

# Contracts and Specifications for Functional Logic Programming

Sergio Antoy  
Computer Science Dept.  
Portland State University, Oregon, U.S.A.  
antoy@cs.pdx.edu

Michael Hanus  
Institut für Informatik  
CAU Kiel, D-24098 Kiel, Germany  
mh@informatik.uni-kiel.de

## Abstract

The expressive power of functional logic languages supports high-level specifications as well as efficient implementations of problems in the same language. If specifications are executable, they can be used both as initial prototypical implementations and as contracts for checking the reliable execution of implementations expected to satisfy the specification. In this paper, we propose a concrete framework to support this general approach to coding. We discuss the notions of specifications and contracts for functional logic programming and present a tool that supports the development of declarative programs based on these notions.

## 1 Introduction

Functional logic programming languages [3, 15] support a wide spectrum of programming styles. One can apply logic programming features like nondeterminism and logic variables to specify the basic knowledge about a problem and let the run-time system search for appropriate solutions. On the other hand, one can also use a deterministic (functional) programming style to implement sophisticated and efficient algorithms [23].

The combination of both can be exploited to develop more reliable software: high-level (“obviously correct”) specifications can be formulated as functional logic programs. Since they are executable, they can serve as initial prototypical implementations. Executable specifications are useful to run experiments and increase the confidence that a specification captures the intent. If the direct execution of the specification is too inefficient, one can choose more efficient data structures (e.g., balanced search trees instead of lists) or better algorithms. In this case, the initial specification remains valuable since one can use it as an oracle to test the implementation on a large set of test data [8, 13] or to check, via run-time assertions, that the implementation behaves as intended on particular executions.

In this paper we show the feasibility of this idea by making the necessary notions, like specifications, contracts, or assertions, precise, showing some important relations between them, and providing tool support based on this development. The concrete language for our presentation is the multi-paradigm declarative language Curry [17]. We demonstrate that Curry can be used as a wide-spectrum language [5] for software development using a tool for Curry programs that transforms specifications into contracts attached to the corresponding implementations, if present.

Although we assume familiarity with the general concepts of functional logic programming [3, 15], we review in the next section the concepts crucial for this paper. Section 3 presents the fundamental notions of our framework. The corresponding tool support is sketched in Section 4 together with some examples.

## 2 Functional Logic Programming and Curry

The declarative multi-paradigm language Curry [17] extends non-strict functional programming languages such as Haskell [24] with logic programming features, e.g., nondeterminism and equa-

tional constraints. Consequently, Curry has a Haskell-like syntax<sup>1</sup> extended by the possible inclusion of free (logic) variables in conditions and right-hand sides of defining rules. The operational semantics is based on an optimal evaluation strategy [1] which is a conservative extension of lazy functional programming and (concurrent) logic programming.

*Expressions* in Curry programs contain *operations* (defined functions), *constructors* (introduced in data type declarations), and *variables* (arguments of operations or free variables). The goal of a computation is to obtain a value of some expression, where a *value* is an expression that does not contain any operation. Note that in a functional logic language expressions might have more than one value due to nondeterministically defined operations. For instance, Curry contains a *choice* operation defined by:

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

Thus, the expression “0 ? 1” has two values: 0 and 1. If expressions have more than one value, these values are typically constrained by conditions in the rules defining operations according to the program intent. A *rule* has the form “ $f\ t_1 \dots t_n \mid c = e$ ” where  $c$  is a *constraint*, i.e., an expression of the built-in type `Success`. For instance, the trivial constraint `success` is a value of type `Success` that denotes the always satisfiable constraint. An *equational constraint*  $e_1 =: e_2$  is satisfiable if both sides  $e_1$  and  $e_2$  are reducible to unifiable values. Furthermore, if  $c_1$  and  $c_2$  are constraints,  $c_1 \& c_2$  denotes their concurrent conjunction (i.e., both constraints are concurrently evaluated) and  $c_1 \&> c_2$  denotes their sequential conjunction (i.e.,  $c_2$  is evaluated after the successful evaluation of  $c_1$ ).

Nondeterministic expressions could cause a semantical ambiguity when bound to variables. Consider the operations

```
coin = 0 ? 1
double x = x + x
```

Standard term rewriting produces, among others, the derivation

```
double coin → coin + coin → 0 + coin → 0 + 1 → 1
```

whose result is unintended. Therefore, González-Moreno et al. [14] proposed the rewriting logic CRWL as a logical foundation for declarative programming with non-strict and nondeterministic operations. This logic specifies the *call-time choice* semantics [18] where values of the arguments of an operation are determined before the operation is evaluated. In a lazy strategy, this is naturally obtained by sharing. For instance, the two occurrences of `coin` in the derivation above are shared so that “`double coin`” has only the results: 0 or 2. Since standard term rewriting does not conform to the intended call-time choice semantics, other notions of rewriting have been proposed to formalize this idea, like graph rewriting [11, 12] or let rewriting [20]. For our purposes, it is sufficient to use a simple reduction relation that we sketch without giving all details (which can be found in [20]).

To cover non-strict computations, expressions can also contain the special symbol  $\perp$  to represent *undefined or unevaluated values*. A *partial value* is a value containing occurrences of  $\perp$ . A *partial constructor substitution* is a substitution that replaces variables by partial values. A *context*  $\mathcal{C}[\cdot]$  is an expression with some “hole”. Then the reduction relation we use throughout this paper is defined as follows (conditional rules are not considered for the sake of simplicity):

---

<sup>1</sup>Variables and function names usually start with lowercase letters and the names of type and data constructors start with an uppercase letter. The application of  $f$  to  $e$  is denoted by juxtaposition (“ $f\ e$ ”).

$$\begin{aligned} \mathcal{C}[f \sigma(t_1) \dots \sigma(t_n)] &\rightarrow \mathcal{C}[\sigma(r)] && f \ t_1 \dots t_n \rightarrow r \text{ program rule,} \\ &&& \sigma \text{ partial constructor substitution} \\ \mathcal{C}[e] &\rightarrow \mathcal{C}[\perp] \end{aligned}$$

The first rule models call-time choice: if a rule is applied, the actual arguments of the operation must have been evaluated to partial values. The second rule models non-strictness by allowing the evaluation of any subexpression to an undefined value (which is intended if the value of this subexpression is not demanded). As usual,  $\rightarrow^*$  denotes the reflexive and transitive closure of this reduction relation. The equivalence of this rewrite relation and CRWL is shown in [20].

Sometimes we use `let`-expressions to enforce the call-time choice semantics. In order to avoid the explicit handling of `let`-expressions in the reduction relation (as done in [20]), we consider `let`-expressions as syntactic sugar for auxiliary functions. For instance, the definition

```
f x = let z = coin*x in z+coin
```

is syntactic sugar for

```
f x = g (coin*x)
g z = z+coin
```

where `g` is a fresh name.

In nondeterministic programming, it is sometimes useful to examine the set of all the values of some expression. A “set-of-values” operation applied to an arbitrary argument might produce results that depend on the degree of evaluation of the argument (see [6] for a detailed discussion). *Set functions* overcome this problem [2]. For each defined function  $f$ ,  $f_S$  denotes the corresponding set function.  $f_S$  encapsulates the nondeterminism of  $f$ , but excludes the potential nondeterminism of the arguments to which  $f$  is applied. For instance, consider the operation `decOrInc` defined by:

```
negOrPos x = -x ? x
```

Then “`negOrPos_S 2`” evaluates to the set  $\{-2, 2\}$ , i.e., the nondeterminism originating from `negOrPos` is encapsulated into a set. However, “`negOrPos_S (1?2)`” evaluates to two different sets  $\{-1, 1\}$  and  $\{-2, 2\}$  due to its nondeterministic argument, i.e., the nondeterminism originating from the argument produces different sets. The type *set* is abstract, i.e., the implementation is hidden but there are operations, e.g., to determine whether a set is empty, `isEmpty`, or an element belongs to a set.

