

Reactive Imperative Programming with Dataflow Constraints

Camil Demetrescu

Dept. of Computer and System Sciences
Sapienza University of Rome
demetres@dis.uniroma1.it

Irene Finocchi

Dept. of Computer Science
Sapienza University of Rome
finocchi@di.uniroma1.it

Andrea Ribichini

Dept. of Computer and System Sciences
Sapienza University of Rome
ribichini@dis.uniroma1.it

Abstract

Dataflow languages provide natural support for specifying constraints between objects in dynamic applications, where programs need to react efficiently to changes of their environment. Researchers have long investigated how to take advantage of dataflow constraints by embedding them into procedural languages. Previous mixed imperative/dataflow systems, however, require syntactic extensions or libraries of *ad hoc* data types for binding the imperative program to the dataflow solver. In this paper we propose a novel approach that smoothly combines the two paradigms without placing undue burden on the programmer.

In our framework, programmers can define ordinary statements of the imperative host language that enforce constraints between objects stored in special memory locations designated as “reactive”. Differently from previous approaches, reactive objects can be of any legal type in the host language, including primitive data types, pointers, arrays, and structures. Statements defining constraints are automatically re-executed every time their input memory locations change, letting a program behave like a spreadsheet where the values of some variables depend upon the values of other variables. The constraint solving mechanism is handled transparently by altering the semantics of elementary operations of the host language for reading and modifying objects. We provide a formal semantics and describe a concrete embodiment of our technique into C/C++, showing how to implement it efficiently in conventional platforms using off-the-shelf compilers. We discuss common coding idioms and relevant applications to reactive scenarios, including incremental computation, observer design pattern, and data structure repair. The performance of our implementation is compared to *ad hoc* problem-specific change propagation algorithms, as well as to language-centric ap-

proaches such as self-adjusting computation and subject/observer communication mechanisms, showing that the proposed approach is efficient in practice.

Categories and Subject Descriptors D.3.3 [Programming Languages]: Language Constructs and Features—Constraints

General Terms Algorithms, design, experimentation, languages.

Keywords Reactive programming, dataflow programming, imperative programming, constraint solving, incremental computation, observer design pattern, data structure repair.

1. Introduction

A one-way, dataflow constraint is an equation of the form $y = f(x_1, \dots, x_n)$ in which the formula on the right side is automatically re-evaluated and assigned to the variable y whenever any variable x_i changes. If y is modified from outside the constraint, the equation is left temporarily unsatisfied, hence the attribute “one-way”. Dataflow constraints are recognized as a powerful programming methodology in a variety of contexts because of their versatility and simplicity [46]. The most widespread application of dataflow constraints is perhaps embodied by spreadsheets [2, 36]. In a spreadsheet, the user can specify a cell formula that depends on other cells: when any of those cells is updated, the value of the first cell is automatically recalculated. Rules in a `Makefile` are another example of dataflow constraints: a rule sets up a dependency between a target file and a list of input files, and provides shell commands for rebuilding the target from the input files. When the `Makefile` is run, if any input file in a rule is discovered to be newer than the target, then the target is rebuilt. The dataflow principle can also be applied to software development and execution, where the role of a cell/file is replaced by a program variable. This approach has been widely explored in the context of interactive applications, multimedia animation, and real-time systems [13, 15, 29, 35, 41, 47].

Since the values of program variables are automatically recalculated upon changes of other values, the dataflow computational model is very different from the standard imperative model, in which the memory store is changed explicitly by the program via memory assignments. The execution

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

OOPSLA'11, October 22–27, 2011, Portland, Oregon, USA.
Copyright © 2011 ACM 978-1-4503-0940-0/11/10...\$10.00

flow of applications running on top of a dataflow environment is indeed data-driven, rather than control-driven, providing a natural ground for automatic change propagation in all scenarios where programs need to react to modifications of their environment. Implementations of the dataflow principle share some common issues with self-adjusting computation, in which programs respond to input changes by updating automatically their output [3–5, 30].

