



Faculty of Science



Modular Tree Automata

Deriving Modular Recursion Schemes from Tree Automata

Patrick Bahr

University of Copenhagen,
Department of Computer Science
paba@diku.dk

11th International Conference on
Mathematics of Program Construction
Madrid, Spain, June 25 - 27, 2012



Goals

Syntax-directed computations on ASTs

- program **analysis**
- complex program **transformations**
- **compiler construction** in general



Goals

Syntax-directed computations on ASTs

- program **analysis**
- complex program **transformations**
- **compiler construction** in general

Desired properties

- extensibility
- modularity
- reusability
- build complex programs by **combining** simple ones



Goals

Syntax-directed computations on ASTs

- program **analysis**
- complex program **transformations**
- **compiler construction** in general

Desired properties

- extensibility
- modularity
- reusability
- build complex programs by **combining** simple ones

Embed the solution into Haskell.



How do we achieve these goals?



How do we achieve these goals?

Locality

simple syntax-directed functions are local in nature



How do we achieve these goals?

Locality

simple syntax-directed functions are local in nature

Compositionality

syntax-directed functions can be combined and composed



How do we achieve these goals?

Locality

simple syntax-directed functions are local in nature

Compositionality

syntax-directed functions can be combined and composed

Contextuality

syntax-directed functions may depend on (the result of) others



How do we achieve these goals?

Locality

simple syntax-directed functions are local in nature

Compositionality

syntax-directed functions can be combined and composed

Contextuality

syntax-directed functions may depend on (the result of) others

- **NB:** This breaks locality and has to be carefully restricted!
- But it is convenient/necessary for
 - ▶ compositionality
 - ▶ expressivity



Locality

Tree automata

- Computation according to a set of **rules**.
- **Applicability** of rules depend only on “**local**” information.
- The **effect** of a rule application is **locally restricted**.



Locality

Tree automata

- Computation according to a set of **rules**.
- **Applicability** of rules depend only on “**local**” information.
- The **effect** of a rule application is **locally restricted**.



$$f(x_1, x_2, \dots, x_n) \longrightarrow t[x_1, x_2, \dots, x_n]$$



Locality

Tree automata

- Computation according to a set of **rules**.
- **Applicability** of rules depend only on “**local**” information.
- The **effect** of a rule application is **locally restricted**.



$$f(q_1(x_1), q_2(x_2), \dots, q_n(x_n)) \quad \longrightarrow \quad q(t[x_1, x_2, \dots, x_n])$$



Compositionality

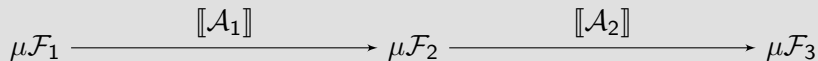
We shall compose tree automata along **3 different dimensions**.



Compositionality

We shall compose tree automata along **3 different dimensions**.

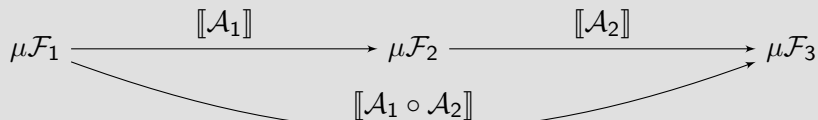
sequential composition: a.k.a. deforestation



Compositionality

We shall compose tree automata along 3 different dimensions.

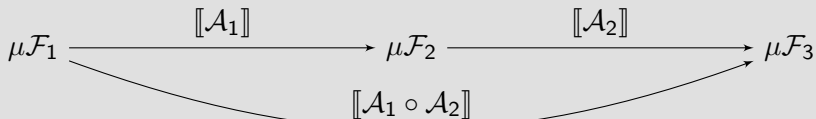
sequential composition: a.k.a. deforestation



Compositionality

We shall compose tree automata along 3 different dimensions.

sequential composition: a.k.a. deforestation



input signature: the type of the AST

$$[[\mathcal{A}_1]] : \mu\mathcal{F} \rightarrow R$$

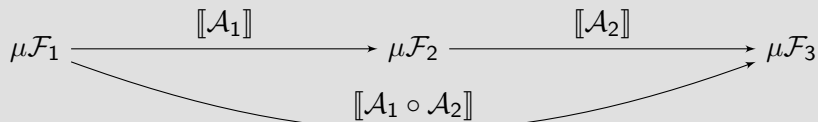
$$[[\mathcal{A}_2]] : \mu\mathcal{G} \rightarrow R$$



Compositionality

We shall compose tree automata along **3 different dimensions**.

sequential composition: a.k.a. deforestation



input signature: the type of the AST

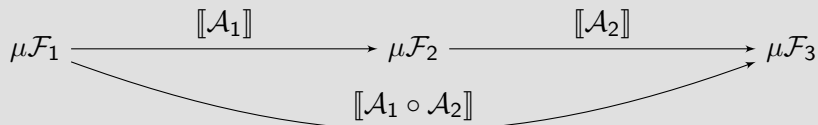
$$\begin{array}{l}
 \llbracket \mathcal{A}_1 \rrbracket : \mu\mathcal{F} \rightarrow R \\
 \llbracket \mathcal{A}_2 \rrbracket : \mu\mathcal{G} \rightarrow R
 \end{array}
 \quad \Longrightarrow \quad
 \llbracket \mathcal{A}_1 + \mathcal{A}_2 \rrbracket : \mu(\mathcal{F} + \mathcal{G}) \rightarrow R$$



