



Adding Data to Curry

Michael Hanus^(✉)  and Finn Teegen 

Institut für Informatik, CAU Kiel, 24098 Kiel, Germany
{mh,fte}@informatik.uni-kiel.de

Abstract. Functional logic languages can solve equations over user-defined data and functions. Thus, the definition of an appropriate meaning of equality has a long history in these languages, ranging from reflexive equality in early equational logic languages to strict equality in contemporary functional logic languages like Curry. With the introduction of type classes, where the equality operation “=” is overloaded and user-defined, the meaning became more complex. Moreover, logic variables appearing in equations require a different typing than pattern variables, since the latter might be instantiated with functional values or non-terminating operations. In this paper, we present a solution to these problems by introducing a new type class `Data` which is associated with specific algebraic data types, logic variables, and strict equality. We discuss the ideas of this class and its implications on various concepts of Curry, like unification, functional patterns, and program optimization.

1 Introduction

The amalgamation of the main declarative programming paradigms, namely functional and logic programming, has a long history. The advantages of such integrated functional logic languages are manifold. One can use the features of functional programming (e.g., powerful type systems, higher-order functions, lazy evaluation) and logic programming (e.g., non-deterministic search, computing with partial information) in a single language which also leads to new design patterns [3,8]. Compared to logic programming, computations can be more efficient due to the use of optimal evaluation strategies [2].

Early approaches to integrating functional and logic programming (see [15] for a good collection of these proposals) used equational logic programming [19,37] as a unifying framework. From a logic programming point of view, equational logic programming extends the meaning of the standard equality predicate “=” by taking user-defined functions into account before checking the equality of both sides of an equation. Hence, both sides are evaluated before they are unified. If the definition of evaluable functions are considered as axioms for an equational theory, this process is also known as *E-unification* [17]. In order to use logic programming techniques (computing with partial information) also for the evaluation of user-defined functions, one can use narrowing instead of reduction [40], i.e., replace pattern matching by unification when a function call should be reduced. In this way, functional logic languages based on narrowing can be used to *solve equations*.

Example 1. Consider the following definition of Peano numbers and their addition (in Haskell [39] syntax):

```
data Nat = Z | S Nat
add :: Nat → Nat → Nat
add Z   n = n
add (S m) n = S (add m n)
```

In the functional language Haskell, we can only compute the value of expressions, e.g.,

```
> add (S Z) (S Z)
S (S Z)
```

However, if we interpret these definitions as a program written in the (narrowing-based) functional logic language Curry [22,27], we can also solve the equation

```
> add x (S Z) == S (S Z)   where x free
{x=S Z} True
```

Here, “==” denotes equality w.r.t. user-defined operations (see below for more details) and x is declared as a free (logic) variable which is bound to $S Z$ in order to evaluate the equation to `True`.

For the practical applicability of functional logic languages, it is important to reduce the computation space by using specific evaluation strategies. Thus, much work in this area has been devoted to develop appropriate narrowing strategies (see [21] for an early account of this research). In order to provide the advantages of lazy evaluation used in Haskell, e.g., optimal evaluation [29] and modularity [30], later research concentrated on demand-driven strategies. Needed narrowing [2] is an optimal strategy [1] and, thus, the basis of the language Curry.

Demand-driven evaluation strategies, like Haskell’s lazy evaluation or Curry’s needed narrowing, can deal with non-terminating operations that compute infinite data structures [30]. However, this could be in conflict with the equation solving capabilities of functional logic languages discussed above. Standard equality in the mathematical sense is required to be reflexive, i.e., $x = x$ should always hold [37]. Now consider two operations to compute infinite lists of Peano numbers:

```
f1 :: Nat → [Nat]
f1 n = n : f1 (S n)

f2 :: Nat → [Nat]
f2 n = n : S n : f2 (S (S n))
```

By reflexivity, $f1 Z = f1 Z$ should hold. This means that the infinite lists of all Peano numbers are equal. As a consequence, $f1 Z = f2 Z$ should also hold, but it is unclear to verify it during run time. In early equational logic programming, equations are solved by narrowing both sides to normal forms and unifying these normal forms. However, this does not work here since $f1 Z$ and $f2 Z$ have no normal form. Thus, reflexivity is not a feasible property of equations to be

evaluated (more details including issues about semantics are discussed in [18, 36]).

Therefore, contemporary languages interpret equations to be evaluated as *strict equality*, denoted by “`:=`” in Curry: $e_1 := e_2$ is satisfied iff e_1 and e_2 are reducible to a same ground constructor term, i.e., an expression without variables and defined functions. In particular, soundness, completeness, and optimality results are stated w.r.t. strict equality [2]. As a consequence, `f1 Z := f1 Z` does not hold so that it is not a defect that this equation cannot be solved.