Differently from purely declarative constraints [7], dataflow constraints are expressed by means of (imperative) methods whose execution makes a relation satisfied. This programming style is intuitive and readily accessible to a broad range of developers [46]. The ability to smoothly combine different paradigms in a unified framework makes it possible to take advantage of different programming styles in the context of the same application. The problem of integrating imperative and dataflow programming has already been the focus of previous work in the context of specific application domains [11, 40, 41, 46]. Previous mixed imperative/dataflow systems are based on libraries of *ad hoc* data types and functions for representing constraint variables and for binding the imperative program to the constraint solver. One drawback of these approaches is that constraint variables can only be of special data types provided by the runtime library, causing loss of flexibility and placing undue burden on the programmer. A natural question is whether the dataflow model can be made to work with general-purpose, imperative languages, such as C, without adding syntactic extensions and *ad hoc* data types. In this paper we affirmatively answer this question.

Our Contributions. We present a general-purpose framework where programmers can specify generic one-way constraints between objects of arbitrary types stored in *reactive* memory locations. Constraints are written as ordinary statements of the imperative host language and can be added and removed dynamically at run time. Since they can change multiple objects within the same execution, they are multi-output. The main feature of a constraint is its sensitivity to modifications of reactive objects: a constraint is automatically re-evaluated whenever any of the reactive locations it depends on is changed, either by the imperative program, or by another constraint. A distinguishing feature of our approach is that the whole constraint solving mechanism is handled transparently by altering the semantics of elementary operations of the imperative host language for reading and modifying objects. No syntax extensions are required and no new primitives are needed except for adding/removing constraints, allocating/deallocating reactive memory locations, and controlling the granularity of solver activations. Differently from previous approaches, programmers are not forced to use any special data types provided by the language extension, and can resort to the full range of conventional constructs for accessing and manipulating objects offered by the host language. In addition, our framework supports all

the other features that have been recognized to be important in the design of dataflow constraint systems [46], including:

Arbitrary code: constraints consist of arbitrary code that is legal in the underlying imperative language, thus including loops, conditionals, function calls, and recursion.

Address dereferencing: constraints are able to reference variables indirectly via pointers.

Automatic dependency detection: constraints automatically detect the reactive memory locations they depend on during their evaluation, so there is no need for programmers to explicitly declare dependencies, which are also allowed to vary over time.

We embodied these principles into an extension of C/C++ that we called DC. Our extension has the same syntax as C/C++, but operations that read or modify objects have a different semantics. All other primitives, including creating and deleting constraints and allocating and deallocating reactive memory blocks, are provided as runtime library functions.

Our main contributions are reflected in the organization of the paper and can be summarized as follows:

- In Section 2 we abstract our mechanism showing how to extend an elementary imperative language to support one-way dataflow constraints using reactive memory. We formally describe our mixed imperative/dataflow computational model, provide a formal semantics, and discuss convergence properties of our dataflow constraint solver by modeling the computation as an iterative process that aims at finding a common fixpoint for the current set of constraints.
- In Section 3 we describe the concrete embodiment of our technique into C/C++, introducing the main features of DC.
- In Section 4 we give a variety of elementary and advanced programming examples and discuss how DC can improve C/C++ programmability in three relevant application scenarios: incremental computation, implementation of the observer software design pattern, and data structure checking and repair. To the best of our knowledge, these applications have not been explored before in the context of dataflow programming.
- In Section 5 we describe how DC can be implemented using off-the-shelf compilers on conventional platforms via a combination of runtime libraries, hardware/operating system support, and dynamic code patching, without requiring any source code preprocessing.
- In Section 6 we discuss the results of an extensive experimental analysis of DC in a variety of settings, showing that our implementation is effective in practice. We consider both interactive applications and computationally demanding benchmarks that manipulate lists, grids, trees, and graphs. We assess the performances of DC

against conventional C-based implementations as well as against competitors that can quickly react to input changes, i.e., *ad hoc* dynamic algorithms [20, 44], incremental solutions realized in CEAL [30] (a state-of-the-art C-based framework for self-adjusting computation), and Qt’s signal-slot implementation of the subject/observer communication mechanism [26].