Compositionality

We shall compose tree automata along **3 different dimensions**.

sequential composition: a.k.a. deforestation



input signature: the type of the AST

$$\begin{array}{l} [[\mathcal{A}_1]] : \mu\mathcal{F} \rightarrow R \\ [[\mathcal{A}_2]] : \mu\mathcal{G} \rightarrow R \end{array} \quad \Longrightarrow \quad [[\mathcal{A}_1 + \mathcal{A}_2]] : \mu(\mathcal{F} + \mathcal{G}) \rightarrow R$$

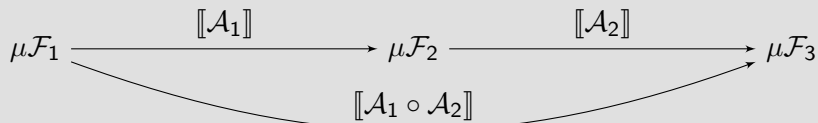
output type: tupling / product automaton construction

$$\begin{array}{l} [[\mathcal{A}_1]] : \mu\mathcal{F} \rightarrow R \\ [[\mathcal{A}_2]] : \mu\mathcal{F} \rightarrow S \end{array}$$

Compositionality

We shall compose tree automata along **3 different dimensions**.

sequential composition: a.k.a. deforestation



input signature: the type of the AST

$$\begin{aligned} [[\mathcal{A}_1]] : \mu\mathcal{F} \rightarrow R \\ [[\mathcal{A}_2]] : \mu\mathcal{G} \rightarrow R \end{aligned} \quad \Longrightarrow \quad [[\mathcal{A}_1 + \mathcal{A}_2]] : \mu(\mathcal{F} + \mathcal{G}) \rightarrow R$$

output type: tupling / product automaton construction

$$\begin{aligned} [[\mathcal{A}_1]] : \mu\mathcal{F} \rightarrow R \\ [[\mathcal{A}_2]] : \mu\mathcal{F} \rightarrow S \end{aligned} \quad \Longrightarrow \quad [[\mathcal{A}_1 \times \mathcal{A}_2]] : \mu\mathcal{F} \rightarrow R \times S$$

Contextuality

tupling / product automaton construction

$$\begin{array}{l} \llbracket \mathcal{A}_1 \rrbracket : \mu\mathcal{F} \rightarrow R \\ \llbracket \mathcal{A}_2 \rrbracket : \mu\mathcal{F} \rightarrow S \end{array} \quad \Longrightarrow \quad \llbracket \mathcal{A}_1 \times \mathcal{A}_2 \rrbracket : \mu(\mathcal{F}) \rightarrow R \times S$$



Contextuality

tupling / product automaton construction

$$\begin{array}{l} \mathcal{A}_1 : \mathcal{F} \rightarrow R \\ \mathcal{A}_2 : \mathcal{F} \rightarrow S \end{array} \quad \Longrightarrow \quad \mathcal{A}_1 \times \mathcal{A}_2 : \mathcal{F} \rightarrow R \times S$$



Contextuality

tupling / product automaton construction

$$\begin{array}{l} \mathcal{A}_1 : \mathcal{F} \rightarrow R \\ \mathcal{A}_2 : \mathcal{F} \rightarrow S \end{array} \quad \Longrightarrow \quad \mathcal{A}_1 \times \mathcal{A}_2 : \mathcal{F} \rightarrow R \times S$$

mutumorphisms / dependent product automata

$$\begin{array}{l} \mathcal{A}_1 : \mathcal{F} \rightarrow R \\ \mathcal{A}_2 : R \Rightarrow \mathcal{F} \rightarrow S \end{array}$$



Contextuality

tupling / product automaton construction

$$\begin{array}{l} \mathcal{A}_1 : \mathcal{F} \rightarrow R \\ \mathcal{A}_2 : \mathcal{F} \rightarrow S \end{array} \quad \Longrightarrow \quad \mathcal{A}_1 \times \mathcal{A}_2 : \mathcal{F} \rightarrow R \times S$$

mutumorphisms / dependent product automata

$$\begin{array}{l} \mathcal{A}_1 : \mathcal{F} \rightarrow R \\ \mathcal{A}_2 : R \Rightarrow \mathcal{F} \rightarrow S \end{array} \quad \Longrightarrow \quad \mathcal{A}_1 \times \mathcal{A}_2 : \mathcal{F} \rightarrow R \times S$$



