

Institute of Information and Computing Sciences
Utrecht University
P.O. Box 80.089
3508 TB Utrecht
The Netherlands
<http://www.generic-haskell.org>

The Generic HASKELL User's Guide

Version 0.99 — Amber release

The **Generic HASKELL** Team

Dæv Clarke
Ralf Hinze
Johan Jeuring
Andres Löh
Jan de Wit

info@generic-haskell.org

Contents

| | | |
|----------|--|-----------|
| 1 | What is Generic HVSHELL? | 4 |
| 1.1 | Generic programming | 4 |
| 1.2 | Generic HVSHELL overview | 4 |
| 2 | Installation | 5 |
| 2.1 | System requirements | 5 |
| 2.2 | Installing the binary distribution | 5 |
| 2.3 | Building from source | 6 |
| 2.4 | Running <code>gh</code> | 6 |
| 2.5 | Compiling and running the generated code | 6 |
| 3 | Generic HVSHELL: The Language | 7 |
| 3.1 | Special Parentheses | 7 |
| 3.2 | Kind-indexed types | 7 |
| 3.3 | Type-indexed values | 8 |
| 3.4 | Generic application | 9 |
| 3.5 | Generic abstraction | 10 |
| 3.6 | Type-indexed types | 10 |
| 3.7 | Generated function naming | 11 |
| 3.8 | Haskell compatibility | 11 |
| 4 | Library | 12 |
| 4.1 | Introduction | 12 |
| 4.2 | Module <code>Bounds</code> | 12 |
| 4.3 | Module <code>Collect</code> | 12 |
| 4.4 | Module <code>Compare</code> | 13 |
| 4.5 | Module <code>Eq</code> | 13 |
| 4.6 | Module <code>Map</code> | 13 |
| 4.7 | Module <code>MapM</code> | 13 |
| 4.8 | Module <code>ReadShow</code> | 14 |
| 4.9 | Module <code>Reduce</code> | 14 |
| 4.10 | Module <code>Table</code> | 15 |
| 4.11 | Module <code>ZipWith</code> | 15 |
| 5 | Future Work | 16 |

| | | |
|----------|--------------------------------------|-----------|
| 6 | Meta-information | 17 |
| 6.1 | Contact | 17 |
| 6.2 | Caveats | 17 |
| 6.3 | Known bugs and limitations | 17 |
| 6.4 | Change log | 18 |
| 6.5 | Acknowledgements | 18 |
| 6.6 | Copyright information | 18 |

1 What is Generic HASKELL?

1.1 Generic programming

Software development often consists of designing datatypes around which functionality is added. Some functionality is datatype specific, whereas other functionality is defined on almost all datatypes in such a way that it depends only on the structure of the datatype. A function that works on many datatypes in this way is called a *generic* (or polytypic) *function*. Examples of generic functionality include storing a value in a database, editing a value, comparing two values for equality, and pretty-printing a value.

Since datatypes often change and new datatypes are introduced, we have developed **Generic HASKELL** which supports generic definitions to save the programmer from (re)writing instances of generic functions. **Generic HASKELL** extends the functional programming language Haskell [5] with, among other things, a construct for defining type-indexed values with kind-indexed types, based on recent work by Hinze [2]. These values can be specialised to all Haskell datatypes, facilitating wider application of generic programming than provided by earlier systems such as PolyP [4].

1.2 Generic HASKELL overview

Generic HASKELL extends Haskell with a number of features.

- *type-indexed values* are defined as a value indexed over the various Haskell type constructors (unit, primitive types, sums, products, and user-defined type constructors). The resulting type-indexed value can be specialised to any type.
- *kind-indexed types* are types indexed over kinds, defined by giving a case for both $*$ and $\kappa \rightarrow \kappa'$. Instances are obtained by applying the kind-indexed type to a kind.
- generic definitions can be used by applying them to a type or kind. This is called *generic application*. The result is a type or value, depending on which sort of generic definition is applied.
- *generic abstraction* allows generic definitions to take a type parameter. Such a definition can apply only to types of a single kind.
- *type-indexed types* are types which are indexed over the type constructors. These can be used to give types to more involved generic values. The resulting type-indexed types can be specialised to any type.

