# Generalizing Generic Functions

Andres Löh

7 July 2004



#### Motivation

Despite "dependency style" Generic Haskell, generic functions have a number of restrictions:

- only one type argument
- no higher-order type-indexed functions
- only flat type patterns
- complicated types for generic functions of higher arity
- no inference of type arguments



#### Motivation

Despite "dependency style" Generic Haskell, generic functions have a number of restrictions:

- only one type argument
- no higher-order type-indexed functions
- only flat type patterns
- complicated types for generic functions of higher arity
- no inference of type arguments

Not all of the restrictions pose difficult problems, but all of them are "remaining work".



#### Motivation

Despite "dependency style" Generic Haskell, generic functions have a number of restrictions:

- only one type argument
- no higher-order type-indexed functions
- only flat type patterns
- complicated types for generic functions of higher arity
- no inference of type arguments

Not all of the restrictions pose difficult problems, but all of them are "remaining work".

Type classes (+extensions) solve many of these problems. Arjan has shown how to encode "dependency style" using type classes.



#### Getting Rid of Type Classes

Andres Löh

7 July 2004



#### More motivation

There are many similarities between type classes and type-indexed functions.

But type-indexed functions are better because:

- Type classes create a separate programming language on top of Haskell.
- Type classes seem to have the need of several extensions to acquire their full power.
- ► Type classes are not first-class either. They are "fixed".
- ► Type classes force implicit passing of dictionaries.



#### Long-term goals

- Extend Haskell language with a type abstraction and type application construct, and a typecase.
- Type-indexed types take the role of functional dependencies.
- Type system and translation are similar to "dependency style" and type classes: use of qualified types, dictionary passing.
- ► Type arguments can be inferred in special cases.
- Type arguments can always be specified explicitly.
- Typecases can be open and closed.
- ► Type-indexed functions are first class.



#### Long-term goals

- Extend Haskell language with a type abstraction and type application construct, and a typecase.
- Type-indexed types take the role of functional dependencies.
- Type system and translation are similar to "dependency style" and type classes: use of qualified types, dictionary passing.
- ► Type arguments can be inferred in special cases.
- Type arguments can always be specified explicitly.
- Typecases can be open and closed.
- ► Type-indexed functions are first class.
- ▶ Generic functions come (almost) for free.



#### Long-term goals

- Extend Haskell language with a type abstraction and type application construct, and a typecase.
- Type-indexed types take the role of functional dependencies.
- Type system and translation are similar to "dependency style" and type classes: use of qualified types, dictionary passing.
- Type arguments can be inferred in special cases.
- ► Type arguments can always be specified explicitly.
- ► Typecases can be open and closed.
- ► Type-indexed functions are first class.
- ▶ Generic functions come (almost) for free.
- ► This talk: a few small steps.



# Pattern Matching for Type-indexed Functions Andres Löh

7 July 2004



## Current situation (Dependency-style)

#### Patterns are flat.

$$x \langle T \alpha_1 \dots \alpha_k \rangle = e$$

#### Examples:

```
 \begin{vmatrix} x \langle [\alpha] \rangle & = \dots \\ x \langle Fix \varphi \rangle & = \dots \\ x \langle GRose \varphi \alpha \rangle = \dots \end{vmatrix}
```

#### Forbidden:

#### Historical reasons (MPC-style)

In MPC-style, type patterns are (unapplied) type constructors:

```
\begin{array}{ccc}
x & \langle [] \rangle & = \dots \\
x & \langle Fix \rangle & = \dots \\
x & \langle GRose \rangle & = \dots
\end{array}
```

#### corresponds to

$$\begin{vmatrix} x \langle [\alpha] \rangle &= \dots \\ x \langle Fix \varphi \rangle &= \dots \\ x \langle GRose \varphi \alpha \rangle &= \dots \end{vmatrix}$$

in Dependency-style.