### 3 Specifications and Contracts

Our framework to support the development of reliable declarative programs is based on the idea of using a single language for specifications, contracts, and implementations. A functional logic language such as Curry is appropriate for this purpose because it is nondeterministic and it has equation-solving capabilities.

Using the same language makes specifications and implementations similar. In fact, a specification is like any other operation but with a specific tag so that the specification is more versatile:

- If there is only a specification but no implementation of an operation, the specification is used as an initial implementation for this operation.
- If there are both a specification and an implementation of an operation, the specification can be used to check the implementation in two different ways:

**Dynamic checking:** If the implementation computes some result when the operation is executed, test whether this result conforms to the specification.

**Static checking:** If one formally proves that the implementation is correct w.r.t. the specification, run-time checking is not necessary.

Hence, we distinguish between a specification and a contract for an operation. A *specification* describes precisely the intended meaning of an operation. However, a *contract* describes conditions that must be satisfied by the implementation. These conditions can be weaker than a specification. Contracts have been introduced in the context of imperative and object-oriented programming languages [22] to improve the quality of software. Typically, a contract consists of both a pre- and a postcondition. The *precondition* is an obligation for the arguments of an operation application. The *postcondition* is an obligation for both the arguments of an operation application and the result of the operation application to those arguments. Intuitively, the application or call to each operation must satisfy its precondition, and, if both the precondition is satisfied and the operation returns a result, this result must satisfy the postcondition. When a contract is checked at run-time, the pre- and postcondition are called *assertions*.

Specifications, preconditions, and postconditions are independent notions separately useful for software development. A precondition for an operation states general restrictions on arguments that must be satisfied in order to apply this operation. Hence, a specification is intended only for inputs satisfying the precondition. Likewise, a postcondition must only be satisfied for these inputs. In a strongly typed language, a type restriction on arguments can be considered a precondition. In general, one is interested in preconditions that are more expressive than a traditional type system. For instance, a precondition for a factorial function could require the argument to be non-negative. A postcondition is some requirement on all results of an operation. It could be a type restriction, but it could also be much stronger. For instance, a postcondition for an operation to sort a list of values could state that the length of the output list is identical to the length of the input list. If a postcondition specifies all and only the intended results of an operation, it can be considered a specification. As we will see later, we can exploit the logic programming features of our language to execute a postcondition as a prototypical implementation by generating result values satisfying the postcondition.

The following definition fixes the notions discussed so far. For the sake of simplicity, we formally define our notions only for unary operations, but the extension to operations with several arguments is straightforward and, thus, it will be used in the subsequent examples.

**Definition 1** (Specification, Contract). *Let  $f$  be an operation of type  $\tau \rightarrow \tau'$ . A specification for  $f$  is an operation  $f^{spec}$  of type  $\tau \rightarrow \tau'$ . A precondition for  $f$  is an operation  $f^{pre}$  of type  $\tau \rightarrow \text{Bool}$ . A postcondition for  $f$  is an operation  $f^{post}$  of type  $\tau \rightarrow \tau' \rightarrow \text{Bool}$ . A precondition and postcondition pair is also called a contract for the operation. If a precondition is not explicitly defined, the most general precondition “ $f^{pre} \_ = \text{True}$ ” is assumed.*

Similarly to other proposals for assertions or contracts for functional (logic) programs (e.g., [7, 9, 16]), we define pre- and postconditions as Boolean functions. An exception is [4] where constraints are used as conditions which was motivated by the use of postconditions as specifications instead of an explicit specification as in this work.

As an example, consider an operation, `sort`, to sort a list of integers. The type of `sort` is:

```
sort :: [Int] → [Int]
```

Since we have no further requirements on arguments (apart from its type), our precondition for

`sort` is the constant operation<sup>2</sup>

```
sort'pre :: [Int] → Bool
sort'pre _ = True
```

As an example for a postcondition, we require that the length of the input and output lists must be equal:

```
sort'post :: [Int] → [Int] → Bool
sort'post xs ys = length xs == length ys
```

However, a precise specification states that the result of `sort` is a permutation in ascending order of its input:

```
sort'spec :: [Int] → [Int]
sort'spec xs | sorted ys = ys where ys = perm xs
```

This specification requires the definition of permutations and sorted lists which are easily formalized in Curry (“`<=`” denotes the less-or-equal constraint):

```
perm [] = []
perm (x:xs) = ndinsert x (perm xs)
  where
    ndinsert x ys = x : ys
    ndinsert x (y:ys) = y : ndinsert x ys

sorted [] = success
sorted [_] = success
sorted (x:y:ys) = x<=:y & sorted (y:ys)
```

We can use the specification `sort'spec` to sort lists since it is a Curry program and, thus, executable. Obviously, it is inefficient for larger lists so that we implement it more efficiently using the well-known quicksort algorithm:

```
sort :: [Int] → [Int]
sort [] = []
sort (x:xs) = sort (filter (<x) xs) ++ [x] ++
              sort (filter (>x) xs)
```

If we apply our tool, `DSDCurry`, to this program, the specification is transformed into an additional postcondition and all existing pre- and postconditions are attached to the `sort` operation for dynamic assertion checking. Therefore, the execution of this transformed program reveals an error in our implementation:

```
SortC> sort [5,1,2,6,5,3]
ERROR: Postcondition of operation 'sort' violated for:
[5,1,2,6,5,3] → [1,2,3,5,6]
```

If we correct the error, by replacing the condition `(>x)` with `(>=x)`, the transformed program behaves as expected.

Before discussing some details about our tool, we have to define the precise meaning of correct implementations and violated assertions. In imperative or strict functional languages, this seems obvious. However, in a functional logic language like Curry, operations might have multiple results or reduce to infinite structures (i.e., their evaluation does not terminate). In order to support contract checking also in these situations, we have to prepare an appropriate setup.

---

<sup>2</sup>Note that in the concrete syntax we use in our tool (see below) we write `f'pre` instead of `fpre` (and similarly for postconditions and specifications).

First, we consider the possible violation of contracts. Obviously, a precondition  $f^{pre}$  is violated for some expression  $e$  if  $f^{pre} e$  is reducible to **False**, since we want to avoid that any calls on operations where the argument does not satisfy the precondition. For postconditions, the situation is less clear for nondeterministic functions. Consider a value  $v$  such that  $f^{pre} v$  is reducible to **True**,  $f v \xrightarrow{*} v_1$ ,  $f v \xrightarrow{*} v_2$ , and  $f^{post} v v_1 \xrightarrow{*} \mathbf{True}$ , but  $f^{post} v v_2 \xrightarrow{*} \mathbf{False}$ , i.e., one result,  $v_1$ , satisfies the postcondition but another result for the same input,  $v_2$ , does not satisfy the postcondition. In a complete implementation, all results of an operation could be produced. Therefore, we propose the strong view in which *all* the results of a function must satisfy the function's postcondition.

**Definition 2** (Violation). *Let  $f$  be an operation of type  $\tau \rightarrow \tau'$ ,  $f^{pre}$  and  $f^{post}$  be pre- and postconditions for  $f$ , and  $e$  an expression of type  $\tau$ . A violation of the precondition  $f^{pre}$  of  $f$  at  $e$  is a derivation of  $f^{pre} e$  to **False**. A violation of the postcondition  $f^{post}$  of  $f$  at  $e$  is a derivation of (where  $x$  is a fresh variable)*

```
let x = e in not (fpre x) || fpost x (f x)
```

to **False**.

The definition of a postcondition violation considers the fact that a violation should be reported only if the precondition holds for the given argument. Note that the `let`-expression is reasonable for nondeterministic arguments since the condition “`not (fpre e) || fpost e (f e)`” is different from the one given in the above definition if  $e$  is nondeterministic. For instance, consider

```
f'post x y = x==y
f x = x
e = 0 ? 1
```

Then “`f'post e (f e)`” reduces to **True** or **False** whereas

```
let x = e in f'post x (f x)
```

cannot reduce to **False** due to the call-time choice semantics. The intent is that the postcondition should be satisfied for the same values used in the precondition, thus, our definition captures this demand.