Related work is discussed in Section 7 and directions for future research are sketched in Section 8. An extended version of this paper, including proofs and additional experiments, can be found in the accompanying technical report [23].

2. Abstract Model

To describe our approach, we consider an elementary imperative language and we show how to extend it to support one-way dataflow constraints. We start from WHILE [43], an extremely simple language of commands including a sub-language of expressions. Although WHILE does not support many fundamental features of concrete imperative languages (including declarations, procedures, dynamic memory allocation, type checking, etc.), it provides all the building blocks for a formal description of our mechanism, abstracting away details irrelevant for our purposes. We discuss how to modify the semantics of WHILE to integrate a *dataflow constraint solver*. We call the extended language DWHILE. DWHILE is identical to WHILE except for a different semantics and additional primitives for adding/deleting constraints dynamically and for controlling the granularity of solver activations. As we will see in Section 3, these primitives can be supported in procedural languages as runtime library functions.

2.1 The DWHILE Language

The abstract syntax of DWHILE is shown in Figure 1. The language distinguishes between commands (i.e., statements) and expressions. We use c, c_1, c_2 as meta-variables ranging over the set of commands $Comm$, and e, e_1, e_2 as meta-variables ranging over the set of expressions Exp . Canonical forms of expressions are either storage locations $\ell \in Loc$, or storable values v over some arbitrary domain Val . Expressions can also be obtained by applying to sub-expressions any primitive operations defined over domain Val (e.g., plus, minus, etc.). Commands include:

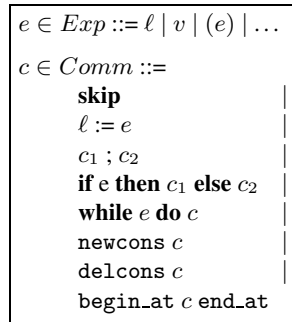


Figure 1. Abstract syntax of DWHILE.

- Assignments of values to storage locations ($\ell := e$). These commands are the basic state transformers.

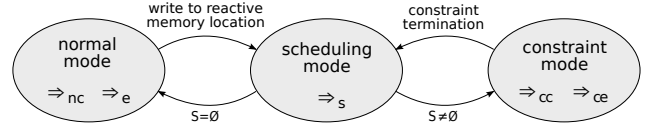


Figure 2. Transitions between different execution modes.

- Constructs for sequencing, conditional execution, and iteration, with the usual meaning.
- Two new primitives, `newcons` and `delcons`, for adding and deleting constraints dynamically. Notice that a constraint in DWHILE is just an ordinary command.
- An atomic block construct, `begin_at c end_at`, that executes a command c atomically so that any constraint evaluation is deferred until the end of the block. This offers fine-grained control over solver activations.

In Section 3 we will show a direct application of the concepts developed in this section to the C/C++ programming languages.

2.2 Memory Model and Execution Modes

Our approach hinges upon two key notions: *reactive memory locations* and *constraints*. Reactive memory can be read and written just like ordinary memory. However, differently from ordinary memory:

1. If a constraint c reads a reactive memory location ℓ during its execution, a dependency (ℓ, c) of c from ℓ is added to a set D of dependencies.
2. If the value stored in a reactive memory location ℓ is changed, all constraints depending on ℓ (i.e., all constraints c such that $(\ell, c) \in D$) are automatically re-executed.

Point 2 states that constraints are sensitive to modifications of the contents of the reactive memory. Point 1 shows how to maintain dynamically the set D of dependencies needed to trigger the appropriate constraints upon changes of reactive memory locations. We remark that re-evaluating a constraint c may completely change the set of its dependencies: prior to re-execution, all the old dependencies $(-, c) \in D$ are discarded, and new dependencies are logged in D during the re-evaluation of c .