### Deep patterns are useful

```
show \langle [Char] \rangle x = "\" + x + "\" "
show \langle [\alpha] \rangle x = "[" + concat (intersperse ", " (map show \langle \alpha \rangle x)) + "]"
```

```
flatten \langle [[\alpha]] \rangle x = [flatten \langle [\alpha] \rangle concat x]
flatten \langle [\alpha] \rangle x = x
```



### Deep patterns are useful

$$show \langle [Char] \rangle x = "\" + x + "\" \\ show \langle [\alpha] \rangle x = "[" \\ + concat (intersperse ", " (map show \langle \alpha \rangle x)) \\ + "]"$$

flatten 
$$\langle [[\alpha]] \rangle x = [flatten \langle [\alpha] \rangle concat x]$$
  
flatten  $\langle [\alpha] \rangle x = x$ 

The order of cases becomes relevant (currently irrelevant):

$$\begin{array}{ccc} x \langle (Int, \alpha) \rangle & = 1 \\ x \langle (\alpha, Int) \rangle & = 2 \end{array}$$



### The plan

First, we liberalize the notion of dependencies. Then, we present a translation of a type-indexed function with deep patterns to

- multiple type-indexed functions
- using only flat patterns
- with fallthrough cases (new)
- possibly with multiple type arguments (new)



# Liberalized dependencies

Dependencies are currently fixed *per function*. We want to track dependencies *by function case*. Example (from my thesis):

Only one case (for functions) depends on *enum*, but the whole function depends on it.



# Liberalized dependencies – contd.

Currently, this means that a local redefinition for *equal* must redefine *enum* as well:

```
let equal \langle \alpha \rangle x y = toUpper x == toUpper y
enum \langle \alpha \rangle = enum \langle Char \rangle
in equal \langle [\alpha] \rangle "laMBdA" "Lambda".
```

- Liberalized dependencies make dependencies variable from case to case.
- ▶ In the above redefinition, *enum* would not be needed.
- Only if *equal* is called on function types, *enum* dependencies are passed.
- ► This is very similar to type classes, which can have different context for different instances.



## Liberalized dependencies – contd.

Liberalized dependencies have disadvantages as well:

- ► Type signatures are needed for every case (modulo type inference, which is future work as well).
- ► The qualified type of a function call depends on all dependencies of all cases, whereas now one need only know the type signature of the function.

#### Nested pattern example: flatten

Usage:

flatten 
$$\langle [[[Int]]] \rangle [[[1,2,3],[4,5,6]],[[7,8,9]]]$$
  
 $\rightarrow [[[1,2,3,4,5,6,7,8,9]]]$ 

A more interesting variant that always returns a list of depth 1 could be written using a type-indexed type.



#### becomes

Note the fallthrough case in *flatten*<sub>1</sub>.



#### New concept: Fallthrough cases

- ▶ We allow a single dependency variable as a type pattern.
- ► For a fallthrough case, one component is generated, as for any other case.
- ► A fallthrough case matches always.
- ► The translation is similar to the one for generic abstractions.
- ► In fact, fallthrough cases can be seen as integrating generic abstractions with typecase-based generic definitions.



#### Fallthrough cases – contd.

#### becomes

```
cp(flatten_1, []) cp(flatten, \beta) cp(flatten_1, \beta) x = ... cp(flatten_1, Any) cp(flatten, \beta) cp(flatten_1, \beta) x = x
```

The call  $flatten_1 \langle Char \rangle$  is translated to

cp(flatten<sub>1</sub>, Any) cp(flatten, Char) cp(flatten<sub>1</sub>, Char)



