

# Strongly Normalising Cyclic Data Computation by Iteration Categories of Second-Order Algebraic Theories

Makoto Hamana

Department of Computer Science,  
Gunma University, Japan

FSCD'16

Porto University, Portuguese.

# Computing with Cyclic Datatypes

- ▷ Cyclic data structures in functional programming
- ▷ In Haskell, one can define a cyclic list by

```
clist = 2:1:clist
```

- ▷ Why is this technique correct? — lazy evaluation

```
head clist  ~>  2
```

- ▷ **Not** completely safe

- ▷ E.g. what is the sum of all elements of `clist`?

```
sum clist  ~>  non-termination
```

- ▷ The “encoding” does **not** give **complete understanding** of cyclic datatypes

## This Work

- ▷ Complete axiomatisation using
  - second-order algebraic theory [Fiore et al.'10]
  - iteration category [Bloom-Esik'93]
- ▷ Do not rely on lazy evaluation
- ▷ Strongly normalising fold combinator
- ▷ A framework for syntax and semantics of cyclic datatypes

# Our Framework

ctype `CList` where

`[] : CList`

`:: : CNat, CList → CList`

with axioms `AxCy`

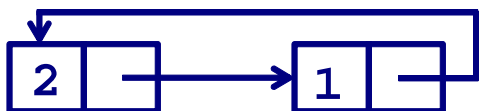
ctype `CNat` where

`0 : CNat`

`S : CNat → CNat`

with axioms `AxCy`

`cy(x.2::1::x)`



`clist`

## Specification

`sum : CList → CNat`

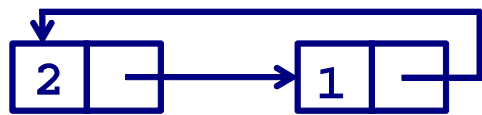
`spec sum [] = 0`

`sum (k::t) = plus(k, sum t)`

`fun sum t = fold (0, k, x.plus(k, x)) t`

# Example

[I]  $cy(x.2::1::x)$



clist

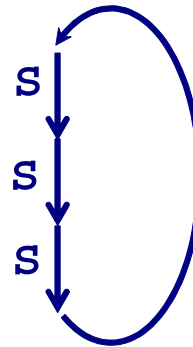
[III]

fold



sum

cyclic natural number



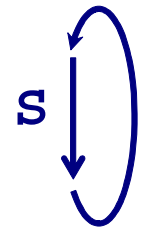
[II]

$AxCy$

=

equational theory  
for bisimulation

“infinity”



```
fun sum t = fold (0, k,x.plus(k,x)) t
```

```
sum(cy(x.S2(0)::S(0)::x)) →+ cy(x.sum(x.S2(0)::S(0)::x)@x)
→+ cy(x.S(S(S(x))))
```

▷ How to understand?

▷ Second-order equational logic and axioms  $AxCy$   
to equate cyclic data

▷  $\Leftrightarrow$  decidable **bisimulation**

# Cartesian Second-Order Algebraic Theory

▷ Base types  $a, b, c, \dots$

▷ Signature  $\Sigma$

$$f : (\vec{a}_1 \rightarrow \vec{b}_1), \dots, (\vec{a}_m \rightarrow \vec{b}_m) \rightarrow c_1, \dots, c_n.$$

▷ Metavariables  $(T, S, \dots)$   $\dots$  variables of at most first-order

$$T : \vec{a} \rightarrow b$$

▷ Variables  $\dots$  variables of order 0

▷ Syntax of meta-terms (Aczel [1978], Klop's CRS [1980])

$$t ::= x \mid x.t \mid f(t_1, \dots, t_n) \mid T[t_1, \dots, t_n]$$

# Cyclic Datatypes

## Default constructors

Tuple  $\langle -, \dots, - \rangle : (\vec{c}_1), \dots, (\vec{c}_n) \rightarrow \vec{c}_1, \dots, \vec{c}_n$

Cycle constructor  $\text{cy} : (\vec{c} \rightarrow \vec{c}) \rightarrow \vec{c}$

Composition  $\diamond : (\vec{a} \rightarrow \vec{c}), \vec{a} \rightarrow \vec{c}$

E.g.