Next we have to define the correctness of an implementation w.r.t. a given specification. A simple approach could require that the values of the specification are all and only the values of the implementation. However, this is not reasonable for non-strict languages. For instance, consider

```
nums'spec n = n : nums'spec (n+1)
```

Since `nums'spec` has no value as a result (its evaluation to some value does not terminate), any other operation (of the same type) without a result value would be correct w.r.t. this specification, e.g.:

```
nums n = n : nums n
```

Obviously, this is not intended. If we put the specification and the implementation in an identical context (e.g., applying “`take 2`” to `nums'spec` and `nums`), then we might obtain different results. This motivates the following definition.

**Definition 3** (Equivalence, Correctness). *Let  $f_1, f_2$  be operations of type  $\tau \rightarrow \tau'$ .  $f_1$  is equivalent to  $f_2$  iff, for any expression  $E_1$ ,  $E_1 \xrightarrow{*} v$  iff  $E_2 \xrightarrow{*} v$ , where  $v$  is a value and  $E_2$  is obtained from  $E_1$  by replacing any occurrence of  $f_1$  with  $f_2$ . An implementation  $f$  is correct w.r.t. a specification  $f^{spec}$  if  $f$  and  $f^{spec}$  are equivalent when applied to expressions satisfying  $f^{pre}$ .*

The correctness of an implementation w.r.t. a specification imposes an equality of two sets of result values. The implementation could produce a value more or less times than the specification in the sense that the same expression has distinct derivations to the same value. Furthermore, we do not require that equivalent operations must have the same failure or non-termination behavior. For instance, the evaluation of one operation could diverge where an equivalent operation might terminate with a failure or some exception, e.g., due to black hole detection [19], on the same input.

Intuitively, two operations are equivalent if it is impossible to detect any difference between them in any context. If operations do not produce values or produce some values as well as failures, the consideration of a context is important. For instance, consider the following alternative implementation of sorting a list based on an operation `idSorted` that is the identity on sorted lists:

```
sort' xs = idSorted (perm xs)
  where idSorted [] = []
        idSorted [x] = [x]
        idSorted (x:y:ys) | x<=y = x : idSorted (y:ys)
```

Although this implementation only returns values that are sorted lists, it is not a correct w.r.t. the specification `sort'spec`. For instance, consider the context operation `head` that returns the first element of the list. Then there is a derivation

$$\begin{aligned} \text{head (sort' [2,3,1])} &\xrightarrow{*} \text{head (idSorted [2,3,1])} \\ &\xrightarrow{*} \text{head (2 : idSorted [3,1])} \\ &\xrightarrow{*} 2 \end{aligned}$$

whereas “`head (sort'spec [2,3,1])`” cannot be reduced to 2. The implementation `sort'` is incorrect with respect to the specification of `sort`: if we want to compute the minimum of a list by sorting the list and taking the first element, the previous derivation shows that we obtain an unintended result.

Specifications can be used to verify programs. This is a complex task that could be supported by proof systems which is an interesting topic for future work. In this paper we exploit the property that specifications are executable so that we can use them to detect an incorrect execution of the implementation. For this purpose, we use a specification as a contract for an implementation. Thus, if we detect a violation at run-time, we can deduce that the implementation is not correct. This demands for a postcondition that is generated from a specification. In a naive approach, we could try to define such a postcondition as

$$f^{post} \ x \ y = y \in f_S^{spec} \ x$$

i.e., the postcondition checks whether the actual result is in the set of all the results according to the specification. Unfortunately, this simple definition does not work as intended due to the following problems:

1. For partially defined operations, this postcondition could be violated even though the implementation is correct. For instance, consider

```
head'spec (x:_) = x
head (x:_) = x
```

Obviously, `head` is correct w.r.t. `head'spec`. However, the set `head'specS []` is empty so that the condition “`head [] ∈ head'specS []`” could reduce to `False`. Therefore, this condition should be checked only if the actual result is a value and not a failure. However, the implementation of “`∈`” may not require the evaluation of its left argument when its

right argument is empty.

2. The membership test requires the decision that two entities are equal. Since in functional logic languages, this test is evaluated by strict equality on (finite) values, the test will never be successful for operations delivering infinite structures.

The first problem can be handled by the addition of an equality test “ $y=y$ ”. Since the equality “ $=$ ” compares *values*, the test is successful only if  $y$  is a value. This has the consequence that postconditions are not checked for failure cases. From a conceptual point of view, it would be better to exclude such cases by appropriate preconditions. Since the test for such an exclusion is undecidable in general, we add this sufficient condition to the postcondition.

The second problem can be handled in part by avoiding the comparison of complete results, and comparing only some computed parts, instead. For this purpose, we define a postcondition that is parametric w.r.t. some observation operation  $g$ .

**Definition 4.** Let  $f^{spec}$  be a specification of type  $\tau \rightarrow \tau'$  and  $g$  an operation of type  $\tau' \rightarrow \tau''$ . The postcondition  $f_g^{post}$  generated from  $f^{spec}$  w.r.t.  $g$  is defined by

$$f_g^{post} \ x \ y = \text{let } z = g \ y \\ \quad g' \ a = g \ (f^{spec} \ a) \\ \text{in } z==z \ \&\& \ z \in g'_S \ x$$

If we use  $g = \text{id}$  (the identity function), the generated postcondition checks whether a result  $y$  is a value and it is contained in the set of all the results according to the specification. For instance, consider

$$f' \ \text{spec} = 0 \ ? \ 1 \\ f = 1 \ ? \ 0$$

The generated postcondition  $f_{\text{id}}^{post}$  requires that each value of the implementation  $f$  is contained in the set  $\{0, 1\}$ .

If we know that a specification is deterministic, i.e., it yields at most one result for a given input, then we can provide a simpler postcondition without using an observation operation and set functions:

$$f^{post} \ x \ y = y == f^{spec} \ x$$

Although this formulation does not support the detection of violations for failed computations, it might report violations when computing infinite structures, if the equality is checked in a demand-driven manner. Hence, this optimized formulation is supported by our tool.

The use of a postcondition generated from a specification to check an implementation is justified by the following propositions. The first proposition shows that equivalent operations have the same violations.

**Proposition 1.** Let  $f^{post}$  be a postcondition for  $f$ . If  $f$  is equivalent to  $f'$  and there is a violation of the postcondition  $f^{post}$  for  $f$  at  $e$ , then there is also a violation of the postcondition  $f^{post}$  for  $f'$  at  $e$ .

*Proof.* Suppose that there is a violation of  $f^{post}$  for  $f$  at  $e$ . By definition of a postcondition violation,

$$\text{let } x = e \text{ in not } (f^{pre} \ x) \ || \ f^{post} \ x \ (f \ x)$$

is reducible to **False**. We have to show that

$$\text{let } x = e \text{ in not } (f^{pre} \ x) \ || \ f^{post} \ x \ (f' \ x)$$



is also reducible to **False**. The second expression is equal to the first one except for the replacement of  $f$  with  $f'$ . Since  $f$  and  $f'$  are assumed to be equivalent, by definition, the two expressions produce the same value(s).  $\square$

The next proposition shows that any postcondition  $f^{post}$  derived from a specification  $f^{spec}$  cannot cause any violation when  $f^{post}$  is used to check an execution of  $f^{spec}$ .

