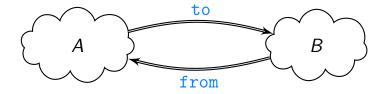
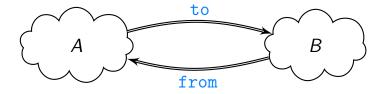
# Programming Language Approaches to Bidirectional Transformation

Janis Voigtländer

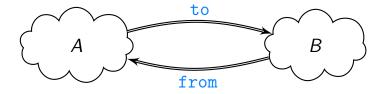
University of Bonn

LDTA'12

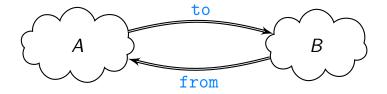




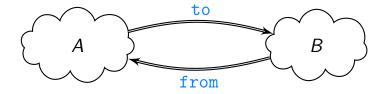
concrete syntax  $\Leftrightarrow$  abstract syntax



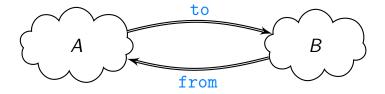
database source  $\Leftrightarrow$  materialized view



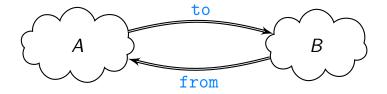
#### document representation $\Leftrightarrow$ screen visualization



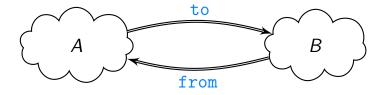
#### software model $\Leftrightarrow$ code



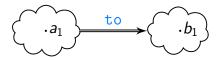
#### abstract datatype $\Leftrightarrow$ actual implementation

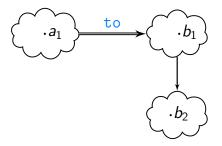


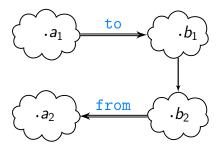
#### program input $\Leftrightarrow$ program output

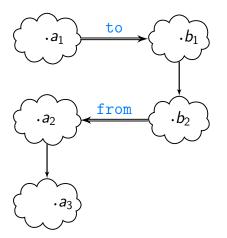


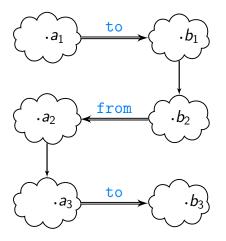
abstract syntax concrete syntax  $\Leftrightarrow$ database source materialized view  $\Leftrightarrow$ document representation screen visualization  $\Leftrightarrow$ software model code  $\Leftrightarrow$ actual implementation abstract datatype  $\Leftrightarrow$ program input program output  $\Leftrightarrow$ 

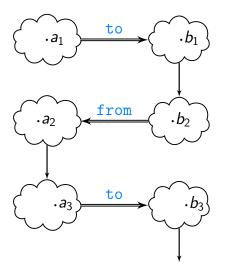


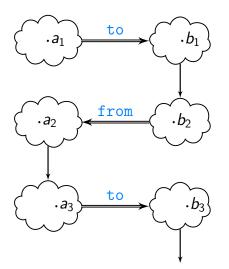




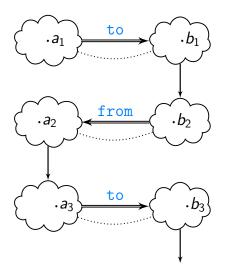






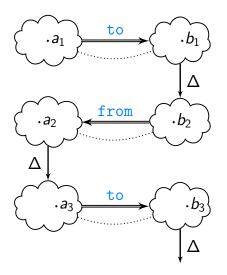


unless bijective, typically additional information needed/useful



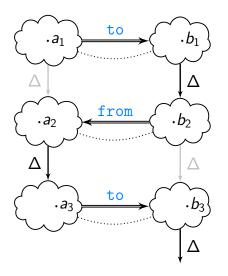
unless bijective, typically additional information needed/useful:

 about connections between A and B (objects)