$\text{cy}(x.S(0) :: x)$

## Axioms AxCy for Cycles

$$(sub) \quad (\overrightarrow{y} \cdot T[\overrightarrow{y}]) \diamond \langle s_1, \dots, s_n \rangle = T[s_1, \dots, s_n]$$

$$(SP) \quad \langle (\overrightarrow{y} \cdot y_1) \diamond T, \dots, (\overrightarrow{y} \cdot y_n) \diamond T \rangle = T$$

$$(dinat) \quad cy(\overrightarrow{x} \cdot s[T[\overrightarrow{x}]]) = (\overrightarrow{z} \cdot s[\overrightarrow{z}]) \diamond cy(\overrightarrow{z} \cdot T[s[\overrightarrow{z}]])$$

$$(Beki\bar{c}) \quad cy(\overrightarrow{x}, \overrightarrow{y} \cdot \langle \hat{T}, \hat{S} \rangle) = \langle cy(\overrightarrow{x} \cdot (\overrightarrow{y} \cdot \hat{T}) \diamond cy(\overrightarrow{y} \cdot \hat{S})), \\ cy(\overrightarrow{y} \cdot (\overrightarrow{x} \cdot \hat{S}) \diamond cy(\overrightarrow{x} \cdot (\overrightarrow{y} \cdot \hat{T}) \diamond cy(\overrightarrow{y} \cdot \hat{S}))) \rangle$$

$$(CI) \quad cy(\overrightarrow{y} \cdot \langle T[\rho_1], \dots, T[\rho_m] \rangle) = \langle cy(\overrightarrow{y} \cdot \tilde{T}), \dots, cy(\overrightarrow{y} \cdot \tilde{T}) \rangle$$

### Note

▷ “cy” is a (Conway) parameterised fixed point operator

$$(fix) \quad cy(\overrightarrow{x} \cdot s[\overrightarrow{x}]) = (\overrightarrow{z} \cdot s[\overrightarrow{z}]) \diamond cy(\overrightarrow{z} \cdot s[\overrightarrow{z}])$$



# Example (1): Cyclic Lists modulo Bisimulation

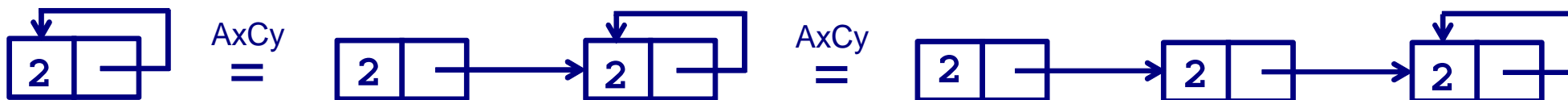
▷ ctype `CList` where

`[]` : `CList`

`::` : `CNat, CList`  $\rightarrow$  `CList`

with axioms `AxCy`

▷ `AxCy` captures the intended meaning of cyclic lists



$$\text{cy}(x.2 :: x) = 2 :: \text{cy}(x.2 :: x) = 2 :: 2 :: \text{cy}(x.2 :: x)$$

# Example (2): Cyclic Sharing Trees modulo Bisimulation

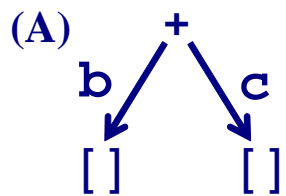
ctype CTree where

$f : \text{CTree} \rightarrow \text{CTree}$  ( $f = a, b, c, p, q, \dots$ )