**Proposition 2.** *If  $f_g^{post}$  is the postcondition generated from  $f^{spec}$  w.r.t. some operation  $g$ , then there is no  $e$  such that there is a violation of the postcondition  $f_g^{post}$  for  $f^{spec}$  at  $e$ .*

*Proof.* Assume the contrary, i.e., there is some  $e$  such that there is a violation of the postcondition  $f_g^{post}$  for  $f^{spec}$  at  $e$ . By definition of violation,

$$\text{let } x = e \text{ in not } (f^{pre} \ x) \ || \ f_g^{post} \ x \ (f^{spec} \ x)$$

is reducible to **False**. Hence,  $\text{let } x = e \text{ in } f_g^{post} \ x \ (f^{spec} \ x) \xrightarrow{*} \text{False}$ . Due to the reduction relation for the call-time choice semantics, there is a derivation  $e \xrightarrow{*} t$ , where  $t$  is a partial value, such that  $f_g^{post} \ t \ (f^{spec} \ t) \xrightarrow{*} \text{False}$ . By definition of  $f_g^{post}$ , there must be some partial value  $t'$  with  $f^{spec} \ t \xrightarrow{*} t'$  and

$$\begin{aligned} \text{let } z = g \ t' \\ g' \ a = g \ (f^{spec} \ a) \\ \text{in } z==z \ \&\& \ z \in g'_S \ t \end{aligned}$$

is reducible to **False**. Thus, there is some partial value  $t''$  with  $g \ t' \xrightarrow{*} t''$  and  $t''==t'' \ \&\& \ t'' \in g'_S \ t$  is reducible to **False**. Since  $t''==t''$  cannot be reducible to **False**, it must be reducible to **True** and  $t'' \in g'_S \ t \xrightarrow{*} \text{False}$ . Since  $g' \ t \rightarrow g \ (f^{spec} \ t) \xrightarrow{*} g \ t' \xrightarrow{*} t''$  where  $t''$  is a value (since the strict equality relation “==” is only true on values),  $t''$  must be an element of  $g'_S \ t$  by the definition of set functions [2] so that  $t'' \in g'_S \ t$  cannot be reducible to **False**.  $\square$

As a consequence, we can use the postcondition generated from  $f^{spec}$  to detect an incorrect implementation:

**Corollary 1.** *Let  $f_g^{post}$  be the postcondition generated from  $f^{spec}$  w.r.t. some operation  $g$ . If there is a violation of  $f_g^{post}$  for  $f$  at  $e$ , then  $f$  is not correct w.r.t.  $f^{spec}$ .*

*Proof.* Assume the contrary, i.e., there is a violation of the postcondition  $f_g^{post}$  for  $f$  at  $e$  but  $f$  is correct w.r.t.  $f^{spec}$ . Since  $f$  is equivalent to  $f^{spec}$ ,  $f_g^{post}$  is also violated for  $f^{spec}$  at  $e$  by Proposition 1. This is a contradiction to Proposition 2.  $\square$

Similarly to testing, the correctness of an implementation cannot be determined by individual executions of a program. Nevertheless, we can infer from a satisfied postcondition which is generated from  $f^{spec}$  and an observation operation  $g$  that the observed part of the computation is correct w.r.t. the specification:

**Proposition 3.** *Let  $f_g^{post}$  be the postcondition generated from  $f^{spec}$  w.r.t. some operation  $g$  and  $e$  an expression such that  $f_g^{post} \ e \ (f \ e) \xrightarrow{*} \text{True}$ . Then there is a value  $s$  with  $g \ (f \ e) \xrightarrow{*} s$  and  $g \ (f^{spec} \ e) \xrightarrow{*} s$ .*

*Proof.* Since  $f_g^{post}$  is defined by

$$\begin{aligned} f_g^{post} \ x \ y = \text{let } z = g \ y \\ g' \ a = g \ (f^{spec} \ a) \\ \text{in } z==z \ \&\& \ z \in g'_S \ x \end{aligned}$$

$f_g^{post} e (f e) \overset{*}{\rightarrow} \text{True}$  implies that there are partial values  $t$  and  $t'$  with  $e \overset{*}{\rightarrow} t$ ,  $(f e) \overset{*}{\rightarrow} t'$ , and

```
let z = g t'
    g' a = g (fspec a)
in z==z && z ∈ g'_S t
```

is reducible to  $\text{True}$ . Hence, there is a partial value  $s$  with  $g t' \overset{*}{\rightarrow} s$  and  $s==s \ \&\& \ s \in g'_S t \overset{*}{\rightarrow} \text{True}$ . Therefore,  $s$  is a value and, by definition of set functions,  $g' t \overset{*}{\rightarrow} s$ . By definition of  $g'$ ,  $g (f^{spec} t) \overset{*}{\rightarrow} s$ . Altogether, we have  $g (f^{spec} e) \overset{*}{\rightarrow} g (f^{spec} t) \overset{*}{\rightarrow} s$  and  $g (f e) \overset{*}{\rightarrow} g t' \overset{*}{\rightarrow} s$ .  $\square$

Now we are ready to put this theoretical framework into a tool to support the development of reliable declarative programs.

## 4 Tool Support

In this section we discuss a tool, `DSDCurry`<sup>3</sup>, based on the ideas described in the previous sections. Basically, the tool transforms a Curry module  $M$  containing specifications, pre- and/or postconditions for some operations into a new Curry module  $MC$  providing the same interface, but where some operations are checked against the provided specifications and/or contracts. Providing specifications and/or contracts is not mandatory. However, when they are provided, they are used as follows in the transformed module:

- If there is a specification  $f^{spec}$ , then a corresponding postcondition is generated according to Definition 4 (if an observation operation is not provided by the programmer, the identity function `id` is used for  $g$ ). If there is also a user-defined postcondition, it is combined with the generated postcondition by conjunction.
- If there is only a specification  $f^{spec}$  but no implementation<sup>4</sup> of operation  $f$  is provided, then an implementation for  $f$  is generated by the rule  $f = f^{spec}$ .
- If there is neither a specification nor an implementation but a postcondition  $f^{post}$  for some operation  $f$ , the postcondition is used as a (weak) specification for  $f$ , i.e., an initial implementation is generated for  $f$  by the following definition:

$$f \ x \mid f^{post} \ x \ y = y \ \text{where } y \ \text{free}$$

Since postconditions typically accept more results than specifications, one could also add the restriction that only one solution of the postcondition should be taken as an implementation. This issue is discussed below in more detail.

- If there is a contract  $f^{pre}/f^{post}$  for some operation  $f$ , the implementation of  $f$  is replaced by

$$\begin{aligned} f \ x \mid & \text{checkPre "f" } (f^{pre} \ x) \ \&\gt; \ \text{checkPost "f" } (f^{post} \ x \ y) \\ & = y \\ & \text{where } y = f' \ x \\ & \quad f' \ \dots \end{aligned}$$

<sup>3</sup>The tool together with more examples is available at:  
<http://www.informatik.uni-kiel.de/~pakcs/dsdcurry/>.

<sup>4</sup>An operation defined by the rule “ $f = \text{unknown}$ ” is considered as undefined. Such a vacuous definition might be necessary if  $f$  is referenced in the definition of other operations in  $M$ .

where “ $f'$  ...” contains the original definition of  $f$  with every occurrence of  $f$  replaced by  $f'$ . Thus, the original interface of any function is preserved by DSDCurry. The auxiliary operations `checkPre` and `checkPost` just check whether their second argument is true and produce an error message if this is not the case. For instance, `checkPre` is defined by:<sup>5</sup>

```
checkPre fname checkresult =
  if checkresult then success else
    error ("Precondition of operation '++fname++' violated!")
```

The postcondition checker, `checkPost`, is similarly defined. Note that the pre- and postcondition checkers are constraints rather than Boolean operations. This is useful for lazy assertion checking [16] since constraints can be concurrently evaluated.

We demonstrate the development of a simple program using DSDCurry. Consider the specification `sort'spec` and the contract `sort'pre/sort'post` for sorting a list as shown in Section 3. According to Definition 4, the specification and postcondition are combined into a new postcondition of the form

```
sort'post x y = sort'post'org x y && y == y && y ∈ sort'specS x
  where sort'post'org xs ys = length xs == length ys
```

where `sort'post'org` is the original, user-supplied postcondition. If we do not provide any implementation of the operation `sort`, an implementation is generated from its specification where contract checking is added:

```
sort x | checkPre "sort" (sort'pre x)
      &> checkPost "sort" (sort'post x y)
      = y
  where y = sort'spec x
```

In principle, postcondition checking should be superfluous for specifications since any user-defined postcondition should be a logical consequence of the specification. Nevertheless, it is included since this entailment is not checked at compile time by our tool.