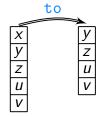
unless bijective, typically additional information needed/useful:

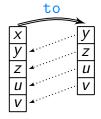
- about connections between A and B (objects)
- about the updates on either side



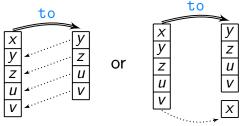
unless bijective, typically additional information needed/useful:

- about connections between A and B (objects)
- about the updates on either side

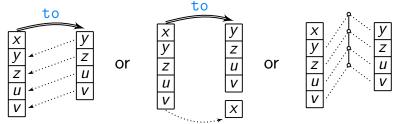




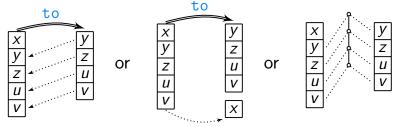
A closer look at representing  $a_{i_1,...}, b_i$  connections. For example:



A closer look at representing  $a_{i_1,...}, b_i$  connections. For example:



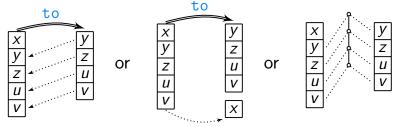
A closer look at representing  $a_{i_1} \cdot b_i$  connections. For example:



Why is it not enough to look just at the data?

Ζ

A closer look at representing  $a_{i_1} b_i$  connections. For example:



Why is it not enough to look just at the data?

Ζ

Ζ

U

 $\frac{x}{x}$ 

x

- What artefacts need to be specified?
  - both to and from
  - only one of them, the other derived
  - ▶ a more abstract artefact, from which both derivable

- What artefacts need to be specified?
  - both to and from
  - only one of them, the other derived
  - ▶ a more abstract artefact, from which both derivable
- How are they specified, manipulated, analyzed?

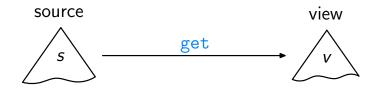
- What artefacts need to be specified?
  - both to and from
  - only one of them, the other derived
  - ▶ a more abstract artefact, from which both derivable
- How are they specified, manipulated, analyzed?
- What properties are they expected to have?

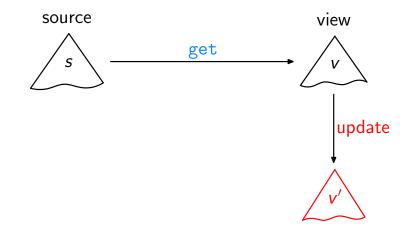
- What artefacts need to be specified?
  - both to and from
  - only one of them, the other derived
  - ▶ a more abstract artefact, from which both derivable
- How are they specified, manipulated, analyzed?
- What properties are they expected to have?
- What influence does a user, modeller, programmer have?

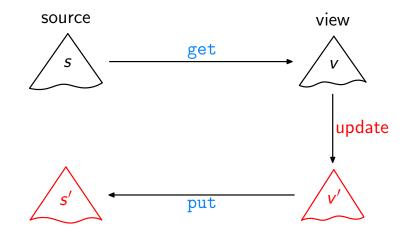
Some further relevant aspects:

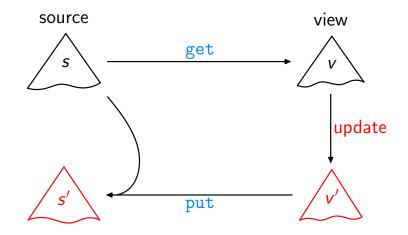
- What artefacts need to be specified?
  - both to and from
  - only one of them, the other derived
  - ▶ a more abstract artefact, from which both derivable
- How are they specified, manipulated, analyzed?
- What properties are they expected to have?
- What influence does a user, modeller, programmer have?