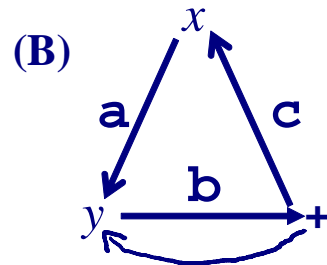
$[] : \text{CTree}$

$+ : \text{CTree}, \text{CTree} \rightarrow \text{CTree}$

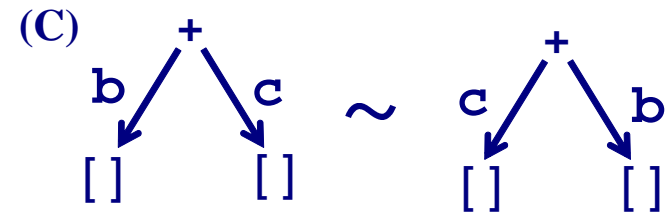
with axioms AxCy, AxBr



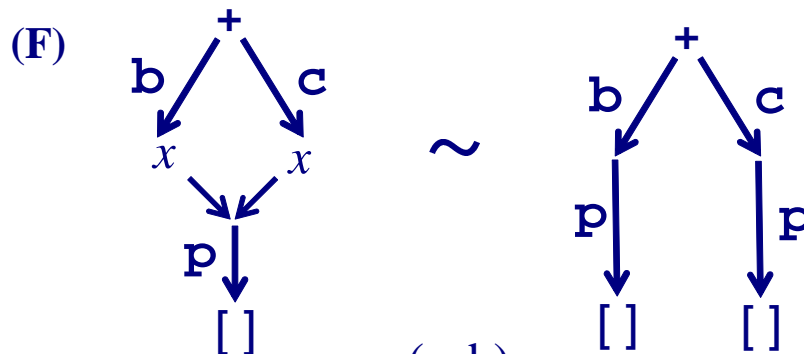
$b([]) + c([])$



$cy(x.a(cy(y.b(c(x+y)))))$



$b([]) + c([]) = c([]) + b([])$

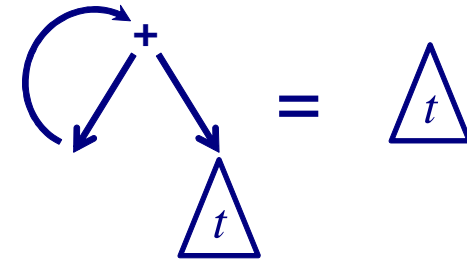


(sub)

$(x.b(x) + c(x)) \circ p([]) = b(p([])) + c(p([]))$

# Axioms AxBr for branching

$$\begin{array}{lcl}
 (\text{del}) & \text{cy}(x^c.T + x) & = \quad T \\
 (\text{unitL}) & [] + T & = \quad T \\
 (\text{unitR}) & T + [] & = \quad T \\
 (\text{assoc}) & (s + T) + U & = \quad s + (T + U) \\
 (\text{comm}) & s + T & = \quad T + s \\
 (\text{degen}) & T + T & = \quad T
 \end{array}$$



axiom (del)

**Prop. (Completeness for bisimulation)** [H. FICS'15]

$\Gamma \vdash s = t : \text{CTree}$  is derivable from AxCy and AxBr

. iff

if  $s$  and  $t$  are bisimilar.

# Categorical Semantics of Cyclic Datatypes

- ▷ Why categorical semantics? — to obtain fold combinator
- ▷ **Categorical structure**
  - Iteration category
  - = traced cartesian category [Joyal et al.'96][Hasegawa-Hyland'97]
  - + Bloom-Ésik's Commutative Identities axiom
- ▷ **Reference**
  - Bloom, Ésik. *Iteration Theories*, EATCS Monographs. 1993.