 $flatten \langle [[[Int]]] \rangle x$ 



```
 \begin{array}{l} \textit{flatten} \; \langle [[Int]]] \rangle \; x \\ == \{ \text{expansion of type application } \} \\ \text{let} \; \{ \textit{flatten} \; \langle \beta \rangle = \textit{flatten} \; \langle [[Int]] \rangle; \textit{flatten}_1 \; \langle \beta \rangle = \textit{flatten}_1 \; \langle [[Int]] \rangle \} \\ \text{in} \; \textit{flatten} \; \; \langle [\beta] \rangle \; x \\ \end{array}
```





```
flatten \ \langle [[[Int]]] \rangle \ x \\ == \{ \ expansion \ of \ type \ application \ \} \\ \ let \ \{ flatten \ \langle \beta \rangle = flatten \ \langle [[Int]] \rangle; flatten_1 \ \langle \beta \rangle = flatten_1 \ \langle [[Int]] \rangle \} \\ \ in \ flatten \ \langle [\beta] \rangle \ x \\ == \{ flatten \ \langle [\beta] \rangle == flatten_1 \ \langle \beta \rangle \} \\ \ flatten_1 \ \langle [[Int]] \rangle \ x \\ == \{ \ expansion \ of \ type \ application \ \} \\ \ let \ flatten \ \langle \beta \rangle = flatten \ \langle [Int] \rangle \\ \ flatten_1 \ \langle \beta \rangle = flatten_1 \ \langle [Int] \rangle \\ \ in \ flatten_1 \ \langle [\beta] \rangle \ x \\ \end{cases}
```



```
flatten \langle [[[Int]]] \rangle x
== \{ expansion of type application \}
      let { flatten \langle \beta \rangle = flatten \langle [[Int]] \rangle; flatten<sub>1</sub> \langle \beta \rangle = flatten<sub>1</sub> \langle [[Int]] \rangle }
      in flatten \langle [\beta] \rangle x
 = \{\mathit{flatten}\ \langle [\beta] \rangle = \mathit{flatten}_1\ \langle \beta \rangle \}
     flatten_1 \langle [[Int]] \rangle x
 == { expansion of type application }
      let flatten \langle \beta \rangle = flatten \langle [Int] \rangle
            flatten_1 \langle \beta \rangle = flatten_1 \langle [Int] \rangle
      in flatten<sub>1</sub> \langle [\beta] \rangle x
 = \{ flatten_1 \langle [\beta] \rangle \ x = [flatten \langle [\beta] \rangle \ (concat \ x)] \}
      let flatten \langle \beta \rangle = flatten \langle [Int] \rangle
            flatten_1 \langle \beta \rangle = flatten_1 \langle [Int] \rangle
       in [flatten \langle [\beta] \rangle (concat x)]
```

```
flatten \ \langle [[[Int]]] \rangle \ x
= \{ \text{previous slide } \}
\mathbf{let} \ flatten \ \ \langle \beta \rangle = flatten \ \ \langle [Int] \rangle
flatten_1 \ \ \langle \beta \rangle = flatten_1 \ \ \langle [Int] \rangle
\mathbf{in} \ \ [flatten \ \ \langle [\beta] \rangle \ \ (concat \ x)]
= \{ flatten \ \ \langle [\beta] \rangle = flatten_1 \ \ \langle \beta \rangle \}
[flatten_1 \ \ \langle [Int] \rangle \ \ (concat \ x)]
```

```
flatten \langle [[Int]] \rangle x
== {previous slide }
let flatten \langle \beta \rangle = flatten \langle [Int] \rangle
                    flatten_1 \langle \beta \rangle = flatten_1 \langle [Int] \rangle
in [flatten \langle [\beta] \rangle (concat x)]

== {flatten \langle [\beta] \rangle == flatten<sub>1</sub> \langle \beta \rangle}

[flatten<sub>1</sub> \langle [Int] \rangle (concat x)]

== {expansion of type application }
           \begin{array}{ll} \textbf{let} \ \textit{flatten} & \langle \beta \rangle = \textit{flatten} & \langle \textit{Int} \rangle \\ \textit{flatten}_1 & \langle \beta \rangle = \textit{flatten}_1 & \langle \textit{Int} \rangle \end{array}
            in flatten<sub>1</sub> \langle [\beta] \rangle x
```

```
flatten \langle [[Int]] \rangle x
== {previous slide }
    let flatten \langle \beta \rangle = flatten \langle \lceil Int \rceil \rangle
          flatten_1 \langle \beta \rangle = flatten_1 \langle [Int] \rangle
    in [flatten \langle [\beta] \rangle (concat x)]
= \{ flatten \langle [\beta] \rangle = flatten_1 \langle \beta \rangle \}
     [flatten<sub>1</sub> \langle [Int] \rangle (concat x)]
== { expansion of type application }
    let flatten \langle \beta \rangle = flatten \langle Int \rangle
          flatten_1 \langle \beta \rangle = flatten_1 \langle Int \rangle
    in flatten<sub>1</sub> \langle [\beta] \rangle x
= \{ flatten_1 \ \langle [\beta] \rangle \ x = [flatten \ \langle [\beta] \rangle \ (concat \ x)] \}
    let flatten \langle \beta \rangle = flatten \langle Int \rangle
          flatten_1 \langle \beta \rangle = flatten_1 \langle Int \rangle
    in [[flatten \langle [\beta] \rangle (concat (concat x))]]
```

```
flatten \ \langle [[Int]]] \rangle \ x
= \{ \text{previous slide } \}
\mathbf{let} \ flatten \ \langle \beta \rangle = flatten \ \langle Int \rangle
flatten_1 \ \langle \beta \rangle = flatten_1 \ \langle Int \rangle
\mathbf{in} \ [[flatten \ \langle [\beta] \rangle \ (concat \ (concat \ x))]]
```



```
flatten \langle [[[Int]]] \rangle x

== { previous slide }

let flatten \langle \beta \rangle = flatten \ \langle Int \rangle

flatten<sub>1</sub> \langle \beta \rangle = flatten_1 \ \langle Int \rangle

in [[flatten \langle [\beta] \rangle \ (concat \ (concat \ x))]]

== { flatten \langle [\beta] \rangle = flatten_1 \ \langle \beta \rangle \}

[[flatten_1 \ \langle Int \rangle \ (concat \ (concat \ x))]]
```