answers/approaches vary with field

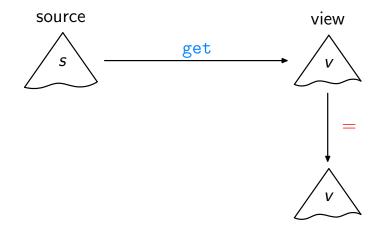






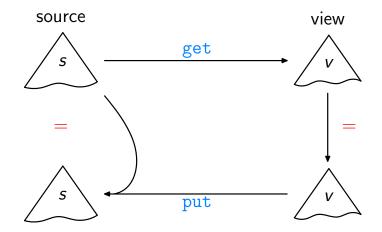


A specific (asymmetric) setting:



GetPut law

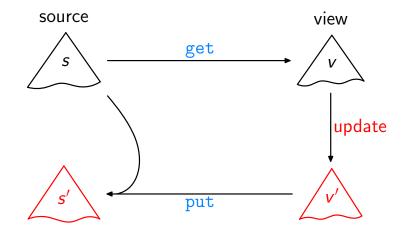
A specific (asymmetric) setting:



GetPut law

### **Bidirectional Transformations**

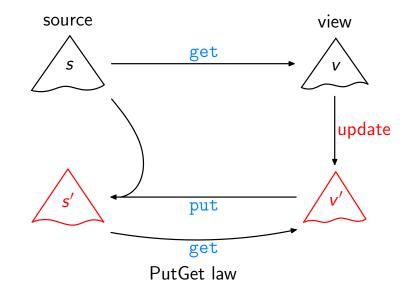
A specific (asymmetric) setting:



PutGet law

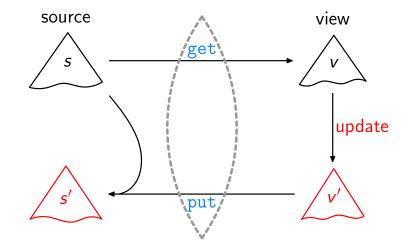
## **Bidirectional Transformations**

A specific (asymmetric) setting:



# **Bidirectional Transformations**

A specific (asymmetric) setting:

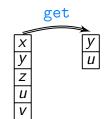


A simple example:

$$\begin{array}{l} \texttt{get} :: [\alpha] \to [\alpha] \\ \texttt{get} [] &= [] \\ \texttt{get} [x] &= [] \\ \texttt{get} (x: y: zs) = y: (\texttt{get} zs) \end{array}$$

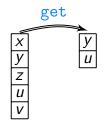
A simple example:

$$\begin{array}{l} \texttt{get} :: [\alpha] \to [\alpha] \\ \texttt{get} [] &= [] \\ \texttt{get} [x] &= [] \\ \texttt{get} (x: y: zs) = y: (\texttt{get} zs) \end{array}$$



A simple example:

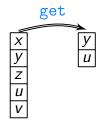
$$\begin{array}{l} \texttt{get} :: [\alpha] \to [\alpha] \\ \texttt{get} [] &= [] \\ \texttt{get} [x] &= [] \\ \texttt{get} (x : y : zs) = y : (\texttt{get} zs) \end{array}$$



One possible backwards transformation:

A simple example:

$$\begin{array}{l} \texttt{get} :: [\alpha] \to [\alpha] \\ \texttt{get} [] &= [] \\ \texttt{get} [x] &= [] \\ \texttt{get} (x : y : zs) = y : (\texttt{get} zs) \end{array}$$



One possible backwards transformation:

#### **Programming Language Approaches**

There has been, and is ongoing, great work in the "lenses" PL/DSLs tradition [Foster et al., ACM TOPLAS'07, ...]. Not covered today.

# **Programming Language Approaches**

There has been, and is ongoing, great work in the "lenses" PL/DSLs tradition [Foster et al., ACM TOPLAS'07, ...]. Not covered today.