Note that Haskell also offers the operation “`==`” intended to compare expressions. Although standard textbooks on Haskell define this operation as “equality” [11, 31, 41], its actual implementation can be different since, as a member of the type class `Eq`, it can be defined with a behavior different than equality on concrete type instances. Actually, the documentation of the type class `Eq`¹ denotes “`==`” as “equality” but also contains the remark: “`==` is customarily expected to implement an equivalence relationship where two values comparing equal are indistinguishable by “public” functions.” Thus, it is intended that $e_1 == e_2$ evaluates to `True` even if e_1 and e_2 have not the same but only equivalent values. On the other hand, the documentation requires that the reflexivity property

```
x == x = True
```

holds for any implementation, but this is not true even for the standard integer equality (choose “`last [1..] :: Int`” for `x`).

This discussion shows that the precise treatment of equality, which is essential for functional logic languages, might have some pitfalls when type classes are used. As long as “`==`” is defined in the standard way (by the use of “`deriving Eq`”), “`==`” conforms with strict equality. With the introduction of type classes to Curry, one has to be more careful. For instance, consider the “classical” functional logic definition of the operation `last` to compute the last element of a list by exploiting list concatenation (“`++`”) and equation solving [21, 24]:

```
last xs | _ ++ [e] == xs = e
  where e free
```

If “`==`” denotes equivalence rather than strict equality, `last` might not return the last element of a list but one (or more than one) value which is equivalent to the last element.

In this paper, we propose a solution to these problems by distinguishing between strict equality and equivalence. For this purpose, we propose a new type class `Data` which is associated with specific algebraic data types. We will see that this type class can also be used for a better characterization of the meaning of logic variables and the Curry’s unification operator “`:=`”.

This paper is structured as follows. In the next section, we review some aspects of functional logic programming and Curry. After motivating the problem this paper tackles in Sect. 3, we propose in Sect. 4 a new standard type class for Curry, namely `Data`, as a solution to the problem. In Sects. 5, 6, and Sect. 7,

¹ <http://hackage.haskell.org/package/base-4.12.0.0/docs/Data-Eq.html>.

we discuss how the proposed `Data` type class affects logic variables, optimization of equality constraints, and non-left-linear rules and functional patterns, respectively. Finally, Sect. 8 discusses related work before we conclude in Sect. 9.

2 Functional Logic Programming and Curry

We briefly review some aspects of functional logic programming and Curry that are necessary to understand the contents of this paper. More details can be found in surveys on functional logic programming [7, 24] and in the language report [27].

Curry is a declarative multi-paradigm language intended to combine the most important features from functional and logic programming. The syntax of Curry is close to Haskell [39] but also allows *free (logic) variables* in conditions and right-hand sides of rules. Thus, *expressions* in Curry programs contain *operations* (defined functions), *constructors* (introduced in data type declarations), and *variables* (arguments of operations or free variables). Function calls with free variables are evaluated by a possibly non-deterministic instantiation of demanded arguments [2]. This corresponds to narrowing [40], but Curry narrows with possibly non-most-general unifiers to ensure the optimality of computations [2]. In contrast to Haskell, rules with overlapping left-hand sides are non-deterministically (rather than sequentially) applied.

Example 2. The following simple program shows the functional and logic features of Curry. It defines the well-known list concatenation and an operation that returns some element of a list having at least two occurrences:

```
(++) :: [a] → [a] → [a]
[]    ++ ys = ys
(x:xs) ++ ys = x : (xs ++ ys)

someDup :: [a] → a
someDup xs | xs == _ ++ [x] ++ _ ++ [x] ++ _ = x
  where x free
```

Since “++” can be called with free variables in arguments, the condition in the rule of `someDup` is solved by instantiating `x` and the anonymous free variables “_” to appropriate values before reducing the function calls. As already mentioned in the introduction, “==” denotes strict equality, i.e., the condition of `someDup` is satisfied if both sides are reduced to a same ground constructor term. In order to avoid the enumeration of useless values, “==” is implemented as unification: if `y` and `z` are free (unbound) variables, `y == z` is evaluated (to `True`) by binding `y` and `z` (or vice versa) instead of non-deterministically binding `y` and `z` to identical ground constructor terms. This can be interpreted as an optimized implementation by delaying the bindings to ground constructor terms [10]. Due to this implementation, “==” is also called an *equational constraint* (rather than Boolean equality).

We already used the logic programming features of Curry in the definition of `last` shown in Sect. 1. In contrast to `last`, `someDup` is a *non-deterministic operation*

since it could yield more than one result for a given argument, e.g., the evaluation of `someDup [1,2,2,1]` yields the values 1 and 2. Non-deterministic operations, which can formally be interpreted as mappings from values into sets of values [20], are an important feature of contemporary functional logic languages. Hence, Curry has also a predefined *choice* operation:

```
x ? _ = x
_ ? y = y
```

Thus, the expression “0 ? 1” evaluates to 0 and 1 with the value non-deterministically chosen.