This prototypical implementation is not efficient because it does not exploit any knowledge about sorting algorithms developed over decades of research in computer science. We improve the efficiency of this implementation by adopting one of these algorithms known as *straight selection sort*. Informally, a list is sorted by selecting its smallest element, sorting the remaining elements, and placing the smallest element in front of the sorted remaining elements. If we know how to select the smallest element of a list, the implementation of this sort method is straightforward by a case distinction on the form of the input list:

```
sort [] = []
sort (x:xs) = min : sort rest
  where (min,rest) = minRest (x:xs)
```

Here, we assume that the essential operation of selecting the smallest element is encoded by the operation `minRest` that, for a non-empty input list, returns both the smallest element and the remaining elements. Since finding the smallest element is a non-trivial task, we define a contract for `minRest`:

```
minRest'pre = not . null
minRest'post xs (min,rest) = (min:rest) ∈ permS xs && all (>= min) xs
```

---

<sup>5</sup>The actual implementation provides more information, e.g., about the concrete arguments of the pre- and postcondition.

The precondition requires that `minRest` is only applied to non-empty lists. Since there might be different methods to select a minimal element and return the remaining ones, we do not put any requirements on the order of the remaining elements in the postcondition, hence `(min:rest)` is some permutation of the input list. This is also the reason why it would be too restrictive to provide a specification of `minRest`. However, we can use the postcondition as an initial implementation.<sup>6</sup> This has the undesirable effect that such an implementation of `minRest` produces many values, i.e., the minimal element together with all permutations of the remaining elements. We can either restrict this implementation to return only one value and ignore the others (for this reason, `DSDCurry` has an option to enforce this behavior), or provide a more informed implementation of the operation `minRest` as follows.

A direct implementation of `minRest` could be obtained via two auxiliary operations, `min` and `del`, that return the minimal element of a list and delete an occurrence of an element in a list, respectively:

```
minRest (x:xs) = let m = min x xs
                 in (m, del m (x:xs))

where
  min x [] = x
  min x (y:ys) = if x<=y then min x ys else min y ys
  del x (y:ys) = if x==y then ys else y : del x ys
```

If we transform this augmented program with `DSDCurry`, it works as intended without any contract violation. We observe that our implementation of `minRest`, in the worst case, performs two traversals of the input list, whereas it is possible to compute the minimal element and the remaining elements with a single traversal. To improve the performance, we re-code `minRest` as

```
minRest (x:xs) = mr x [] xs
where mr m r [] = (m,r)
      mr m r (y:ys) = if m<=y then mr m (y:r) ys else mr y (m:r) ys
```

This implementation is more efficient, but also more complicated and its correctness is not as apparent as before. Thus, we apply again our transformation tool to integrate the contract into this implementation and execute the program to increase our confidence in its correctness. Now that we are satisfied with the implementation, we could attempt a formal correctness proof of this implementation. However, this is outside the scope of this paper.

As a further example, consider a program to compute the infinite list, `fibs`, of all the Fibonacci numbers. The specification maps the operation, `fib`, to compute the  $n$ -th Fibonacci number defined by the immediate recursive definition, onto the list of all naturals:

```
fibs'spec = map fib [0..]
where fib n | n == 0 = 0
           | n == 1 = 1
           | otherwise = fib (n-1) + fib (n-2)
```

The application of `DSDCurry` immediately gives us a correct implementation of `fibs` from this specification, e.g., the expression “take 10 fibs” reduces to `[0,1,1,2,3,5,8,13,21,34]`. Since each number in the list is computed by applying operation `fib`, the implementation is quite inefficient due to the exponential complexity of `fib`. Hence, we improve the implementation and construct the list (in linear time) by creating the next element by adding the two previous ones:

```
fibs = fiblist 0 1 where fiblist x y = x : fiblist (x+y) y
```

---

<sup>6</sup>In the current implementation, this requires replacing the Boolean operation “ $\in$ ” into a corresponding constraint since the equality test implicitly performed by “ $\in$ ” suspends on free variables [17].

When we execute “take 10 fibs” again after transforming our program with DSDCurry, a violation is reported for the third element, 2, of the result list. We made a typical error in iterative definitions by swapping some arguments. If we correct the program to

```
fibs = fiblist 0 1 where fiblist x y = x : fiblist y (x+y)
```

and transform and run it again, no more violations are reported.

Contract checking in the presence of infinite structures requires the lazy evaluation of assertions. Thus, our simple implementation where the contract is completely checked in the condition of an operation would lead to an infinite loop in the transformed fibs operation. In general, the eager or strict checking of assertions might influence the execution behavior of a program. To avoid this problem, Chitil et al. [7] proposed *lazy assertions*. Lazy assertions do not evaluate their arguments, but check them when they become evaluated by the application program. Thus, as long as every assertion is satisfied, program executions with or without lazy assertion checking deliver the same results.

On the other hand, lazy assertion checking might not detect contract violations if the assertion arguments are not sufficiently evaluated by the main program. Thus, it is debatable whether full assertion checking should be avoided in order to preserve the behavior of programs [9, 16]. Lazy assertions do not modify the behavior, but a lazily computed result cannot be trusted as long as some assertion has not been checked. As a compromise between these conflicting goals, *enforceable assertions* are proposed in [16]. These assertions behave like lazy assertions, but they can also be checked upon an explicit request of the programmer, e.g., at the end of a program run or at key intermediate execution points.

Making the appropriate choice might be dependent on the application or require some sophisticated program analysis. Therefore, DSDCurry supports strict, lazy, and enforceable assertions by some transformation options so that it can be easily adapted to future insights.

## 5 Static Contract Checking

In the previous sections we have developed a framework to improve the reliable execution of declarative programs. If a specification for some operation is provided, this specification is used to check the execution of this operation during each program run. Obviously, the improved reliability comes with a price. If the execution of the specification needs a considerable amount of time, contract checking slows down the overall execution. This could be partially avoided by exploiting multi-core computers and performing contract checking in parallel to the main program, e.g., as proposed in [10]. However, this does not help if the computational complexity of contract checking is high. Moreover, assertion checking ensures only the reliable execution of a particular execution but implies no guarantees for all inputs. Therefore, it would be much better to check contracts statically. For instance, if one can prove that the implementation of an operation is equivalent to its specification, there is no need to check the postcondition generated from the specification at run time. Since there is no fully automatic method to prove such equivalences, we propose in this section sufficient criteria for proving such equivalences and apply them to some examples.

The definition of equivalence as stated in Def. 3 covers the intuition that equivalent operations can be interchanged at any place in an expression without changing its value. Proving such a general form of equivalence could be difficult. Therefore, we define another form of equivalence that is based on an operation to observe the computed results of the corresponding operations.

**Definition 5** (Observable equivalence). *Let  $f_1, f_2$  be operations of type  $\tau \rightarrow \tau'$ .  $f_1$  is observable equivalent to  $f_2$  iff, for all operations  $g$  of type  $\tau' \rightarrow \tau''$ , all expressions  $e$  and values  $v$ ,  $g (f_1 e) \xrightarrow{*} v$  iff  $g (f_2 e) \xrightarrow{*} v$ .*

One can expect that proving observable equivalence is easier than equivalence since one has to consider only a single (but arbitrary) observation operation instead of all expressions. Fortunately, the next theorem shows that proving observable equivalence is sufficient in general.

**Theorem 1.** *Let  $f_1, f_2$  be operations of type  $\tau \rightarrow \tau'$ .  $f_1$  and  $f_2$  are equivalent iff they are observable equivalent.*

*Proof.* It is trivial that equivalence implies observable equivalence. Hence, we assume that  $f_1$  and  $f_2$  are observable equivalent, i.e., for all operations  $g$  of type  $\tau' \rightarrow \tau''$ , all expressions  $e$  and values  $v$ ,  $g (f_1 e) \xrightarrow{*} v$  iff  $g (f_2 e) \xrightarrow{*} v$ . We show by induction on the number  $n$  of occurrences of the symbol  $f_1$  the following claim:

If  $E_1$  is an expression with  $n$  occurrences of  $f_1$ ,  $E_2$  is obtained from  $E_1$  by replacing any occurrence of  $f_1$  with  $f_2$ , and  $v$  is a value, then  $E_1 \xrightarrow{*} v$  iff  $E_2 \xrightarrow{*} v$ .