We will mention/look at:

- syntactic program transformation
- semantic/type-based transformation
- benefits of higher-order types and abstraction
- search-based program synthesis (if time permits, otherwise see PEPM'12 short paper)

Given

$$\texttt{get}::S\to V$$

Given

$$\texttt{get}::S\to V$$

define a C and

 $\texttt{res}::S\to C$ 

Given

$$\texttt{get}::S\to V$$

define a C and

 $\texttt{res}::S\to C$ 

such that

$$paired = \lambda s \rightarrow (get \ s, res \ s)$$

is injective

Given

$$\texttt{get}::S\to V$$

define a C and

 $\texttt{res}::S\to C$ 

such that

$$paired = \lambda s \rightarrow (get s, res s)$$

is injective and has an inverse  $inv :: (V, C) \rightarrow S$ .

Given

$$\texttt{get}::S\to V$$

define a C and

 $\texttt{res}::S\to C$ 

such that

$$\texttt{paired} = \lambda s \rightarrow (\texttt{get } s, \texttt{res } s)$$

is injective and has an inverse  $inv :: (V, C) \rightarrow S$ .

Then:

put :: 
$$V \rightarrow S \rightarrow S$$
  
put  $v' \ s = inv \ (v', res \ s)$ 

Given

$$\texttt{get}::S\to V$$

define a C and

**res** ::  $S \rightarrow C$ 

such that

$$extsf{paired} = \lambda s 
ightarrow ( extsf{get} \ s, extsf{res} \ s)$$

is injective and has an inverse inv ::  $(V, C) \rightarrow S$ . has to be effective! Then:

put :: 
$$V \rightarrow S \rightarrow S$$
  
put  $v' \ s = inv \ (v', res \ s)$ 

Guarantees "very-well-behavedness":

• put (get s) s = s

• get (put 
$$v' s$$
) =  $v'$ 

• put v'' (put v' s) = put v'' s

Guarantees "very-well-behavedness":

• put (get s) s = s

• get (put 
$$v' s$$
) =  $v'$ 

 $\blacktriangleright \text{ put } v'' \text{ (put } v' \text{ s)} = \text{put } v'' \text{ s}$ 

Example:

 $\begin{array}{l} \texttt{get} :: \mathsf{Nat} \to \mathsf{Nat} \\ \texttt{get} \ n = n \ \texttt{`div'} \ 2 \end{array}$ 

Guarantees "very-well-behavedness":

• put (get s) s = s

• get (put 
$$v' s$$
) =  $v'$ 

 $\blacktriangleright \text{ put } v'' \text{ (put } v' \text{ s)} = \text{put } v'' \text{ s}$ 

Example:

Guarantees "very-well-behavedness":

• put (get s) s = s

• get (put 
$$v' s$$
) =  $v'$ 

 $\blacktriangleright \text{ put } v'' \text{ (put } v' \text{ s)} = \text{put } v'' \text{ s}$ 

Example:

$$\texttt{inv} :: (\mathsf{Nat}, \mathsf{Nat}_2) \to \mathsf{Nat}$$
  
 $\texttt{inv} (\mathbf{v}', \mathbf{c}) = 2 * \mathbf{v}' + \mathbf{c}$ 

Example:

$$inv(v', c) = 2 * v' + c$$

Example:

Then:

put :: Nat 
$$\rightarrow$$
 Nat  $\rightarrow$  Nat  
put  $v' \ s = inv \ (v', res \ s)$ 

Example:

Then:

put :: Nat 
$$\rightarrow$$
 Nat  $\rightarrow$  Nat  
put v'  $s = inv (v', res s)$   
 $= 2 * v' + s \mod 2$ 

Example:

$$\begin{array}{ccc} \texttt{get} :: \mathsf{Nat} \to \mathsf{Nat} & \texttt{res} :: \mathsf{Nat} \to \mathsf{Nat}_2 \\ \texttt{get} & n = n \; \texttt{`div'} \; 2 & \texttt{res} \; n = n \; \texttt{`mod'} \; 2 \\ & \texttt{inv} :: (\mathsf{Nat}, \mathsf{Nat}_2) \to \mathsf{Nat} \\ & \texttt{inv} \; (v', c) = 2 * v' + c \end{array} \; \begin{array}{c} \texttt{other choices} \\ \texttt{possible, and} \\ \texttt{give different} \end{array}$$