As shown in Figure 2, at any point in time the execution can be in one of three modes: *normal execution*, *constraint execution*, or *scheduling*. As we will see more formally later in this section, different instructions (such as reading a reactive memory location or assigning it with a value) may have different semantics depending on the current execution mode.

We assume *eager* constraint evaluation, i.e., out-of-date constraints are brought up-to-date as soon as possible. This choice is better suited to our framework and, as previous experience has shown, lazy and eager evaluators typically de-

liver comparable performance in practice [46]. Eager evaluation is achieved as follows. A scheduler maintains a data structure S containing constraints to be first executed or re-evaluated. As an invariant property, S is guaranteed to be empty during normal execution. As soon as a reactive memory location ℓ is written, the scheduler queries the set D of dependencies and adds to S all the constraints depending on ℓ . These constraints are then run one-by-one in constraint execution mode, and new constraints may be added to S throughout this process. Whenever S becomes empty, normal execution is resumed.

An exception to eager evaluation is related to atomic blocks. The execution of an atomic block c is regarded as an uninterruptible operation: new constraints created during the evaluation of c are just added to S . When c terminates, for each reactive memory location ℓ whose value has changed, all the constraints depending on ℓ are also added to S , and the solver is eventually activated. Constraint executions are uninterruptible as well.

We remark that any scheduling mechanism may be used for selecting from S the next constraint to be evaluated: in this abstract model we rely on a function `pick` that implements any appropriate scheduling strategy.

2.3 Configurations

Given a set X , we denote by 2^X its power set. A configuration of our system is a six-tuple

$$(\rho, a, \sigma, D, S, c_{self}) \in \mathcal{R} \times Bool \times \Sigma \times Dep \times 2^{Cons} \times Cons$$

where:

- \mathcal{R} is a set of Boolean store attributes of the form $\rho : Loc \rightarrow \{\text{normal}, \text{reactive}\}$, where function $\rho \in \mathcal{R}$ specifies which memory locations are reactive.
- $a \in \{\text{true}, \text{false}\}$ is a Boolean flag that is `true` inside atomic blocks and is used for deferring solver activations.
- $\Sigma = \{\sigma : Loc \rightarrow Val\}$ is a set of stores mapping storage locations to storable values.
- $D \subseteq Loc \times Cons$ is a set of dependencies, $Cons$ is a set of constraints, and $Dep = 2^{Loc \times Cons}$. A constraint can be any command in DWHILE, i.e., $Cons = Comm$. We use different names for the sake of clarity.
- $S \subseteq Cons$ is the scheduling data structure.
- c_{self} is a meta-variable that denotes the *current* constraint (i.e., the constraint that is being evaluated) in constraint execution mode, and is undefined otherwise. If the scheduler were deterministic, c_{self} may be omitted from the configuration, but we do not make this assumption in this paper.

2.4 Operational Semantics

Most of the operational semantics of the DWHILE language can be directly derived from the standard semantics

of WHILE. The most interesting aspects of our extension include reading and writing the reactive memory, adding and deleting constraints, executing commands atomically, and defining the behavior of the scheduler and its interactions with the other execution modes. Rules for these aspects are given in Figure 4 and are discussed below.

Let $\Rightarrow_e \subseteq (\Sigma \times Exp) \times Val$ and $\Rightarrow_c \subseteq (\Sigma \times Comm) \times \Sigma$ be the standard big-step transition relations used in the operational semantics of the WHILE language [43]. Besides \Rightarrow_e and \Rightarrow_c , we use additional transition relations for expression evaluation in constraint mode (\Rightarrow_{ce}), command execution in normal mode (\Rightarrow_{nc}), command execution in constraint mode (\Rightarrow_{cc}), and constraint solver execution in scheduling mode (\Rightarrow_s), as defined in Figure 3. Notice that expression evaluation in normal mode can be carried on directly by means of transition relation \Rightarrow_e of WHILE. As discussed below, relation \Rightarrow_{ce} is obtained by appropriately modifying \Rightarrow_e . Similarly, relations \Rightarrow_{nc} and \Rightarrow_{cc} are obtained by appropriately modifying \Rightarrow_c . All the rules not reported in Figure 4 can be derived in a straightforward way from the corresponding rules in the standard semantics of WHILE [43].