2 Installation

2.1 System requirements

Generic HVSHELL is written in Haskell 98. Minor parts require rank-2-polymorphism, as implemented, for example, by GHC and Hugs. The package has been tested with GHC and comes with a `Makefile` suitable to build it using GHC. Although parts of the system are written using the Utrecht University's attribute grammar system `ag` and Ralf Hinze's `frown` `:-` parser generator, these tools are not required to build the compiler. The code generated by **Generic HVSHELL** is Haskell 98 compliant, except that functions generated for higher kinded types require rank- n -polymorphism. (Fortunately higher-kinded types are not particular common for most applications.)

Two kinds of distribution are available.

The *binary* distribution includes a `gh` compiler binary and can be used together with any Haskell system, as it translates **Generic HVSHELL** input files into ordinary Haskell files. We currently provide binaries for Linux, Solaris, and Windows. In principle, we can provide binaries for any platform supported by GHC.

The *source* distribution includes the Haskell code for the `gh` compiler which has been generated from our compiler source using both `ag` and `frown` `:-`. GHC is required to compile this distribution. Configuration files are provided.

At present, the compiler source (including both `ag` and `frown` `:-` sources) may only be obtained by emailing info@generic-haskell.org.

2.2 Installing the binary distribution

Installation is simple. Instructions for Unix users are as follows. Windows users must consult the `INSTALL` file.

```
./configure --prefix=install_path
```

`--prefix=install_path` is optional and defaults to `/usr/local/`.

Next run (GNU make is required):

```
make install
```

This will result in the binary `gh` being installed in the directory `${prefix}/bin`.

2.3 Building from source

Building from source requires exactly the same command sequence as for installing the binary distribution. The difference of course is the amount of work which is subsequently done.

2.4 Running gh

The **Generic HASKELL** compiler is called `gh`. There are essentially two modes for running it.

If you call `gh` without supplying a file to process, you will be asked for a file name. Specify an input file relative to your working directory, and the compiler will process the file and generate an output file with the same basename as the input file, but the extension `.hs`.

Alternatively, input files can be specified on the command line.

A number of command line options are also available:

```
Usage: gh [options...] files...
  -v      --verbose          (number of v's controls the verbosity)
  -V      --version, --release show version info
  -h, -?  --help            show help
  -L DIR  --library=DIR     add DIR to search path
```

The first level of verbosity (no `-v` flag) produces only error messages. The second level (`-v`) in addition provides some diagnostic information and warnings. The third level (`-vv`) in addition provides debugging information.

2.5 Compiling and running the generated code

The **Generic HASKELL** compiler generates ordinary Haskell code. This can be run using GHC, Hugs, or any other Haskell compiler. Ensure that you include the path to `GHPrelude.hs`, which can be found in the `lib` subdirectory, in your compiler's search path.

3 Generic HASKELL: The Language

The **Generic HASKELL** compiler implements a number of extensions to Haskell. These are described briefly here. Further information can also be found by consulting the literature [2, 1, for example] and in the library included in the distribution (in `lib/*.ghs`).

3.1 Special Parentheses

Kind-indexed and type-indexed definitions take a (single) kind or type argument which is wrapped by special parentheses. The parentheses $\{\llbracket _ \rrbracket\}$ (i.e., $\{\llbracket _ \rrbracket\}$) wraps a kind argument, whereas $\{\llbracket _ \rrbracket\}$ (i.e., $\{\llbracket _ \rrbracket\}$) wraps a type argument.

3.2 Kind-indexed types

Type-indexed values (may) possess kind-indexed types. Kind-indexed types are defined in **Generic HASKELL** using a new top-level declaration which has the following syntax:

$$\begin{aligned} \mathbf{type} \langle \mathit{Conid} \rangle \{\ast\} t_1 \dots t_n &= \langle \mathit{type} \rangle \\ \mathbf{type} \langle \mathit{Conid} \rangle \{k \rightarrow l\} t_1 \dots t_n &= \langle \mathit{type} \rangle \end{aligned}$$

A case is defined for both kind \ast and for higher kinds $k \rightarrow l$. To a certain degree the $k \rightarrow l$ case is predetermined, depending on the number of arguments [2]. This is exemplified by the $k \rightarrow l$ case in the following example.