Then:

put :: Nat 
$$\rightarrow$$
 Nat  $\rightarrow$  Nat  
put v'  $s = inv (v', res s)$   
 $= 2 * v' + s \mod 2$ 

behavior

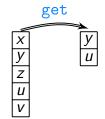
Let:

$$get :: [\alpha] \rightarrow [\alpha]$$
  

$$get [] = []$$
  

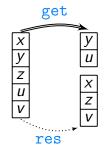
$$get [x] = []$$
  

$$get (x : y : zs) = y : (get zs)$$



Let:

$$\begin{array}{l} \texttt{get} :: [\alpha] \to [\alpha] \\ \texttt{get} [] &= [] \\ \texttt{get} [x] &= [] \\ \texttt{get} (x: y: zs) = y: (\texttt{get} zs) \end{array}$$



A syntactically derived complement function:

$$\begin{array}{l} \texttt{res} \left[ \right] &= \mathsf{C}_1 \\ \texttt{res} \left[ x \right] &= \mathsf{C}_2 \ x \\ \texttt{res} \left( x : y : zs \right) = \mathsf{C}_3 \ x \ (\texttt{res} \ zs) \end{array}$$

A syntactically derived complement function:

$$\begin{array}{l} \operatorname{res} \left[ \right] &= \mathsf{C}_1 \\ \operatorname{res} \left[ x \right] &= \mathsf{C}_2 \ x \\ \operatorname{res} \left( x : y : zs \right) = \mathsf{C}_3 \ x \ (\operatorname{res} \ zs) \end{array}$$

Syntactic pairing:

# Syntactic pairing:

Syntactic inversion:

inv ([] , 
$$C_1$$
) = []  
inv ([] ,  $C_2 x$ ) = [x]  
inv (y: v,  $C_3 x c$ ) = x : y : zs  
where  $zs = inv (v, c)$ 

#### Syntactic inversion:

$$inv([], C_1) = []$$
  
 $inv([], C_2 x) = [x]$   
 $inv(y: v, C_3 x c) = x: y: zs$   
where  $zs = inv(v, c)$ 

Then,

put 
$$v' \ s = inv \ (v', res \ s)$$

#### Syntactic inversion:

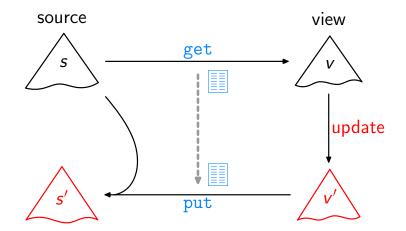
$$inv([], C_1) = []$$
  
 $inv([], C_2 x) = [x]$   
 $inv(y: v, C_3 x c) = x: y: zs$   
where  $zs = inv(v, c)$ 

Then,

put 
$$v' \ s = inv \ (v', res \ s)$$

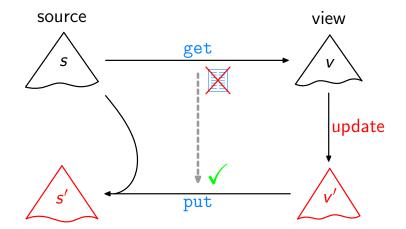
corresponds to (the earlier seen):

## **Automatic Bidirectionalization**



Syntactic Bidirectionalization [Matsuda et al., ICFP'07]

### **Automatic Bidirectionalization**



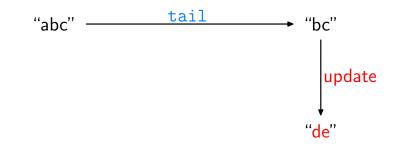
Semantic Bidirectionalization [V., POPL'09]

Aim: Write higher-order function bff<sup>†</sup> such that any get and bff get satisfy GetPut, PutGet, ....

