cong \coprod_{w \in A^{(n \bmod m)}} \prod_{i \leq \lfloor \frac{n}{m} \rfloor} A^m$$

DECOMPOSITION LEMMA

For $k, \iota, \xi \geq 0$ *and* $n \geq m \geq 2$ *,*

$$\Lambda_n^{\alpha}(k, \iota, \xi) \cong \coprod_{E \in \mathcal{B}_m^n} \prod_{i \leq \mathrm{shn}(E)} U_{E.i}^m$$

some set of second-order contexts (for $\mathcal{B}_m^n$)

the number of holes $[\![\,]\!]$ *in E* (for $\mathrm{shn}(E)$)

cf.

$$A^n \cong \coprod_{w \in A^{(n \bmod m)}} \prod_{i \leq \lfloor \frac{n}{m} \rfloor} A^m$$

DECOMPOSITION LEMMA

For $k, \iota, \xi \geq 0$ *and* $n \geq m \geq 2$ *,*

$$\Lambda^{\alpha}_{n}(k, \iota, \xi) \cong \coprod_{E \in \mathcal{B}^{n}_{m}} \prod_{i \leq \mathrm{shn}(E)} U^{m}_{E.i}$$

some set of second-order contexts

the number of holes $[\![\,]\!]$ *in E*

the set of "good" contexts that can be filled in the i-th hole of E.

cf. $$A^{n} \cong \coprod_{w \in A^{(n \bmod m)}} \prod_{i \leq \lfloor \frac{n}{m} \rfloor} A^{m}$$

DECOMPOSITION LEMMA

For $k, \iota, \xi \geq 0$ *and* $n \geq m \geq 2$ *,*

$$\Lambda_n^\alpha(k, \iota, \xi) \cong \coprod_{E \in \mathcal{B}_m^n} \prod_{i \leq \mathrm{shn}(E)} U_{E.i}^m$$

Each decomposed part $U_{E.i}^m$ **DOES depend** on the residual part E (and also on the index i)

Each decomposed part A^m **does NOT depend** on the residual part w

cf. $$A^n \cong \coprod_{w \in A^{(n \bmod m)}} \prod_{i \leq \lfloor \frac{n}{m} \rfloor} A^m$$

OUTLINE

- Introduction
- Our result
- Proof of our result
  - Idea
  - Infinite Monkey Theorem
  - Decomposition of terms
  - Sketch of the proof
- Conclusion

PROOF OF PARAMETERISED INFINITE MONKEY THOREM FOR TERMS

For any family of contexts $(C_n)_n$ *of* $\Lambda_n^\alpha(k,\iota,\xi)$ *such that* $|C_n| = \lceil \log^{(2)}(n) \rceil$,

$$\lim_{n\to\infty} \frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k,\iota,\xi)\} \mid C_n \preceq t\}}{\#\Lambda_n^\alpha(k,\iota,\xi)} = 1.$$

if $k,\iota,\xi \geq 2$.

$\because$ It is suffice to show that

$$\frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k,\iota,\xi) \mid C_n \not\preceq t\}}{\#\Lambda_n^\alpha(k,\iota,\xi)} \to 0 \quad (n\to\infty)$$

PROOF OF PARAMETERISED INFINITE MONKEY THOREM FOR TERMS

$\because$

$$\frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k,\iota,\xi) \mid C_n \not\leq t\}}{\#\Lambda_n^\alpha(k,\iota,\xi)} \overset{?}{\to} 0 \quad (n \to \infty)$$

cf. $$\begin{array}{ccccc} \Lambda_n^\alpha(k,\iota,\xi) \cong & \displaystyle\coprod_{E \in \mathcal{B}_n^{\log^{(2)}(n)}} & & \displaystyle\prod_{i \leq \mathrm{shn}(E)} U_{E.i}^{\log^{(2)}(n)} \\ \Cup & \Cup & & \Cup \\ [t]_\alpha & \xmapsto{\Phi_{\log^{(2)}(n)}} & E \,\&\, \bigl(u_1, u_2, \cdots, u_{\mathrm{shn}(E)}\bigr) \end{array}$$