The evaluation of a DWHILE program is started by rule EVAL, which initializes the atomic flag a to `false` and both the scheduling queue S and the set D of dependencies to the empty set.

Writing Memory. Assigning an ordinary memory location in normal execution mode (rule ASGN-N1) just changes the store as in the usual semantics of WHILE. This is also the case when the new value of the location to be assigned equals its old value or inside an atomic block. Otherwise, if the location ℓ to be assigned is reactive, the new value differs from the old one, and execution is outside atomic blocks (rule ASGN-N2), constraints depending on ℓ are scheduled in S and are evaluated one-by-one. As we will see, the transition relation \Rightarrow_s guarantees S to be empty at the end of the constraint solving phase. In conformity with the atomic execution of constraints, assignment in constraint mode (rule ASGN-C) just resorts to ordinary assignment in WHILE for both normal and reactive locations. We will see in rule SOLVER-2, however, that constraints can be nevertheless scheduled by other constraints if their execution changes the contents of reactive memory locations.

Reading Memory. Reading an ordinary memory location in constraint execution mode (rule DEREf-C1) just evaluates the location to its value in the current store: this is specified by transition relation \Rightarrow_e of the WHILE semantics. If the location ℓ to be read is reactive (rule DEREf-C2), a new dependency of the active constraint c_{self} from ℓ is also added to the set D of dependencies.

Executing Atomic Blocks. To execute an atomic block in normal mode (rule BEGINEND-N2), the uninterruptible command c is first evaluated according to the rules defined by transition \Rightarrow_{nc} . If the content of some reactive loca-

$\Rightarrow \subseteq (\mathcal{R} \times \Sigma \times \text{Comm}) \times \Sigma$	$\langle \rho, \sigma, c \rangle \Rightarrow \sigma'$
$\Rightarrow_{ce} \subseteq (\mathcal{R} \times \Sigma \times \text{Cons} \times \text{Dep} \times \text{Exp}) \times (\text{Dep} \times \text{Val})$	$\langle \rho, \sigma, c_{self}, D, e \rangle \Rightarrow_{ce} \langle D', v \rangle$
$\Rightarrow_{nc} \subseteq (\mathcal{R} \times \text{Bool} \times \Sigma \times \text{Dep} \times 2^{\text{Cons}} \times \text{Comm}) \times (\Sigma \times \text{Dep} \times 2^{\text{Cons}})$	$\langle \rho, a, \sigma, D, S, c \rangle \Rightarrow_{nc} \langle \sigma', D', S' \rangle$
$\Rightarrow_{cc} \subseteq (\mathcal{R} \times \Sigma \times \text{Dep} \times 2^{\text{Cons}} \times \text{Cons} \times \text{Comm}) \times (\Sigma \times \text{Dep} \times 2^{\text{Cons}})$	$\langle \rho, \sigma, D, S, c_{self}, c \rangle \Rightarrow_{cc} \langle \sigma', D', S' \rangle$
$\Rightarrow_s \subseteq (\mathcal{R} \times \Sigma \times \text{Dep} \times 2^{\text{Cons}}) \times (\Sigma \times \text{Dep})$	$\langle \rho, \sigma, D, S \rangle \Rightarrow_s \langle \sigma', D' \rangle$

Figure 3. Transition relations for DWHILE program evaluation (\Rightarrow), expression evaluation in constraint mode (\Rightarrow_{ce}), command execution in normal mode (\Rightarrow_{nc}), command execution in constraint mode (\Rightarrow_{cc}), and constraint solver execution in scheduling mode (\Rightarrow_s).