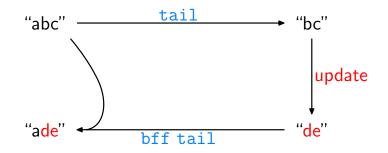
Aim: Write higher-order function bff<sup>†</sup> such that any get and bff get satisfy GetPut, PutGet, ....



Aim: Write higher-order function bff<sup>†</sup> such that any get and bff get satisfy GetPut, PutGet, ....



Aim: Write higher-order function bff<sup>†</sup> such that any get and bff get satisfy GetPut, PutGet, ....



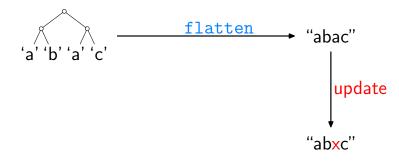
Aim: Write higher-order function bff<sup>†</sup> such that any get and bff get satisfy GetPut, PutGet, ....



### Semantic Bidirectionalization

Aim: Write higher-order function bff<sup>†</sup> such that any get and bff get satisfy GetPut, PutGet, ....

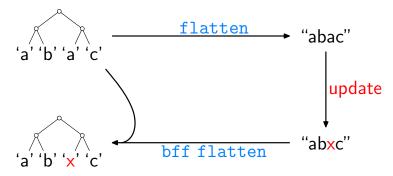
Examples:



#### Semantic Bidirectionalization

Aim: Write higher-order function bff<sup>†</sup> such that any get and bff get satisfy GetPut, PutGet, ....

Examples:



<sup>†</sup> "Bidirectionalization for free!"

Assume we are given some  $\texttt{get} :: [\alpha] \to [\alpha]$  How can we, or <code>bff</code>, analyze it without access to its source code?

Assume we are given some get ::  $[\alpha] \rightarrow [\alpha]$ 

How can we, or **bff**, analyze it without access to its source code?

Idea: How about applying get to some input?

Assume we are given some  $\texttt{get}::[\alpha] \to [\alpha]$ 

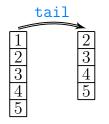
How can we, or **bff**, analyze it without access to its source code?

Idea: How about applying get to some input?

Like:  
get 
$$[1..n] = \begin{cases} [2..n] & \text{if get} = \texttt{tail} \\ [n..1] & \text{if get} = \texttt{reverse} \\ [1..(\min 5 n)] & \text{if get} = \texttt{take 5} \\ \vdots \end{cases}$$

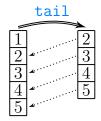
Like:

$$get [1..n] = \begin{cases} [2..n] & \text{if get} = \texttt{tail} \\ [n..1] & \text{if get} = \texttt{reverse} \\ [1..(\texttt{min 5 } n)] & \text{if get} = \texttt{take 5} \\ \vdots \end{cases}$$



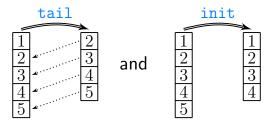
Like:

$$get [1..n] = \begin{cases} [2..n] & \text{if get} = \texttt{tail} \\ [n..1] & \text{if get} = \texttt{reverse} \\ [1..(\texttt{min 5 } n)] & \text{if get} = \texttt{take 5} \\ \vdots \end{cases}$$



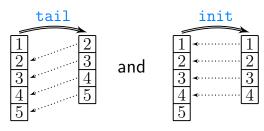
Like:

$$get [1..n] = \begin{cases} [2..n] & \text{if get} = \texttt{tail} \\ [n..1] & \text{if get} = \texttt{reverse} \\ [1..(\texttt{min 5 } n)] & \text{if get} = \texttt{take 5} \\ \vdots \end{cases}$$