Contextuality

tupling / product automaton construction

$$\begin{array}{l} \mathcal{A}_1 : \mathcal{F} \rightarrow R \\ \mathcal{A}_2 : \mathcal{F} \rightarrow S \end{array} \quad \Longrightarrow \quad \mathcal{A}_1 \times \mathcal{A}_2 : \mathcal{F} \rightarrow R \times S$$

mutomorphisms / dependent product automata

$$\begin{array}{l} \mathcal{A}_1 : S \Rightarrow \mathcal{F} \rightarrow R \\ \mathcal{A}_2 : R \Rightarrow \mathcal{F} \rightarrow S \end{array} \quad \Longrightarrow \quad \mathcal{A}_1 \times \mathcal{A}_2 : \mathcal{F} \rightarrow R \times S$$



Outline

- 1 Introduction
- 2 State Transition Functions
 - Composing State Spaces
 - Compositional Signatures
- 3 Tree Transducers
 - Bottom-Up Tree Transducers
 - Decomposing Tree Transducers
- 4 Conclusions



Terms in Haskell

Data types as fixed points of functors

```
data Term f = In (f (Term f))
```



Terms in Haskell

Data types as fixed points of functors

```
data Term f = In (f (Term f))
```

Functors

```
class Functor f where
```

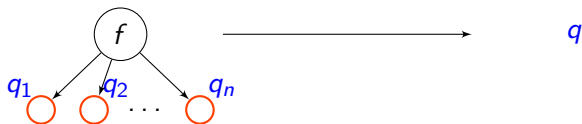
```
  fmap :: (a → b) → f a → f b
```



Bottom-Up State Transitions in Haskell



Bottom-Up State Transitions in Haskell



Bottom-Up State Transitions in Haskell



Bottom-Up State Transitions in Haskell

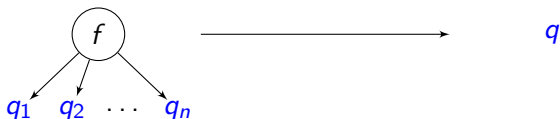


Bottom-up state transition rules as algebras

```
type UpState f q = f q → q
```



Bottom-Up State Transitions in Haskell



Bottom-up state transition rules as algebras

type $UpState\ f\ q = f\ q \rightarrow q$

$runUpState\ ::\ Functor\ f \Rightarrow UpState\ f\ q \rightarrow Term\ f \rightarrow q$

$runUpState\ \phi\ (In\ t) = \phi\ (fmap\ (runUpState\ \phi)\ t)$



Bottom-Up State Transitions in Haskell



Bottom-up state transitions a.k.a. catamorphism / fold

type $UpState\ f\ q = f\ q \rightarrow q$

$runUpState :: Functor\ f \Rightarrow UpState\ f\ q \rightarrow Term\ f \rightarrow q$

$runUpState\ \phi\ (In\ t) = \phi\ (fmap\ (runUpState\ \phi)\ t)$



Composing State Spaces – Motivating Example

A simple expression language

data *Sig* $e = \text{Val Int} \mid \text{Plus } e e$



Composing State Spaces – Motivating Example

A simple expression language

data *Sig e = Val Int | Plus e e*

Task: writing a code generator

type *Addr = Int*

data *Instr = Acc Int | Load Addr | Store Addr | Add Addr*

type *Code = [Instr]*



Composing State Spaces – Motivating Example

A simple expression language

```
data Sig e = Val Int | Plus e e
```

Task: writing a code generator

```
type Addr = Int
```

```
data Instr = Acc Int | Load Addr | Store Addr | Add Addr
```

```
type Code = [Instr]
```

The problem

```
codeSt :: UpState Sig Code
```

```
codeSt (Val i)    = [Acc i]
```

```
codeSt (Plus x y) = x ++ [Store a] ++ y ++ [Add a]
```

```
  where a = ...
```



Composing State Spaces – Motivating Example

A simple expression language

```
data Sig e = Val Int | Plus e e
```

Task: writing a code generator

```
type Addr = Int
```

```
data Instr = Acc Int | Load Addr | Store Addr | Add Addr
```

```
type Code = [Instr]
```

Sig Code \rightarrow Code

The problem

```
codeSt :: UpState Sig Code
```

```
codeSt (Val i) = [Acc i]
```

```
codeSt (Plus x y) = x ++ [Store a] ++ y ++ [Add a]
```

```
  where a = ...
```

Tupling

Tuple the code with an address counter

$codeAddrSt :: UpState \text{ Sig } (Code, Addr)$

$codeAddrSt (Val i) = ([Acc i], 0)$

$codeAddrSt (Plus (x, a') (y, a)) = (x \# [Store a] \# y \# [Add a],$
 $1 + \max a a')$



Tupling

Tuple the code with an address counter

$codeAddrSt :: UpState Sig (Code, Addr)$

$codeAddrSt (Val i) = ([Acc i], 0)$

$codeAddrSt (Plus (x, a') (y, a)) = (x \# [Store a] \# y \# [Add a],$
 $1 + \max a a')$

