A lightweight approach to datatype-generic rewriting

THOMAS VAN NOORT
Institute for Computing and Information Sciences, Radboud University Nijmegen,
P.O. Box 9010, 6500 GL Nijmegen, The Netherlands
(e-mail: thomas@cs.ru.nl)

ALEXEY RODRIGUEZ YAKUSHEV and STEFAN HOLDERMANS
Vector Fabrics, Paradijslaan 28, 5611 KN Eindhoven, The Netherlands
(e-mail: {alexey, stefan}@vectorfabrics.com)

JOHAN JEURING
Department of Information and Computing Sciences, Utrecht University,
P.O. Box 80.089, 3508 TB Utrecht, The Netherlands
and
School of Computer Science, Open University of the Netherlands,
P.O. Box 2960, 6401 DL Heerlen, The Netherlands
(e-mail: johanj@cs.uu.nl)

BASTIAAN HEEREN
School of Computer Science, Open University of the Netherlands,
P.O. Box 2960, 6401 DL Heerlen, The Netherlands
(e-mail: bastiaan.heeren@ou.nl)

JOSE´ PEDRO MAGALHÃES
Department of Information and Computing Sciences, Utrecht University,
P.O. Box 80.089, 3508 TB Utrecht, The Netherlands
(e-mail: jpm@cs.uu.nl)

Abstract

Term-rewriting systems can be expressed as generic programs parameterised over the shape of the terms being rewritten. Previous implementations of generic rewriting libraries require users to either adapt the datatypes that are used to describe these terms or to specify rewrite rules as functions. These are fundamental limitations: the former implies a lot of work for the user, while the latter makes it hard if not impossible to document, test, and analyze rewrite rules. In this article, we demonstrate how to overcome these limitations by making essential use of type-indexed datatypes. Our approach is lightweight in that it is entirely expressible in Haskell with GADTs and type families and can be readily packaged for use with contemporary Haskell distributions.

1 Introduction

Consider a Haskell datatype Prop for representing formulae of propositional logic,
\[
data \text{Prop} = \text{Var String} \mid T \mid F \mid \text{Not Prop} \mid \text{Prop} : \land : \text{Prop} \mid \text{Prop} : \lor : \text{Prop},\]

and suppose we wish to simplify such formulae using the principle of contradiction:

\[ p \land \neg p \rightarrow \bot. \]

Ideally, our formulation of this rewrite rule as an executable program is neither much longer nor much more complicated than this rule itself.

One approach is to encode the rule as a function and then to apply it to individual formulae using some bottom-up traversal combinator \textit{transform}:

\[
simplify :: \text{Prop} \rightarrow \text{Prop} \\
simplify = \text{transform contradiction}
\]

\[
\text{where}
\]

\[
\text{contradiction} (p : \land: \text{Not } q) \mid p \equiv q = F
\]

\[
\text{contradiction } p = p.
\]

Although this implementation is relatively straightforward, encoding rules by functions has a number of drawbacks. To start with, rules cannot be concise one-line definitions as we have to provide a catch-all case in order to avoid pattern-matching failures at run time. Second, pattern guards (such as \( p \equiv q \) in our example) are needed to deal with multiple occurrences of variables, cluttering the definition. Lastly, rules cannot be analyzed easily since it is hard to inspect functions.

A way to overcome these drawbacks is to provide specialised rewriting functionality. That is, we can define a datatype for representing rewrite rules on formulae and implement the machinery required for rewriting (e.g., functions for matching formulae against rules and substituting formulae for metavariables) on top of this datatype. While this does overcome the drawbacks mentioned above, this approach comes with a serious disadvantage: it requires a large amount of datatype-specific code. If our next task is to rewrite, say, arithmetic expressions, we have to define a new datatype for representing rewrite rules and a new implementation of all the rewriting machinery.

However, both the datatype for representing rules and the associated rewriting machinery can be determined from the type that is used to describe the terms being rewritten. Hence, there is an excellent opportunity for datatype-generic programming here. In this article, we seize this opportunity and present a rewriting library that is generic in the type of terms being rewritten. Using our library, the example above can be written as

\[
simplify :: \text{Prop} \rightarrow \text{Prop} \\
simplify = \text{transform (rewriteWith contradiction)}
\]

\[
\text{where}
\]

\[
\text{contradiction } p = p : \land: \text{Not } p \rightarrow F.
\]

The library provides \textit{rewriteWith} and \( \mapsto \), which are generic and, in this case, instantiated with the type of propositional formulae \text{Prop}. A noticeable aspect of our approach is that metavariables in rewrite rules, such as \( p \) in our example, are introduced through ordinary function abstraction in Haskell, allowing the user to define her rules in terms of the term type \text{Prop} rather than some dedicated type for representing rules over \text{Prop}. The body of the function \text{contradiction} is now a fairly
direct transcript of the rule $p \land \neg p \rightarrow \bot$. As we will see, rewrite rules constructed with our library neither suffer from the drawbacks of the approach that uses pattern matching nor require large amounts of datatype-specific boilerplate code.

More specifically, the contributions of this article are the following:

- We present a library for term rewriting that is implemented using a simple design pattern (Section 4) for datatype-generic programming in Haskell extended with type families (Chakravarty et al. 2005a, 2005b; Schrijvers et al. 2008). As such, our library is “lightweight” and can be used readily with recent versions of the Glasgow Haskell Compiler (GHC).¹

- To represent rewrite rules our library needs to extend the type that is used to describe the terms being rewritten internally with an extra constructor for metavariables (Section 5.2). This extension is constructed generically using a type-indexed datatype (Hinze et al. 2004). Distinct metavariables in a single rewrite rule can, in our approach, range over rewritable terms of different type (Section 5.1).

- Internally, the library implements rewriting in terms of generic functions for pattern matching (Section 5.4) and substitution (Section 5.3) over generically extended datatypes. These datatypes are, however, completely hidden from the user, who writes her rewrite rules using the constructors of the types of terms that are to be rewritten (Section 6).

- We compare the efficiency of our library to that of other approaches to term rewriting in Haskell (Section 10).

This article is based on a paper presented at the 2008 Workshop on Generic Programming (Van Noort et al. 2008). The present article includes several improvements over this previous work. Most prominently, while the library described in the WGP paper could only be used to generically rewrite values of regular datatypes, we now support generic rewriting for a strictly larger class of datatypes, including types from families of mutually recursive datatypes. Furthermore, we now detect ill-formed rewrite rules (Section 7) and facilitate guarded rewrite rules (Section 8) as well as heterogeneously typed metavariables (Section 5.1).

### 1.1 Road map

The remainder of this article is structured as follows. In Section 2, we discuss the two fundamental approaches to representing rewrite rules in Haskell. In Section 3, we present our proposal for a datatype-generic library for term rewriting from a user’s perspective.

Sections 4 to 6 deal with the implementation of our library’s main functionality. Section 4 showcases, through an example generic function, how datatype-generic functions are implemented in our library. Section 5 discusses how generic rewriting functionality is composed from more elementary generic functions for pattern matching and substitution and shows how these functions are implemented. In

¹ The library is dubbed guarded-rewriting and available on Hackage.
Section 6, we demonstrate how the not so programmer-friendly representation of rewrite rules, used internally by the generic functions from Section 5, is hidden from the users of our library.

Sections 7 and 8 discuss additions to the core functionality. In Section 7, it is shown how nonsensical rewrite rules can be detected statically, i.e., without applying them. In Section 8, the library is extended with support for rewrite rules that have preconditions associated with them.

Section 9 discusses, as a case study, the use of our library in a realistic application. Section 10 presents the results of two performance benchmarks. Section 11 discusses related work; Section 12 concludes.

2 Representing rewrite rules

Before we present our approach to datatype-generic rewriting in Section 3, let us first have a more in-depth look at the two fundamental approaches to representing rewrite rules in Haskell that were already briefly discussed in the introduction: the extensional approach (Section 2.1) and the intensional approach (Section 2.2).

2.1 Extensional representations

The extensional approach to representing rewrite rules encodes rules as Haskell functions, using pattern matching to check whether the argument term matches the left-hand side of the rule. If this is indeed the case, the right-hand side of the rule is returned; otherwise, the argument term is returned unchanged. For example, the rule

\[ \neg (p \land q) \rightarrow \neg p \lor \neg q \]

that is derived from one of De Morgan’s laws is extensionally encoded as

\[
demorgan :: \text{Prop} \rightarrow \text{Prop} \\
demorgan (\text{Not} (p :\land: q)) = \text{Not} \ p :\lor: \text{Not} \ q \\
demorgan \ p = p.
\]

Note that the last line ensures that prevents arguments that do not match the pattern \( \neg (p \land q) \) from causing run-time errors.

As Haskell lacks support for nonlinear patterns, rewrite rules containing metavariables with multiple left-hand-side occurrences cannot be written as functions directly. Instead, such variables are encoded by means of so-called pattern guards. For instance, a rule for the principle of the excluded middle

\[ p \lor \neg p \rightarrow \top \]

in which the metavariable \( p \) occurs twice at the left-hand side is implemented by

\[
excludedmiddle :: \text{Prop} \rightarrow \text{Prop} \\
excludedmiddle (p :\lor: \text{Not} \ q) \mid p \equiv q = T \\
excludedmiddle \ p = p
\]

where the second occurrence of \( p \) is replaced by an occurrence of a fresh variable \( q \) and equality of \( p \), and \( q \) is enforced through the guard \( p \equiv q \). Note that this encoding requires equality to be defined for values of type \( \text{Prop} \).
In some applications of rewriting, it is useful to know whether or not a rewrite rule was applied successfully. This information can be made available, at the expense of some additional notational overhead, by wrapping the rewriting result in a `Maybe` value.

\[
\text{excludedMiddleM} :: \text{Prop} \rightarrow \text{Maybe Prop} \\
\text{excludedMiddleM} (p : \lor : \text{Not } q) | p \equiv q = \text{Just } T \\
\text{excludedMiddleM} \ p = \text{Nothing}.
\]

Encoding rewrite rules in terms of Haskell functions allows for function-parameterised traversal combinators to be used directly in rewriting applications. As an example, the Uniplate library (Mitchell & Runciman 2007) provides, among others, the combinator `transform`

\[
\text{transform} :: \text{Uniplate } \alpha \Rightarrow (\alpha \rightarrow \alpha) \rightarrow \alpha \rightarrow \alpha
\]

which applies its argument function in a bottom-up fashion to all recursive positions in a tree. Given a suitable Uniplate-instance for the type `Prop`, it is straightforward to use this combinator to remove certain classes of tautological clauses from propositional formulae:

\[
\text{removeTautologies} :: \text{Prop} \rightarrow \text{Prop} \\
\text{removeTautologies} = \text{transform excludedMiddle}.
\]

However, even though Haskell’s pattern-matching facilities enable a more or less direct encoding of rewrite rules as functions and the interaction with traversal libraries comes almost for free, the extensional approach to representing rewrite rules raises some issues.

- Extensionally, represented rules cannot be easily observed as in Haskell it is not possible to inspect functions. Still, there are several reasons why it is desirable to have observable rewrite rules:

  **Documentation:** If rules are observable, they can be pretty-printed in order to generate documentation for a rewrite system.

  **Static checking:** Observability of rules allows for checking whether a given set of rewrite rules constitutes a confluent and terminating rewrite system.

  **Automated testing:** In most applications, a rule is expected to preserve the semantics of the term being rewritten. One way to test this property is to randomly generate terms, to rewrite these, and then to check whether the rewritten terms indeed have the same semantics as the original terms. However, a rewrite rule with a nontrivial left-hand side will most likely not match successfully against a randomly generated term. Hence, such rules are in danger of not getting tested sufficiently. If left-hand sides of rules are inspectable, term generation can be directed to produce matching terms more often, effectively improving test coverage.

  **Associativity- and commutativity-aware rewriting:** Many domains, such as that of logical propositions, have associative and commutative operators. If the
rewriting infrastructure is aware of this fact, rewrite rules can be specified more concisely and repetition can be avoided. With an intensional approach, this can be implemented by making the matching algorithm return all possible substitutions. In an extensional approach, the behavior of pattern matching is fixed and cannot be made aware of these operators.

**Inversion:** If the left-hand side and right-hand side of a rewrite rule can be accessed, these can be exchanged, resulting in the inverse of the rule.

**Tracing:** When a sequence of rewrite steps leads to an unexpected result, one may want to learn which rules were applied in which order.