3 Equality vs. Equivalence

Type classes are an important feature to express ad-hoc polymorphism in a structured manner [42]. In the context of Curry, it is also useful to restrict the application of some operations to unintended expressions. For instance, in the definition of Curry without type classes [27], the type of the unification operator is defined as

```
(=:=) :: a → a → Bool
```

This implies that we could unify values of any type, including defined functions. However, the meaning of equality on functions is not well defined. The Curry implementation PAKCS [26], which compiles Curry programs into Prolog programs, uses an intensional meaning, i.e., functions are equal if they have the same name. This means that PAKCS evaluates

```
not =:= not
```

to `True` but it fails on

```
not =:= (\x → not x)
```

(since the lambda abstraction will be lifted into a new top-level function). Moreover, the Curry implementation KiCS2 [12], which compiles Curry programs into Haskell programs, produces an internal error for these expressions.

It would be preferable to forbid the application of “=:=” to functional values at compile time. This is similar to the requirement on Haskell’s operator “==”. Haskell uses the type class `Eq` in order to express that “==” is not parametric polymorphic but overloaded for some (but not all) types. The type class `Eq` contains two operations (we omit the default implementations):

```
class Eq a where
  (==) :: a → a → Bool
  (/=) :: a → a → Bool
```

Hence, the operator “==” cannot be applied to any type but only to types defining instances of this class. We can use this operator to check whether an element occurs in a list:

```
elem :: Eq a => a → [a] → Bool
```

```
elem _ [] = False
elem x (y:ys) = x==y || elem x ys
```

Although type classes express type restrictions in an elegant manner, they might also cause unexpected behaviors if they are not carefully used. For instance, we can define a data type for values indexed by a unique number:

```
data IVal a = IVal Int a
```

Since the index is assumed to be unique, we define the comparison of index values by just comparing the indices:

```
instance Eq a => Eq (IVal a) where
  IVal i1 _ == IVal i2 _ = i1 == i2
```

With this definition, the operation `elem` defined above could yield surprising results:

```
> elem (IVal 1 'b') [IVal 1 'a']
True
```

This is not intended since the element (first argument) does not occur in the list. Actually, the Haskell documentation² about `elem` contains the explanation “Does the element occur in the structure?” which ignores the fact that some instances of `Eq` are only equivalences rather than identities.

This unusual behavior could also influence logic-oriented computations in a surprising manner. If the operation `last` is defined as shown in Sect. 1, we obtain the following answer when computing the last element of a given `IVal` list (here, “_” denotes a logical variable of type `Char`):

```
> last [IVal 1 'a']
IVal 1 _
```

Hence, instead of the last element, we get a rather general representation of it.

The next section presents our proposal to solve these problems.

4 Data

As discussed above, type classes are an elegant way to express type restrictions. On the other hand, it is not a good idea to allow user-defined instance definitions of important operations like strict equality. Therefore, we propose the introduction of a specific type class where only standard instances can be derived so that all instances satisfy the intended meaning. This type class is called `Data` and has the following definition:

```
class Data a where
  aValue :: a
  (===) :: a -> a -> Bool
```

Thus, any instance of this class provides two operations:

- The non-deterministic operation `aValue` returns some value, i.e., the complete evaluation of `aValue` yields all values of type `a`.

² <http://hackage.haskell.org/package/base-4.12.0.0/docs/Prelude.html>.

- The operation “`===`” implements the standard equality on values, i.e., it returns `True` or `False` depending on whether the argument values are identical or not.

The following definition specifies how to automatically derive a `Data` instance for any algebraic datatype.

Definition 1. *If T is an algebraic datatype declared by*

$$\mathit{data} \ T \ a_1 \ \dots \ a_k \ = \ C_1 \ b_{1k_1} \ \dots \ b_{1k_1} \ / \ \dots \ / \ C_n \ b_{nk_n} \ \dots \ b_{nk_n}$$

the standard derived `Data` instance has the following form:

```
instance cx => Data (T a1 ... ak) where
  aValue = C1 aValue ... aValue ? ... ? Cn aValue ... aValue
  C1 x1 ... xk1 === C1 y1 ... yk1 = x1 === y1 &&& ... &&& xk1 === yk1
  ⋮
  Cn x1 ... xkn === Cn y1 ... ykn = x1 === y1 &&& ... &&& xkn === ykn
  Ci _ ... _ === Cj _ ... _ = False ∀i, j ∈ {1, ..., n} with i ≠ j
```

In the instance declaration above, the context cx consists of `Data` constraints ensuring that `Data` b_{ij} holds for each type b_{ij} with $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, k_i\}$.

Example 3. For the type of Peano numbers (see Example 1), the `Data` instance can be defined as follows:

```
instance Data Nat where
  aValue = Z ? S aValue
  Z === Z = True
  S m === S n = m === n
  Z === S _ = False
  S _ === Z = False
```

A `Data` instance for lists requires a `Data` instance for its elements:

```
instance Data a => Data [a] where
  aValue = [] ? aValue : aValue
  [] === [] = True
  (x:xs) === (y:ys) = x === y &&& xs === ys
  [] === (_:_) = False
  (_:_) === [] = False
```

The operation `aValue` is useful when a value of some data type should be guessed, e.g., for testing [25]. The obvious relation to logic variables will be discussed later.