# Axioms AxCy

$$(sub) \quad (\overrightarrow{y} \cdot T[\overrightarrow{y}]) \diamond \langle s_1, \dots, s_n \rangle = T[s_1, \dots, s_n]$$

$$(SP) \quad \langle (\overrightarrow{y} \cdot y_1) \diamond T, \dots, (\overrightarrow{y} \cdot y_n) \diamond T \rangle = T$$

$$(dinat) \quad cy(\overrightarrow{x} \cdot S[T[\overrightarrow{x}]]) = (\overrightarrow{z} \cdot S[\overrightarrow{z}]) \diamond cy(\overrightarrow{z} \cdot T[S[\overrightarrow{z}]])$$

$$(Beki\check{c}) \quad cy(\overrightarrow{x}, \overrightarrow{y} \cdot \langle \hat{T}, \hat{S} \rangle) = \langle cy(\overrightarrow{x} \cdot (\overrightarrow{y} \cdot \hat{T}) \diamond cy(\overrightarrow{y} \cdot \hat{S})), \\ cy(\overrightarrow{y} \cdot (\overrightarrow{x} \cdot \hat{S}) \diamond cy(\overrightarrow{x} \cdot (\overrightarrow{y} \cdot \hat{T}) \diamond cy(\overrightarrow{y} \cdot \hat{S}))) \rangle$$

$$(CI) \quad cy(\overrightarrow{y} \cdot \langle T[\rho_1], \dots, T[\rho_m] \rangle) = \langle cy(\mathbf{y} \cdot \tilde{T}), \dots, cy(\mathbf{y} \cdot \tilde{T}) \rangle$$

where  $\rho_i$  is a permutation of  $\overrightarrow{y}$ ,  $\tilde{T} = T[\mathbf{y}, \dots, \mathbf{y}]$

# Categorical Semantics and Completeness

**Thm.**  $\Gamma \vdash s = t : \vec{c}$  is derivable from  $\mathcal{E}$  iff  $[[s]]_{\mathcal{C}}^M = [[t]]_{\mathcal{C}}^M$  holds for all iteration categories  $\mathcal{C}$  and all  $(\Sigma, \mathcal{E})$ -structures in  $\mathcal{C}$ .

## Category $\mathbf{Tm}(\mathcal{E})$

- ▷ object: sequence of types  $b_1, \dots, b_n$
- ▷ arrow: equivalence class  $[\Gamma \vdash t : \vec{c}]_{\mathcal{E}}$  of judgments modulo axioms  $\mathcal{E} = \text{AxCy} \cup \text{AxBr}$

# Universality

**Thm.** For a  $(\Sigma, \mathcal{E})$ -structure  $M$  in an iteration category  $\mathcal{C}$ , there exists a unique iteration functor  $\Psi^M : \mathbf{Tm}(\mathcal{E}) \longrightarrow \mathcal{C}$  that preserves  $(\Sigma, \mathcal{E})$ -structures.

$$\begin{array}{ccc} \mathbf{Tm} & \xrightarrow{[[ - ]^U} & \mathbf{Tm}(\mathcal{E}) \\ \downarrow [[ - ]^M & & \swarrow \Psi^M \\ \mathcal{C} & & \end{array}$$

## Extract “Fold” from Universality

**Thm.** For a  $(\Sigma, \mathcal{E})$ -structure  $M$  in an iteration category  $\mathcal{C}$ , there exists a unique iteration functor  $\Psi^M : \mathbf{Tm}(\mathcal{E}) \longrightarrow \mathcal{C}$  that preserves  $(\Sigma, \mathcal{E})$ -structures.

$$\begin{array}{ccc}
 \mathbf{Tm} & \xrightarrow{\llbracket - \rrbracket^U} & \mathbf{Tm}(\mathcal{E}) \ni x_1 : \mathbf{c}, \dots, x_n : \mathbf{c} \vdash t : \mathbf{c} \\
 \downarrow \llbracket - \rrbracket^M & \searrow \Psi^M & \downarrow \\
 \mathbf{Tm}(\mathcal{E}) & & \ni x_1 : b, \dots, x_n : b \vdash u : b
 \end{array}$$

- ▷ where the  $(\Sigma, \mathcal{E})$ -structure  $M$  interprets a type  $\mathbf{c}$  as type  $b$
- ▷  $\Psi^M$  gives a translation from the cyclic datatype  $\mathbf{c}$  to  $b$  (e.g.  $\mathbf{CList}$  to  $\mathbf{CNat}$ )
- ▷ Extract the law of fold by formalising  $\Psi^M$