- It is tedious to have to specify a catch-all case when rules are encoded as functions. All rule definitions require this extra case.
- The lack of nonlinear pattern matching in Haskell becomes a nuisance if left-hand sides of rules contain many occurrences of the same variables.
- As Haskell lacks first-class pattern matching, the user cannot easily abstract over commonly occurring structures in the left-hand sides of rewrite rules.

These issues can be overcome by switching to an intensional representation instead.

### 2.2 Intensional representations

In the intensional approach, rewrite rules are not encoded as functions, but as values of a datatype, so that the left- and right-hand sides of rules become observable:

```plaintext
data Rule α = Rule {lhs :: α, rhs :: α}.
```

Values of type `Rule α` are used to encode rewrite rules with left- and right-hand sides of type `α`. For example, rewrite rules for formulae of propositional logic can be expressed as values of type `Rule EProp`, where `EProp` is an extended version of the datatype `Prop` of propositional formulae with an extra constructor `Metavar` to represent metavariable occurrences in rewrite rules:

```plaintext
```

With values of type `Rule EProp` in place, we need to define rewrite functions that interpret these values as functions over propositions represented by `Prop`:

```plaintext
rewritePropWith :: Rule EProp → Prop → Prop.
```

Here, we do not give an implementation of `rewritePropWith`, but note that its type (and thus its implementation) is specific to propositional formulae. If we want to implement rewrite functionality that works on different datatypes, then we have to define new rewrite functions for these types.

With the proposition-specific rewrite function `rewritePropWith`, rules over propositional formulae can be written and used as in

```plaintext
removeTautologies :: Prop → Prop
removeTautologies = transform (rewritePropWith excludedMiddle)
where
    excludedMiddle = Rule {lhs = Metavar "p" :\(\oplus\): ENot (Metavar "p"), rhs = ET}.
```
An apparent inconvenience of this style of defining rules is that we cannot reuse the type \( Prop \) of terms being rewritten and its constructors \( Not \) and \( \vee \). Instead, to provision for metavariables, we have to use the extended representation \( EProp \) and its constructors \( ENot \) and \( :\& : \).

### 3 Datatype-generic rewriting

In this section, we present the interface to our library for datatype-generic rewriting. In Sections 4 to 6, we zoom in at the concrete implementation of this interface.

The rewrite system that we present in this paper uses intensionally represented rewrite rules. As observed in the previous section, straightforward implementations of such rewrite systems suffer from two drawbacks: (1) they require a significant amount of datatype-specific code and (2) rewrite rules need to be expressed in terms of a new datatype obtained by extending the original datatype with a constructor for metavariables. Our system, however, is carefully designed to circumvent these drawbacks: (1) we provide a single implementation of rewriting that is generic in the type of terms being rewritten and (2) we completely hide the internal representation of rewrite rules from the user of our library.

More specifically, in our approach rewrite rules are specified in terms of templates:

\[
\begin{align*}
\text{closedWorldTemplate} & :: \text{Template} \ Prop \\
\text{contradictionTemplate} & :: \text{Prop} \rightarrow \text{Template} \ Prop \\
\text{deMorganTemplate} & :: \text{Prop} \rightarrow \text{Prop} \rightarrow \text{Template} \ Prop \\
\text{closedWorldTemplate} & = \text{Not} \ T \mapsto F \\
\text{contradictionTemplate} & \ p = p :\& : \text{Not} \ p \mapsto F \\
\text{deMorganTemplate} & \ p \ q = \text{Not} \ (p :\& : q) \mapsto \text{Not} \ p :\vee : \text{Not} \ q.
\end{align*}
\]

Templates are constructed by means of an operator \( \mapsto \),

\[(\mapsto) :: \alpha \rightarrow \alpha \rightarrow \text{Template} \ \alpha,
\]

which takes a left-hand side and a right-hand side of a type \( \alpha \) and produces a template for rewrite rules on \( \alpha \). Note that both sides of a template are just values of the type of terms being rewritten. In particular, templates are expressed without need for an additional datatype providing for metavariables. Instead, metavariables are encoded as ordinary Haskell function arguments. The template for the De Morgan rule from the example above, for instance, uses two metavariables, which are introduced through function arguments \( p \) and \( q \).

To prepare templates for use in our rewrite system, the user needs to synthesise rules from these. To this end, the library provides an overloaded function \( \text{synthesize} \) (defined in Section 6.3) that takes templates or functions producing templates for rewrite rules on some type \( \alpha \) to values of type \( \text{Rule} \ \alpha \):

\[
\begin{align*}
\text{closedWorld} , \text{contradiction} , \text{deMorgan} & :: \text{Rule} \ Prop \\
\text{closedWorld} & = \text{synthesize} \ \text{closedWorldTemplate} \\
\text{contradiction} & = \text{synthesize} \ \text{contradictionTemplate} \\
\text{deMorgan} & = \text{synthesize} \ \text{deMorganTemplate}.
\end{align*}
\]
Here, values of type Rule α (with an implementation that differs slightly from the one given above; see Section 5) form the internal representation of rewrite rules on α in our library.

The generic rewrite functionality is now exposed through a pair of rewrite functions \texttt{rewriteWith} and \texttt{rewriteWithM}. The first

\[
\text{rewriteWith :: Rewritable α} \Rightarrow \text{Rule α} \rightarrow \text{α} \rightarrow \text{α},
\]

takes as arguments a rule over some rewritable type α (see Section subsec:making-terms-rewritable) and a value of type α and attempts to apply the rule to the value. For example, we have that

\[
\text{rewriteWith closedWorld (Not T)} \quad \text{yields } F.
\]

If the second argument to \texttt{rewriteWith} does not match the left-hand side of its first argument, the value to be rewritten is returned unmodified; for instance,

\[
\text{rewriteWith contradiction (Var "x" :∧: Not (Var "y"))}
\]

\[
\quad \text{yields } Var "x" :∧: Not (Var "y")
\]

as the argument term does not match a contradictory formula. To make a failed attempt at rewriting explicit in the value returned, the second generic rewrite function,

\[
\text{rewriteWithM :: (Rewritable α, Monad μ) ⇒ Rule α} \rightarrow \text{α} \rightarrow \mu \text{α},
\]

wraps its result in a monad μ. For example, instantiating μ with the \texttt{Maybe}-monad,

\[
\text{rewriteWithM deMorgan (T :∧: F)} \quad \text{yields } \text{Nothing},
\]

while

\[
\text{rewriteWithM deMorgan (Not (T :∧: F))} \quad \text{yields } \text{Just (Not T :∨: Not F)}.
\]

As with other lightweight approaches to generic rewriting, such as Scrap Your Boilerplate (Lämmel & Peyton Jones 2003) and Uniplate (Mitchell & Runciman 2007), a small effort is required from the users of our library in order to prepare their datatypes for generic rewriting. In particular, they must describe the structure of their datatypes (Section 3.1) and make these datatypes instances of the type class \texttt{Rewritable} (Section 3.2).

### 3.1 Representing the structure of datatypes

In our library, the structure of datatypes is described through instances of a type class \texttt{Representable}:

```haskell
class Representable α where
    type Rep α :: *
    from :: α → Rep α
    to :: Rep α → α.
```

Here, \texttt{Rep} is a so-called associated type synonym (Chakravarty et al. 2005a). A type α is representable if it is isomorphic to its generic representation type \texttt{Rep α}; the
isomorphism is witnessed by a pair of functions \textit{from} and \textit{to} that convert between the type and its generic representation.

Base types, such as \texttt{Int}, \texttt{Float}, and \texttt{Char} form their own generic representations:

\begin{verbatim}
instance Representable Int where type Rep Int = Int ; from = id ; to = id
instance Representable Float where type Rep Float = Float ; from = id ; to = id
instance Representable Char where type Rep Char = Char ; from = id ; to = id.
\end{verbatim}

Further generic representation types are composed from a fixed set of structure constructors. These include the nullary type constructor \texttt{Nil} and the binary type constructors \texttt{:+} and \texttt{:::}, defined as:

\begin{verbatim}
infixr 6 :+:;
infixr 5 :::
data Nil = Nil
data \(\alpha \leftarrow \beta\) = Inl \(\alpha\) \mid Inr \(\beta\)
data \(\alpha \rightleftharpoons\beta\) = \(\alpha\leftarrow Nil\) \leftarrow \(\beta\leftarrow Nil\).
\end{verbatim}

A given datatypes representation type follows immediately from its structure. Choice among data constructors is encoded in terms of right-nested sums constructed by \texttt{:+}. A data constructor itself is represented as a type-level list of its argument types, constructed by \texttt{:::} and \texttt{Nil}. Note that, instead of the more common \textit{sums-of-products} representation of datatypes (Jansson & Jeuring 1997; Backhouse et al. 1999; Hinze 2000), we use a list-like representation (Holdermans et al. 2006) as we want to make sure that constructor arguments are always encoded as the first operand of the constructor \texttt{:::}. For example, Haskell’s \texttt{Maybe}-type, given by

\begin{verbatim}
data Maybe \(a\) = Nothing \mid Just \(a\)
\end{verbatim}

is represented by the type \(\texttt{Nil} \leftarrow (\texttt{a} \rightleftharpoons \texttt{Nil})\) and we can write

\begin{verbatim}
instance Representable (Maybe \(a\)) where
type Rep (Maybe \(a\)) = \(\texttt{Nil} \leftarrow (\texttt{a} \rightleftharpoons \texttt{Nil})\)
from Nothing = Inl \(\texttt{Nil}\)
from (Just \(x\)) = Inr \((\texttt{x} \rightleftharpoons \texttt{Nil}))\)
to (Inl \(\texttt{Nil}\)) = \texttt{Nothing}
to (Inr \((\texttt{x} \rightleftharpoons \texttt{Nil}))\)) = Just \(x\).
\end{verbatim}

The type-class methods \textit{from} and \textit{to} form a so-called \textit{embedding-projection pair} and are supposed to witness the isomorphism between a type and its generic representation “modulo undefinedness”, i.e., it should hold that \(\textit{to} \circ \textit{from} = \textit{id}\) and \(\textit{from} \circ \textit{to} \subseteq \textit{id}\) (Hinze 2000).

The functional programmer’s all-time favorite datatype, i.e., the type of cons-lists,

\begin{verbatim}
data \([\texttt{a}]\) = [\texttt{\_}] \mid \texttt{a} : [\texttt{a}]
\end{verbatim}

is in our approach represented by \(\texttt{Nil} \leftarrow (\texttt{a} \rightleftharpoons [\texttt{a}] \rightleftharpoons \texttt{Nil})\), yielding the declaration
instance Representable \([\alpha]\) where

\[
\text{type Rep } [\alpha] = \text{Nil} :+ (\alpha ::: [\alpha] ::: \text{Nil})
\]

from [] = Inl Nil
from (x : xs) = Inr (x ::: xs ::: \text{Nil})

to (Inl Nil) = []
to (Inr (x ::: xs ::: \text{Nil}))) = x : xs.

Note that the generic representation types of recursive datatypes are themselves nonrecursive: from only converts the top-level constructor of a value into its generic representation and leaves all subtrees untouched.