Example The kind-indexed type for the generic *map* function is defined as:

$$\begin{aligned} \mathbf{type} \mathit{Map}\{\ast\} t_1 t_2 &= t_1 \rightarrow t_2 \\ \mathbf{type} \mathit{Map}\{k \rightarrow l\} t_1 t_2 &= \mathbf{forall} u v . \mathit{Map}\{k\} u v \rightarrow \mathit{Map}\{l\} (t_1 u) (t_2 v) \end{aligned}$$

Note that both cases have the same number of arguments and an equal number of type variables introduced by the **forall**.

| | | |
|------------------------------------|-----|--|
| $map\{Label\} l\} m (Label\ l\ b)$ | $=$ | $Label\ l\ (m\ b)$ |
| $map\{(\rightarrow)\}$ | $=$ | $error\ "cannot_map_over_function_type"$ |
| $map\{Int\} i$ | $=$ | i |
| $map\{Char\} c$ | $=$ | c |

This function can also be implemented in a point-free style (see [1]).

Generics defined over $\langle tycon \rangle$ Type-indexed values (and later types) can be defined over (possibly user-defined) type constructors. This covers the case for *Int* and *Char*, as illustrated above. Additional cases such as the following are also possible (though in this case superfluous):

| | | |
|-------------------------------|-----|-------------------------------------|
| $map\{List\} m\ Nil$ | $=$ | Nil |
| $map\{List\} m (Cons\ a\ as)$ | $=$ | $Cons\ (m\ a)\ (map\{List\} m\ as)$ |

for the user defined type

data $List\ a = Nil \mid Cons\ a\ (List\ a)$

Notice the call $map\{List\} m\ as$ on the right-hand side of this definition. This is required since *List* is a recursive type.

3.4 Generic application

A type-indexed value can be specialised to a value by applying it to a type. Generic application extends the syntax of expressions ($\langle aexp \rangle$) as follows:

| | | |
|------------------------|-------|---|
| $\langle aexp \rangle$ | $::=$ | \dots |
| | | $\mid \langle Varid \rangle \{ \langle type \rangle \}$ |

Similarly, a kind-indexed type can be specialised to a type by supplying the kind at which the definition is to be applied. The syntax of type expressions ($\langle gtycon \rangle$) is thus extended as follows:

| | | |
|--------------------------|-------|---|
| $\langle gtycon \rangle$ | $::=$ | \dots |
| | | $\mid \langle Conid \rangle \{ \{ \langle kind \rangle \} \}$ |

Example Given the datatype:

data $BinTree\ a = Empty \mid Node\ a\ (BinTree\ a)\ (BinTree\ a)$

The *map* function for *BinTree* is $map\{BinTree\}$. The type of $map\{BinTree\}$ is $Map\{* \rightarrow *\} BinTree\ BinTree$, which is $(a \rightarrow b) \rightarrow (BinTree\ a \rightarrow BinTree\ b)$.

3.5 Generic abstraction

A type variable (of fixed kind) can be abstracted generically from an expression using a kind of generic abstraction. Declarations take the following form:

$$\begin{aligned} \langle \text{Varid} \rangle \{ t :: \langle \text{kind} \rangle \} &:: \langle \text{type} \rangle \\ \langle \text{Varid} \rangle \{ t \} \dots &= \langle \text{exp} \rangle \end{aligned}$$

Here t is a type variable of the given kind, where $\langle \text{kind} \rangle$ ranges over grounded kinds (i.e., those without kind variables).