Base case ( $n = 0$ ): Since  $E_1$  contains no occurrence of  $f_1$ ,  $E_2 = E_1$  and the claim is trivially satisfied.

Inductive case ( $n > 0$ ): Assume the claim holds for  $n - 1$  and  $E_1$  contains  $n$  occurrences of  $f_1$  and  $E_1 \xrightarrow{*} v$  for some value  $v$ . We have to show that  $E_2 \xrightarrow{*} v$  (the opposite direction is symmetric) where  $E_2$  is obtained from  $E_1$  by replacing any occurrence of  $f_1$  with  $f_2$ . Let  $p$  be a position in  $E_1$  with  $E_1|_p = f_1 e$  and  $e$  does not contain any occurrence of  $f_1$ . Since  $E_1 \xrightarrow{*} v$ , by definition of  $\xrightarrow{*}$ , there is a partial value  $t_1$  with  $f_1 e \xrightarrow{*} t_1$  and  $E_1[t_1]_p \xrightarrow{*} v$ . We define a new operation  $g$  by

$$g x = E_1[x]_p$$

where  $x$  is a variable that does not occur in  $E_1$ . Hence  $g (f_1 e) \xrightarrow{*} g t_1 \rightarrow E_1[t_1]_p \xrightarrow{*} v$ . Our assumption implies  $g (f_2 e) \xrightarrow{*} v$ . By definition of  $\xrightarrow{*}$ , there is a partial value  $t_2$  with  $g (f_2 e) \xrightarrow{*} g t_2 \rightarrow E_1[t_2]_p \xrightarrow{*} v$ . Since  $E_1[t_2]_p$  contains  $n - 1$  occurrences of  $f_1$ , the induction hypothesis implies that  $E_2[t_2]_p \xrightarrow{*} v$ . Therefore,  $E_2 = E_2[f_2 e]_p \xrightarrow{*} E_2[t_2]_p \xrightarrow{*} v$ .  $\square$

A proof that two operations are observable equivalent could still be difficult since one has to take all possible observation operations into account. However, the next result shows that it is sufficient to verify that two operations yield always the same partial values on identical inputs.

**Theorem 2.** *Let  $f_1, f_2$  be operations of type  $\tau \rightarrow \tau'$ . If for all expressions  $e$  and partial values  $t$ ,  $f_1 e \xrightarrow{*} t$  iff  $f_2 e \xrightarrow{*} t$ , then  $f_1$  and  $f_2$  are equivalent.*

*Proof.* By Theorem 1 it is sufficient to show that observable equivalence of  $f_1$  and  $f_2$ . Hence, let  $g$  be an operation of type  $\tau' \rightarrow \tau''$ ,  $e$  an expression and  $v$  a value with  $g (f_1 e) \xrightarrow{*} v$ . We have to show that  $g (f_2 e) \xrightarrow{*} v$  (the other direction is symmetric). By definition of  $\xrightarrow{*}$ , there is some partial value  $t$  with  $f_1 e \xrightarrow{*} t$  and  $g t \xrightarrow{*} v$ . By the assumption of the theorem,  $f_2 e \xrightarrow{*} t$ . Hence,  $g (f_2 e) \xrightarrow{*} g t \xrightarrow{*} v$ .  $\square$

Note that the consideration of all *partial* result values is essential to establish equivalence. For instance, consider the operation `sort'` and the specification `sort'spec` defined in Section 3. Although `sort'` and `sort'spec` compute the same *values*, we have that `sort' [2,3,1]  $\xrightarrow{*}$  2:⊥` but

`sort'spec [2,3,1]` cannot be derived to  $2:\perp$ . Actually, we have seen that `sort'` and `sort'spec` are not equivalent.

The next corollary is useful to avoid the consideration of all argument expressions in equivalence proofs.

**Corollary 2.** *Let  $f_1, f_2$  be operations of type  $\tau \rightarrow \tau'$ . If, for all partial values  $t$  and  $t'$ ,  $f_1 t \overset{*}{\rightarrow} t'$  iff  $f_2 t \overset{*}{\rightarrow} t'$ , then  $f_1$  and  $f_2$  are equivalent.*

*Proof.* Assume that  $f_1 t \overset{*}{\rightarrow} t'$  iff  $f_2 t \overset{*}{\rightarrow} t'$  holds for all partial values  $t$  and  $t'$ . Consider an expression  $e$  and a partial value  $t$  such that  $f_1 e \overset{*}{\rightarrow} t_1$ . By definition of  $\overset{*}{\rightarrow}$ , there is a partial value  $t_0$  with  $e \overset{*}{\rightarrow} t_0$  and  $f_1 t_0 \overset{*}{\rightarrow} t_1$ . Our assumption implies  $f_2 t_0 \overset{*}{\rightarrow} t_1$ . Hence  $f_2 e \overset{*}{\rightarrow} f_2 t_0 \overset{*}{\rightarrow} t_1$ . Since the other direction is symmetric, Theorem 2 implies the equivalence of  $f_1$  and  $f_2$ .  $\square$

As a concrete example, we want to show the correctness of the implementation of `sort` given in Section 4 (by computing the minimum of a list and the remaining elements in one pass), i.e., we proof that the implemented operation is equivalent to the specification  $\text{sort}^{\text{spec}}$  defined in Section 3. As a first step, we show that the implementation of `minRest` satisfies its postcondition. For this purpose, we proof the following lemma about the auxiliary operation `mr`.

**Lemma 1.** *Let  $x$  be an integer and  $xs$  and  $l$  lists of integers such that all  $(\geq x) xs \overset{*}{\rightarrow} \text{True}$ . Then  $\text{mr } x \ xs \ l \overset{*}{\rightarrow} (m, r)$  with all  $(\geq m) (x : xs++l) \overset{*}{\rightarrow} \text{True}$  and  $\text{perm } (x : xs++l) \overset{*}{\rightarrow} m : r$ .*

*Proof.* We prove the lemma by induction on the length of the list  $l$ .

Base case:  $l = []$ : Since  $\text{mr } x \ xs \ [] \overset{*}{\rightarrow} (x, xs)$ , all  $(\geq x) (x : xs++[]) \overset{*}{\rightarrow} \text{True}$  follows from the assumption all  $(\geq x) xs \overset{*}{\rightarrow} \text{True}$  and  $\text{perm } (x : xs++[]) \overset{*}{\rightarrow} x : xs$  obviously holds.

Inductive case:  $l = y : ys$ : We assume that all  $(\geq x) xs \overset{*}{\rightarrow} \text{True}$  and the claim holds for the list  $ys$ . We consider the two cases occurring in the definition of `mr`.

If  $x \leq y$ , then  $\text{mr } x \ xs \ (y : ys) \overset{*}{\rightarrow} \text{mr } x \ (y : xs) \ ys$ . Since all  $(\geq x) (y : xs) \overset{*}{\rightarrow} \text{True}$ , we can apply the induction hypothesis to the last expression and obtain  $\text{mr } x \ (y : xs) \ ys \overset{*}{\rightarrow} (m, r)$  with all  $(\geq m) (x : y : xs++ys) \overset{*}{\rightarrow} \text{True}$  and  $\text{perm } (x : y : xs++ys) \overset{*}{\rightarrow} m : r$ . Hence  $m : r$  is also a permutation of  $(x : xs++y : ys)$  and all  $(\geq m) (x : xs++y : ys) \overset{*}{\rightarrow} \text{True}$ .