# Second-Order Algebraic Theory for fold

$$\text{fold}(\mathbf{E}, \vec{y}^c \Vdash y_i) = \vec{y}^b \Vdash y_i \quad (\text{for } y_i \in \{\vec{y}\})$$

$$\text{fold}(\mathbf{E}, \vec{y} \Vdash \langle \rangle) = \vec{y} \Vdash \langle \rangle$$

$$\text{fold}(\mathbf{E}, \vec{y} \Vdash \langle s[\vec{y}], T[\vec{y}] \rangle) = \vec{y} \Vdash \langle \text{app}(\text{fold}(\mathbf{E}, \vec{y} \Vdash s[\vec{y}]), \vec{y}), \text{app}(\text{fold}(\mathbf{E}, \vec{y} \Vdash T[\vec{y}]), \vec{y}) \rangle$$

$$\text{fold}(\mathbf{E}, \vec{y} \Vdash \text{cy}(\vec{x}.T[\vec{y}, \vec{x}])) = \vec{y} \Vdash \text{cy}(\vec{x}.\text{app}(\text{fold}(\mathbf{E}, \vec{y}, \vec{x} \Vdash T[\vec{y}, \vec{x}]), \vec{y}, \vec{x}))$$

$$\text{fold}(\mathbf{E}, \vec{y} \Vdash d(\vec{A}, T_1[\vec{y}], \dots, T_n[\vec{y}])) = \vec{y} \Vdash (\vec{x}.\text{Ed}[\vec{A}, \vec{x}]) \diamond \langle \text{app}(\text{fold}(\mathbf{E}, \vec{y} \Vdash T_1[\vec{y}]), \vec{y}), \vec{y} \rangle$$

$$\text{fold}(\mathbf{E}, \vec{x} \Vdash (\vec{y}.T[\vec{y}]) \diamond s[\vec{x}]) = \vec{x} \Vdash \vec{y}.\text{app}(\text{fold}(\mathbf{E}, \vec{y} \Vdash T[\vec{y}]), \vec{y}) \diamond \text{app}(\text{fold}(\mathbf{E}, \vec{x} \Vdash s[\vec{x}]), \vec{x})$$

$$\text{app}(\vec{x} \Vdash s[\vec{x}], z_1, \dots, z_m) = s[z_1, \dots, z_m]$$

To obtain computation rules

- ▷ Include some theorems of AxCy and AxBr for simplification
- ▷ Orient equations