An example is the so-called categorical strength:

$$\begin{aligned} \text{strength} \{ t :: * \rightarrow * \} &:: t \ a \rightarrow b \rightarrow t \ (a, b) \\ \text{strength} \{ t \} \ ta \ b &= \text{map} \{ t \} \ (\lambda x \rightarrow (x, b)) \ ta \end{aligned}$$

3.6 Type-indexed types

Type-indexed types [3] can be defined just as type-indexed values, except that the right-hand side of a definition is a constructor followed by a type. Thus the syntax consists of a collection of definitions, indexed over the type constructors, of the form:

$$\mathbf{type} \ \langle \text{Conid} \rangle \{ \langle \text{stype} \rangle \} \ tv1 \ \dots \ tvn = \langle \text{con} \rangle \ \langle \text{type} \rangle$$

New constructors ($\langle \text{con} \rangle$) must be introduced for each case of such a definition — each case will be compiled into a **newtype** declaration.

A type-indexed type can be specialised to a type by supplying its type argument.

$$\begin{aligned} \langle \text{gtycon} \rangle &::= \dots \\ &| \ \langle \text{Conid} \rangle \{ \langle \text{type} \rangle \} \end{aligned}$$

Example The type-indexed type $FMap$ is defined as follows:

$$\begin{aligned} \mathbf{type} \ FMap \{ \{ \text{Unit} \} \} \ v &= \text{FMU} \ (\text{Maybe } v) \\ \mathbf{type} \ FMap \{ \{ \text{:+:} \} \} \ fma \ fmb \ v &= \text{FMP} \ (fma \ v, fmb \ v) \\ \mathbf{type} \ FMap \{ \{ \text{:*} \} \} \ fma \ fmb \ v &= \text{FMT} \ (fma \ (fmb \ v)) \\ \mathbf{type} \ FMap \{ \{ \text{Con} \} \} \ fm \ v &= \text{FMC} \ (fm \ v) \\ \mathbf{type} \ FMap \{ \{ \text{Label} \} \} \ fm \ v &= \text{FML} \ (fm \ v) \end{aligned}$$

$FMap$ can be used anywhere a type can be by supplying $FMap$ with its type parameter, for example in the following:

$$\begin{aligned} \mathbf{type} \ Lookup \{ \{ * \} \} \ t &= \mathbf{forall} \ v . FMap \{ \{ t \} \} \ v \rightarrow t \rightarrow \text{Maybe } v \\ \mathbf{type} \ Lookup \{ \{ k \rightarrow l \} \} \ t &= \mathbf{forall} \ a . Lookup \{ \{ k \} \} \ a \rightarrow Lookup \{ \{ l \} \} \ (t \ a) \end{aligned}$$

The constructors introduced in the definition of *FMap* can be used in pattern matching:

| | | |
|---|---|---|
| <i>lookup</i> { <i>t</i> :: <i>k</i> } | <i>Lookup</i> { <i>k</i> } | <i>t</i> |
| <i>lookup</i> { <i>Unit</i> } | <i>FMU</i> <i>fm</i> | <i>Unit</i> = <i>fm</i> |
| <i>lookup</i> { <i>+</i> } | <i>LA</i> <i>LB</i> (<i>FMP</i> (<i>fma</i> , <i>fmb</i>)) | (<i>Inl</i> <i>a</i>) = <i>LA fma a</i> |
| <i>lookup</i> { <i>+</i> } | <i>LA</i> <i>LB</i> (<i>FMP</i> (<i>fma</i> , <i>fmb</i>)) | (<i>Inr</i> <i>b</i>) = <i>LB fmb b</i> |
| <i>lookup</i> { <i>*</i> } | <i>LA</i> <i>LB</i> (<i>FMT</i> <i>fma</i>) | (<i>a</i> :: <i>b</i>) = do { <i>fmb</i> ← <i>LA fma a</i> ; <i>LB fmb b</i> } |
| <i>lookup</i> { <i>Con</i> <i>d</i> } | <i>l</i> (<i>FMC</i> <i>fm</i>) | (<i>Con</i> _ <i>b</i>) = <i>l fm b</i> |
| <i>lookup</i> { <i>Label</i> <i>d</i> } | <i>l</i> (<i>FML</i> <i>fm</i>) | (<i>Label</i> _ <i>b</i>) = <i>l fm b</i> |

Note: This feature is experimental and subject to revision.