If  $x > y$ , then  $\text{mr } x \ xs \ (y : ys) \overset{*}{\rightarrow} \text{mr } y \ (x : xs) \ ys$ . Since all  $(\geq y) (x : xs) \overset{*}{\rightarrow} \text{True}$ , we can apply the induction hypothesis to the last expression and obtain  $\text{mr } y \ (x : xs) \ ys \overset{*}{\rightarrow} (m, r)$  with all  $(\geq m) (y : x : xs++ys) \overset{*}{\rightarrow} \text{True}$  and  $\text{perm } (y : x : xs++ys) \overset{*}{\rightarrow} m : r$ . Hence  $m : r$  is also a permutation of  $(x : xs++y : ys)$  and all  $(\geq m) (x : xs++y : ys) \overset{*}{\rightarrow} \text{True}$ .  $\square$

Now we can easily prove that the implementation of `minRest` satisfies the postcondition  $\text{minRest}^{\text{post}}$  defined in Section 3.

**Lemma 2.** *Let  $xs = [x_1, \dots, x_n]$  be a list of integers,  $n > 0$ , and  $\text{minRest } xs \overset{*}{\rightarrow} (m, r)$ . Then  $\text{minRest}^{\text{post}} xs \ (m, r) \overset{*}{\rightarrow} \text{True}$ .*

*Proof.* If the assumptions of the claim hold, then, by definition of `minRest`,

$$\text{minRest } xs \rightarrow \text{mr } x_1 \ [] \ [x_2, \dots, x_n]$$

Since all  $(\geq x_1) [] \overset{*}{\rightarrow} \text{True}$ , Lemma 1 implies that  $\text{mr } x_1 \ [] \ [x_2, \dots, x_n] \overset{*}{\rightarrow} (m, r)$  such that all  $(\geq m) (x_1 : []++[x_2, \dots, x_n]) \overset{*}{\rightarrow} \text{True}$  and  $\text{perm } (x_1 : []++[x_2, \dots, x_n]) \overset{*}{\rightarrow} m : r$ . Hence all  $(\geq m) xs \overset{*}{\rightarrow} \text{True}$  and  $\text{perm } xs \overset{*}{\rightarrow} m : r$  so that  $\text{minRest}^{\text{post}} xs \ (m, r) \overset{*}{\rightarrow} \text{True}$ .  $\square$

Now we can prove the soundness of *sort*, i.e., the fact that each value computed by *sort* can also be derived by *sort<sup>spec</sup>*.

**Lemma 3.** *Let  $l = [x_1, \dots, x_n]$  be a list of integers,  $n \geq 0$ , and  $\text{sort } l \xrightarrow{*} l'$  for some value  $l'$ . Then  $\text{sort}^{\text{spec}} l \xrightarrow{*} l'$ .*

*Proof.* We prove the lemma by induction on the length of the list  $l$ .

Base case:  $l = []$ : Since  $\text{sort } [] \rightarrow []$  and  $\text{sort}^{\text{spec}} [] \rightarrow []$ , the claim obviously holds.

Inductive case:  $l = x : xs$ : By definition of *sort*,  $\text{sort } l \xrightarrow{*} m : ys$  with  $\text{minRest } l \xrightarrow{*} (m, r)$  and  $\text{sort } r \xrightarrow{*} ys$ . Lemma 2 implies (1)  $\text{perm } (x : xs) \xrightarrow{*} m : r$  and (2)  $\text{all } (\geq m) (x : xs) \xrightarrow{*} \text{True}$ . Since  $(m : r)$  is a permutation of  $(x : xs)$ ,  $m$  is an element of  $l$  and  $r$  is a list shorter than  $l$ . Thus, we can apply the induction hypothesis to  $r$  and we obtain  $\text{sort}^{\text{spec}} r \xrightarrow{*} ys$ . Hence, by definition of *sort<sup>spec</sup>*, (3)  $\text{perm } r \xrightarrow{*} ys$  and (4)  $\text{sorted } ys \xrightarrow{*} \text{success}$ . Properties (1) and (3) imply that  $(m : ys)$  is a permutation of  $(x : xs)$ . Together with (2),  $m$  is smaller or equal to all elements of  $(m : ys)$  so that, due to (4),  $\text{sorted } (m : ys) \xrightarrow{*} \text{success}$ . Therefore,  $\text{sort}^{\text{spec}} (x : xs) \xrightarrow{*} m : ys$ .  $\square$

Next we prove the completeness of *sort*, i.e., the fact that each value derived by *sort<sup>spec</sup>* can also be computed by *sort*.

**Lemma 4.** *Let  $l = [x_1, \dots, x_n]$  be a list of integers,  $n \geq 0$ , and  $\text{sort}^{\text{spec}} l \xrightarrow{*} l'$  for some value  $l'$ . Then  $\text{sort } l \xrightarrow{*} l'$ .*

*Proof.* Assume that  $\text{sort}^{\text{spec}} l \xrightarrow{*} l'$  for some value  $l'$ . Since it is easy to check that both *sort* and *minRest* are totally defined and terminating, there exists a value  $l''$  with  $\text{sort } l \xrightarrow{*} l''$ . By Lemma 3,  $\text{sort}^{\text{spec}} l \xrightarrow{*} l''$ . By definition of *sort<sup>spec</sup>*, both  $l'$  and  $l''$  are permutations of  $l$  that are sorted. To prove the claim, we show  $l' = l''$  by induction on the length  $n$  of  $l$ .

Base case:  $n = 0$ : Then  $l' = [] = l''$  by definition of *sort<sup>spec</sup>*.

Inductive case:  $n > 0$ : Since both  $l'$  and  $l''$  are permutations of  $l$ , they must be non-empty so that they have the form  $l' = x : xs$  and  $l'' = y : ys$ . Since  $l'$  and  $l''$  are sorted,  $x$  and  $y$  are smallest elements of  $l$ . Thus  $x = y$  and  $xs$  and  $ys$  are sorted and permutation of  $l$  without element  $x$ . Hence, we can apply the induction hypothesis.  $\square$

Now we can finally prove that our sort implementation is equivalent to its specification.

**Corollary 3.**  *$\text{sort}$  and  $\text{sort}^{\text{spec}}$  are equivalent.*

*Proof.* First, consider an argument list  $l$  of integers and a partial value  $t$  with  $\text{sort } l \xrightarrow{*} t$ . Since both *sort* and *minRest* are totally defined, orthogonal and terminating, there exists a unique value  $l'$  with  $\text{sort } l \xrightarrow{*} l'$ . Hence,  $t$  must be smaller than  $l'$  w.r.t. information ordering and  $l' \xrightarrow{*} t$  (by definition of  $\xrightarrow{*}$ ). By Lemma 3,  $\text{sort}^{\text{spec}} l \xrightarrow{*} l'$ . Hence,  $\text{sort}^{\text{spec}} l \xrightarrow{*} l' \xrightarrow{*} t$ .

If  $\text{sort}^{\text{spec}} l \xrightarrow{*} t$  for some partial value  $t$ , either  $t = \perp$  (then  $\text{sort } l \rightarrow t$  trivially holds) or  $\text{sort}^{\text{spec}} l \xrightarrow{*} l' \xrightarrow{*} t$  for some value  $l'$  (since a constructor occurs at the root only if the condition of *sort<sup>spec</sup>* has been proved which implies that the argument is a value). By Lemma 4,  $\text{sort } l \xrightarrow{*} l'$  so that  $\text{sort } l \xrightarrow{*} t$ .

Now consider a partial value  $t$  that is not a value as an argument, i.e.,  $t$  contains occurrences of  $\perp$ . Since both *sort<sup>spec</sup>* and *sort* consider their complete list argument,  $\text{sort}^{\text{spec}} t \xrightarrow{*} \perp$  and  $\text{sort } t \xrightarrow{*} \perp$  are the only computable results.

Altogether, we have shown that  $\text{sort}^{\text{spec}} t \xrightarrow{*} t'$  iff  $\text{sort } t \xrightarrow{*} t'$  for all partial values  $t$  and  $t'$ . Therefore, *sort<sup>spec</sup>* and *sort* are equivalent by Corollary 2.  $\square$



## 6 Conclusions and Related Work

We have discussed some notions that are essential for a methodology intended to develop reliable declarative programs. Specifications are executable so that they can be used as initial prototypes as well as contracts for implementations that might later be developed. We have shown some relationships between these notions that are the basis of a transformation tool to support this development. Our tool, DSDCurry, transforms a specification into an initial implementation, if an implementation is not provided, otherwise it transforms the specification into a contract that checks the results computed by the implementation. Furthermore, our tool supports various forms of contract checking, such as eager, lazy, or enforceable assertions.