# Second-Order Rewrite System FOLDr

## Fold

$$\begin{aligned}
 \text{fold}(\mathbf{E}, \overrightarrow{\mathbf{y}}^{\text{Var}_c}. v(\mathbf{y}_i)) &\rightarrow \overrightarrow{\mathbf{y}}^{\text{Var}_b}. v(\mathbf{y}_i) \\
 \text{fold}(\mathbf{E}, \overrightarrow{\mathbf{y}}. \langle \rangle) &\rightarrow \overrightarrow{\mathbf{y}}. \langle \rangle \\
 \text{fold}(\mathbf{E}, \overrightarrow{\mathbf{y}}. \langle s[\overrightarrow{\mathbf{y}}], \mathsf{T}[\overrightarrow{\mathbf{y}}] \rangle) &\rightarrow \overrightarrow{\mathbf{y}}. \langle \text{fold}(\mathbf{E}, \overrightarrow{\mathbf{y}}. s[\overrightarrow{\mathbf{y}}]) @ \overrightarrow{\mathbf{y}}, \text{fold}(\mathbf{E}, \overrightarrow{\mathbf{y}}. \mathsf{T}[\overrightarrow{\mathbf{y}}]) @ \overrightarrow{\mathbf{y}} \rangle \\
 \text{fold}(\mathbf{E}, \overrightarrow{\mathbf{y}}. \text{cy}^1(x. \mathsf{T}[\overrightarrow{\mathbf{y}}, x])) &\rightarrow \overrightarrow{\mathbf{y}}. \text{cy}^1(x. \text{fold}(\mathbf{E}, \overrightarrow{\mathbf{y}}, x. \mathsf{T}[\overrightarrow{\mathbf{y}}, x]) @ \overrightarrow{\mathbf{y}}, x) \\
 \text{fold}(\mathbf{E}, \overrightarrow{\mathbf{y}}. d(\overrightarrow{\mathbf{A}}, \mathsf{T}_1[\overrightarrow{\mathbf{y}}], \dots, \mathsf{T}_n[\overrightarrow{\mathbf{y}}])) &\rightarrow \overrightarrow{\mathbf{y}}. (\overrightarrow{\mathbf{x}}. \text{Ed}[\overrightarrow{\mathbf{A}}, \overrightarrow{\mathbf{x}}]) \diamond \langle \text{fold}(\mathbf{E}, \overrightarrow{\mathbf{y}}. \mathsf{T}_1[\overrightarrow{\mathbf{y}}]) @ \overrightarrow{\mathbf{y}}, \dots \rangle \\
 \text{fold}(\mathbf{E}, \overrightarrow{\mathbf{x}}. (\overrightarrow{\mathbf{y}}. \mathsf{T}[\overrightarrow{\mathbf{y}}]) \diamond s[\overrightarrow{\mathbf{x}}]) &\rightarrow \overrightarrow{\mathbf{x}}. (\overrightarrow{\mathbf{y}}. \text{fold}(\mathbf{E}, \overrightarrow{\mathbf{y}}. \mathsf{T}[\overrightarrow{\mathbf{y}}]) @ \overrightarrow{\mathbf{y}}) \diamond \text{fold}(\mathbf{E}, \overrightarrow{\mathbf{x}}. s[\overrightarrow{\mathbf{x}}]) @ \overrightarrow{\mathbf{x}}
 \end{aligned}$$

## Bekič and cycle cleaning

$$\begin{aligned}
 \text{cy}^{m+n}(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}}. \langle \hat{\mathbf{T}}, \hat{\mathbf{S}} \rangle) &\rightarrow \langle \text{cy}^m(\overrightarrow{\mathbf{x}}. (\overrightarrow{\mathbf{y}}. \hat{\mathbf{T}}) \diamond \text{cy}^n(\overrightarrow{\mathbf{y}}. \hat{\mathbf{S}})), \\
 &\quad \text{cy}^n(\overrightarrow{\mathbf{y}}. (\overrightarrow{\mathbf{x}}. \hat{\mathbf{S}}) \diamond \text{cy}^m(\overrightarrow{\mathbf{x}}. (\overrightarrow{\mathbf{y}}. \hat{\mathbf{T}}) \diamond \text{cy}^n(\overrightarrow{\mathbf{y}}. \hat{\mathbf{S}}))) \rangle \\
 \text{cy}(\overrightarrow{\mathbf{y}}. \mathsf{T}) &\rightarrow \mathsf{T} \\
 \text{cy}(x. v(x)) &\rightarrow [] \\
 \text{cy}(x. \mathsf{T} + v(x)) &\rightarrow \mathsf{T}
 \end{aligned}$$

## Composition

$$\begin{aligned}
 (\overrightarrow{\mathbf{y}}. v(\mathbf{y}_i)) \diamond \langle \overrightarrow{\mathbf{S}} \rangle &\rightarrow \mathsf{S}_i \\
 (\overrightarrow{\mathbf{y}}. d(\overrightarrow{\mathbf{x}}_1. \mathsf{T}_1[\overrightarrow{\mathbf{y}}, \overrightarrow{\mathbf{x}}_1], \dots)) \diamond \langle \overrightarrow{\mathbf{S}} \rangle &\rightarrow d(\overrightarrow{\mathbf{x}}. (\overrightarrow{\mathbf{y}}. \mathsf{T}_1[\overrightarrow{\mathbf{y}}, \overrightarrow{\mathbf{x}}_1]) \diamond \langle \overrightarrow{\mathbf{S}} \rangle, \dots, (\overrightarrow{\mathbf{y}}. \mathsf{T}_n[\overrightarrow{\mathbf{y}}, \overrightarrow{\mathbf{x}}_n]) \diamond \langle \overrightarrow{\mathbf{S}} \rangle)
 \end{aligned}$$