3.7 Generated function naming

The **Generic HVSHELL** programmer must (unfortunately) be aware of the names which the compiler uses in the generated Haskell source in order to avoid name clashes.

For a type-indexed value with name *poly*, the name of the code generated for this function for type *Type* is *polyType*; this is the name of the function *poly*{*Type*}. Thus *map* for *BinTree* or *map*{*BinTree*} is called *mapBinTree*. There are a number of exceptions: the suffixes for type constructors [], (,), (, ,) and (→) are *LIST*, *TUPLE2*, *TUPLE3*, *Fun*, respectively. The generated *map* functions for these types are *mapLIST*, *mapTUPLE2*, and so forth.

With this in mind, a programmer may link ordinary Haskell code with code generated by the **Generic HVSHELL** compiler by using the names just described.

3.8 Haskell compatibility

Generic HVSHELL parses all Haskell programs, except when one of the following problems occur:

- The token **forall** is an additional keyword in **Generic HVSHELL**. As this is already the case in the extensions provided by many Haskell implementations, it should hopefully not cause too much trouble.
- The special parentheses for type and kind arguments, i.e. {}, [], {}, [], are all handled as a single token. Unfortunately, some pieces of regular Haskell code can trick the lexer and result in parse errors. For example, in

do { [*x*] ← *action*; *return* *x* }

the initial { [is treated as a single token { [rather than the two tokens { and [which an Haskell programmer would expect. In other instances, sequences such as +|} are considered as the operator +| followed by a }, since | may occur in operators, whereas { |+ is considered as the token { | followed by +.

The required fix in both cases is to insert a space in the appropriate place, for example, by writing instead **do** { [*x*] ← *action*; *return* *x* }.

4 Library

4.1 Introduction

Provided with the **Generic HVSHELL** system is a library of useful generic functions. These are summarised below; for the details, consult the library itself (in subdirectory `lib`). We give the types of the generic functions for kind `*` and `* → *`, and usually a short description.

Naming conventions When generic functions defined in the **Generic HVSHELL** library have an equivalent in the Haskell Prelude or libraries, the name of the generic function is prefixed with a ‘g’.

4.2 Module Bounds

$$\begin{aligned} gminBound, gmaxBound\{t :: *\} &:: t \\ gminBound, gmaxBound\{t :: * \rightarrow *\} &:: a \rightarrow t \ a \end{aligned}$$

These are slight generalizations of the `minBound` and `maxBound` members of the `Bounded` class. They have the property that for all types `t` of kind `*`:

$$\forall a :: t. gminBound\{t\} \leq a \leq gmaxBound\{t\}$$

However, these functions are also defined for types for which `Bounded` is not derivable; i.e. types which are not enumerations or simple product types (see [5, Appendix D]).

4.3 Module Collect

The functions in this module collect information about types and values of these types.

$$\begin{aligned} constructorOf\{t :: *\} &:: t \rightarrow ConDescr \\ constructorOf\{t :: * \rightarrow *\} &:: (a \rightarrow ConDescr) \rightarrow t \ a \rightarrow ConDescr \end{aligned}$$

`constructorOf` returns a description of the topmost constructor in a value.

$$\begin{aligned} constructors\{t :: *\} &:: [ConDescr] \\ constructors\{t :: * \rightarrow *\} &:: [ConDescr] \rightarrow [ConDescr] \end{aligned}$$

constructors returns a list of descriptions of all topmost constructors used in a datatype.

$$\begin{aligned} \text{labelsOf}\{t :: *\} &:: t \rightarrow [\text{LabelDescr}] \\ \text{labelsOf}\{t :: * \rightarrow *\} &:: (a \rightarrow \text{LabelDescr}) \rightarrow t \ a \rightarrow \text{LabelDescr} \end{aligned}$$

labelsOf returns a list of descriptions of labels in a value, or the empty list when the current type constructor has no labels.

$$\begin{aligned} \text{constructorsAndLabels}\{t :: *\} &:: [(\text{ConDescr}, [\text{LabelDescr}])] \\ \text{constructorsAndLabels}\{t :: * \rightarrow *\} &:: [(\text{ConDescr}, [\text{LabelDescr}])] \\ &\rightarrow [(\text{ConDescr}, [\text{LabelDescr}])] \end{aligned}$$

constructorsAndLabels combines the above information: it returns a list of all constructors, paired with the labels present in the given type constructor.