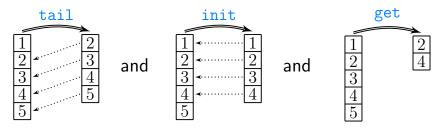
Like:

$$get [1..n] = \begin{cases} [2..n] & \text{if get} = \texttt{tail} \\ [n..1] & \text{if get} = \texttt{reverse} \\ [1..(\texttt{min 5 } n)] & \text{if get} = \texttt{take 5} \\ \vdots \end{cases}$$



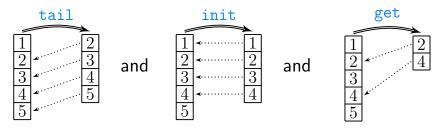
Like:

$$get [1..n] = \begin{cases} [2..n] & \text{if get} = \texttt{tail} \\ [n..1] & \text{if get} = \texttt{reverse} \\ [1..(\texttt{min 5 } n)] & \text{if get} = \texttt{take 5} \\ \vdots \end{cases}$$



Like:

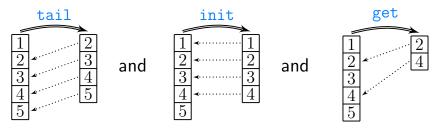
$$get [1..n] = \begin{cases} [2..n] & \text{if get} = \texttt{tail} \\ [n..1] & \text{if get} = \texttt{reverse} \\ [1..(\texttt{min 5 } n)] & \text{if get} = \texttt{take 5} \\ \vdots \end{cases}$$



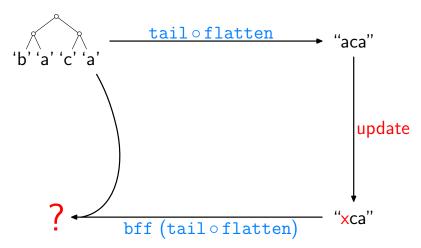
Like:

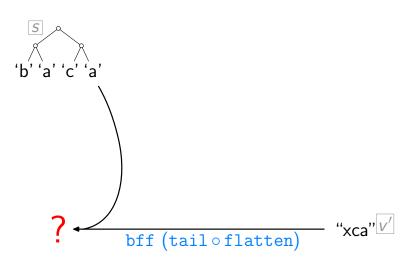
$$get [1..n] = \begin{cases} [2..n] & \text{if get} = \texttt{tail} \\ [n..1] & \text{if get} = \texttt{reverse} \\ [1..(\texttt{min 5 } n)] & \text{if get} = \texttt{take 5} \\ \vdots \end{cases}$$

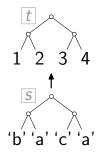
Indeed, this gives us traceability for free:



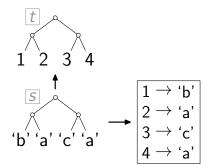
Then transfer the gained insights to arbitrary lists!