For the type Prop of propositional formulae,

\[
data \text{Prop} = \text{Var String} | T | F | \text{Not Prop} | \text{Prop} : \land : \text{Prop} | \text{Prop} : \lor : \text{Prop}
\]

we have

\[
\begin{align*}
type \text{Var} & = \text{String} ::: \text{Nil} \\
type \text{T} & = \text{Nil} \\
type \text{F} & = \text{Nil} \\
type \text{Not} & = \text{Prop} ::: \text{Nil} \\
type \text{And} & = \text{Prop} ::: \text{Prop} ::: \text{Nil} \\
type \text{Or} & = \text{Prop} ::: \text{Prop} ::: \text{Nil}\end{align*}
\]

as abbreviations for the generic representations of the alternatives and then

instance Representable Prop where

\[
\text{type Rep Prop} = \text{Var} :+ \text{T} :+ \text{F} :+ \text{Not} :+ \text{And} :+ \text{Or}
\]

from (Var x) = Inl (x ::: \text{Nil})
from T = Inr (Inl \text{Nil})
from F = Inr (Inr (Inl \text{Nil}))
from (Not p) = Inr (Inr (Inl (p ::: \text{Nil}))))
from (p :\land : q) = Inr (Inr (Inl (Inr (Inl (p ::: q ::: \text{Nil}))))))
from (p :\lor : q) = Inr (Inr (Inr (Inr (Inr (p ::: q ::: \text{Nil}))))))

to (Inl (x ::: \text{Nil})) = \text{Var x}
to (Inr (Inl \text{Nil})) = T

to (Inr (Inr (Inl \text{Nil}))) = F

to (Inr (Inr (Inr (Inl (p ::: \text{Nil})))))) = \text{Not p}

to (Inr (Inr (Inr (Inl (p ::: q ::: \text{Nil})))))) = p :\land : q

to (Inr (Inr (Inr (Inr (Inr (p ::: q ::: \text{Nil}))))))) = p :\lor : q.
\]

Instance declarations of Representable can be quite verbose, as in the case for Prop. However, these declarations are completely determined by the structure of the represented datatypes and can easily be derived automatically, for example by means of a Template Haskell program (Sheard & Peyton Jones 2002). Moreover, all that needs to be done to use our library on a user-defined datatype, such as Prop, is declaring it an instance of Representable, Typeable, and Rewritable – and, as we will see next, instances of the latter two can be given almost effortlessly.
3.2 Making terms rewritable

The class `Rewritable` of types with rewritable values is given by

```haskell
class (Representable α, Typeable α, Eq (Rep α), Extensible (Rep α), Matchable (Rep α), Substitutable (Rep α), Sampleable (Rep α), Diffable (Rep α)) ⇒ Rewritable α.
```

As this class does not have any methods or associated types, it is only introduced for its superclass constraints. These constraints encode the conditions that need to be fulfilled by a term type in order for its values to be rewritable.

Not only do we need an instance of `Representable`, we also require term types to be in the class `Typeable` that was originally introduced for use with the Scrap Your Boilerplate-library (Lämmel & Peyton Jones 2003). Currently, `Typeable` is Haskell's *de facto* standard API for reifying types at the value level and as such it is included in the base libraries that ship with the GHC. Recent versions of the GHC even provide support for automatically deriving instances of `Typeable` for user-defined datatypes.

The remaining superclass constraints on `Rewritable` place restrictions on the generic representations of term types and make specific parts of the generic rewriting machinery available for all instances of `Rewritable`. More specifically, each of these constraints accounts for one generic function. As representation types are built from a limited set of type constructors, these constraints imply no additional burden on the user of our generic rewriting library. That is, all needed instances for the base types `Int`, `Float`, and `Char` and the representation constructors `Nil`, `:+:`, and `:::` are already provided by the library. The details behind these instances are discussed in the next sections: in Section 4, we give instances of the standard class `Eq` for our generic representation types; in Section 5 we give the definitions and instances of the custom classes `Extensible`, `Matchable`, and `Substitutable`, while Section 6 covers `Sampleable` and `Diffable`.

For now, we observe that, with the appropriate instances of `Representable` and `Typeable` in place, putting a term type in the class `Rewritable` reduces to a mere one-liner:

```haskell
instance Rewritable Int
instance Rewritable Float
instance Rewritable Char
instance Rewritable α ⇒ Rewritable (Maybe α)
instance Rewritable α ⇒ Rewritable [α]
instance Rewritable Prop.
```

4 Generic equality

The previous section introduced the interface to our library for datatype-generic rewriting. Let us now turn to the concrete implementation of this interface.
In this section, we present an implementation of a type-indexed equality function. In the next section, this generic function is used in our implementation of generic pattern matching, but here it also serves as a neat example of the design pattern for lightweight type-indexed functions that we employ for all generic functions in our library. The general pattern for implementing generic functions is that we overload a given function $f$ for all generic representation types and then derive a generic version $f'$ that “ties the knot” and works for all types in Rewritable.

In our implementation, we rely on the class `Eq` from Haskell’s Standard Prelude to provide an interface for overloaded equality:

```haskell
class Eq α where
    (≡), (̸≡) :: α → α → Bool
    x ≡ y = ¬ (x ̸≡ y)
    x ̸≡ y = ¬ (x ≡ y).
```

As the class `Rewritable` requires the generic representation types of all its instances to be in the class `Eq`, we can directly define an equality operator $≡'$ that works for all types of rewritable terms:

```haskell
(≡') :: Rewritable α ⇒ α → α → Bool
x ≡' y = from x ≡ from y.
```

To test two equally typed rewritable terms for equality, we convert them to their generic representations and then test these for equality.

It remains to declare instances of `Eq` for the types that appear in generic representations. The case for `Nil` is straightforward:

```haskell
instance Eq Nil where
    Nil ≡ Nil = True.
```

For sums, we require the summands to be instances of `Eq` and test whether both generic representations have their origins in the same alternative. If so, both values are compared recursively; otherwise, we produce `False`:

```haskell
instance (Eq α, Eq β) ⇒ Eq (α :+: β) where
    Inl x ≡ Inl y = x ≡ y
    Inr u ≡ Inr v = u ≡ v
    _         ≡ _         = False.
```

In the case for `:::` we make use of that, in our encoding of a datatype’s structure, the second type argument of `:::` is always another type-level list, and so we can assume that this type argument is itself in `Eq` as well. The first type argument, however, can be any type and, hence, we cannot just assume it to be an instance of `Eq`. Instead, we require this type argument to be in `Rewritable`, so that we can use the operator $≡'$ defined above to compare values of this type:

```haskell
instance (Rewritable α, Eq β) ⇒ Eq (α ::: β) where
    (x ::: xs) ≡ (y ::: ys) = x ≡' y ∧ xs ≡ ys.
```
5 Matching and substituting

In the previous section, we demonstrated how generic functions are implemented in our library. We continue our exploration of the internals of the library by discussing the core functionality of our library: the implementation of the function rewriteWith and its monadic companion rewriteWithM.

These are implemented in terms of two generic functions match' and substitute',

\[
\text{match'} :: (\text{Rewritable } \alpha, \text{Mappable } \Gamma, \text{Monad } \mu) \Rightarrow \text{Pattern } \Gamma \alpha \rightarrow \alpha \rightarrow \mu (\text{Substitution } \Gamma)
\]

\[
\text{substitute'} :: (\text{Rewritable } \alpha, \text{Monad } \mu) \Rightarrow \text{Substitution } \Gamma \rightarrow \text{Pattern } \Gamma \alpha \rightarrow \mu \alpha.
\]

The type Pattern \( \Gamma \alpha \) (see Section 5.2) is used in our library for the intensional representation of the left- and right-hand sides of rewrite rules over a term type \( \alpha \). Its type argument \( \Gamma \) is a so-called metavariable environment: a type-level list that encodes the types of the metavariables in a rewrite rule. Successfully matching a term against a left-hand-side pattern results in a substitution (Section 5.3) for the metavariables that occur in the pattern. As pattern matching may fail, the function match' returns its result in a monad \( \mu \). This function requires the metavariable environment \( \Gamma \) involved to be in the class Mappable (defined in Section 5.1), which simply means that an empty substitution can be produced for \( \Gamma \). Substitutions are partial maps from metavariables drawn from a given environment to matched subterms. Given such a substitution and a right-hand-side pattern, the generic function substitute' attempts to construct a new term value. This construction fails if the substitution is not defined for all metavariables that occur in the right-hand-side pattern, which explains the monadic result type of substitute'.

As metavariable environments are only of interest to the internals of our library, they are hidden from the user by wrapping the left- and right-hand-side patterns that constitute a rewrite rule in an existential type:

\[
\text{data Rule} :: \star \rightarrow \star \text{ where}
\]

\[
\text{Rule} :: \text{Mappable } \Gamma \Rightarrow \text{Pattern } \Gamma \alpha \rightarrow \text{Pattern } \Gamma \alpha \rightarrow \text{Rule } \alpha.
\]

Here, the existential type Rule is defined using the syntax of a so-called generalised algebraic datatypes or GADTs (Xi et al. 2003; Peyton Jones et al. 2006).

Given the existential Rule and suitable definitions of match' and substitute', the monadic rewrite function rewriteWithM can be written as

\[
\text{rewriteWithM} :: (\text{Rewritable } \alpha, \text{Monad } \mu) \Rightarrow \text{Rule } \alpha \rightarrow \alpha \rightarrow \mu \alpha
\]

\[
\text{rewriteWithM} \ (\text{Rule lhs rhs}) \ x = \text{do}
\]

\[
\text{let } s \leftarrow \text{match'} \text{ lhs x}
\]

\[
\text{substitute'} \ s \ \text{rhs}.
\]

That is, the term \( x \) is matched against the left-hand side \( \text{lhs} \) of a given rewrite rule. If the match is successful, the resulting substitution \( s \) is applied to the right-hand side \( \text{rhs} \) of the rewrite rule in order to produce the result term. An implementation for the nonmonadic rewrite function rewriteWith is obtained by instantiating the
5.1 Typed metavariables

In our intensional representation of rewrite rules, we encode metavariables by De Bruijn indices (De Bruijn, 1972). Our implementation allows different metavariables to range over differently typed subterms. To enforce a type-safe use of metavariables, we adopt the approach of Pasalic & Linger (2004) and implement metavariables as values of the GADT \texttt{Ref} of typed references:

\begin{verbatim}
data Ref :: \texttt{\star} \rightarrow \texttt{\star} \rightarrow \texttt{\star} where
  RZero :: Ref (\texttt{\alpha} ::: \Gamma) \texttt{\alpha}
  RSucc :: Ref \Gamma \texttt{\alpha} \rightarrow Ref (\beta ::: \Gamma) \texttt{\alpha}.
\end{verbatim}

Here, we use as metavariable environments \(\Gamma\) the heterogeneous lists constructed from \texttt{Nil} and \texttt{::=} that we also use in generic representations. A value of type \texttt{Ref} \(\Gamma \texttt{\alpha}\) then carries the Peano encoding of an index for an \(\texttt{\alpha}\)-typed position in a heterogeneous list of type \(\Gamma\). Note that such a value can never refer to an empty list, simply because the constructor types dictate that the lists contain at least one value.

As an example of the use of \texttt{Ref}, consider the function \texttt{deref} for dereferencing a typed reference to a value in a heterogeneously typed list:

\begin{verbatim}
deref :: Ref \Gamma \texttt{\alpha} \rightarrow \Gamma \rightarrow \texttt{\alpha}
deref RZero (x ::: xs) = x
deref (RSucc r) (x ::: xs) = deref r xs.
\end{verbatim}

In the implementation of \texttt{match'} and \texttt{substitute'}, typed references are used as indices into heterogeneously typed partial maps:

\begin{verbatim}
data PMap :: \texttt{\star} \rightarrow \texttt{\star} where
  PNil :: PMap Nil
  PCons :: Rewritable \texttt{\alpha} \Rightarrow Maybe \texttt{\alpha} \rightarrow PMap \Gamma \rightarrow PMap (\texttt{\alpha} ::: \Gamma).
\end{verbatim}

Values of type \texttt{PMap} \(\Gamma\) are partial maps from \(\Gamma\)-typed references to rewritable terms. Looking up a value in a partial map is implemented through the function \texttt{lookup},

\begin{verbatim}
lookup :: Monad \mu \Rightarrow Ref \Gamma \texttt{\alpha} \rightarrow PMap \Gamma \rightarrow \mu \texttt{\alpha}
lookup RZero (PCons Nothing s) = fail "unbound variable"
lookup RZero (PCons (Just x) s) = return x
lookup (RSucc r) (PCons mb s) = lookup r s
\end{verbatim}
that returns its result in a monad $\mu$ to provide for the case in which looking up fails. Since the types of the $RZero$ and $RSucc$ constructors ensure that the referenced partial map is nonempty, the definition of $lookup$ does not require a case for $PNil$.