Run the automaton

$code :: Term Sig \rightarrow (Code, Addr)$

$code = runUpState codeAddrSt$



Tupling

Tuple the code with an address counter

$$\begin{aligned}
 \text{codeAddrSt} &:: \text{UpState Sig} \ (\text{Code}, \text{Addr}) \\
 \text{codeAddrSt} \ (\text{Val } i) &= ([\text{Acc } i], 0) \\
 \text{codeAddrSt} \ (\text{Plus } (x, a') \ (y, a)) &= (x \# \# [\text{Store } a] \# \# y \# \# [\text{Add } a], \\
 &\quad 1 + \max a \ a')
 \end{aligned}$$

Run the automaton

$$\begin{aligned}
 \text{code} &:: \text{Term Sig} \rightarrow (\text{Code}, \text{Addr}) \\
 \text{code} &= \text{fst} \ . \ \text{runUpState} \ \text{codeAddrSt}
 \end{aligned}$$


Tupling

Tuple the code with an address counter

```
codeAddrSt :: UpState Sig (Code, Addr)
codeAddrSt (Val i)           = ([Acc i], 0)
codeAddrSt (Plus (x, a') (y, a)) = (x ++ [Store a] ++ y ++ [Add a],
                                     1 + max a a')
```

Run the automaton

```
code :: Term Sig → Code
code = fst . runUpState codeAddrSt
```



Product Automata

Deriving projections

class $a \in b$ **where**

$pr :: b \rightarrow a$



Product Automata

Deriving projections

class $a \in b$ **where**

$pr :: b \rightarrow a$

$a \in b$ iff

- b is of the form $(b_1, (b_2, \dots b_n))$ and
- $a = b_i$ for some i



Product Automata

Deriving projections

class $a \in b$ **where**

$pr :: b \rightarrow a$

$a \in b$ iff

- b is of the form $(b_1, (b_2, \dots b_n))$ and
- $a = b_i$ for some i

For example: $Addr \in (Code, Addr)$



Product Automata

Deriving projections

class $a \in b$ **where**

$pr :: b \rightarrow a$

$a \in b$ iff

- b is of the form $(b_1, (b_2, \dots b_n))$ and
- $a = b_i$ for some i

For example: $Addr \in (Code, Addr)$

Dependent state transition functions

type $UpState$ f $q =$

f $q \rightarrow q$



Product Automata

Deriving projections

class $a \in b$ **where**

$pr :: b \rightarrow a$

$a \in b$ iff

- b is of the form $(b_1, (b_2, \dots b_n))$ and
- $a = b_i$ for some i

For example: $Addr \in (Code, Addr)$

Dependent state transition functions

type $UpState$ f $q =$ f $q \rightarrow q$

type $DUpState$ f p $q = (q \in p) \Rightarrow f$ $p \rightarrow q$



Product Automata

Deriving projections

class $a \in b$ **where**

$pr :: b \rightarrow a$

$a \in b$ iff

- b is of the form $(b_1, (b_2, \dots b_n))$ and
- $a = b_i$ for some i

For example: $Addr \in (Code, Addr)$

Dependent state transition functions

type $UpState$ f $q =$ f $q \rightarrow q$

type $DUpState$ f p $q = (q \in p) \Rightarrow f$ $p \rightarrow q$



Product Automata

Deriving projections

class $a \in b$ **where** $a \in b$ iff

$pr :: b \rightarrow a$

- b is of the form $(b_1, (b_2, \dots b_n))$ and
- $a = b_i$ for some i

For example: $Addr \in (Code, Addr)$

Dependent state transition functions

type $UpState$ f $q =$ f $q \rightarrow q$

type $DUpState$ f p $q = (q \in p) \Rightarrow f$ $p \rightarrow q$



Product Automata

Deriving projections

class $a \in b$ **where** $a \in b$ iff

$pr :: b \rightarrow a$

- b is of the form $(b_1, (b_2, \dots b_n))$ and
- $a = b_i$ for some i

For example: $Addr \in (Code, Addr)$

Dependent state transition functions

type $UpState\ f\ q = f\ q \rightarrow q$

type $DUpState\ f\ p\ q = (q \in p) \Rightarrow f\ p \rightarrow q$

Product state transition

$(\otimes) :: (p \in c, q \in c) \Rightarrow DUpState\ f\ c\ p \rightarrow DUpState\ f\ c\ q$
 $\rightarrow DUpState\ f\ c\ (p, q)$