The definition of “`===`” is identical to “`==`” if the definition of the latter is automatically derived (by a “`deriving Eq`” clause). As discussed above, it is also possible to define other instances of `Eq` that leads to unintended results. To ensure that “`===`” always denotes equality on values, *it is not allowed to define explicit*

Data instances as shown above. Such instances can only be generated by adding a “*deriving Data*” clause to a data definition. Note that an instance derivation requires that all arguments of all data constructors have *Data* instances. In particular, if some argument has a functional type, e.g.,

```
data IntRel = IntRel (Int → Bool)
```

then a *Data* instance can not be derived.

For ease of use, one could always derive *Data* instances for data declarations whenever it is possible (i.e., functional values do not occur in arguments), or provide a language option to turn this behavior on or off.

With the introduction of the class *Data*, we can specify a more precise type to Curry’s strict equality operation “*:=:*”. As discussed in [10], the meaning of “*:=:*” is the “positive” part of “*===*”, i.e., its semantics can be defined by

$$x \text{ :=: } y = \text{solve } (x \text{ === } y) \quad (1)$$

where *solve* is an operator that enforces positive evaluations for Boolean expressions:

$$\text{solve True} = \text{True} \quad (2)$$

Since expressions of the form $e_1 \text{ :=: } e_2$ might return *True* but never *False*, “*:=:*” can be implemented by unification, as already discussed in Sect. 2. Such an optimized implementation is justified by the definition (1) above. However, if the semantics of “*:=:*” is defined by

$$x \text{ :=: } y = \text{solve } (x == y) \quad (3)$$

as suggested before the introduction of type classes to Curry [9], an implementation of “*:=:*” by unification would not be correct since unification might put stronger requirements on expressions to be compared than actually defined by *Eq* instances.

As a spin-off of definition (1), we obtain a more restricted type of “*:=:*”:

$$(\text{:=:}) :: \text{Data } a \Rightarrow a \rightarrow a \rightarrow \text{Bool} \quad (4)$$

This avoids the problems with the application of “*:=:*” to functional values sketched at the beginning of Sect. 3.

5 Logic Variables

When a function call with free variables in arguments is evaluated by narrowing, the free variables are instantiated to values so that the function call becomes reducible. Conceptually, a free variable denotes possible values so that a computation can pick one in order to proceed. With the definition of the type class *Data* and the non-deterministic operation *aValue*, we make the notion of “possible value” explicit. Actually, it has been shown that non-deterministic operations and logic variables have the same expressive power [5, 14] since one can replace logic variables occurring in a functional logic program by non-deterministic value generators.

Example 4. Consider the addition on Peano numbers shown in Example 1 which is exploited to define subtraction:

```
sub :: Nat → Nat → Nat
sub x y | add y z === x = z
  where z free
```

We can replace the logic variable `z` by a value generator:

```
sub x y | add y z === x = z
  where z = aValue
```

The equivalence of logic variables and non-deterministic value generators can be exploited when Curry is implemented by translation into a target language without support for non-determinism and logic variables. For instance, KiCS2 [12] compiles Curry into Haskell by adding only a mechanism to handle non-deterministic computations. Therefore, KiCS2 is able to evaluate a logic variable to all its values. Thus, KiCS2 could exploit this fact by using the following alternative definition for `aValue`:

$$\text{aValue} = _ \tag{5}$$

This equivalence also sheds some new light on the type of logic variables. Currently, logic variables without any constraints on their types are considered to have a polymorphic type. For instance, the inferred type of `aValue` as defined in (5) is

```
aValue :: a
```

However, this type does not really describe the intent of this operation, since `aValue` does not yield functional values. For instance, consider the definition

```
f x = y where y free
```

The type currently inferred is

```
f :: a → b
```

However, it is meaningless to use the result of some application of `f` in contexts where a function is required. For instance, the evaluation of the expression

$$\text{map (f True) [0,1]} \tag{6}$$

suspends in PAKCS and produces a run-time error in KiCS2 (very similar to the examples described at the beginning of Sect. 3). Furthermore, the inferred type of the definition

```
g x = g x
```

is

```
g :: a → b
```

Thus, it looks very similar to the type of `f` although `g` has a quite different meaning: in contrast to `f`, an application of `g` never returns a value.

All these problems can be avoided by a simple fix: logic variables are considered as equivalent to the operation `aValue` of type class `Data` so that a logic variable without any constraints on its type has type `a` where `a` is constrained with the type class context `Data a`. With this change, the inferred type of `f` is

```
f :: Data b => a -> b
```

As a consequence, expression (6) will be rejected by the type checker since functions have no `Data` instance.