PROOF OF PARAMETERISED INFINITE MONKEY THOREM FOR TERMS

$\because$

$$\frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k,\iota,\xi) \mid C_n \not\preceq t\}}{\#\Lambda_n^\alpha(k,\iota,\xi)}$$

$$\leq \frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k,\iota,\xi) \mid C_n \not\preceq u_i \text{ for every } i\}}{\#\Lambda_n^\alpha(k,\iota,\xi)}$$

cf. $\Lambda_n^\alpha(k,\iota,\xi) \cong \coprod_{E \in \mathcal{B}_n^{\log^{(2)}(n)}} \prod_{i \leq \mathrm{shn}(E)} U_{E.i}^{\log^{(2)}(n)}$

$\quad\ \ \rotatebox{90}{$\in$} \qquad\qquad\qquad \rotatebox{90}{$\in$} \qquad\qquad\qquad \rotatebox{90}{$\in$}$

$$[t]_\alpha \xmapsto{\Phi_{\log^{(2)}(n)}} E \ \&\ (u_1, u_2, \cdots, u_{\mathrm{shn}(E)})$$

PROOF OF PARAMETERISED INFINITE MONKEY THOREM FOR TERMS

$\because$

$$\frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k,\iota,\xi) \mid C_n \not\leq t\}}{\#\Lambda_n^\alpha(k,\iota,\xi)}$$

$$\leq \frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k,\iota,\xi) \mid C_n \not\leq u_i \text{ for every } i\}}{\#\Lambda_n^\alpha(k,\iota,\xi)}$$

$$= \frac{\# \coprod_{E \in \mathcal{B}_n^{\lceil \log^{(2)}(n) \rceil}} \prod_{i \leq \mathrm{shn}(E)} \left\{u_i \in U_{E.i}^{\lceil \log^{(2)}(n) \rceil} \mid C_n \not\leq u_i\right\}}{\#\Lambda_n^\alpha(k,\iota,\xi)}$$

PROOF OF PARAMETERISED INFINITE MONKEY THOREM FOR TERMS

$$\vdots$$

$$\frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k,\iota,\xi) \mid C_n \not\leq t\}}{\#\Lambda_n^\alpha(k,\iota,\xi)}$$

$$\leq \frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k,\iota,\xi) \mid C_n \not\leq u_i \text{ for every } i\}}{\#\Lambda_n^\alpha(k,\iota,\xi)}$$

$$= \frac{\# \coprod_{E \in \mathcal{B}_n^{\lceil \log^{(2)}(n) \rceil}} \prod_{i \leq \mathrm{shn}(E)} \left\{ u_i \in U_{E.i}^{\lceil \log^{(2)}(n) \rceil} \mid C_n \not\leq u_i \right\}}{\#\Lambda_n^\alpha(k,\iota,\xi)}$$

$$\leq \left(1 - 1/c\gamma^{2\lceil \log^{(2)}(n) \rceil}\right)^{n/4\lceil \log^{(2)}(n) \rceil}$$

PROOF OF PARAMETERISED INFINITE MONKEY THOREM FOR TERMS

$\vdots$

Lemma

$\mathrm{shn}(E) \geq n/4\lceil \log^{(2)}(n) \rceil$ for any $E \in \mathcal{B}_n^{\lceil \log^{(2)}(n) \rceil}$

$$\leq \frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k,\iota,\xi) \mid C_n \not\leq u_i \text{ for every } i\}}{\#\Lambda_n^\alpha(k,\iota,\xi)}$$

$$= \frac{\# \coprod_{E \in \mathcal{B}_n^{\lceil \log^{(2)}(n) \rceil}} \prod_{i \leq \mathrm{shn}(E)} \left\{ u_i \in U_{E.i}^{\lceil \log^{(2)}(n) \rceil} \mid C_n \not\leq u_i \right\}}{\#\Lambda_n^\alpha(k,\iota,\xi)}$$

$$\leq \left(1 - 1/c\gamma^{2\lceil \log^{(2)}(n) \rceil}\right)^{n/4\lceil \log^{(2)}(n) \rceil}$$

PROOF OF PARAMETERISED INFINITE MONKEY THOREM FOR TERMS

Lemma

$\mathrm{shn}(E) \geq n/4\lceil \log^{(2)}(n) \rceil$ for any $E \in \mathcal{B}_n^{\lceil \log^{(2)}(n) \rceil}$

$$\leq \frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k,\iota,\xi) \mid C_n \not\leq u_i \text{ fo[r] every } i\}}{\ldots}$$