$\frac{\langle \rho, a, \sigma, D, S, c \rangle \Rightarrow_{nc} \langle \sigma', D', S \rangle}{\langle \rho, \sigma, c \rangle \Rightarrow \sigma'}$ <p style="text-align: center;">(EVAL)</p>	<p>where: $\begin{cases} a = \text{false} \\ D = \emptyset \\ S = \emptyset \end{cases}$</p> $\frac{\langle \rho, \sigma, c_{self}, D, e \rangle \Rightarrow_{ce} \langle D', v \rangle \quad \sigma' = \sigma _{\ell \mapsto v}}{\langle \rho, \sigma, D, S, c_{self}, \ell := e \rangle \Rightarrow_{cc} \langle \sigma', D', S \rangle}$ <p style="text-align: center;">(ASGN-C)</p>
$\frac{\sigma \vdash e \Rightarrow_e v \quad \sigma' = \sigma _{\ell \mapsto v}}{\langle \rho, a, \sigma, D, S, \ell := e \rangle \Rightarrow_{nc} \langle \sigma', D, S \rangle}$ <p style="text-align: center;">if $\rho(\ell) = \text{normal}$ or $\sigma'(\ell) = \sigma(\ell)$ or $a = \text{true}$</p> <p style="text-align: center;">(ASGN-N1)</p>	$\frac{S = \emptyset \quad S' = \{c \mid (\ell, c) \in D\}}{\sigma \vdash e \Rightarrow_e v \quad \sigma' = \sigma _{\ell \mapsto v} \quad \langle \rho, \sigma', D, S' \rangle \Rightarrow_s \langle \sigma'', D' \rangle}$ $\langle \rho, a, \sigma, D, S, \ell := e \rangle \Rightarrow_{nc} \langle \sigma'', D', S \rangle$ <p style="text-align: center;">if $\rho(\ell) = \text{reactive}$ and $\sigma'(\ell) \neq \sigma(\ell)$ and $a = \text{false}$</p> <p style="text-align: center;">(ASGN-N2)</p>
$\frac{\sigma \vdash \ell \Rightarrow_e v}{\langle \rho, \sigma, c_{self}, D, \ell \rangle \Rightarrow_{ce} \langle D, v \rangle}$ <p style="text-align: center;">if $\rho(\ell) = \text{normal}$</p> <p style="text-align: center;">(DEREF-C1)</p>	$\frac{\sigma \vdash \ell \Rightarrow_e v \quad D' = D \cup \{(\ell, c_{self})\}}{\langle \rho, \sigma, c_{self}, D, \ell \rangle \Rightarrow_{ce} \langle D', v \rangle}$ <p style="text-align: center;">if $\rho(\ell) = \text{reactive}$</p> <p style="text-align: center;">(DEREF-C2)</p>
$\frac{\langle \rho, \sigma, D, S, c_{self}, c \rangle \Rightarrow_{cc} \langle \sigma', D', S' \rangle}{\langle \rho, \sigma, D, S, c_{self}, \text{begin_at } c \text{ end_at} \rangle \Rightarrow_{cc} \langle \sigma', D', S' \rangle}$ <p style="text-align: center;">(BEGINEND-C)</p>	$\frac{\langle \rho, a, \sigma, D, S, c \rangle \Rightarrow_{nc} \langle \sigma', D, S' \rangle}{\langle \rho, a, \sigma, D, S, \text{begin_at } c \text{ end_at} \rangle \Rightarrow_{nc} \langle \sigma', D, S' \rangle}$ <p style="text-align: center;">if $a = \text{true}$</p> <p style="text-align: center;">(BEGINEND-N1)</p>
$\frac{S = \emptyset \quad a' = \text{true} \quad \langle \rho, a', \sigma, D, S, c \rangle \Rightarrow_{nc} \langle \sigma', D, S' \rangle}{S'' = S' \cup \{c \mid (\ell, c) \in D \wedge \sigma(\ell) \neq \sigma'(\ell)\} \quad \langle \rho, \sigma', D, S'' \rangle \Rightarrow_s \langle \sigma'', D' \rangle}$ $\langle \rho, a, \sigma, D, S, \text{begin_at } c \text{ end_at} \rangle \Rightarrow_{nc} \langle \sigma'', D', S \rangle$ <p style="text-align: center;">if $a = \text{false}$</p> <p style="text-align: center;">(BEGINEND-N2)</p>	