For the construction of partial maps of type $PMap \Gamma$, we require that $\Gamma$ is a type-level list of rewritable-term types, so that $PNil$ and $PCons$ can be used to produce an initial, empty map. To this end, we make the list constructors $\text{Nil}$ and $:::$ instances of a class $\text{Mappable}$ that provides an empty-map constructor:

\[
\begin{align*}
\text{class Mappable } \Gamma \text{ where} \\
\text{empty} :: \text{PMap } \Gamma
\end{align*}
\]

\[
\begin{align*}
\text{instance Mappable Nil where} \\
\text{empty} = \text{PNil}
\end{align*}
\]

\[
\begin{align*}
\text{instance (Rewritable } \alpha, \text{Mappable } \Gamma \text{) } \Rightarrow \text{Mappable } (\alpha ::: \Gamma) \text{ where} \\
\text{empty} = \text{PCons Nothing empty}
\end{align*}
\]

Updating a rewritable term in a partial map involves destructing a typed reference and traversing the map until the appropriate position has been reached:

\[
\begin{align*}
\text{update} :: \text{Ref } \Gamma \alpha \to \alpha \to \text{PMap } \Gamma \to \text{PMap } \Gamma \\
\text{update } RZero \ x \ (\text{PCons } mb \ s) &= \text{PCons } (\text{Just } x) \ s \\
\text{update } (RSucc \ r) \ x \ (\text{PCons } mb \ s) &= \text{PCons } mb \ (\text{update } r \ x \ s).
\end{align*}
\]

Singleton mappings are then constructed by updating a single term in an empty map:

\[
\begin{align*}
\text{singleton} :: (\text{Rewritable } \alpha, \text{Mappable } \Gamma) \Rightarrow \text{Ref } \Gamma \alpha \to \alpha \to \text{PMap } \Gamma \\
\text{singleton } r \ x = \text{update } r \ x \ \text{empty}
\end{align*}
\]

Finally, two maps for the same environment $\Gamma$ can be merged if they agree on their codomain:

\[
\begin{align*}
\oplus :: \text{Monad } \mu \Rightarrow \text{PMap } \Gamma \to \text{PMap } \Gamma \to \mu (\text{PMap } \Gamma) \\
\text{PNil} \oplus \text{PNil} &= \text{return } \text{PNil} \\
\text{PCons Nothing } s \oplus \text{PCons Nothing } s' &= \text{liftM } (\text{PCons Nothing}) (s \oplus s') \\
\text{PCons Nothing } s \oplus \text{PCons } (\text{Just } y) \ s' &= \text{liftM } (\text{PCons } (\text{Just } y)) (s \oplus s') \\
\text{PCons } (\text{Just } x) \ s \oplus \text{PCons Nothing } s' &= \text{liftM } (\text{PCons } (\text{Just } x)) (s \oplus s') \\
\text{PCons } (\text{Just } x) \ s \oplus \text{PCons } (\text{Just } y) \ s' \\
\mid x \equiv' y &= \text{liftM } (\text{PCons } (\text{Just } x)) (s \oplus s') \\
\mid \text{otherwise} &= \text{fail } "\text{merging failed}".
\end{align*}
\]

Here, $\text{liftM}$,

\[
\begin{align*}
\text{liftM} :: \text{Monad } \mu \Rightarrow (\alpha \to \beta) \to \mu \alpha \to \mu \beta,
\end{align*}
\]

is the function from Haskell's standard libraries that lifts a given unary function into an arbitrary monad. If, for at least one reference, the arguments of the monadic merge operator $\oplus$ produce different terms, merging fails. As all terms contained in a partial map are of types in the class $\text{Rewritable}$, equality of terms can be tested by means of the generic equality test $\equiv'$.
5.2 Generic patterns

Recall from the definition of the GADT Rule that the left- and right-hand sides of rewrite rules are represented by values of the type Pattern $\Gamma \alpha$, where $\alpha$ is the type of terms to be rewritten and $\Gamma$ is a metavariable environment. The idea is to derive the definition of Pattern $\Gamma \alpha$ from the definition of $\alpha$, much like in Section 2.2 the definition of EProp was derived from the definition of Prop, but without requiring the user to explicitly declare the pattern type. As pattern types are supposed to hold the same values as their corresponding term types, but additionally allow each subterm to be replaced by a metavariable, Pattern can be elegantly defined in terms of a so-called type-indexed datatype. A type-indexed datatype (Hinze et al. 2004) is a datatype that is defined by induction over the structure of generically representable types.

Here, we encode type-indexed datatypes as datatype families (Schrijvers et al. 2008). That is, we define a datatype family Extended,

\[
\text{data family} \quad \text{Extended} \quad \alpha :: \star \rightarrow \star
\]

with index $\alpha$. A type Extended $\alpha \Gamma$ is to be interpreted as the type that is obtained from extending $\alpha$ with metavariables from $\Gamma$.

Instances of Extended are given for all representation constructors. These instances recursively introduce metavariable alternatives in all subterm positions in the generic representation of a type’s structure, while duplicating the remainder of the structure. A pattern is then defined as either a duplicate of a type’s structure with metavariable alternatives for all subterm positions or otherwise a metavariable of the appropriate type:

\[
\text{type} \quad \text{Pattern} \quad \Gamma \quad \alpha = \text{Extended} \quad \text{(Rep} \quad \alpha \text{)} \quad \Gamma \quad ::+\text{: Ref} \quad \Gamma \\quad \alpha.
\]

As values of base types do not contain subterms, the extension of these types amounts to mere duplication:

\[
\text{newtype instance} \quad \text{Extended} \quad \text{Int} \quad \Gamma = \text{Int}' \quad \text{Int}.
\]

The cases for Float and Char are analogous; in the sequel, we provide instance declarations for Int as representatives for all base types.

Note that in our library subterm positions in a type are encoded as elements of type-level lists. Hence, sums and lists are extended recursively with metavariable alternatives inserted for all list elements:

\[
\text{data instance} \quad \text{Extended} \quad \text{Nil} \quad \Gamma = \text{Nil}'
\]

\[
\text{data instance} \quad \text{Extended} \quad (\alpha ::+\beta) \quad \Gamma = \text{Inl}' \quad (\text{Extended} \quad \Gamma \quad \alpha) \mid \text{Inr}' \quad (\text{Extended} \quad \Gamma \quad \beta)
\]

\[
\text{data instance} \quad \text{Extended} \quad (\alpha :::\beta) \quad \Gamma = \text{Pattern} \quad \Gamma \quad \alpha :::+\beta \quad \Gamma.
\]

Because extended types contain at least the values of the original representation types (modulo renaming of constructors and redirections into sum types), converting from terms to patterns is straightforward. First, we declare a class Extensible of types that can be lifted into their extended counterparts,

\[
\text{class} \quad \text{Extensible} \quad \alpha \quad \text{where}
\]

\[
\text{extend} :: \alpha \rightarrow \text{Extended} \quad \alpha \quad \text{Nil}
\]
and then we define a generic extension function $extend'$ for constructing patterns from terms:

$$
\begin{align*}
    extend' :: & \text{Rewritable } \alpha \Rightarrow \alpha \rightarrow \text{Pattern } \alpha \text{ Nil} \\
    extend' x & = \text{Inl } (extend \ (\text{from } x)).
\end{align*}
$$

Note that a value of a type $\text{Pattern } \alpha \text{ Nil}$, due to the empty metavariable environment, is guaranteed to not contain any metavariables.

Lifting base types reduces to wrapping values in extension constructors:

- **instance** Extensible Int where $extend = \text{Int'}$. 

Extension of the empty list involves converting from $\text{Nil}$ to $\text{Nil'}$

- **instance** Extensible Nil where $extend \text{ Nil} = \text{Nil'}$ 

while sums are extended recursively:

- **instance** (Extensible $\alpha$, Extensible $\beta$) $\Rightarrow$ Extensible ($\alpha \oplus: \beta$) where 
  - $extend \ (\text{Inl } x) = \text{Inl'} \ (extend \ x)$
  - $extend \ (\text{Inr } y) = \text{Inr'} \ (extend \ y)$. 

For $:::$, we require the first type argument to be rewritable, so that subterms can be lifted generically:

- **instance** (Rewritable $\alpha$, Extensible $\beta$) $\Rightarrow$ Extensible ($\alpha :::\beta$) where 
  - $extend \ (x ::: x) = extend' \ x :::' extend \ x$. 

The conversion from terms to patterns is used in Section 6 for the synthesis of rewrite rules from functions over term types.

### 5.3 Generic substitutions

Substitutions are just partial maps over a given metavariable environment:

- **type** Substitution $\Gamma = \text{PMap } \Gamma$. 

Applying a substitution then involves traversing a value of an extended type and replacing all metavariable occurrences by subterms drawn from the partial map in order to obtain a term representation:

- **class** Substitutable $\alpha$ where 
  - $\text{substitute :: Monad } \mu \Rightarrow \text{Substitution } \Gamma \rightarrow \text{Extended } \alpha \text{ } \Gamma \rightarrow \mu \alpha$.

As looking up metavariables in partial maps may fail, $\text{substitute}$ returns its result in a monad $\mu$. To apply a substitution to a pattern, we distinguish between values of extended types and metavariables. In the former case, we use $\text{substitute}$ to yield a representation and then convert this representation to a term by means of $\text{to}$. In the latter case, the metavariable is looked up in the partial map that represents the substitution:

- $\text{substitute'} :: (\text{Rewritable } \alpha, \text{Monad } \mu) \Rightarrow \text{Substitution } \Gamma \rightarrow \text{Pattern } \Gamma \alpha \rightarrow \mu \alpha$
- $\text{substitute'} \ s \ (\text{Inl } e) = \text{liftM to } (\text{substitute } s \ e)$
- $\text{substitute'} \ s \ (\text{Inr } r) = \text{lookup } r \ s$
Substitutions over extended base types are performed by stripping off the extension constructors:

\[
\text{instance } \text{Substitutable } \text{Int where} \quad \text{substitute } s \ (\text{Int}' \ n) = \text{return } n.\\
\]

Similarly, for the empty lists of constructor arguments, we have

\[
\text{instance } \text{Substitutable } \text{Nil where} \quad \text{substitute } s \ \text{Nil}' = \text{return } \text{Nil}.\\
\]

Extended sum values are processed recursively, and the obtained values are reinjected into the appropriate side of the original sum type:

\[
\text{instance } (\text{Substitutable } \alpha, \text{Substitutable } \beta) \Rightarrow \text{Substitutable } (\alpha + \beta) \quad \text{where}\\
\quad \text{substitute } s \ (\text{Inl} \ e) = \text{liftM} \ \text{Inl} \ (\text{substitute } s \ e)\\
\quad \text{substitute } s \ (\text{Inr} \ e) = \text{liftM} \ \text{Inr} \ (\text{substitute } s \ e).\\
\]

The instance for `:::` once more requires all elements in a list to be in the class \text{Rewritable} and invokes the generic function `substitute'` to apply substitutions to patterns:

\[
\text{instance } (\text{Rewritable } \alpha, \text{Substitutable } \beta) \Rightarrow \text{Substitutable } (\alpha ::\beta) \quad \text{where}\\
\quad \text{substitute } s \ (\text{pat} ::\prime \text{es}) = \text{liftM2} \ (\ (\text{substitute' } s \ \text{pat}) (\text{substitute } s \ \text{es}).\\
\]

To lift the list constructor `:::` into a monad, this instance uses the standard function `liftM2`:

\[
liftM2 :: \text{Monad } \mu \Rightarrow (\alpha \rightarrow \beta \rightarrow \gamma) \rightarrow \mu \alpha \rightarrow \mu \beta \rightarrow \mu \gamma\\
\]

for turning binary functions into monadic operations.

### 5.4 Generic pattern matching

Finally, let us consider how substitutions are constructed, namely, by generically matching term values against patterns. The required machinery consists of a class \text{Matchable} of representation types, which can be matched against their recursively extended counterparts,

\[
\text{class Matchable } \alpha \text{ where}\\
\quad \text{match} :: (\text{Mappable } \Gamma, \text{Monad } \mu) \Rightarrow \text{Extended } \alpha \Gamma \rightarrow \alpha \rightarrow \mu (\text{Substitution } \Gamma)\\
\]

and a top-level generic function `match'` for matching terms against either an extended representation or otherwise a top-level metavariable:

\[
\text{match'} :: (\text{Rewritable } \alpha, \text{Mappable } \Gamma, \text{Monad } \mu) \Rightarrow\\
\quad \text{Pattern } \Gamma \alpha \rightarrow \alpha \rightarrow \mu (\text{Substitution } \Gamma)\\
\quad \text{match'} (\text{Inl } e) \ x = \text{match } e \ (\text{from } x)\\
\quad \text{match'} (\text{Inr } r) \ x = \text{return} \ (\text{singleton } r \ x).\\
\]

If a term `x` is to be matched against an extended representation `e`, `x` is itself converted to a generic representation `from x` and matched by means of `match`. If `x` is matched against a metavariable `r`, a singleton substitution is constructed that maps `r` to `x`. Pattern-match failures are dealt with monadically.
Matching values of base type against extended base values requires an equality test. If this test succeeds, an empty substitution is produced; otherwise, a mismatch is reported:

```haskell
instance Matchable Int where
  match (Int' n) n' |
    n ≡ n' = return empty
  | otherwise = fail "pattern mismatch".
```

Provided that both the extended representation and the term representation are completely defined (i.e., do not diverge), matching is always successful for empty lists:

```haskell
instance Matchable Nil where
  match Nil' Nil = return empty.
```

For values of sum types, we check whether the extended representation and the term representation encode the same alternative. If so, we proceed recursively; otherwise, matching fails:

```haskell
instance (Matchable α, Matchable β) ⇒ Matchable (α :+: β) where
  match (Inl' e) (Inl x) = match e x
  match (Inr' e) (Inr y) = match e y
  match _ _ = fail "pattern mismatch".
```

For nonempty lists, we attempt to match the head `x` against a pattern `pat` by means of a call to the generic function `match'` and the tail `xs` against extended representations `es` through a recursive call to `match`. If both `x` and `xs` are matched successfully, the resulting substitutions are merged with the operator `⊕` from Section 5.1:

```haskell
instance (Rewritable α, Matchable β) ⇒ Matchable (α ::: β) where
  match (pat :::' es) (x ::: xs) = join (liftM2 (⊕) (match' pat x) (match es xs)).
```

As both matching and merging may fail, this gives rise to a nested monadic structure, which we flatten with a call to the function `join`,

```haskell
join :: Monad μ ⇒ μ (μ α) ⇒ μ α
```

from the standard libraries.