6 Equality Optimization

Choosing the appropriate kind of equality might not be obvious to the programmer. The difference between identity and equivalence is semantically relevant so that the decision between “`===`” and “`==`” is not avoidable. However, “`:=`” can be considered as an optimization of “`===`” so that it is not obvious when it should be applied. In order to simplify this situation, it has been argued in [9,10] that the programmer should always use strict equality (i.e., “`===`”) and the selection of “`:=`” should be done by an optimization tool. This tool analyzes the required values of Boolean expressions. If an application of strict equality requires only the result value `True`, e.g., in guards of conditional rules or in arguments of `solve`, see (2), then one can safely replace the equality operator by the unification operator “`:=`” (see [10] for details). For instance, if `last` is defined by

```
last xs | _ ++ [e] === xs
       = e  where e free
```

then it can be transformed into

```
last xs | _ ++ [e] := xs
       = e  where e free
```

As shown in [10], this transformation can have a big impact on the execution time.

Up to now, this tool (which is part of the compilation chain of Curry systems) considered the optimization of calls to “`==`”. Since this might lead to incompleteness, as discussed above, it has to consider calls to “`===`” when the type class `Data` is introduced. However, for backward compatibility and better optimizations, one can extend the optimizer also to calls of the form $e_1 == e_2$: if the types of the arguments e_1, e_2 are monomorphic and the `Eq` instances of these types are derived with the default scheme (by `deriving` annotations), the semantics of “`==`” is identical to the semantics of “`===`” so that one can replace $e_1 == e_2$ by $e_1 === e_2$ and apply the optimization sketched above.

7 Non-left-Linear Rules and Functional Patterns

The proposed introduction of the type class `Data` together with the adjusted type of the unification operator “`:=`” has also some influence on language constructs where unification is implicitly used. We discuss this in more detail in this section.

In contrast to Haskell, Curry allows non-left-linear rules, i.e., defining rules with multiple occurrences of a variable in the patterns of the left-hand side. For instance, this function definition is valid in Curry:

```
f x x = x
```

Multiple occurrences of variables in the left-hand side are considered as an abbreviation for equational constraints between these occurrences [27], i.e., the definition above is expanded to

```
f x y | x := y = x
```

This feature of Curry is motivated by logic programming where multiple variable occurrences in rule heads are also solved by unification. However, in Curry the situation is a bit more complex due to the inclusion of functions and infinite data structures. As a matter of fact, our refined type of “`:=`” makes the status of non-left-linear rules clearer. According to the type shown in (4), the type inferred for the definition above is

```
f :: Data a => a → a → a
```

Hence, `f` can not be called with functional values as arguments. This even increases the compatibility with logic programming where unification is applied to Herbrand terms, i.e., algebraic data.

Another feature of Curry, where equational constraints are implicitly used, are functional patterns. *Functional patterns* are proposed in [4] as an elegant way to describe pattern matching with an infinite set of patterns. For instance, consider the definition of `last` shown above. Since the equational condition requires the complete evaluation of the input list, an expression like `last [failed,3]` (where `failed` is an expression that has no value) can not be evaluated to some value. Now, consider that `last` is defined by the following (infinite) set of rules:

```
last [x] = x
last [x1,x] = x
last [x1,x2,x] = x
⋮
```

Then the expression above is reduced to the value `3` by applying the second rule. This set of rules can be abbreviated by a single rule:

$$\text{last } (_ ++ [x]) = x \tag{7}$$

Since the argument contains the defined operation “`++`”, it is called a *functional pattern*. Conceptually, a functional pattern denotes all constructor terms to which it can be evaluated (by narrowing). In this case, these are the patterns

shown above. Operationally, pattern matching with functional patterns can be implemented by a specific unification procedure which evaluates the functional pattern in a demand-driven manner [4]. Functional patterns are useful to express pattern matching at arbitrary depths in a compact manner. For instance, they can be exploited for a compact and declarative approach to process XML documents [23].

A delicate point of functional patterns are non-linear patterns, i.e., if a functional pattern is evaluated to some constructor term containing multiple occurrences of a variable. For instance, consider the function

```
dup :: a → (a,a)
dup x = (x,x)
```

and its use in a functional pattern:

```
whenDup (dup x) = x
```

By the semantics of functional patterns, the latter rule is equivalent to the definition

```
whenDup (x,x) = x
```

Due to the non-linear left-hand side, the type of `whenDup` is

```
whenDup :: Data a => (a,a) → a
```

Now, consider the operation `const` defined by

```
const :: a → b → a
const x _ = x
```

and its use in a functional pattern:

$$g(\text{const } x \ x) = x \tag{8}$$

By the semantics of functional pattern, the definition of `g` is equivalent to

```
g x = x
```

so that a correct type is

```
g :: a → a
```

Hence, the type context `Data a` is not required, although the variable `x` has a multiple occurrence in (8). This example shows that, if functional patterns are used, the requirement for a `Data` context depends on the linearity of the constructor terms to which the functional patterns evaluate. Since this property is undecidable in general, a safe approximation is to add a `Data` constraint to the result type of the functional pattern. This has the consequence that the type of `last`, when defined as in (7), is inferred as

```
last :: Data a => [a] → a
```

Basically, this type is the same as we would obtain when defining `last` with an equational constraint, but it could be done better: since the functional pattern `(_ ++ [x])` always yields a linear term, the type class constraint `Data a` is not

necessary. Hence, one can make the type checking for operations defined with functional patterns more powerful by approximating the linearity property of the functional pattern. Such an approximation has already been used in [4] to improve the efficiency of the unification procedure for functional patterns. However, a significant drawback would be the fact that the inferred type of a function would depend on the quality of the approximation. As a consequence, the principal type of a function [13,28] would become ambiguous under certain circumstances and would depend on a function’s implementation.