The translation of *flatten* depends on *flatten* $_1$ . What happens with local redefinitions?

```
let flatten \langle \alpha \rangle x = reverse x in flatten \langle [\alpha] \rangle [[[1,2,3],[4,5,6]],[[7,8,9]]]
```

The translation of *flatten* depends on *flatten* $_1$ . What happens with local redefinitions?

```
let flatten \langle \alpha \rangle x = reverse x in flatten \langle [\alpha] \rangle [[[1,2,3],[4,5,6]],[[7,8,9]]]
```

This is translated to:

```
let flatten \langle \alpha \rangle x = reverse \ x flatten<sub>1</sub> \langle \alpha \rangle x = x in flatten \langle [\alpha] \rangle [[[1,2,3],[4,5,6]],[[7,8,9]]]
```

The fallthrough case of  $flatten_1$  is added. The result is



## New concept: Multiple type arguments

In the general case, we need multiple type arguments.

```
\begin{array}{lll} poly & \langle Int, Int \rangle & (x,y) = x+y \\ poly & \langle Int, Char \rangle & (x,\_) = x \\ poly & (\langle \alpha, [Int] \rangle & (\_,ys) = maximum \ ys \\ poly & \langle Int, \alpha \rangle & (x,y) = x+poly \ \langle \alpha \rangle \ y \\ poly & \langle Char \rangle & x = ord \ x \end{array}
```

# New concept: Multiple type arguments

In the general case, we need multiple type arguments.

#### becomes

```
 \begin{array}{llll} & poly & \langle (\alpha,\beta) \rangle & = poly_1 & \langle \alpha \rangle & \langle \beta \rangle \\ & poly & \langle Char \rangle & x & = ord & x \\ & poly_1 & \langle Int \rangle & \langle Int \rangle & (x,y) & = x+y \\ & poly_1 & \langle Int \rangle & \langle Char \rangle & (x,\_) & = x \\ & poly_1 & \langle \alpha \rangle & \langle [\beta] \rangle & = poly_2 & \langle \alpha \rangle & \langle \beta \rangle \\ & poly_1 & \langle \alpha \rangle & \langle \beta \rangle & = poly_3 & \langle \alpha \rangle & \langle \beta \rangle \\ & poly_2 & \langle \alpha \rangle & \langle Int \rangle & (\_,ys) & = maximum & ys \\ & poly_2 & \langle \alpha \rangle & \langle \beta \rangle & = poly_3 & \langle \alpha \rangle & \langle [\beta] \rangle \\ & poly_3 & \langle Int \rangle & \langle \alpha \rangle & (x,ys) & = x+poly & \langle \alpha \rangle & y \\ \end{array}
```

Universiteit Utrecht



How do multiple type arguments work?



How do multiple type arguments work? In each case of the definition,

- each of the type patterns must be flat,
- all type variables of all patterns must be distinct.