$(sp \otimes sq)\ t = (sp\ t, sq\ t)$

Running Dependent State Transition Functions

The types

type *UpState* $f\ q = f\ q \rightarrow q$

type *DUpState* $f\ p\ q = (q \in p) \Rightarrow f\ p \rightarrow q$



Running Dependent State Transition Functions

The types

type *UpState* $f\ q = f\ q \rightarrow q$

type *DUpState* $f\ p\ q = (q \in p) \Rightarrow f\ p \rightarrow q$

Running dependent state transitions

runDUpState $:: \text{Functor } f \Rightarrow \text{DUpState } f\ q\ q \rightarrow \text{Term } f \rightarrow q$

runDUpState $f = \text{runUpState } f$



Running Dependent State Transition Functions

The types

type *UpState* $f\ q = f\ q \rightarrow q$

type *DUpState* $f\ p\ q = (q \in p) \Rightarrow f\ p \rightarrow q$

Running dependent state transitions

runDUpState $:: \text{Functor } f \Rightarrow \text{DUpState } f\ q\ q \rightarrow \text{Term } f \rightarrow q$

runDUpState $f = \text{runUpState } f$

From state transition to dependent state transition

dUpState $:: \text{Functor } f \Rightarrow \text{UpState } f\ q \rightarrow \text{DUpState } f\ p\ q$

dUpState $st = st . \text{fmap } pr$



The Code Generator Example

The code generator

```
codeSt :: (Int ∈ q) ⇒ DUpState Sig q Code
codeSt (Val i)    = [Acc i]
codeSt (Plus x y) = pr x ++ [Store a] ++ pr y ++ [Add a]
  where a = pr y
```



The Code Generator Example

The code generator

```
codeSt :: (Int ∈ q) ⇒ DUpState Sig q Code
codeSt (Val i)    = [Acc i]
codeSt (Plus x y) = pr x ++ [Store a] ++ pr y ++ [Add a]
  where a = pr y
```

Generating fresh addresses

```
heightSt :: UpState Sig Int
heightSt (Val _)    = 0
heightSt (Plus x y) = 1 + max x y
```



The Code Generator Example

The code generator

```

codeSt :: (Int ∈ q) ⇒ DUpState Sig q Code
codeSt (Val i)    = [Acc i]
codeSt (Plus x y) = pr x ++ [Store a] ++ pr y ++ [Add a]
  where a = pr y
  
```

Generating fresh addresses

```

heightSt :: UpState Sig Int
heightSt (Val _)    = 0
heightSt (Plus x y) = 1 + max x y
  
```

Combining the components

```

code :: Term Sig → Code
code = fst . runUpState (codeSt ⊗ dUpState heightSt)
  
```

The Code Generator Example

The code generator

```
codeSt :: (Int ∈ q) ⇒ DUpState Sig q Code
codeSt (Val i)    = [Acc i]
codeSt (Plus x y) = pr x ++ [Store a] ++ pr y ++ [Add a]
  where a = pr y
```

Generating fresh addresses

```
heightSt :: UpState Sig Int
heightSt (Val _)    = 0
heightSt (Plus x y) = 1 + max (heightSt x) (heightSt y)
```

$(Code \in q, Int \in q) \Rightarrow DUpState Sig q (Code, Int)$

Combining the components

```
code :: Term Sig → Code
code = fst . runUpState (codeSt ⊗ dUpState heightSt)
```


Outline

- 1 Introduction
- 2 State Transition Functions
 - Composing State Spaces
 - Compositional Signatures
- 3 Tree Transducers
 - Bottom-Up Tree Transducers
 - Decomposing Tree Transducers
- 4 Conclusions



Combining Signatures

Signatures & automata may be combined in the style of “Data types à la carte” [Swierstra 2008].



Combining Signatures

Signatures & automata may be combined in the style of “Data types à la carte” [Swierstra 2008].

Coproduct of signatures

data $(f \oplus g) e = \text{Inl } (f e) \mid \text{Inr } (g e)$



Combining Signatures

Signatures & automata may be combined in the style of “Data types à la carte” [Swierstra 2008].

Coproduct of signatures

data $(f \oplus g) e = \text{Inl } (f e) \mid \text{Inr } (g e)$

Example

data $\text{Inc } e = \text{Inc } e$

type $\text{Sig}' = \text{Inc} \oplus \text{Sig}$



Combining Signatures

Signatures & automata may be combined in the style of “Data types à la carte” [Swierstra 2008].

Coproduct of signatures

```
data  $(f \oplus g)$  e = Inl (f e) | Inr (g e)
```

Example

```
data Inc e = Inc e
type Sig' = Inc  $\oplus$  Sig
```

Subsignature type class

```
class  $f \preceq g$  where
  inj :: f a  $\rightarrow$  g a
```

Combining Signatures

Signatures & automata may be combined in the style of “Data types à la carte” [Swierstra 2008].

Coproduct of signatures

data $(f \oplus g) e = \text{Inl } (f e) \mid \text{Inr } (g e)$

Example

data $\text{Inc } e = \text{Inc } e$
type $\text{Sig}' = \text{Inc} \oplus \text{Sig}$

Subsignature type class

class $f \preceq g$ **where**
 $\text{inj} :: f a \rightarrow g a$

$f \preceq g$ iff

- $g = g_1 \oplus g_2 \oplus \dots \oplus g_n$ and
- $f = g_i, \quad 0 < i \leq n$

Combining Signatures

Signatures & automata may be combined in the style of “Data types à la carte” [Swierstra 2008].