This completes our implementation of generic matching and substitution.

### 6 Synthesising rewrite rules

In the previous section, we have demonstrated how rewrite rules are intensionally represented in terms of the type synonym `Pattern` and the type-indexed datatype `Extended`. Implementing patterns through generic types frees the user of the library from the burden of defining separate datatypes for representing the left- and right-hand sides of rewrite rules for various term types, but still allows us to enjoy the benefits of observable rules.
However, this use of generic types raises the question how the user is supposed to define her rewrite rules. She could write the rewrite rule derived from the principle of contradiction, for example, as

\[
\text{contradiction :: Rule Prop} \\
\text{contradiction = Rule lhs rhs} \\
\text{where} \\
\text{lhs = Inl (Inr (Inr (Inr (Inl (Inr (Inl (Inr (Inr \text{RZero ::: Nil})))))) ::: Nil))))} \\
\text{rhs = Inl (Inr (Inl \text{Nil'))})}
\]

but clearly this style of definition is tedious and, moreover, error-prone. Of course, the definition of so-called smart constructors, such as

\[
(\land') :: \text{Extended Prop} \to \text{Extended Prop} \to \text{Extended Prop} \\
p \land' q = \text{Inl (Inr (Inr (Inr (Inl (Inr (p ::: q ::: Nil'))))))}
\]

make take away some of the burden, but these smart constructors then need to be defined for all types of rewriteable terms, defeating the very purpose of datatype-generic programming. Instead, our library allows for rewrite rules to be defined in terms of the real constructors of the type of terms that are to be rewritten. A library function \textit{synthesize} then takes care of translating the terms in rewrite rules into generic representations. The rule above, for example, can conveniently and concisely be written as

\[
\text{contradiction :: Rule Prop} \\
\text{contradiction = synthesize (\lambda p \to p : \land: Not p \to F).}
\]

That is, rewrite rules are synthesised from functions that take placeholders for metavariables as arguments and produce values of the type of rewriteable terms – in this case, \textit{Prop}. This way, rewrite rules are specified in the same way for different term types, while the internal representation of the rules remains hidden from the user.

To synthesise rules from functions, we develop some more generic machinery. The idea is to instantiate each function parameter twice – each time with distinct term values – and compare the resulting values. This approach restricts us to function parameters of types that have at least two values; but note that this restriction is by no means essential as rules with metavariables that range over types that have only one value are not meaningful.

For the function \(\lambda p \to p : \land: Not p \to F\) that we used above, we could instantiate the parameter \(p\) first with the value \(T\) and then with the value \(F\). The first instantiation then yields the left-hand side \(T : \land: Not T\) and the right-hand \(F\); the second instantiation yields \(F : \land: Not F\) and \(F\). Next, we compare the obtained pairs of left- and right-hand sides to determine where metavariables are to be inserted. As, in our example, the produced left-hand sides differ in the left operand of \(\land\): and in the argument of \textit{Not}, an occurrence of some metavariable is inserted in these locations. The two right-hand sides are identical, so no metavariable occurrence will show up there.
In this section, we implement this scheme of producing rewrite rules generically. We first show how to generate pairs of distinct values for term types (Section 6.1). Then, we present a generic \texttt{diff} function that localises the positions in which metavariables are to be inserted (Section 6.2). Finally, a class of synthesiser types is given (Section 6.3).

### 6.1 Generic sampling

To produce pairs of distinct values for types in \texttt{Rewritable}, we define a class \texttt{Sampleable},

\begin{verbatim}
class Sampleable α where
  left :: α
  right :: α.
\end{verbatim}

Instances of \texttt{Sampleable} are supposed to have their methods \texttt{left} and \texttt{right} produce values that differ in their top-level constructors. With instances of \texttt{Sampleable} declared for all generic representation types, functions \texttt{left}' and \texttt{right}' can be defined generically for all types of rewritable terms:

\begin{verbatim}
left', right' :: Rewritable α => α
left' = to left
right' = to right.
\end{verbatim}

As always, appropriate instances for base types are straightforward to produce:

\begin{verbatim}
instance Sampleable Int where left = 0; right = 1.
\end{verbatim}

For \texttt{Nil}, it is not possible to produce distinct \texttt{left} and \texttt{right} values. Still, we have to provide an instance declaration to allow for types that contain \texttt{Nil}-values to be in \texttt{Sampleable}:

\begin{verbatim}
instance Sampleable Nil where
  left = Nil
  right = Nil.
\end{verbatim}

As a result, no metavariables are ever introduced in rules over types with only a single nonbottom value. (Note that this is by no means a fundamental limitation as meaningful rewrite rules for such types cannot be given anyway.) Sum types are easy: here we have the opportunity to actually produce values that are distinct in their top-level constructor. For \texttt{left}, we choose \texttt{Inl}, while for \texttt{right}, \texttt{Inr} is selected:

\begin{verbatim}
2 There is a minor caveat associated with the given Sampleable-declaration for sum types. Our library requires the values for \texttt{left} and \texttt{right} to be finite as infinite values will lead to nontermination of the generic \texttt{diff} function in Section 6.2. To guarantee termination, we require the lefmost constructor of a datatype to be nonrecursive, such that \texttt{left} always produces a finite value for this constructor. Note that this may require an implicit reordering of constructors when defining or generating generic representations and precludes types that have no finite values. Then, we can use a single \texttt{left}-produced value in the definitions of \texttt{left} and \texttt{right} for \texttt{+:} as the top-level constructors \texttt{Inl} and \texttt{Inr} already distinguish the values.
\end{verbatim}
instance (Sampleable α, Sampleable β) ⇒ Sampleable (α ::: β) where
  left = Inl left
  right = Inr left.

For :::, we have only one constructor at our disposal, so a distinction in top-level constructors is to be made at a deeper level:

instance (Rewritable α, Sampleable β) ⇒ Sampleable (α ::: β) where
  left = left' ::: left
  right = right' ::: right.

6.2 Generic diff

To determine at which positions in a pattern metavariables are to be introduced, we require the ability to generically compute a “diff” between two patterns. If such a position is found, it depends on the type of the metavariable to be introduced whether or not a new pattern can be distilled from the differences between the pattern values compared. To this end, we require term types to be in the class Typeable, so that their types can be compared at run-time.

The class Typeable comes with an operation gc

\[ gc :: (\text{Typeable } \alpha, \text{Typeable } \beta) \Rightarrow \varphi \alpha \rightarrow \text{Maybe } (\varphi \beta) \]

that allows values of type \( \varphi \alpha \) to be cast into values of type \( \varphi \beta \) if and only if \( \alpha \) and \( \beta \) are the same type. In our implementation of a generic diff, we attempt to cast values of type Pattern \( \Gamma \alpha \) into values of type Pattern \( \Gamma \beta \) with both \( \alpha \) and \( \beta \) in Typeable.

We define a class Diffable of representation types for which a diff can be computed:

\[
\text{class Diffable } \alpha \text{ where } \\
\quad \text{diff} :: \text{Typeable } \beta \Rightarrow \\
\quad \quad \text{Extended } \alpha \Gamma \rightarrow \text{Extended } \alpha \Gamma \rightarrow \text{Maybe } (\text{Extended } \alpha (\beta ::: \Gamma)).
\]

For each generic representation type \( \alpha \), the overloaded function \( \text{diff} \) takes two values of type \( \text{Extended } \alpha \Gamma \) for some environment \( \Gamma \) and attempts to introduce a new \( \beta \)-typed metavariable at the deeper locations in which the two values differ. If the two values differ at top-level or at an inappropriately typed location, \( \text{diff} \) fails and produces Nothing. Note that values generated by \( \text{left} \) and \( \text{right} \) for \( \alpha \) can only ever differ at top level.

Diffs for rewritable terms can now be computed by means of a generic function \( \text{diff}' \):

\[
\text{diff}' :: (\text{Rewritable } \alpha, \text{Typeable } \beta) \Rightarrow \\
\quad \text{Pattern } \Gamma \alpha \rightarrow \text{Pattern } \Gamma \alpha \rightarrow \text{Maybe } (\text{Pattern } (\beta ::: \Gamma) \alpha)
\]

\[
\text{diff}' (\text{Inl } e) (\text{Inl } e') = \\
\quad \text{case } \text{diff } e e' \text{ of } \text{Nothing } \rightarrow gc; \text{Just } e'' \rightarrow \text{Just } (\text{Inl } e'')
\]

\[
\text{diff}' (\text{Inr } r) (\text{Inr } r') | r \equiv r' = \text{Just } (\text{Inr } (\text{RSucc } r))
\]

\[
\text{diff}' _ _ = \text{Nothing}.
\]
This generic function takes patterns over a type \( \alpha \) as argument. If both patterns consist of values \( e \) and \( e' \) of an extended type, the overloaded \( \text{diff} \) function is used to compare \( e \) and \( e' \). If \( \text{diff} \) successfully computes a combined value \( e'' \) of extended type, this value is wrapped in a pattern \( \text{Inl} \ e'' \) and returned. If \( \text{diff} \) fails, we attempt to insert a metavariable \( \text{RZero} \) of type \( \beta \) at top-level. As insertion of such a metavariable is only allowed if \( \alpha \) and \( \beta \) are the same type, we use \( \text{gcast} \) to compare \( \alpha \) and \( \beta \) at run-time. If both patterns are metavariable alternatives \( \text{Inr} \ r \) and \( \text{Inr} \ r' \), we require \( r \) and \( r' \) to be the same metavariable and construct a corresponding metavariable \( \text{Inr} \ (\text{RSucc} \ r) \) in the extended environment \( \beta :: \Gamma \). If \( r \) and \( r' \) are not equal, or if the two patterns are constructed from different alternatives, we produce \( \text{Nothing} \).