▷ **Strongly normalising** by [Blanqui RTA'00]'s the General Schema

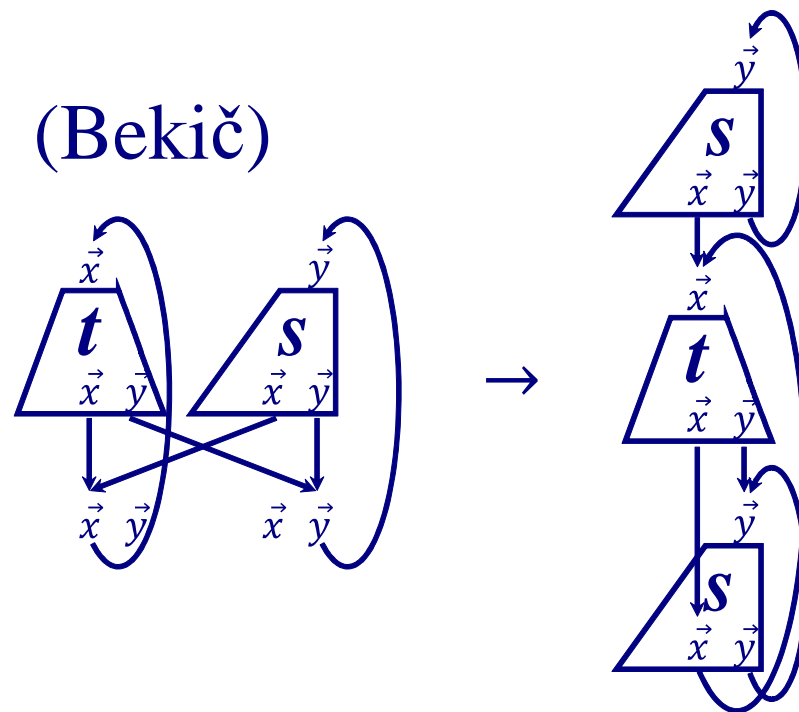
# Fixed point law

We **do not include**

$$(\text{fix}) \quad \text{cy}(\vec{x}.s[\vec{x}]) \rightarrow (\vec{z}.s[\vec{z}]) \diamond \text{cy}(\vec{z}.s[\vec{z}])$$

in the rewrite system FOLDr

But include



# Example: Tail of Cyclic List



`ctail : CList → CList`

`spec ctail ([]) = []`

`ctail (k::t) = t`

`ctail (cy(x. t)) = ??`

# Example: Tail of Cyclic List



`ctail : CList → CList`

`spec ctail ([]) = []`      ▶ Primitive recursion by fold  
`ctail (k::t) = t`      . “paramorphism”

`ctail (cy(x. t)) = ??`

`fun ctail(t) = π1 ◊ fold (<[], []>, k.x.y.<y, k::y>) t`

`ctail(cy(x.1::2::x)) →+ π1 ◊ cy(x.y. <2::y, 1::2::y>)`

`→+ π1 ◊ <cy(x.2::cy(y.1::2::y)), cy(y.1::2::y)>`

`→ cy(x.2::cy(y.1::2::y))`

`→ 2::cy(y.1::2::y)`      (Normal form)

▷ The step `→+` uses Bekič law

▷ Fixed point law is not used

▶ **Not** infinite unfolding

# Summary

- I. A framework for **cyclic datatypes** using second-order algebraic theories
- II. Semantics: iteration categories
- III. Extracted **SN** fold by rewrite system FOLDr

## Further results

- ▷ FOLDr: Confluent on plain terms
- ▷ FOLDr: Confluent modulo bisimulation  $\sim$  via Huet [1980]'s theorem using *local coherence modulo  $\sim$*
- ▷ A complete model of graph database language UnCAL [Buneman'96] (with Matsuda and Asada) **draft available**