Coproduct of signatures

data $(f \oplus g) e = \text{Inl } (f e) \mid \text{Inr } (g e)$

Example

data $\text{Inc } e = \text{Inc } e$

type $\text{Sig}' = \text{Inc} \oplus \text{Sig}$

Subsignature type class

class $f \preceq g$ **where**

$\text{inj} :: f a \rightarrow g a$

For example: $\text{Inc} \preceq \text{Sig}'$

$f \preceq g$ iff

- $g = g_1 \oplus g_2 \oplus \dots \oplus g_n$ and
- $f = g_i, \quad 0 < i \leq n$

Combining Automata

Making the height compositional

class *HeightSt* *f* **where**

heightSt :: *DUpState f q Int*

instance (*HeightSt f*, *HeightSt g*) \Rightarrow *HeightSt (f \oplus g)* **where**

heightSt (Inl x) = *heightSt x*

heightSt (Inr x) = *heightSt x*



Combining Automata

Making the height compositional

class *HeightSt* *f* **where**

heightSt :: *DUpState f q Int*

instance (*HeightSt f*, *HeightSt g*) \Rightarrow *HeightSt* (*f* \oplus *g*) **where**

heightSt (*Inl* *x*) = *heightSt* *x*

heightSt (*Inr* *x*) = *heightSt* *x*

Defining the height on Sig

instance *HeightSt Sig* **where**

heightSt (*Val* _) = 0

heightSt (*Plus* *x y*) = 1 + *max* *x y*



Combining Automata

Making the height compositional

class *HeightSt* *f* **where**

heightSt :: *DUpState f q Int*

instance (*HeightSt f*, *HeightSt g*) \Rightarrow *HeightSt (f \oplus g)* **where**

heightSt (Inl x) = *heightSt x*

heightSt (Inr x) = *heightSt x*

Defining the height on Sig

instance *HeightSt Sig* **where**

heightSt (Val _) = 0

heightSt (Plus x y) = 1 + max x y

Defining the height on Inc

instance *HeightSt Inc* **where**

heightSt (Inc x) = 1 + x

Outline

- 1 Introduction
- 2 State Transition Functions
 - Composing State Spaces
 - Compositional Signatures
- 3 **Tree Transducers**
 - **Bottom-Up Tree Transducers**
 - Decomposing Tree Transducers
- 4 Conclusions



Bottom-Up Tree Transducers



Bottom-Up Tree Transducers



Bottom-Up Tree Transducers



From terms to contexts

data *Term* $f = \text{In}(f(\text{Term } f))$

data *Context* $f a = \text{In}(f(\text{Context } f a)) \mid \text{Hole } a$



Bottom-Up Tree Transducers



type *Term* *f* = *Context* *f* *Empty*

From terms to contexts

data *Term* *f* = *In* (*f* (*Term* *f*))

data *Context* *f* *a* = *In* (*f* (*Context* *f* *a*)) | *Hole* *a*



Bottom-Up Tree Transducers



From terms to contexts

data *Term* $f = \text{In} (f (\text{Term } f))$

data *Context* $f a = \text{In} (f (\text{Context } f a)) \mid \text{Hole } a$

Representing transduction rules, [Hasuo et al. 2007]

type *UpTrans* $f q g = \forall a. f (q, a) \rightarrow (q, \text{Context } g a)$

Bottom-Up Tree Transducers



From terms to contexts

data *Term* $f = \text{In} (f (Term \ f \))$

data *Context* $f \ a = \text{In} (f (Context \ f \ a)) \mid \text{Hole } a$

Representing transduction rules, [Hasuo et al. 2007]

type *UpTrans* $f \ q \ g = \forall \ a. f (q, a) \rightarrow (q, Context \ g \ a)$

Tree Homomorphisms

type $UpTrans\ f\ q\ g = \forall a . f\ (q, a) \rightarrow (q, Context\ g\ a)$



Tree Homomorphisms

type *UpTrans* *f* *g* = $\forall a . f \quad a \rightarrow \quad$ *Context* *g* *a*



Tree Homomorphisms

type *Hom* f $g = \forall a . f \quad a \rightarrow \text{Context } g \ a$



Tree Homomorphisms

type $\text{Hom } f \ g = \forall a . f \ a \rightarrow \text{Context } g \ a$

Example (Desugaring)

class $\text{DesugHom } f \ g$ **where**

$\text{desugHom} :: \text{Hom } f \ g$

$\text{desugar} :: (\text{Functor } f, \text{Functor } g, \text{DesugHom } f \ g) \Rightarrow \text{Term } f \rightarrow \text{Term } g$