It remains to give instances of \( \text{Diffable} \) for our generic representation constructors. As values of base types contain no subterms and can thus only differ at top-level, an implementation of \( \text{diff} \) for these types reduces to testing for equality:

\[
\text{instance} \ \text{Diffable} \ \text{Int} \ \text{where} \\
\begin{align*}
\text{diff} \ (\text{Int'} \ n) \ (\text{Int'} \ n') & \\
| n \equiv n' & = \text{Just} \ (\text{Int'} \ n) \\
| \text{otherwise} & = \text{Nothing}.
\end{align*}
\]

The extension of \( \text{Nil} \) holds only a single value \( \text{Nil}' \), so \( \text{diff} \) for empty lists cannot fail:

\[
\text{instance} \ \text{Diffable} \ \text{Nil} \ \text{where} \\
\begin{align*}
\text{diff} \ \text{Nil'} \ \text{Nil}' & = \text{Just} \ \text{Nil}'.
\end{align*}
\]

For values of sum type, we compare the top-level constructors. If these are different, we produce \( \text{Nothing} \); otherwise, comparison proceeds recursively:

\[
\text{instance} \ (\text{Diffable} \ \alpha, \text{Diffable} \ \beta) \Rightarrow \text{Diffable} \ (\alpha ::\beta) \ \text{where} \\
\begin{align*}
\text{diff} \ (\text{Inl'} \ e) \ (\text{Inl'} \ e') & = \text{case} \ \text{diff} \ e \ e' \ \text{of} \ \text{Nothing} \rightarrow \text{Nothing} ; \ \text{Just} \ e'' \rightarrow \text{Just} \ (\text{Inl'} \ e'') \\
\text{diff} \ (\text{Inr'} \ e) \ (\text{Inr'} \ e') & = \text{case} \ \text{diff} \ e \ e' \ \text{of} \ \text{Nothing} \rightarrow \text{Nothing} ; \ \text{Just} \ e'' \rightarrow \text{Just} \ (\text{Inr'} \ e'') \\
\text{diff} \ _ & = \text{Nothing}.
\end{align*}
\]

Similarly, for \( :::: \), the comparison of two values \( \text{pat} ::::' \text{es} \) and \( \text{pat}' ::::' \text{es}' \) continues recursively underneath the constructor \( ::::' \):

\[
\text{instance} \ (\text{Rewritable} \ \alpha, \text{Diffable} \ \beta) \Rightarrow \text{Diffable} \ (\alpha ::\beta) \ \text{where} \\
\begin{align*}
\text{diff} \ (\text{pat} ::::' \text{es}) \ (\text{pat}' ::::' \text{es}') & = \\
\text{case} \ (\text{diff}' \ \text{pat} \ \text{pat}', \text{diff} \ \text{es} \ \text{es}') \ \text{of} \\
(\text{Just} \ \text{pat}'', \text{Just} \ \text{es}'') & \rightarrow \text{Just} \ (\text{pat}'' ::::' \text{es}'') \\
_ & \rightarrow \text{Nothing}.
\end{align*}
\]

### 6.3 Generic synthesis

With generic sampling and generic \( \text{diff} \) defined, we can now implement the synthesis of rewrite rules from functions over term types. These functions wrap the left- and right-hand sides of rules in values of a type \( \text{Template} \),

\[
\text{data} \ \text{Template} \ \alpha = \text{Template} \ \alpha \ \alpha
\]
of which the values simply constitute pairs of terms. For the concise definition of templates, we introduce an operator $\mapsto$:

\[
\text{infix 1} \mapsto \\
(\mapsto) :: x \rightarrow x \rightarrow \text{Template } x \\
\text{lhs} \mapsto \text{rhs} = \text{Template lhs rhs}.
\]

Next, we define a class $\text{Synthesiser}$ of types of which the values can be used to synthesise rewrite rules:

\[
\text{class Rewritable (Term } x) \Rightarrow \text{Synthesiser } x \text{ where} \\
\text{type Term } x :: * \\
\text{type Env } x :: * \\
\text{patterns :: } x \rightarrow (\text{Pattern (Env } x) (\text{Term } x), \text{Pattern (Env } x) (\text{Term } x)).
\]

Each instance $x$ of $\text{Synthesiser}$ has an associated type synonym $\text{Term } x$ that gives the type of terms that are rewritten by a synthesised rewrite rule. Similarly, the associated type synonym $\text{Env } x$ gives the term types over which the metavariables of a synthesised rule range. For example, a rewrite rule synthesised from a function of a type $x \rightarrow \beta \rightarrow \text{Template } \gamma$ has two metavariables, ranging over values of types $x$ and $\beta$, and is used to rewrite terms of type $\gamma$. Operationally, a value $x$ of a type from $\text{Synthesiser}$ can be used to produce a pair $\text{patterns } x$ that contains the left- and right-hand-side components of a rewrite rule. Synthesis then reduces to combining these components in a $\text{Rule}$-value:

\[
\text{synthesize :: (Synthesiser } x, \text{Mappable (Env } x)) \Rightarrow x \rightarrow \text{Rule (Term } x) \\
\text{synthesize } x = \text{let } (\text{lhs, rhs}) = \text{patterns } x \text{ in } \text{Rule lhs rhs}.
\]

Instances of the class $\text{Synthesiser}$ are defined inductively over the structure of function types. As a base case, we have an instance for $\text{Template } x$ for any type $x$ of rewritable terms:

\[
\text{instance Rewritable } x \Rightarrow \text{Synthesiser (Template } x) \text{ where} \\
\text{type Term (Template } x) = x \\
\text{type Env (Template } x) = \text{Nil} \\
\text{patterns (Template lhs rhs)} = (\text{extend' lhs, extend' rhs}).
\]

Rewrite rules that are synthesised directly from templates over $x$ operate on terms of type $x$ and contain no metavariables. Left- and right-hand sides for these rules can be obtained simply by lifting template components into the type $\text{Pattern Nil } x$ of patterns over $x$ without variables, for which we use the generic function $\text{extend'}$ defined in Section 5.2.

In the inductive step, we require, in order for a function type $x \rightarrow \beta$ to be in the class $\text{Synthesiser}$, $x$ to be a type of rewritable terms and $\beta$ to be a type of
A lightweight approach to datatype-generic rewriting

instance (Rewritable α, Synthesiser β) ⇒ Synthesiser (α → β) where
  type Term (α → β) = Term β
  type Env (α → β) = α ::: Env β

patterns f =
  let (lhs, rhs) = patterns (f left')
  (lhs', rhs') = patterns (f right')
  in case (diff' lhs lhs', diff' rhs rhs') of
    (Just lhs'', Just rhs'') → (lhs'', rhs'')
    _ → error "synthesis failure".

Function abstraction over α adds an α-typed metavariable to the environment Env β, but does not alter the type Term β of terms the synthesised rule operates on. Patterns of the left- and right-hand sides of the rewrite rule are constructed by applying the function twice (once to the value produced by left' and once to the value produced by right') and then computing diffs from the obtained components, possibly introducing occurrences of a new metavariable that ranges over terms of type α. If diffs cannot be computed, synthesis fails with a run-time error – an issue to be discussed in more detail in the next section.

7 Detecting ill-formed rewrite rules

In the previous sections, we have shown the implementation of our library’s core functionality. In particular, we have shown how, although we use an intensional representation of rewrite rules internally, we allow the user to define rules in terms of functions over domain-specific types. Due to this sugarcoating, additional verification of rewrite rules is required.

Consider, for example, the following rewrite rule over propositional formulae,

funny :: Rule Prop
funny = synthesise (λn → f n ↦ T)

where f is some function taking Int-values to values of type Prop:

f :: Int → Prop.

It is unclear what the semantics of such a rewrite rule should be. That is, in a well-formed rewrite rule, we expect metavariables to exclusively occur as constructor arguments, not as arguments to arbitrary functions. Using Haskell’s variables as placeholders for our metavariables means, however, that we cannot preclude such ill-formed rules and that we have to rely on the user not to construct nonsensical rules as the one above.

Another class of meaningless rewrite rules can be excluded by equipping our library with functionality for detecting their ill-formedness. Consider, for instance, the rule

unbound :: Rule Prop
unbound = synthesise (λp → T ↦ T : ∨ : p)
in which the metavariable \( p \) on the right-hand side is not bound on the left-hand side of the rewrite rule, and,

\[
\text{superfluous :: Rule Prop}
\]
\[
\text{superfluous} = \text{synthesise} \ (\lambda p \ q \to p : \forall : p \mapsto p)
\]
in which the metavariable \( q \) is superfluous since it is “declared” but not used at all in the rewrite rule. In general, we consider a rewrite rule well-formed if and only if all of its declared metavariables are bound in its left-hand side – and, interestingly, this notion of well-formedness can be checked for statically, i.e., without applying the rule.

To this end, we extend the library with a function \( \text{validate} \) that provides the user with an opportunity to verify the use of declared metavariables in rewrite rules:

\[
\text{validate :: Rewritable } \alpha \Rightarrow \text{Rule } \alpha \to \text{Bool}.
\]

This function is intended to be applied just after rule synthesis.

Validation is achieved by constructing a \( \text{use record} \) with a field for each metavariable, denoting its presence in the left-hand side of the rewrite rule:

\[
\text{data Record :: } \star \to \star \text{ where}
\]
\[
\text{RNil :: Record Nil}
\]
\[
\text{RCons :: Bool } \to \text{ Record } \Gamma \to \text{ Record } (\alpha :: : \Gamma).
\]

An initial blank record is created by setting each presence to \( False \):

\[
\text{class Recordable } \Gamma \text{ where}
\]
\[
\text{blank :: Record } \Gamma
\]
\[
\text{instance Recordable Nil where}
\]
\[
\text{blank} = \text{RNil}
\]
\[
\text{instance Recordable } \Gamma \Rightarrow \text{Recordable } (\alpha :: : \Gamma) \text{ where}
\]
\[
\text{blank} = \text{RCons False blank}.
\]

We now require environments to be instances of the type class \( \text{Recordable} \) and, hence, a constraint is added to the constructor \( \text{Rule} \) from Section 5:

\[
\text{data Rule :: } \star \to \star \text{ where}
\]
\[
\text{Rule :: (\cdots, \text{Recordable } \Gamma) } \Rightarrow \text{Pattern } \Gamma \alpha \to \text{Pattern } \Gamma \alpha \to \text{Rule } \alpha.
\]

A use record is updated by traversing the left-hand side of a rewrite rule and checking off each metavariable encountered:

\[
\text{class Validateable } \alpha \text{ where}
\]
\[
\text{record :: Extended } \alpha \Gamma \to \text{Record } \Gamma \to \text{Record } \Gamma.
\]

Recall from Section 5 that a \( \text{Pattern} \) is either a value of a corresponding extended type or else a metavariable. In the former case, we traverse the extended term recursively, looking for metavariable occurrences; in the latter case we check off the
metavariable in the use record:

\[
\begin{align*}
record' ::& \text{Rewritable } \alpha \Rightarrow \text{Pattern } \Gamma \ x \rightarrow \text{Record } \Gamma \rightarrow \text{Record } \Gamma \\
record' (\text{Inl } e) \ rec = & \ \text{record } e \ rec \\
record' (\text{Inr } \text{RZero}) (\text{RCons } b \ rec) = & \ \text{RCons True } rec \\
record' (\text{Inr } (\text{RSucc } r)) (\text{RCons } b \ rec) = & \ \text{RCons } b \ (\text{record'} (\text{Inr } r) \ rec).
\end{align*}
\]

Traversing base-type values results in no change to the use record as base values cannot contain metavariables:

\[
\text{instance Validateable } \text{Int where } \text{record } (\text{Int' } n) = \id.
\]

Similarly, traversing Nil values results in the original record:

\[
\text{instance Validateable } \text{Nil where } \\
\text{record Nil' } = \id.
\]

Values of sum types are traversed by stripping their top-level constructor:

\[
\text{instance } (\text{Validateable } \alpha, \text{Validateable } \beta) \Rightarrow \text{Validateable } (\alpha +: \beta) \text{ where } \\
\text{record } (\text{Inl' } e) = \text{record } e \\
\text{record } (\text{Inr' } e) = \text{record } e.
\]

For :::, we update the record by traversing the subterms and the pattern in sequence:

\[
\text{instance } (\text{Rewritable } \alpha, \text{Validateable } \beta) \Rightarrow \text{Validateable } (\alpha ::: \beta) \text{ where } \\
\text{record } (\text{pat :::'} es) = \text{record'} \ \text{pat} \circ \text{record } es.
\]

Note that since the record is only used to check off metavariable use, the order of the calls to record' and record plays no rôle.

Next, we add a superclass constraint for Validateable to the declaration of the class Rewritable from Section 3.2,

\[
\text{class } (\cdots, \text{Validateable } (\text{Rep } \alpha)) \Rightarrow \text{Rewritable } \alpha,
\]

and define a top-level function for validating rules:

\[
\text{validate :: Rewritable } \alpha \Rightarrow \text{Rule } \alpha \rightarrow \text{Bool} \\
\text{validate } (\text{Rule } \text{lhs } \text{rhs}) = \text{check } (\text{record'} \ \text{lhs} \ \text{blank}) \\
\text{where } \\
\text{check } \text{RNil} = \text{True} \\
\text{check } (\text{RCons } b \ \text{rec}) = b \land \text{check } \text{rec}.
\]

Starting with a blank record, validate records all occurrences of metavariables on the left-hand side of a rewrite rule and then verifies that all metavariables in the environment of the rule are checked off in the updated record.