8 Related Work

We already discussed in the previous sections some work related to the interpretation and use of equality in declarative languages. In the following, we focus on some additional work related to our proposal.

The necessity to distinguish different equalities in the context of functional logic programming and to define their exact semantics has been recognized before. In [16], the authors introduce several equality (and disequality) operations, among others also an operation for strict equality. However, no explicit distinction between equality and equivalence is made as only the former is discussed. Note also that some of these operations became obsolete with [9].

In [33], the author discusses the addition of Haskell-like overloading to Curry. In doing so, a new type class `Equal` that contains the unification operation “`:=`” is proposed. The intent is to restrict this operation similarly to the equivalence operation “`==`” so that it is only applicable to certain types. In contrast to our proposal, it is not enforced that instances of the `Equal` type class should always have the same form. In the same work, another type class `Narrowable` containing a method called `narrow` is proposed in order to restrict the type of logical variables against the background of higher-rank types. The method `narrow` is very similar to our method `aValue`. But aside from a few downsides of the introduction of such a method, e.g., a possibly fixed order when enumerating solutions, no further consequences for the language itself are discussed in that work.

The idea to use a type relation to restrict the type of logical variables has also been introduced in [35] for a better characterization of free theorems. In [34], a type class `Data` is used for the same reason, but the class is only used as a marker (as in [35]) so that the type class does not contain any methods.

On a side note, there is also a `Data` type class in Haskell. However, this particular type class is used for generic programming in Haskell and shares nothing but the name with our type class [32].

9 Conclusions

In this paper we presented a solution to various problems w.r.t. equality and logic variables in functional logic programs by introducing a new type class `Data`. Instances of this class support a generator operation `aValue` for values and a strict equality operation “`===`” on these values. In contrast to other classes,

instances of this class can only be derived in a standard manner and cannot be defined by the programmer. This decision ensures a reasonable semantics: if $e_1 === e_2$ evaluates to `True`, then the expressions e_1 and e_2 have an identical value. Although this is the notion of strict equality proposed for a long time, Haskell-like overloading of the class `Eq` and its operation “`==`” allows to specify that “some expressions are more equal than others” [38].

At a first glance, it might be unnecessary to add a further equality operator and base type class to a declarative language. The advantage is that this supports a clear documentation for all functions depending on equality, as it makes a huge difference in functional logic programming whether one imposes equality or equivalence in a function’s implementation. If a programmer is interested in identical values, she or he has to use “`===`”.³ If only equivalence is relevant, “`==`” is the right choice. For instance, consider the operation `elem` to check whether an element occurs in a list. The type

```
elem :: Data a => a -> [a] -> Bool
```

indicates that this operation succeeds if the element actually occurs in the list, whereas the type

```
elem :: Eq a => a -> [a] -> Bool
```

indicates that it succeeds if some equivalent element is contained in the list.

Unfortunately, these details are often not taken into account. As discussed in this paper, many textbooks and program documentations simply ignore such differences or are not formally precise in their statements.

We showed that our proposal is also useful to type logic variables in a more meaningful way. The type of a logic variable is required to be an instance of `Data` so that one can enumerate the possible values of this variable. Although logic variables are often instantiated by narrowing or unification to appropriate values, there are situations where an explicit enumeration is necessary to ensure completeness. For instance, consider the encapsulation of non-deterministic computations in order to reason about the various outcomes. Set functions [6] are a declarative, i.e., evaluation-independent, encapsulation approach. If f is a (unary) function, its set function f_S returns the set of all results computed by f for a given argument. For instance, `someDupS xs` returns the set of all duplicate elements (see Example 2) occurring in the list `xs`. An important property of a set function is that it encapsulates only the non-determinism caused by the function’s definition and not by the arguments. Hence, `someDupS ([1,1] ? [2])` yields two different sets: `{1}` and `{}`. This property of set functions is important to ensure their declarative semantics. It has the consequence that arguments must be evaluated *outside* the set function. Hence, to evaluate the expression

```
let x free in ... (fS x) ...
```

³ As discussed in Sect. 6, the unification operator “`:=`” does not need to be used by the programmer since it is an optimization of “`===`”.

it is not allowed to bind x inside the evaluation of f . As a consequence, x must be instantiated outside in order to proceed a computation where f demands its argument. This can easily be obtained by the use of the operation `aValue`:

```
let x = aValue in ...(fS x)...
```