$\text{desugar} = \text{runHom } \text{desugHom}$



Tree Homomorphisms

type *Hom* f $g = \forall a . f \quad a \rightarrow \text{Context } g \quad a$

Example (Desugaring)

class *DesugHom* f g **where**

desugHom :: *Hom* f g

desugar :: (*Functor* f , *Functor* g , *DesugHom* f g) \Rightarrow *Term* $f \rightarrow$ *Term* g

desugar = *runHom* *desugHom*

instance (*Sig* \preceq g) \Rightarrow *DesugHom* *Inc* g **where**

desugHom (*Inc* x) = *Hole* x 'plus' *val* 1

instance (*Functor* g , $f \preceq$ g) \Rightarrow *DesugHom* f g **where**

desugHom = *simpCxt* . *inj*



Tree Homomorphisms

type *Hom* *f g* = $\forall a . f \ a \rightarrow \text{Context } g \ a$

Example (Desugaring)

class *DesugHom* *f g* **where**

desugHom :: *Hom* *f g*

desugar :: (*Functor* *f*, *Functor* *g*, *DesugHom* *f g*) \Rightarrow *Term* *f* \rightarrow *Term* *g*

desugar = *runHom* *desugHom*

instance (*Sig* \preceq *g*) = *simpCxt* :: *Functor* *g* \Rightarrow *g* *a* \rightarrow *Context* *g* *a*
desugHom (*Inc* *x*) = *simpCxt* *t* = *In* (*fmap* *Hole* *t*)

instance (*Functor* *g*, *f* \preceq *g*) \Rightarrow *DesugHom* *f g* **where**

desugHom = *simpCxt* . *inj*



Stateful Tree Homomorphisms

Decomposing tree transducers

type *Hom* $f \ g = \forall a . f \ a \rightarrow \text{Context } g \ a$

type *UpState* $f \ q = f \ q \rightarrow q$

type *UpTrans* $f \ q \ g = \forall a . f \ (q, a) \rightarrow (q, \text{Context } g \ a)$



Stateful Tree Homomorphisms

Decomposing tree transducers

type $Hom\ f\ g = \forall a. f\ a \rightarrow Context\ g\ a$

type $UpState\ f\ q = f\ q \rightarrow q$

type $UpTrans\ f\ q\ g = \forall a. f\ (q, a) \rightarrow (q, Context\ g\ a)$

Making homomorphisms dependent on a state

type $QHom\ f\ q\ g = \forall a. f\ a \rightarrow Context\ g\ a$



Stateful Tree Homomorphisms

Decomposing tree transducers

type $Hom\ f\ g = \forall a. f\ a \rightarrow Context\ g\ a$

type $UpState\ f\ q = f\ q \rightarrow q$

type $UpTrans\ f\ q\ g = \forall a. f\ (q, a) \rightarrow (q, Context\ g\ a)$

Making homomorphisms dependent on a state

type $QHom\ f\ q\ g = \forall a. f\ (q, a) \rightarrow Context\ g\ a$



Stateful Tree Homomorphisms

Decomposing tree transducers

type $Hom\ f\ g = \forall a. f\ a \rightarrow Context\ g\ a$

type $UpState\ f\ q = f\ q \rightarrow q$

type $UpTrans\ f\ q\ g = \forall a. f\ (q, a) \rightarrow (q, Context\ g\ a)$

Making homomorphisms dependent on a state

type $QHom\ f\ q\ g = \forall a. (a \rightarrow q) \rightarrow f\ a \rightarrow Context\ g\ a$



Stateful Tree Homomorphisms

Decomposing tree transducers

type $Hom\ f\ g = \forall a. f\ a \rightarrow Context\ g\ a$

type $UpState\ f\ q = f\ q \rightarrow q$

type $UpTrans\ f\ q\ g = \forall a. f\ (q, a) \rightarrow (q, Context\ g\ a)$

Making homomorphisms dependent on a state

type $QHom\ f\ q\ g = \forall a. (a \rightarrow q) \rightarrow f\ a \rightarrow Context\ g\ a$

From stateful homomorphisms to tree transducers

$upTrans :: (Functor\ f, Functor\ g) \Rightarrow$

$UpState\ f\ q \rightarrow QHom\ f\ q\ g \rightarrow UpTrans\ f\ q\ g$

$upTrans\ st\ hom\ t = (q, c)$ **where**

$q = st\ (fmap\ fst\ t)$

$c = fmap\ snd\ (hom\ fst\ t)$

An Example

Extending the signature with let bindings

type *Name* = *String*

data *Let e* = *LetIn Name e e* | *Var Name*

type *LetSig* = *Let* \oplus *Sig*



An Example

Extending the signature with let bindings

```
type Name = String
data Let e = LetIn Name e e | Var Name
type LetSig = Let  $\oplus$  Sig
```

```
type Vars = Set Name
class FreeVarsSt f where
  freeVarsSt :: UpState f Vars
```

An Example

Extending the signature with let bindings

```
type Name = String
data Let e = LetIn Name e e | Var Name
type LetSig = Let  $\oplus$  Sig
```

```
type Vars = Set Name
class FreeVarsSt f where
  freeVarsSt :: UpState f Vars
instance FreeVarsSt Sig where
  freeVarsSt (Plus x y) = x 'union' y
  freeVarsSt (Val _)    = empty
instance FreeVarsSt Let where
  freeVarsSt (Var v)      = singleton v
  freeVarsSt (LetIn v e s) = if v 'member' s then e 'union' delete v s
                               else s
```