\section{Guarded rewriting}

In the previous section, we have added some infrastructure for statically validating rewrite rules to the core functionality of our library. In this section, we further extend the library and add support for rewrite rules guarded by \textit{preconditions}. 
As an example, consider the datatype Lam of lambda-expressions,
\[
\text{data Lam} = \text{Var String} | \text{Abs String Lam} | \text{App Lam Lam}
\]
and an accompanying function \(fv\) that produces the variables that appear free in a given lambda-expression:
\[
fv :: \text{Lam} \to [\text{String}].
\]
Now, suppose that we want to define a rewrite rule that implements eta-reduction:
\[
\lambda x. e \ x \to e, \quad \text{if } x \text{ not free in } e.
\]
That is, eta-reduction applies to expressions that match the pattern \(\lambda x. e \ x\), but only if such an expression additionally fulfills the precondition that the variable \(x\) does not appear free in the expression \(e\). Using the extension presented in this section, such rewrite rules can be written as in
\[
\text{etaReduction} :: \text{Rule Lam}
\]
\[
\text{etaReduction} = \text{synthesize (} \lambda x. e \ x \to \text{Abs x (App e (Var x))} \to e \iff x \notin fv \ e)\).
\]
Here, we synthesise a rule over lambda-expressions from a function that produces a template constructed with the operators \(\mapsto\) and \(\iff\). The latter adds a guard to the rewrite rule, i.e., a boolean expression that may refer to the metavariables abstracted over by the synthesiser function.

In order to implement preconditions, we extend our type Rule of rewrite rules with a component containing a guard:
\[
\text{data Rule :: } * \to * \text{ where } \text{Rule :: (Mappable } \Gamma, \text{Recordable } \Gamma, \text{Testable } \Gamma) \Rightarrow \text{Pattern } \Gamma \ x \to \text{Pattern } \Gamma \ x \to \text{Guard } \Gamma \to \text{Rule } \ x.
\]
In addition to the classes Mappable (cf. Section 5) and Recordable (cf. Section 7), metavariable environments used within rules are restricted to be instances of the class Testable, to be explained below. Guard types are defined inductively over the structure of metavariable environments. That is, we have a type family Guard,
\[
\text{type family Guard } \Gamma :: *,
\]
with instances
\[
\text{type instance Guard Nil} = \text{Bool}
\]
\[
\text{type instance Guard } (\alpha ::: \Gamma) = \alpha \to \text{Guard } \Gamma.
\]
A guard for a rewrite rule without metavariables is just a boolean expression. For rules that do have metavariables, a guard is a function that takes an argument of appropriate type for each metavariable and produces a boolean.

Given a substitution for a metavariable environment \(\Gamma\) (cf. Section 5.3), values of type Guard \(\Gamma\) can be tested in order to obtain a boolean that indicates whether the corresponding precondition is fulfilled. To this end, we define the type class Testable of environments for which guards are testable:
\[
\text{class Testable } \Gamma \text{ where } \text{test :: Guard } \Gamma \to \text{Substitution } \Gamma \to \text{Bool}.
\]
For the empty-environment type Nil, the guard is itself already a value of type \( \text{Bool} \), so testing can just discard the supplied substitution (which can only be constructed by \( \text{PNil} \) anyway):

\[
\text{instance } \text{Testable} \text{ Nil where}
\]
\[
\text{test } b \text{ PNil } = b.
\]

For an environment \( \alpha :: \Gamma \), the guard function is applied to the value that is to be substituted for the metavariable corresponding to \( \alpha \) and the resulting guard for \( \Gamma \) is tested recursively:

\[
\text{instance } \text{Testable} \Gamma \Rightarrow \text{Testable} (\alpha :: \Gamma) \text{ where}
\]
\[
\text{test } f (\text{PCons} (\text{Just } x) s) = \text{test } s (f \ x)
\]
\[
\text{test } f (\text{PCons} \text{Nothing} s) = \text{error } \text{"test failure"}.
\]

If no substitution value is available, the governing rewrite rule was ill-formed (cf. Section 7) and testing fails with a run-time error.

As the GADT Rule now requires all metavariable environments to be testable, enforcing preconditions is straightforward:

\[
\text{rewriteWithM :: (Rewritable } \alpha, \text{Monad } \mu) \Rightarrow \text{Rule } \alpha \rightarrow \alpha \rightarrow \mu \alpha
\]
\[
\text{rewriteWithM } (\text{Rule } \text{lhs } \text{rhs } \text{grd}) x = \text{do}
\]
\[
\text{s } \leftarrow \text{match' lhs } x
\]
\[
\text{if } \text{test } \text{grd } s \text{ then substitute' } s \text{ rhs } \text{else fail } \text{"precondition failure"}.
\]

If, for a given rule \( \text{Rule } \text{lhs } \text{rhs } \text{grd} \) and term \( x \), \( x \) successfully matches against the left-hand side \( \text{lhs} \), the resulting substitution \( s \) is tested against the guard \( \text{grd} \). If the test succeeds, the substitution \( s \) and the right-hand side \( \text{rhs} \) are combined to produce a new term; otherwise, the rule does not apply and rewriting fails.

What remains is to adapt the synthesis of rules from templates and functions producing templates (cf. Section 6). First, we extend templates with a boolean component:

\[
\text{data Template } \alpha = \text{Template } \alpha \alpha \text{ Bool}.
\]

Next, we redefine and introduce the smart constructors \( \mapsto \) and \( \mapsto \), respectively:

\[
\text{infix 1 } \mapsto
\]
\[
\text{infix 0 } \mapsto
\]
\[
(\mapsto) :: \alpha \rightarrow \alpha \rightarrow \text{Template } \alpha
\]
\[
\text{lhs } \mapsto \text{rhs } = \text{Template } \text{lhs } \text{rhs } \text{True}
\]
\[
(\mapsto) :: \text{Template } \alpha \rightarrow \text{Bool } \rightarrow \text{Template } \alpha
\]
\[
\text{Template } \text{lhs } \text{rhs } \mapsto b = \text{Template } \text{lhs } \text{rhs } b.
\]

The class \text{Synthesiser} now gets an additional method \text{guard} that produces, for a synthesised rule, a guard of appropriate type:

\[
\text{class } \cdots \Rightarrow \text{Synthesiser } \alpha \text{ where}
\]
\[
\text{guard } :: \alpha \rightarrow \text{Guard (Env } \alpha)
\]

\[
\text{(1.2.3.4.5)}
\]
For rules synthesised directly from templates, this guard is just the boolean from the template:

\[
\text{instance } \cdots \Rightarrow \text{Synthesiser} \left( \text{Template } \alpha \right) \text{ where} \\
\vdots \\
guard \left( \text{Template } \text{lhs rhs b} \right) = b.
\]

For rules synthesised from functions, the guard is itself a function too:

\[
\text{instance } \cdots \Rightarrow \text{Synthesiser} \left( \alpha \rightarrow \beta \right) \text{ where} \\
\vdots \\
guard f = guard \circ f.
\]

The function \textit{synthesise}, finally, that turns synthesizers into rewrite rules simply puts guards in the right places in rules:

\[
\text{synthesise } :: \left( \text{Synthesiser } \alpha, \text{Mappable} \left( \text{Env } \alpha \right), \text{Testable} \left( \text{Env } \alpha \right) \right) \Rightarrow \alpha \rightarrow \text{Rule} \left( \text{Term } \alpha \right) \\
\text{synthesise } x = \text{let } \left( \text{lhs, rhs} \right) = \text{patterns } x \text{ in Rule lhs rhs (guard } x \text{).}
\]

Note that the given implementation of guarded rewrite rules has one obvious drawback: preconditions are encoded extensionally rather than intensionally and are therefore not observable. This reintroduces some of the problems mentioned in Section 2. Most prominently, when pretty-printing rewrite rules, the rendering of preconditions will pose a problem. The other issues listed in Section 2 are, however, of lesser importance. To what extent rewrite rules are still suitable for automated testing, strongly depends on how often preconditions apply: only if preconditions are rarely fulfilled, the generation of appropriate test data may be problematic. For inversion and tracing, nonobservability of preconditions plays no limiting rôle.

9 A case study: solving arithmetic equations

The previous section completed our exploration of our generic library for term rewriting. In this and the next section, we evaluate our approach. In this section, we present (part of) a small case study of a more or less realistic use of our library: solving arithmetic equations using term rewriting. In this case study, we use some of the more advanced features of our library, such as heterogeneously typed metavariables and guarded rewrite rules. In Section 10, we discuss some results obtained from benchmarking.

Consider the problem of solving the equation

\[
1 + \frac{8}{(x - 3)^2} = 3.
\]

To solve such an equation with a single variable, we use the so-called cover-up method, which is based on covering up the part of the equation that contains the variable. We can define cover-up rewrite rules for addition, subtraction, multiplication, division, and exponentiation operations; with these rules we solve the example equation in
the following sequence of steps:

\[
1 + \frac{8}{(x - 3)^2} = 3
\]
\[
\iff \frac{8}{(x - 3)^2} = 2
\]
\[
\iff (x - 3)^2 = 4
\]
\[
\iff x - 3 = 2 \lor x - 3 = -2
\]
\[
\iff x = 5 \lor x = 1.
\]

The domain of interest is represented by a variation of the datatype Prop from Section 1, that allows for formulae to be expressed over atoms of different types,

\[
\textbf{data Prop } α = \text{Var } α \mid T \mid F \mid \text{Not (Prop } α) \mid \text{Prop } α : \land : \text{Prop } α \mid \text{Prop } α : \lor : \text{Prop } α,
\]
a type of equations,

\[
\textbf{data Equation } α = α : ≡ : α,
\]
and a type Expr of various arithmetic expressions,

\[
\textbf{data Expr } = \text{Const Rational} \mid \text{Varia String} \mid \text{Expr} :: \text{Expr} \mid \text{Expr} : → : \text{Expr} \mid \text{Expr} : * : \text{Expr} \mid \text{Expr} : / : \text{Expr} \mid \text{Expr} : ^ : \text{Expr}.
\]

For each of these datatypes, we need instances of the class Representable (as described in Section 3), the class Typeable (can be derived by the GHC), and Rewritable (one line).

Using the datatypes, the equation \(1 + \frac{8}{(x - 3)^2} = 3\) is represented as

\[
\text{Var ((Const 1} +: (\text{Const 8} :/: ((\text{Varia } "x" : - : \text{Const 3}) :^: \text{Const 2}))) :≡: \text{Const 3}).
\]

The solution to this equation, \(x = 5 \lor x = -1\), is represented as

\[
\text{Var (Varia } "x" :≡: \text{Const 5}) :\lor: \text{Var (Varia } "x" :≡: \text{Const (−1))}.
\]

Our rewrite system consists of simple rules for simplifying propositions, such as

\[
\text{orTrueLeft} :: \text{Rewritable } α \Rightarrow \text{Rule (Prop } α)
\]
\[
\text{orTrueLeft} = \text{synthesise (}\lambda p → T :\lor: p → p\)
\]
and some rules for rewriting additions, which require preconditions,

\[
\text{coverPlusLeft} :: \text{Rule (Equation Expr)}
\]
\[
\text{coverPlusLeft} = \text{synthesise (}\lambda x \ y \ z \rightarrow 
\ x :+: y :≡: z → x :≡: z :≡: y ; \text{hasVaria } x \land \text{noVaria } y)\).
\]

In the rule coverPlusLeft, all metavariables range over expressions. We only want to apply this rule if there are variables in the expression \(x\) and no variables in the expression \(y\), so as to guarantee the isolation of the variables on the left-hand side of the equation. The helper functions \text{hasVaria} and \text{noVaria} test the presence (or absence) of variables in an expression.
Dealing with exponentiation requires a more complex rule:

\[
\text{coverPowerEven :: Rule (Prop (Equation Expr))} \\
\text{coverPowerEven} = \text{synthesize} (\lambda x \ n \ y \rightarrow \\
\quad \text{let } z = y ^ {\text{Const} \ (1 / n)} \\
\quad \text{in \ Var} \ (x ^ {\text{Const}\ n} \equiv y) \mapsto \\
\quad \text{Var} (x \equiv z) \lor \text{Var} (x \equiv \text{Const} 0 :s; z) \\
\quad \text{hasVaria} \ x \land n > 0 \land \text{isEven} \ n).
\]

As this definition illustrates, complex rewrite rules can be become quite verbose, but we can freely use local definitions to keep rules more or less readable. Since our rewrite rules are observable, a pretty-printer would be able to format such rules nicely. Note, however, that guards in rewrite rules are not observable since these are just boolean values, as described earlier in Section 8.

10 Benchmarks

The biggest disadvantage of generic programming techniques is that they can be a source of inefficiency. The introduction of representation types and corresponding conversions to and from the original datatypes generally imposes a penalty on execution time. We have measured the performance of our generic rewriting library to assess how large this penalty is, compared to hand-written code for a specific datatype. We have performed two separate tests of different complexities. The first one deals with logical propositions and uses neither preconditions nor metavariables of different types. The second one deals with arithmetic equations, and uses the full power of our generic rewriting library. Both are bundled with the library for analysis and repeatability.