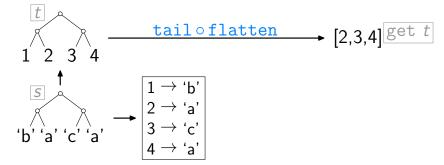




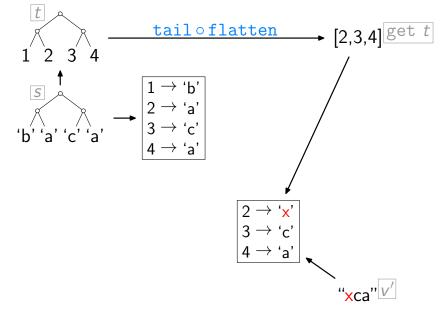


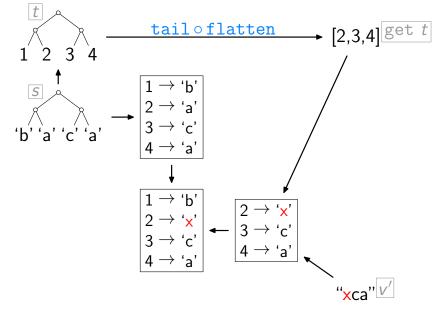


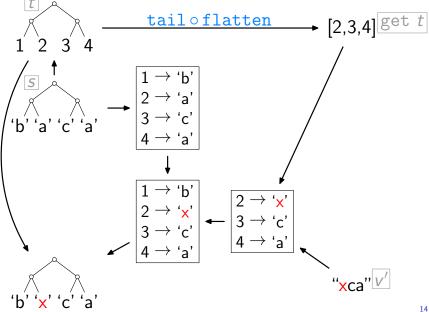


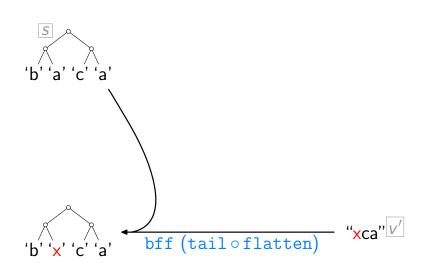


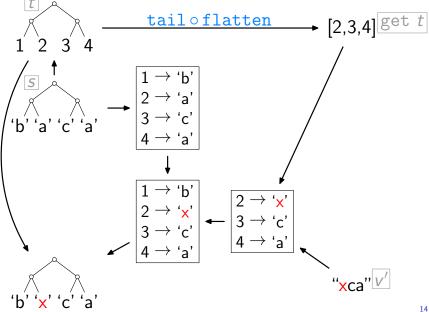












# **Taking Stock of Automatic Bidirectionalization**

[Matsuda et al., ICFP'07]:

- depends on syntactic restraints
- allows (ad-hoc) some shape-changing updates

# **Taking Stock of Automatic Bidirectionalization**

[Matsuda et al., ICFP'07]:

- depends on syntactic restraints
- allows (ad-hoc) some shape-changing updates
- [V., POPL'09]:
  - very lightweight, easy access to bidirectionality
  - essential role: polymorphic function types
  - major problem: rejects shape-changing updates

# Taking Stock of Automatic Bidirectionalization

[Matsuda et al., ICFP'07]:

- depends on syntactic restraints
- allows (ad-hoc) some shape-changing updates
- [V., POPL'09]:
  - very lightweight, easy access to bidirectionality
  - essential role: polymorphic function types
  - major problem: rejects shape-changing updates
- [V. et al., ICFP'10]:
  - synthesis of the two techniques
  - inherits limitations in program coverage from both
  - strictly better in terms of updatability than either

# **References** I

- F. Bancilhon and N. Spyratos.
   Update semantics of relational views.
   ACM Transactions on Database Systems, 6(3):557–575, 1981.
- J.N. Foster, M.B. Greenwald, J.T. Moore, B.C. Pierce, and A. Schmitt.

Combinators for bidirectional tree transformations: A linguistic approach to the view-update problem.

ACM Transactions on Programming Languages and Systems, 29(3):17, 2007.

S. Katayama.

Systematic search for lambda expressions.

In Trends in Functional Programming 2005, Revised Selected Papers, pages 111–126. Intellect, 2007.

# **References II**

E. Kitzelmann and U. Schmid. Inductive synthesis of functional programs: An explanation based generalization approach.

Journal of Machine Learning Research, 7:429–454, 2006.

- K. Matsuda, Z. Hu, K. Nakano, M. Hamana, and M. Takeichi. Bidirectionalization transformation based on automatic derivation of view complement functions. In International Conference on Functional Programming, Proceedings, pages 47–58. ACM Press, 2007.
- J. Voigtländer, Z. Hu, K. Matsuda, and M. Wang. Combining syntactic and semantic bidirectionalization. In International Conference on Functional Programming, Proceedings, pages 181–192. ACM Press, 2010.

# **References III**

## J. Voigtländer.

Bidirectionalization for free!

In Principles of Programming Languages, Proceedings, pages 165-176. ACM Press, 2009.