How do multiple type arguments work? In each case of the definition,

- each of the type patterns must be flat,
- all type variables of all patterns must be distinct.

## When applied,

▶ all type arguments have to be provided.

How do multiple type arguments work? In each case of the definition,

- each of the type patterns must be flat,
- all type variables of all patterns must be distinct.

## When applied,

all type arguments have to be provided.

#### Furthermore,

- ► Multiple type arguments interact with fallthrough cases.
- ▶ Multiple type arguments require per-case dependencies.
- Multiple type arguments allow to get rid of higher-arity generic functions. For instance, map can be written with two type arguments.



# Implementation of multiple type arguments

Once we have liberalized dependencies, they are easy to add.

- ► Each case of the definition is translated to a component.
- Components are parametrized by multiple type constructors now.



# Implementation of multiple type arguments

Once we have liberalized dependencies, they are easy to add.

- ► Each case of the definition is translated to a component.
- Components are parametrized by multiple type constructors now.

#### However:

- Specializations are also parametrized by multiple type constructors.
- Potential explosion of specializations required, bounded by  $d^n$ , where d is the number of datatypes and n is the number of type arguments.
- ► In connection with fallthrough cases, code explosion does not occur.



# Implementation of multiple type arguments – contd.

```
\begin{array}{llll} & poly_1 \; \langle Int \rangle \; \langle Int \rangle & (x,y) & = x+y \\ & poly_1 \; \langle Int \rangle \; \langle Char \rangle & (x,\_) & = x \\ & poly_1 \; \langle \alpha \rangle \; \; \langle [\beta] \rangle & = poly_2 \; \langle \alpha \rangle \; \langle \beta \rangle \\ & poly_1 \; \langle \alpha \rangle \; \; \langle \beta \rangle & = poly_3 \; \langle \alpha \rangle \; \langle \beta \rangle \end{array}
```

#### becomes

$$\begin{array}{llll} & \mathsf{cp}(poly_1, Int \times Int) & (x,y) & = x+y \\ & \mathsf{cp}(poly_1, Int \times Char) & (x,\_) & = x \\ & \mathsf{cp}(poly_1, Any \times [\,]) & \mathsf{cp}(poly_2, \alpha) & (\beta) = \mathsf{cp}(poly_2, \alpha) & (\beta) \\ & \mathsf{cp}(poly_1, Any \times Any) & \mathsf{cp}(poly_3, \alpha) & (\beta) = \mathsf{cp}(poly_3, \alpha) & (\beta) \end{array}$$

#### Call translation:

$$poly_1 \; \langle Int \rangle \; \langle [Char] \rangle \leadsto \mathsf{cp}(poly_1, Any \times [\,]) \; (poly_2 \; \langle Int \rangle \; \langle Char \rangle)$$
 Universiteit Utrecht

### Liberalized dependencies

- make dependency behaviour more similar to type classes
- are necessary to track a large number of dependencies efficiently



### Liberalized dependencies

- make dependency behaviour more similar to type classes
- are necessary to track a large number of dependencies efficiently

### Fallthrough cases

- are an important yet simple to implement extension
- are yet another concept next to generic abstraction (allows higher-kinded abstractions) and default cases (allows redirection of dependencies)

### Liberalized dependencies

- make dependency behaviour more similar to type classes
- are necessary to track a large number of dependencies efficiently

### Fallthrough cases

- are an important yet simple to implement extension
- are yet another concept next to generic abstraction (allows higher-kinded abstractions) and default cases (allows redirection of dependencies)

### Generic functions with multiple type arguments

- are necessary to implement deep patterns
- with liberalized dependencies, allow simplification of type system



### Liberalized dependencies

- make dependency behaviour more similar to type classes
- are necessary to track a large number of dependencies efficiently

### Fallthrough cases

- are an important yet simple to implement extension
- are yet another concept next to generic abstraction (allows higher-kinded abstractions) and default cases (allows redirection of dependencies)

### Generic functions with multiple type arguments

- are necessary to implement deep patterns
- with liberalized dependencies, allow simplification of type system

More to come ... Comments?



# Acknowledgements

Many thanks to Arthur van Leeuwen for taking the time to design this beautiful and (nearly) "huisstijl"-conformant LATEX theme.