In order to evaluate the practical consequences of our proposal, we implemented it in a prototypical manner in our Curry front end that is used by various Curry implementations. The changes in the type checker were minimal (e.g., adding `Data` contexts to the inferred types of logic variables). Concerning libraries, only a single type signature had to be adapted in the standard prelude, one of the largest Curry modules: the type of the “arbitrary value” operation gets a `Data` context:

```
unknown :: Data a => a
unknown = let x free in x
```

In other libraries, only a few types (related to search encapsulation primitives) had to be adapted. With these few changes, even larger Curry applications could be compiled without problems. This demonstrates that our proposal is a viable alternative to the current unsatisfying handling of equality and logic variables in Curry. Usually, no changes are necessary in existing Curry programs. Only in the rare cases of function definitions with polymorphic non-linear left-hand sides or polymorphic logic variables, type signatures have to be adapted.

Acknowledgments. The authors are grateful to Sandra Dylus and Marius Rasch for fruitful discussions during the conception phase of this paper. Furthermore, we thank Kai-Oliver Prott for his efforts in evaluating our proposal by prototypically implementing it.

References

1. Antoy, S.: Optimal non-deterministic functional logic computations. In: Hanus, M., Heering, J., Meinke, K. (eds.) ALP/HOA-1997. LNCS, vol. 1298, pp. 16–30. Springer, Heidelberg (1997). <https://doi.org/10.1007/BFb0027000>
2. Antoy, S., Echahed, R., Hanus, M.: A needed narrowing strategy. *J. ACM* **47**(4), 776–822 (2000)
3. Antoy, S., Hanus, M.: Functional logic design patterns. In: Hu, Z., Rodríguez-Artalejo, M. (eds.) FLOPS 2002. LNCS, vol. 2441, pp. 67–87. Springer, Heidelberg (2002). https://doi.org/10.1007/3-540-45788-7_4
4. Antoy, S., Hanus, M.: Declarative programming with function patterns. In: Hill, P.M. (ed.) LOPSTR 2005. LNCS, vol. 3901, pp. 6–22. Springer, Heidelberg (2006). https://doi.org/10.1007/11680093_2
5. Antoy, S., Hanus, M.: Overlapping rules and logic variables in functional logic programs. In: Etalle, S., Truszczyński, M. (eds.) ICLP 2006. LNCS, vol. 4079, pp. 87–101. Springer, Heidelberg (2006). https://doi.org/10.1007/11799573_9
6. Antoy, S., Hanus, M.: Set functions for functional logic programming. In: Proceedings of the 11th ACM SIGPLAN International Conference on Principles and Practice of Declarative Programming (PPDP 2009), pp. 73–82. ACM Press (2009)

7. Antoy, S., Hanus, M.: Functional logic programming. *Commun. ACM* **53**(4), 74–85 (2010)
8. Antoy, S., Hanus, M.: New functional logic design patterns. In: Kuchen, H. (ed.) WFLP 2011. LNCS, vol. 6816, pp. 19–34. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-22531-4_2
9. Antoy, S., Hanus, M.: Curry without success. In: Proceedings of the 23rd International Workshop on Functional and (Constraint) Logic Programming (WFLP 2014). CEUR Workshop Proceedings, vol. 1335, pp. 140–154. CEUR-WS.org (2014)
10. Antoy, S., Hanus, M.: Transforming Boolean equalities into constraints. *Formal Aspects Comput.* **29**(3), 475–494 (2017). <https://doi.org/10.1007/s00165-016-0399-6>
11. Bird, R.: Introduction to Functional Programming using Haskell, 2nd edn. Prentice Hall, Englewood Cliffs (1998)
12. Braßel, B., Hanus, M., Peemöller, B., Reck, F.: KiCS2: a new compiler from Curry to Haskell. In: Kuchen, H. (ed.) WFLP 2011. LNCS, vol. 6816, pp. 1–18. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-22531-4_1
13. Damas, L., Milner, R.: Principal type-schemes for functional programs. In: Proceedings of the 9th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL 1982), pp. 207–212. ACM, New York (1982)
14. de Dios Castro, J., López-Fraguas, F.J.: Extra variables can be eliminated from functional logic programs. *Electron. Notes Theor. Comput. Sci.* **188**, 3–19 (2007)
15. DeGroot, D., Lindstrom, G. (eds.): Logic Programming, Functions, Relations, and Equations. Prentice Hall, Englewood Cliffs (1986)
16. Gallego Arias, E.J., Mariño Carballo, J., Rey Poza, J.M.: A proposal for disequality constraints in Curry. *Electron. Notes Theor. Comput. Sci.* **177**, 269–285 (2007). Proceedings of the 15th Workshop on Functional and (Constraint) Logic Programming (WFLP 2006)
17. Gallier, J.H., Snyder, W.: Complete sets of transformations for general E-unification. *Theoret. Comput. Sci.* **67**, 203–260 (1989)
18. Giovannetti, E., Levi, G., Moiso, C., Palamidessi, C.: Kernel leaf: a logic plus functional language. *J. Comput. Syst. Sci.* **42**(2), 139–185 (1991)
19. Goguen, J.A., Meseguer, J.: EQLOG: equality, types, and generic modules for logic programming. In: DeGroot, D., Lindstrom, G. (eds.) Logic Programming, Functions, Relations, and Equations, pp. 295–363. Prentice Hall (1986)
20. González-Moreno, J.C., Hortalá-González, M.T., López-Fraguas, F.J., Rodríguez-Artalejo, M.: An approach to declarative programming based on a rewriting logic. *J. Log. Program.* **40**, 47–87 (1999)
21. Hanus, M.: The integration of functions into logic programming: from theory to practice. *J. Log. Program.* **19&20**, 583–628 (1994)
22. Hanus, M.: A unified computation model for functional and logic programming. In: Proceedings of the 24th ACM Symposium on Principles of Programming Languages (Paris), pp. 80–93 (1997)
23. Hanus, M.: Declarative processing of semistructured web data. In: Technical Communications of the 27th International Conference on Logic Programming, vol. 11, pp. 198–208. Leibniz International Proceedings in Informatics (LIPIcs) (2011)
24. Hanus, M.: Functional logic programming: from theory to Curry. In: Voronkov, A., Weidenbach, C. (eds.) Programming Logics. LNCS, vol. 7797, pp. 123–168. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-37651-1_6
25. Hanus, M.: CurryCheck: checking properties of Curry programs. In: Hermenegildo, M.V., Lopez-Garcia, P. (eds.) LOPSTR 2016. LNCS, vol. 10184, pp. 222–239. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-63139-4_13