Again, consult `GHPrelude.hs` for details of *ConDescr* and *LabelDescr*.

4.4 Module Compare

$$\begin{aligned} \text{gcompare}\{t :: *\} &:: t \rightarrow t \rightarrow \text{Ordering} \\ \text{gcompare}\{t :: * \rightarrow *\} &:: (a \rightarrow b \rightarrow \text{Ordering}) \rightarrow t \ a \rightarrow t \ b \rightarrow \text{Ordering} \end{aligned}$$

gcompare is the generic version of *compare* in the *Ord* class.

4.5 Module Eq

$$\begin{aligned} \text{eq}\{t :: *\} &:: t \rightarrow t \rightarrow \text{Bool} \\ \text{eq}\{t :: * \rightarrow *\} &:: (a \rightarrow b \rightarrow \text{Bool}) \rightarrow t \ a \rightarrow t \ b \rightarrow \text{Bool} \end{aligned}$$

eq is the generic version of (`==`) in the *Eq* class.

4.6 Module Map

$$\begin{aligned} \text{gmap}\{t :: *\} &:: t \rightarrow t \\ \text{gmap}\{t :: * \rightarrow *\} &:: (a \rightarrow b) \rightarrow t \ a \rightarrow t \ b \end{aligned}$$

gmap is the generic version of *fmap* in *Functor* class.

4.7 Module MapM

$$\begin{aligned} \text{mapMl}, \text{mapMr}\{t :: *\} &:: (\text{Functor } m, \text{Monad } m) \Rightarrow t \rightarrow m \ t \\ \text{mapMl}, \text{mapMr}\{t :: * \rightarrow *\} &:: (\text{Functor } m, \text{Monad } m) \Rightarrow (a \rightarrow m \ b) \rightarrow t \ a \rightarrow m \ (t \ b) \end{aligned}$$

These are the generic versions of the monadic map $mapM$ in the Prelude. $mapMl$ traverses a data structure from left to right (just like $mapM$) while $mapMr$ traverses from right to left. The *Monad* in the context should also be an instance of class *Functor*, but that's usually not problematic.

4.8 Module ReadShow

$$\begin{aligned} gshowsPrec\{t :: *\} &:: Bool \rightarrow Int \rightarrow t \rightarrow ShowS \\ gshowsPrec\{t :: * \rightarrow *\} &:: (Bool \rightarrow Int \rightarrow a \rightarrow ShowS) \rightarrow \\ &Bool \rightarrow Int \rightarrow t a \rightarrow ShowS \\ \\ greadsPrec\{t :: *\} &:: Bool \rightarrow Int \rightarrow ReadS t \\ greadsPrec\{t :: * \rightarrow *\} &:: (Bool \rightarrow Int \rightarrow ReadS a) \rightarrow \\ &Bool \rightarrow Int \rightarrow ReadS (t a) \end{aligned}$$

The generic versions of *show* and *read* (in classes *Show* and *Read*).

The extra argument of type *Bool* is used internally to specify whether field labels are to be printed (and separated by commas). It should usually be *False*.

Since calling these functions is a bit cumbersome, the following specialisations are provided:

$$\begin{aligned} gshow\{t :: *\} &:: t \rightarrow String \\ gshow1\{t :: * \rightarrow *\} &:: Show a \Rightarrow t a \rightarrow String \\ gread\{t :: *\} &:: String \rightarrow t \\ gread1\{t :: * \rightarrow *\} &:: Read a \Rightarrow String \rightarrow t a \end{aligned}$$