An Example (Cont'd)

class *RemLetHom* *f q g* **where**

remLetHom :: *QHom f q g*

instance (*Vars* ∈ *q*, *Let* ≼ *g*, *Functor g*) ⇒ *RemLetHom Let q g* **where**

remLetHom qOf (*LetIn v _ s*) | ¬ (*v* 'member' *qOf s*) = *Hole s*

remLetHom _ t = *simpCxt (inj t)*

instance (*Functor f*, *Functor g*, *f* ≼ *g*) ⇒ *RemLetHom f q g* **where**

remLetHom _ = *simpCxt . inj*



An Example (Cont'd)

class *RemLetHom* *f q g* **where**

remLetHom :: *QHom f q g*

instance (*Vars* \in *q*, *Let* \preceq *g*, *Functor g*) \Rightarrow *RemLetHom Let q g* **where**

remLetHom qOf (*LetIn* *v _ s*) | \neg (*v* 'member' *qOf s*) = *Hole s*

remLetHom _ t = *simpCxt (inj t)*

instance (*Functor f*, *Functor g*, *f* \preceq *g*) \Rightarrow *RemLetHom f q g* **where**

remLetHom _ = *simpCxt . inj*

Combining state transition and homomorphism

remLet :: (*Functor f*, *FreeVarsSt f*, *RemLetHom f Vars f*)

\Rightarrow *Term f* \rightarrow (*Vars*, *Term f*)

remLet = *runUpHom freeVarsSt remLetHom*



An Example (Cont'd)

class *RemLetHom* *f q g* **where**

remLetHom :: *QHom f q g*

instance (*Vars* ∈ *q*, *Let* ≤ *g*, *Functor g*) ⇒ *RemLetHom Let q g* **where**

remLetHom qOf (*LetIn v _ s*) | ¬ (*v* 'member' *qOf s*) = *Hole s*

remLetHom _ *t* = *simpCxt* (*inj t*)

instance (*Functor f*, *Functor g*, *f* ≤ *g*) ⇒ *RemLetHom f q g* **where**

remLetHom _

runUpHom :: *UpState f q* → *QHom f q g*

→ *Term f* → *Term g*

runUpHom st hom = *runUpTrans* (*upTrans st hom*)

Combining state

remLet :: (*Functor f*, *freeVars f*, *remLetHom f q g*)

⇒ *Term f* → (*Vars*, *Term f*)

remLet = *runUpHom freeVarsSt remLetHom*



An Example (Cont'd)

class *RemLetHom* *f q g* **where**

remLetHom :: *QHom f q g*

instance (*Vars* ∈ *q*, *Let* ≤ *g*, *Functor g*) ⇒ *RemLetHom Let q g* **where**

remLetHom qOf (*LetIn v _ s*) | ¬ (*v* 'member' *qOf s*) = *Hole s*

remLetHom _ t = *simpCxt (inj t)*

instance (*Functor f*, *Functor g*, *f* ≤ *g*) ⇒ *RemLetHom f q g* **where**

remLetHom _ = *simpCxt . inj*

Combining state transition and homomorphism

remLet :: (*Functor f*, *FreeVarsSt f*, *RemLetHom f Vars f*)

⇒ *Term f* → (*Vars*, *Term f*)

remLet = *runUpHom freeVarsSt remLetHom*

remLet :: *Term LetSig* → *Term LetSig*

remLet :: *Term (Inc ⊕ LetSig)* → *Term (Inc ⊕ LetSig)*



Beyond Bottom-Up Tree Automata

What have we seen?

- Bottom-up tree acceptors (a.k.a. folds)
- Bottom-up tree transducers
- “dependent” versions thereof



Beyond Bottom-Up Tree Automata

What have we seen?

- Bottom-up tree acceptors (a.k.a. folds)
- Bottom-up tree transducers
- “dependent” versions thereof

Other Tree recursion schemes

- Top-down tree acceptors
- Top-down tree transducers
- “dependent” versions thereof



Beyond Bottom-Up Tree Automata

What have we seen?

- Bottom-up tree acceptors (a.k.a. folds)
- Bottom-up tree transducers
- “dependent” versions thereof

Other Tree recursion schemes

- Top-down tree acceptors
- Top-down tree transducers
- “dependent” versions thereof
- automata with bidirectional state propagation
- (restricted versions of macro tree transducers)



What have we gained?

Modularity & Reusability

- modularity along **three dimensions** (signature, sequential composition, state space)
- **decoupling** of state propagation and tree transformation
- **operations on automata** (beyond product & sum) allow us to construct new automata from old ones



What have we gained?

Modularity & Reusability

- modularity along **three dimensions** (signature, sequential composition, state space)
- **decoupling** of state propagation and tree transformation
- **operations on automata** (beyond product & sum) allow us to construct new automata from old ones

Interface between tree automata

- dependencies between automata by constraints on the state space
- modularity allows us to replace individual components



Try It Out!

This is part of the **compositional data types** Haskell library `compdata`:

```
> cabal install compdata
```

<http://hackage.haskell.org/package/compdata>