In principle, our method and tool support can be seen as a proposal to use Curry as a wide-spectrum language. In contrast to a wide-spectrum language like CIP-L [5] that supports the development of correct programs by applying a stepwise transformation process to specifications, our approach is more flexible. It does not guarantee correct implementations, but it allows very efficient implementations. The correctness is only checked at each concrete program execution w.r.t. some observation operation.

The use of contracts or assertions to obtain more reliable programs has been proposed for many programming languages and paradigms. Concepts for assertions in strict languages, like imperative, logic, or strict functional languages, are easier to handle than in non-strict languages. For instance, [25] proposes an assertion language for (constraint) logic programming that is combined in [21] with a static verification framework. [10] considered a strict language with side effects and proposed the evaluation of assertions in parallel to the application program to exploit the power of multi-core computers. In non-strict languages, one has the option between lazy assertions [7], which do not change the meaning of a program (apart from reporting violated assertions) but might not report some violations, and strict assertions which could influence the evaluation order. Degen et al. [9] discussed the different approaches and came to the conclusion that there seems no way to satisfy both objectives, meaning preservation and violation reporting, in a non-strict language.

ESC/Haskell [26] is an approach to add pre- and postconditions to Haskell programs that are checked at compile time by sophisticated program transformations. Similarly to our approach, pre- and postconditions are arbitrary Boolean operations implemented in the source language. These conditions are considered as violated if the evaluation of an operation might fail due to incompletely defined operations (e.g., applying the operation `head` to the empty list). Such an interpretation of pre- and postconditions is too restrictive for functional logic languages where failures are used as a programming technique. Moreover, we distinguish between precise specifications and (weak) postconditions. For instance, [26] considers a sorting algorithm as verified if the output is a sorted list. We consider such a property as a weak postcondition whereas a precise specification should additionally require that the output is a permutation of the input list in order to exclude non-intended implementations.

An obvious challenge for future work is to provide proof support for contracts and specifications. If it can be shown at compile time that a contract is always satisfied by the corresponding implementation, its run-time checking can be omitted. This improves the efficiency of reliable software and reduces the need to test the developed software with large sets of test data [8, 13]. Furthermore, a static proof guarantees the correctness of the implementation for all inputs rather than for particular executions.

## References

- [1] S. Antoy, R. Echahed, and M. Hanus. A needed narrowing strategy. *Journal of the ACM*, 47(4):776–822, 2000.
- [2] S. Antoy and M. Hanus. Set functions for functional logic programming. In *Proceedings of the 11th ACM SIGPLAN International Conference on Principles and Practice of Declarative Programming (PPDP'09)*, pages 73–82. ACM Press, 2009.
- [3] S. Antoy and M. Hanus. Functional logic programming. *Communications of the ACM*, 53(4):74–85, 2010.
- [4] S. Antoy and M. Hanus. A transformation tool for functional logic program development. In *Proc. of the 24th Workshop on (Constraint) Logic Programming (WLP 2010)*, pages 23–33. German University of Cairo, 2010.
- [5] F.L. Bauer, M. Broy, R. Gnatz, W. Hesse, B. Krieg-Brückner, H. Partsch, P. Pepper, H. Wössner, and H. Wössner. Towards a wide spectrum language to support program specification and program development. *ACM SIGPLAN Notices*, 13(12):15–24, 1978.
- [6] B. Braßel, M. Hanus, and F. Huch. Encapsulating non-determinism in functional logic computations. *Journal of Functional and Logic Programming*, 2004(6), 2004.
- [7] O. Chitil, D. McNeill, and C. Runciman. Lazy assertions. In *Proceedings of the 15th International Workshop on Implementation of Functional Languages (IFL 2003)*, pages 1–19. Springer LNCS 3145, 2004.
- [8] J. Christiansen and S. Fischer. EasyCheck - test data for free. In *Proc. of the 9th International Symposium on Functional and Logic Programming (FLOPS 2008)*, pages 322–336. Springer LNCS 4989, 2008.
- [9] M. Degen, P. Thiemann, and S. Wehr. True lies: Lazy contracts for lazy languages (faithfulness is better than laziness). In *4. Arbeitstagung Programmiersprachen (ATPS'09)*, pages 370; 2946–2259. Springer LNI 154, 2009.
- [10] C. Dimoulas, R. Pucella, and M. Felleisen. Future contracts. In *Proceedings of the 11th ACM SIGPLAN International Conference on Principles and Practice of Declarative Programming (PPDP'09)*, pages 195–206. ACM Press, 2009.
- [11] R. Echahed and J.-C. Janodet. On constructor-based graph rewriting systems. Research report imag 985-i, IMAG-LSR, CNRS, Grenoble, 1997.
- [12] R. Echahed and J.-C. Janodet. Admissible graph rewriting and narrowing. In *Proc. Joint International Conference and Symposium on Logic Programming (JICSLP'98)*, pages 325–340, 1998.
- [13] S. Fischer and H. Kuchen. Systematic generation of glass-box test cases for functional logic programs. In *Proceedings of the 9th ACM SIGPLAN International Conference on Principles and Practice of Declarative Programming (PPDP'07)*, pages 75–89. ACM Press, 2007.
- [14] J.C. González-Moreno, M.T. Hortalá-González, F.J. López-Fraguas, and M. Rodríguez-Artalejo. An approach to declarative programming based on a rewriting logic. *Journal of Logic Programming*, 40:47–87, 1999.
- [15] M. Hanus. Multi-paradigm declarative languages. In *Proceedings of the International Conference on Logic Programming (ICLP 2007)*, pages 45–75. Springer LNCS 4670, 2007.
- [16] M. Hanus. Lazy and enforceable assertions for functional logic programs. In *Proc. of the 19th International Workshop on Functional and (Constraint) Logic Programming (WFLP 2010)*, pages 84–100. Springer LNCS 6559, 2011.
- [17] M. Hanus (ed.). Curry: An integrated functional logic language (vers. 0.8.2). Available at <http://www.curry-language.org>, 2006.
- [18] H. Hussmann. Nondeterministic algebraic specifications and nonconfluent term rewriting. *Journal of Logic Programming*, 12:237–255, 1992.
- [19] J. Launchbury. A natural semantics for lazy evaluation. In *Proc. 20th ACM Symposium on Principles of Programming Languages (POPL'93)*, pages 144–154. ACM Press, 1993.
- [20] F.J. López-Fraguas, J. Rodríguez-Hortalá, and J. Sánchez-Hernández. A simple rewrite notion for

- call-time choice semantics. In *Proceedings of the 9th ACM SIGPLAN International Conference on Principles and Practice of Declarative Programming (PPDP'07)*, pages 197–208. ACM Press, 2007.
- [21] E. Mera, P. López-García, and M. Hermenegildo. Integrating software testing and run-time checking in an assertion verification framework. In *25th International Conference on Logic Programming (ICLP 2009)*, pages 281–295. Springer LNCS 5649, 2009.
- [22] B. Meyer. *Object-oriented Software Construction*. Prentice Hall, second edition, 1997.
- [23] C. Okasaki. *Purely Functional Data Structures*. Cambridge University Press, 1998.
- [24] S. Peyton Jones, editor. *Haskell 98 Language and Libraries—The Revised Report*. Cambridge University Press, 2003.
- [25] G. Puebla, F. Bueno, and M. Hermenegildo. An assertion language for constraint logic programs. In *Analysis and Visualization Tools for Constraint Programming*, pages 23–62. Springer LNCS 1870, 2000.
- [26] D.N. Xu. Extended static checking for Haskell. In *Proc. of the 36th ACM SIGPLAN Workshop on Haskell (Haskell 2006)*, pages 48–59, 2006.