4.9 Module Reduce

$$\begin{aligned} rreduce\{t :: *\} &:: t \rightarrow b \rightarrow b \\ rreduce\{t :: * \rightarrow *\} &:: (a \rightarrow b \rightarrow b) \rightarrow t a \rightarrow b \rightarrow b \\ lreduce\{t :: *\} &:: b \rightarrow t \rightarrow b \\ lreduce\{t :: * \rightarrow *\} &:: (b \rightarrow a \rightarrow b) \rightarrow b \rightarrow t a \rightarrow b \end{aligned}$$

rreduce is a generic version of *foldr* (note the reversed order of the last two arguments!), while *lreduce* is a generic *foldl*. See [1, section 5.4].

$$crush\{t :: * \rightarrow *\} \quad :: (a \rightarrow a \rightarrow a) \rightarrow a \rightarrow t a \rightarrow a$$

crush is an instance of *lreduce* with a slightly more familiar type.

The following functions are all defined in terms of the above functions, and most have counterparts in the Haskell Prelude:

| | |
|--|---|
| $gsum, gproduct\{t :: * \rightarrow *\}$ | $:: Num\ a \Rightarrow t\ a \rightarrow a$ |
| $gand, gor\{t :: * \rightarrow *\}$ | $:: t\ Bool \rightarrow a$ |
| $flatten\{t :: * \rightarrow *\}$ | $:: t\ a \rightarrow [a]$ |
| $count\{t :: * \rightarrow *\}$ | $:: t\ a \rightarrow Int$ |
| $comp\{t :: * \rightarrow *\}$ | $:: t\ (a \rightarrow a) \rightarrow (a \rightarrow a)$ |
| $gconcat\{t :: * \rightarrow *\}$ | $:: t\ [a] \rightarrow [a]$ |
| $gall, gany\{t :: * \rightarrow *\}$ | $:: (a \rightarrow Bool) \rightarrow t\ a \rightarrow Bool$ |
| $gelem\{t :: * \rightarrow *\}$ | $:: Eq\ a \Rightarrow a \rightarrow t\ a \rightarrow Bool$ |

$flatten$ collects all values of type a in a list, and $comp$ composes all functions contained in a datatype.

4.10 Module Table

The module `Table` provides a type-indexed type and functions for building memo tables of functions. See [1, section 5.6] and the code in `lib/Table.ghs`.

4.11 Module ZipWith

| | |
|------------------------------------|---|
| $gzipWith\{t :: *\}$ | $:: (t, t) \rightarrow Maybe\ t$ |
| $gzipWith\{t :: * \rightarrow *\}$ | $:: ((a, b) \rightarrow Maybe\ c) \rightarrow (t\ a, t\ b) \rightarrow Maybe\ (t\ c)$ |

A generic version of $zipWith$, except that it returns *Nothing* when the two data structures do not have the same shape.

| | |
|--------------------------------------|---|
| $gunzipWith\{t :: *\}$ | $:: t \rightarrow (t, t)$ |
| $gunzipWith\{t :: * \rightarrow *\}$ | $:: (a \rightarrow (b, c)) \rightarrow t\ a \rightarrow (t\ b, t\ c)$ |

$gunzipWith$ is a generic version of $unzip$.

| | |
|----------------------------------|---|
| $gzip\{t :: * \rightarrow *\}$ | $:: t\ a \rightarrow t\ b \rightarrow Maybe\ (t\ (a, b))$ |
| $gunzip\{t :: * \rightarrow *\}$ | $:: t\ (a, b) \rightarrow (t\ a, t\ b)$ |

These functions are more or less direct generalizations of zip and $unzip$ respectively, defined as instances of $gzipWith$ and $gunzipWith$.

5 Future Work

In the future, we plan to continue our work on the compiler. Among the many possible extensions and improvements, we are initially considering:

- support for POPL-style definitions
- full support for modules
- adding a type checker
- a view mechanism (i.e. implicit maps between data types); better support for fixpoints
- improved support for type-indexed data types
- XML, generic traversals, ...

As we have not yet decided how the next major release of the **Generic HASKELL** compiler will look, these topics are subject to change. Any input and feedback is most welcome!

6 Meta-information

6.1 Contact

The Generic HASKELL Project For information regarding the **Generic HASKELL** project send email to `info@generic-haskell.org`.