Lemma

$\#U_{E.i}^{\lceil \log^{(2)}(n) \rceil} = \mathrm{O}(c\gamma^{2\lceil \log^{(2)}(n) \rceil})$

for some constants c and γ

$$= \frac{\ldots (n)\rceil \mid C_n \not\leq u_i\}}{\#\Lambda_n^\alpha(k,\iota,\xi)}$$

$$\leq \left(1 - 1/c\gamma^{2\lceil \log^{(2)}(n) \rceil}\right)^{n/4\lceil \log^{(2)}(n) \rceil}$$

PROOF OF PARAMETERISED INFINITE MONKEY THOREM FOR TERMS

$\vdots$

$$\frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k,\iota,\xi) \mid C_n \not\leq t\}}{\#\Lambda_n^\alpha(k,\iota,\xi)}$$

$$\leq \frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k,\iota,\xi) \mid C_n \not\leq u_i \text{ for every } i\}}{\#\Lambda_n^\alpha(k,\iota,\xi)}$$

$$= \frac{\# \coprod_{E \in \mathcal{B}_n^{\lceil \log^{(2)}(n) \rceil}} \prod_{i \leq \mathrm{shn}(E)} \left\{ u_i \in U_{E.i}^{\lceil \log^{(2)}(n) \rceil} \mid C_n \not\leq u_i \right\}}{\#\Lambda_n^\alpha(k,\iota,\xi)}$$

$$\leq \left(1 - 1/c\gamma^{2\lceil \log^{(2)}(n) \rceil}\right)^{n/4\lceil \log^{(2)}(n) \rceil}$$

PROOF OF PARAMETERISED INFINITE MONKEY THOREM FOR TERMS

$$\because$$

$$\frac{\#\{[t]_\alpha \in \Lambda^\alpha_n(k,\iota,\xi) \mid C_n \not\preceq t\}}{\#\Lambda^\alpha_n(k,\iota,\xi)}$$

$$\leq \frac{\#\{[t]_\alpha \in \Lambda^\alpha_n(k,\iota,\xi) \mid C_n \not\preceq u_i \text{ for every } i\}}{\#\Lambda^\alpha_n(k,\iota,\xi)}$$

$$= \frac{\# \coprod_{E \in \mathcal{B}_n^{\lceil \log^{(2)}(n) \rceil}} \prod_{i \leq \mathrm{shn}(E)} \left\{ u_i \in U_{E.i}^{\lceil \log^{(2)}(n) \rceil} \mid C_n \not\preceq u_i \right\}}{\#\Lambda^\alpha_n(k,\iota,\xi)}$$

$$\leq \left(1 - 1/c\gamma^{2\lceil \log^{(2)}(n) \rceil}\right)^{n/4\lceil \log^{(2)}(n) \rceil} \to 0 \ (n \to \infty) \quad \therefore$$

PROOF OF PARAMETERISED INFINITE MONKEY THOREM FOR TERMS

For any family of contexts $(C_n)_n$ *of* $\Lambda_n^\alpha(k,\iota,\xi)$ *such that* $|C_n| = \lceil \log^{(2)}(n) \rceil$,

$$\lim_{n\to\infty} \frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k,\iota,\xi)\} \mid C_n \preceq t\}}{\#\Lambda_n^\alpha(k,\iota,\xi)} = 1.$$

if $k,\iota,\xi \geq 2$.

$\because$

$$\frac{\#\{[t]_\alpha \in \Lambda_n^\alpha(k,\iota,\xi) \mid C_n \not\preceq t\}}{\#\Lambda_n^\alpha(k,\iota,\xi)} \to 0 \quad (n \to \infty)$$

$\therefore$

SUMMARY OF THE MAIN PROOF

the probability that a term $[t]_\alpha \in \Lambda_n^\alpha(k, \iota, \xi)$ has a β-reduction sequence of length (k-2)-EXP(n)

(∵ explosive property)

$\geq$ the probability that  $^{k}_{\lceil \log^{(2)}(n) \rceil} \preceq t$ *holds*

(∵ Monkey Theorem)

$\to 1 \quad (n \to \infty)$

OUTLINE

- Introduction
- Proof of our result
- Conclusion

CONCLUSION

- Almost every terms of size *n* and order at most *k* has a **β**-reduction sequence of length (*k-2*)-EXP(*n*).
  - The core of our proof is a non-trivial extension of well-known Infinite Monkey Theorem.

FUTURE WORK

- Quantitative analysis of simply typed **λ**-terms in different settings:
  - with an **unbounded** number of variables.
  - with **recursion**.