10.1 Turning propositions into disjunctive normal form

Our first benchmark uses the datatype Prop of propositional formulae from the Introduction, extended with constructors for implication and equivalence. We have defined 16 rewrite rules and used these rules to bring the logical proposition to disjunctive normal form (DNF). This rewrite system is a realistic application of our rewriting library, and is very similar to the system that is used in an exercise assistant for e-learning systems (Heeren et al. 2008). None of the rules has preconditions, and all metavariables are of type Prop.

Conversion to DNF has been tested with four different strategies: such a strategy controls which rewrite rule is tried, and where. The strategies range from naive (i.e., apply some rule somewhere), to more involved strategy specifications that stage the rewriting and use all kinds of traversal combinators. We implemented these combinators in a type-specific fashion. They could also be implemented as generic functions, and not necessarily with the library we present. However, this would add another source of inefficiency to our tests, one that we do not wish to benchmark; hence, our choice for implementing the strategies in a type-specific fashion.

We use QuickCheck (Claessen & Hughes 2000) to generate a sequence of random propositions. The random-number generator is initiated with a fixed seed so that
A lightweight approach to datatype-generic rewriting

Table 1. The strategies benchmarked

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Terms</th>
<th>Rules applied</th>
<th>Rules tried</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>dnf-1</td>
<td>10,000</td>
<td>217,076</td>
<td>113,511,244</td>
<td>0.19%</td>
</tr>
<tr>
<td>dnf-2</td>
<td>50,000</td>
<td>492,114</td>
<td>22,224,222</td>
<td>2.21%</td>
</tr>
<tr>
<td>dnf-3</td>
<td>50,000</td>
<td>487,490</td>
<td>22,467,730</td>
<td>2.17%</td>
</tr>
<tr>
<td>dnf-4</td>
<td>100,000</td>
<td>872,494</td>
<td>18,327,913</td>
<td>4.76%</td>
</tr>
</tbody>
</table>

the same sequence is used for all test runs. We carefully profiled our tests to assure that the computation time was being spent mostly on the rewriting functionality, and not on auxiliary infrastructure such as data generation.

Because the strategy highly influences how many rules are tried, we vary the number of terms that has to be brought to disjunctive normal form depending on the strategy that is used. Table 1 shows for each strategy the number of terms that are normalised, how many rules are successfully applied, and the total number of rules that have been fired. The final column shows the percentage of rules that succeeded: the numbers reflect that the simpler strategies fire more rules.

We compare the execution times of three different implementations for the collection of rewrite rules.

Pattern Matching (PM): The first implementation defines the 16 rewrite rules as functions that use pattern matching. This implementation suffers from all the drawbacks that were mentioned in Section 1, making this version less suitable for an actual application. However, this implementation of the rules is worthwhile to study because Haskell has excellent support for pattern matching, which will likely result in efficient code.

Specialised Rewriting (SR): We have also written a specialised rewriting system that operates on propositions, very much like that described in Section 2.2. The most significant difference is that we have reused the Var constructor for representing metavariables too, thus mixing object variables with metavariables and avoiding the need to introduce an additional, extended datatype of propositions.

Generic Rewriting (GR): Here, we implemented the rules using the generic functions for rewriting that are introduced in this paper. The instance of the Representable type class is similar to the declaration in Section 3, except that it also includes the constructors for equivalence and implication.

All test runs were executed on a machine running Windows XP Professional x64 Edition with SP2 on an Intel Core 2 Duo 3Ghz with 2GB of RAM. The programs were compiled with the GHC (version 6.10.4) with standard optimization level (using the -O1 compiler flag). We do not use optimization level -O2, because we noticed that it sometimes reduced performance. Execution times were measured as the difference of the value returned by the function System.CPUTime.getCPUTime from the base libraries that ship with the GHC, after and before the execution of the test, and averaged over 10 runs.

Table 2 shows the performance for each implementation of the strategies. The absolute figures are given in seconds, and we also show the figures relative to the pattern-matching approach.
Table 2. Benchmark results for the Prop datatype with \texttt{-O1}

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Absolute (s)</th>
<th>Relative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PM</td>
<td>SR</td>
</tr>
<tr>
<td>dnf-1</td>
<td>3.11</td>
<td>10.89</td>
</tr>
<tr>
<td>dnf-2</td>
<td>2.52</td>
<td>4.82</td>
</tr>
<tr>
<td>dnf-3</td>
<td>2.49</td>
<td>4.87</td>
</tr>
<tr>
<td>dnf-4</td>
<td>3.94</td>
<td>7.28</td>
</tr>
</tbody>
</table>

Table 3. Benchmark results for the Prop datatype with increased inlining

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Absolute (s)</th>
<th>Relative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PM</td>
<td>SR</td>
</tr>
<tr>
<td>dnf-1</td>
<td>3.02</td>
<td>10.78</td>
</tr>
<tr>
<td>dnf-2</td>
<td>2.12</td>
<td>4.00</td>
</tr>
<tr>
<td>dnf-3</td>
<td>2.12</td>
<td>4.07</td>
</tr>
<tr>
<td>dnf-4</td>
<td>2.51</td>
<td>4.49</td>
</tr>
</tbody>
</table>

The table shows that PM is significantly faster than the other approaches. The specialised rewriting approach (SR) adds observability of the rewrite rules, at the cost of approximately doubling execution time. The generic approach (GR), when compared to the SR approach, suffers from a slowdown of a factor of about 3. This is probably due to the conversions to and from the structure representation of propositions. We also observe a correlation between strategy ratio of rule application (Table 1) and performance (the higher the ratio, the better the performance). This confirms that the overhead of both the SR and GR approaches is caused by the rewriting infrastructure: the PM approach has little overhead from trying rules as it uses Haskell’s native support for pattern matching.

Inspired by Magalhães et al. (2010), we repeated our benchmark setting compilation flags \texttt{-funfolding-creation-threshold} to 450 and \texttt{-funfolding-use-threshold} to 60. These flags control, respectively, the keenness of the compiler to export function definitions into interface files and to inline them. This has been shown to increase the performance of certain generic functions, since inlining “large” functions such as \texttt{to} and \texttt{from} exposes opportunities for further optimizations. We show the new results in Table 3. Note that the relative figures are still in relation to PM compiled with \texttt{-O1}, as this is the “standard” approach at the “standard” optimization level.

Increased inlining effectively improves the performance. All the approaches benefit from it, but the most pronounced gains are seen in the GR approach, where performance is improved to between 40% and 60% of the original levels. Strategy dnf-4, in particular, shows the highest improvement, now taking only twice as much as the original PM approach.

10.2 Solving arithmetic equations

Our second benchmark is performed on a family of datatypes representing arithmetic equations, as introduced in Section 9. We use 25 rules, some with preconditions and
some using metavariables of different types, therefore testing the full potential of our library in a realistic setting. These rules are applied to isolate variables on the left-hand sides of equations.

Again, we have used QuickCheck for test data generation. We test a single strategy, and use type-specific traversals for its application. We compare our library against a pattern-matching approach (PM) only, and again include figures with standard \texttt{-O1} optimization and with increased inlining as described previously. The results, as an average over 10 runs, are summarised in Table 4. We can conclude that the introduction of preconditions and metavariables of different types does not significantly influence performance. Promoting inlining continues to prove useful to increase the performance of our library.

Our benchmarks confirm that observability of rules comes at the expense of loss in runtime efficiency. Furthermore, generic definitions introduce some additional overhead. The tradeoff between efficiency and genericity depends on the application at hand. For instance, the library would be suitable for the online exercise assistant, because runtime performance is less important in such a context.

We believe that improving the efficiency of generic library code is an interesting area for future research. By inlining and specializing generic definitions, and by applying partial-evaluation techniques, we expect to get code that is more competitive to the hand-written definitions for a specific datatype.

### 11 Related work

Jansson & Jeuring (2000) implement a generic rewriting library in PolyP (Jansson & Jeuring 1997), an extension of Haskell with a special construct for generic programming. Our library differs in a number of aspects. First, we use no extensions of Haskell specific to generic programming. This is a minor improvement, since we expect that Jansson and Jeuring’s library can easily be translated to plain Haskell as well. Second, we use a type-indexed datatype for specifying rules. This is a major difference, since it allows us to generically extend a datatype with metavariables. In Jansson and Jeuring’s library, a datatype either has to be extended by hand, forcing users to introduce a new constructor, or one of the constructors of the original datatype is to be reused for metavariables. Neither solution is very satisfying, since either functions unrelated to rewriting must now handle the new metavariable constructor, or we are forced to introduce a safety problem in the library since an object variable may accidentally be considered a metavariable.

Libraries that provide generic traversal combinators, such as Strafunski (Lämmel & Visser 2002), Scrap Your Boilerplate (Lämmel & Peyton Jones 2003), Uniplate

<table>
<thead>
<tr>
<th>Optimization</th>
<th>Absolute (s)</th>
<th>Relative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PM</td>
<td>GR</td>
</tr>
<tr>
<td>Standard</td>
<td>0.57</td>
<td>2.44</td>
</tr>
<tr>
<td>Increased inlining</td>
<td>0.60</td>
<td>1.87</td>
</tr>
</tbody>
</table>

Table 4. Benchmark results for solving arithmetic equations
(Mitchell & Runciman 2007), Bringert’s “almost compositional” functions (Bringert & Ranta 2006), and probably more, can be used to define extensionally represented rewrite rules. These suffer from the disadvantages described in Section 2, but typically perform better than intensionally represented rules (see Section 10).

Our generic pattern-matching function is a variation on the generic unification functions of Jansson & Jeuring (1998) and Sheard (2001). A generalisation of our library to full unification is possible, but probably hard to keep user-friendly as unification results may contain metavariable occurrences that can then no longer be hidden from the user. Adapting our library to use mutable variables to improve performance, as in Sheard’s work, should be relatively straightforward.

Brown & Sampson (2008) implement generic rewriting using the Scrap Your Boilerplate-library. Patterns are described in a special-purpose datatype that does not depend on the type of values being rewritten. In contrast to our system, rules are not typed and hence ill-typed rules are only detected at runtime.

There exist a number of programming languages built on top of the rewriting paradigm, such as ELAN (Borovanský et al. 2001), OBJ (Goguen & Grant 1997), ASF+SDF (Van Deursen et al. 1996), and Stratego (Bravenboer et al. 2008). Instead of built-in support for rewriting, we focus on how to support rewriting in a mainstream higher order functional programming language by providing a library.

12 Conclusions and further work

We have presented a library for datatype-generic term rewriting. Our library overcomes problems in previous generic rewriting libraries: users do not have to adapt or manually extend the datatypes that are used to represent terms; they do not need knowledge of the internals of the library; and they can document, test, and analyze their rewrite rules. The performance of our library is not as good as that of hand-written, datatype-specific rewrite functions, but we think the loss of performance is acceptable for many applications.

In contrast to rewrite rules that are defined using an extensional representation, our library requires that rule synthesisers do not “cheat” by inspecting their metavariable arguments. Concretely, we do not allow arbitrary function applications in the right-hand side of a rule template, but unfortunately this restriction cannot be enforced statically.

There is ongoing work on generating test data for rewrite rules generically. That is, the left-hand side of a rewrite rule can be used as a template for test-data generation to improve testing coverage. We plan to use this approach in a testing framework that is to be shipped with our library.

Acknowledgments

This work was made possible by the support of the SURF Foundation, the Higher Education and Research Partnership Organisation for Information and Communications Technology (ICT). Please visit http://www.surf.nl/ for more information about SURF.
This work has been partially funded by the Technology Foundation STW through its project on “Demand Driven Workflow Systems” (07729), by the Netherlands Organisation for Scientific Research (NWO) through its projects on “Real-life Datatype-Generic Programming” (612.063.613) and “Scriptable Compilers” (612.063.406), and by the Portuguese Foundation for Science and Technology (FCT) via the SFRH/BD/35999/2007 grant; it was carried out while the second and third author were employed at Utrecht University.

The authors would also like to thank Chris Eidhof and Sebastiaan Visser for their work on testing rewrite rules using generic test-data generation and Andres Löh for productive discussions on this work. Finally, the authors are indebted to Doaitse Swierstra and the anonymous reviewers of the 2008 Workshop on Generic Programming and the present volume for their useful suggestions.

References