Mailing List A low volume mailing list exists. Currently it serves as a place for distributing information relevant to **Generic HASKELL** and for announcing our project meetings. This is the appropriate forum for general language discussions and whatnot. The address is `generic-haskell@generic-haskell.org`. To subscribe to the mailing list send an email to `listadm@generic-haskell.org`.

Bug Reports Bugs can be reported to `bugs@generic-haskell.org`.

6.2 Caveats

The **Generic HASKELL** compiler is a research prototype. Many of its features, especially the more experimental ones, may change as we gain more experience and understanding. It should be noted that the compiler does not perform type checking of the **Generic HASKELL** source language. Thus type errors in **Generic HASKELL** source will often only be discovered when the generated Haskell source is compiled.

6.3 Known bugs and limitations

1. Qualified names are not handled correctly in many parts of the compiler. If you want to have reliable behaviour, then avoid using any qualified names in your **Generic HASKELL** source files.
2. Export/import lists for modules are ignored by the compiler, which will lead to unwanted results if used for **Generic HASKELL** source files.
3. Imported types are not specialised for imported generic definitions. (This is not really a bug or a limitation, but rather prevents a generic function from being specialised to a specific type twice.)

4. The constructor descriptors for user-defined data types that have infix constructors with non-default fixity will be generated incorrectly with the default fixity.
5. Usage of the keyword **forall** in types is a bit tricky. There are places where it is needed, and others where it may cause strange errors. This will be clarified in the future. Let the examples guide you for now.
6. We have not yet decided how to deal with generic definitions which are declared across module boundaries. Any behaviour that these may give is incidental and should not be considered as definitive.
7. Type-indexed data types do not work well with modules. All specialisations have to be generated in the module in which they are defined.
8. A *datatype* (or **newtype**) which has a parameter of higher kinded type that does not appear in the right hand side of the definition produces a bug in the Haskell code generated by **gh**. The datatype Y below exhibits the behaviour.

```

data X a = X a
data Y b = Y (X (Y X))

```

Parameter b has kind $* \rightarrow *$, but it does not occur in $Y (X (Y X))$. We do not imagine that such types are very common, so the bug will remain for now.

9. This list is incomplete.

6.4 Change log

Amber (0.99) Everything is new. This is the first release.

6.5 Acknowledgements

Thanks to Ralf Hinze for **frown** `:-)`, to Arthur Baars and Doaitse Swiestra for **ag**, and to Simon Marlow and Sven Panne for the original Happy Haskell grammar.

6.6 Copyright information

gh – a compiler for **Generic HASKELL**.

Copyright © 2001 The **Generic HASKELL** Team. Utrecht University

This library is free software; you can redistribute it and/or modify it under the terms of the GNU Lesser General Public License as published by the Free Software Foundation; either version 2.1 of the License, or (at your option) any later version.

This library is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU Lesser General Public License for more details.

You should have received a copy of the GNU Lesser General Public License along with this library; if not, write to the Free Software Foundation, Inc., 59 Temple Place, Suite 330, Boston, MA 02111-1307 USA

Bibliography

- [1] Ralf Hinze. *Generic Programs and Proofs*. 2000. Habilitationsschrift, Bonn University. Available from <http://www.informatik.uni-bonn.de/~ralf/>.
- [2] Ralf Hinze. Polytypic values possess polykinded types. In Roland Backhouse and José Nuno Oliveira, editors, *Mathematics of Program Construction*, volume 1837 of *LNCS*, pages 2–27. Springer-Verlag, 2000.
- [3] Ralf Hinze, Johan Jeuring, and Andres Löb. Type-indexed datatypes. Unpublished, 2001. <http://www.cs.uu.nl/~johanj/publications/tidata.ps>.
- [4] P. Jansson and J. Jeuring. PolyP — a polytypic programming language extension. In *Conference Record of POPL '97: The 24th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, pages 470–482. ACM Press, 1997.
- [5] Simon Peyton Jones, John Hughes (editors), et al. Haskell 98 — A non-strict, purely functional language. Available from <http://haskell.org>, February 1999.