26. Hanus, M., et al.: PAKCS: The Portland Aachen Kiel Curry System (2018). <http://www.informatik.uni-kiel.de/~pakcs/>
27. Hanus, M. (ed.): Curry: an integrated functional logic language (vers. 0.9.0) (2016). <http://www.curry-language.org>
28. Hindley, R.: The principal type-scheme of an object in combinatory logic. *Trans. Am. Math. Soc.* **146**, 29–60 (1969)
29. Huet, G., Lévy, J.-J.: Computations in orthogonal rewriting systems. In: Lassez, J.-L., Plotkin, G. (eds.) *Computational Logic: Essays in Honor of Alan Robinson*, pp. 395–443. MIT Press (1991)
30. Hughes, J.: Why functional programming matters. In: Turner, D.A. (ed.) *Research Topics in Functional Programming*, pp. 17–42. Addison Wesley (1990)
31. Hutton, G.: *Programming in Haskell*, 2nd edn. Cambridge University Press, Cambridge (2016)
32. Lämmel, R., Peyton Jones, S.L.: Scrap your boilerplate: a practical design pattern for generic programming. In: *Proceedings of the 2003 ACM SIGPLAN International Workshop on Types in Languages Design and Implementation (TLDI 2003)*, pp. 26–37. ACM Press (2003)
33. Lux, W.: Adding Haskell-style overloading to Curry. In: 25. Workshop der GI-Fachgruppe Programmiersprachen und Rechenkonzepte, Kiel, Germany, pp. 67–76 (2008). Technical report 0811, Institut für Informatik, CAU Kiel
34. Mehner, S.: Tools for reasoning about effectful declarative programs. Ph.D. thesis, Universität Bonn (2015)
35. Mehner, S., Seidel, D., Straßburger, L., Voigtländer, J.: Parametricity and proving free theorems for functional-logic languages. In: *Proceedings of the 16th International Symposium on Principle and Practice of Declarative Programming (PPDP 2014)*, pp. 19–30. ACM Press (2014)
36. Moreno-Navarro, J.J., Rodríguez-Artalejo, M.: Logic programming with functions and predicates: the language babel. *J. Log. Program.* **12**, 191–223 (1992)
37. O’Donnell, M.J.: Equational logic programming. In: Gabbay, D.M., Hogger, C.J., Robinson, J.A. (eds.) *Handbook of Logic in Artificial Intelligence and Logic Programming*, pp. 69–161. Oxford University Press (1998)
38. Orwell, G.: *Animal Farm: A Fairy Story*. Secker and Warburg, London (1945)
39. Peyton Jones, S. (ed.): *Haskell 98 Language and Libraries-The Revised Report*. Cambridge University Press, Cambridge (2003)
40. Reddy, U.S.: Narrowing as the operational semantics of functional languages. In: *Proceedings IEEE International Symposium on Logic Programming*, Boston, pp. 138–151 (1985)
41. Thompson, S.: *Haskell - The Craft of Functional Programming*, 2nd edn. Addison-Wesley, Boston (1999)
42. Wadler, P., Blott, S.: How to make ad-hoc polymorphism less ad hoc. In: *Proceedings POPL 1989*, pp. 60–76 (1989)