$\frac{S = \emptyset \quad S' = \{c\} \quad \langle \rho, \sigma, D, S' \rangle \Rightarrow_s \langle \sigma', D' \rangle}{\langle \rho, a, \sigma, D, S, \text{newcons } c \rangle \Rightarrow_{nc} \langle \sigma', D', S \rangle}$ <p style="text-align: center;">if $a = \text{false}$</p> <p style="text-align: center;">(NEWCONS-N1)</p>	$\frac{S' = S \cup \{c\}}{\langle \rho, a, \sigma, D, S, \text{newcons } c \rangle \Rightarrow_{nc} \langle \sigma, D, S' \rangle}$ <p style="text-align: center;">if $a = \text{true}$</p> <p style="text-align: center;">(NEWCONS-N2)</p>
$\frac{D' = D \setminus \{(\ell, x) \in D \mid x = c\} \quad S' = S \setminus \{c\}}{\langle \rho, a, \sigma, D, S, \text{delcons } c \rangle \Rightarrow_{nc} \langle \sigma, D', S' \rangle}$ <p style="text-align: center;">(DELCONS-N)</p>	$\frac{S' = S \cup \{c\}}{\langle \rho, \sigma, D, S, c_{self}, \text{newcons } c \rangle \Rightarrow_{cc} \langle \sigma, D, S' \rangle}$ <p style="text-align: center;">(NEWCONS-C)</p>
$\frac{D' = D \setminus \{(\ell, x) \in D \mid x = c\} \quad S' = S \setminus \{c\}}{\langle \rho, \sigma, D, S, c_{self}, \text{delcons } c \rangle \Rightarrow_{cc} \langle \sigma, D', S' \rangle}$ <p style="text-align: center;">(DELCONS-C)</p>	$\langle \rho, \sigma, D, S \rangle \Rightarrow_s \langle \sigma, D \rangle$ <p style="text-align: center;">if $S = \emptyset$</p> <p style="text-align: center;">(SOLVER-1)</p>
$\frac{\langle \rho, \sigma, D', S \setminus \{c_{self}\}, c_{self}, c_{self} \rangle \Rightarrow_{cc} \langle \sigma', D'', S' \rangle \quad \langle \rho, \sigma', D', S'' \rangle \Rightarrow_s \langle \sigma'', D''' \rangle}{\langle \rho, \sigma, D, S \rangle \Rightarrow_s \langle \sigma'', D''' \rangle}$ <p style="text-align: center;">(SOLVER-2)</p>	<p>where: $\begin{cases} c_{self} = \text{pick}(S) \\ D' = D \setminus \{(\ell, x) \in D \mid x = c_{self}\} \\ S'' = S' \cup \{c \mid (\ell, c) \in D'' \wedge \sigma(\ell) \neq \sigma'(\ell)\} \end{cases}$ if $S \neq \emptyset$</p>

Figure 4. DWHILE program evaluation.

tion changes due to the execution of c , the solver is then activated at the end of the block. The `begin_at/end_at` command has instead no effect when execution is already atomic, i.e., in constraint mode (rule `BEGINEND-C`) and inside atomic blocks (rule `BEGINEND-N1`), except for executing command c .

Creating and Deleting Constraints. In non-atomic normal execution mode, rule `NEWCONS-N1` creates a new constraint and triggers its first execution as specified by \Rightarrow_s . In atomic normal execution and in constraint mode, rules `NEWCONS-N2` and `NEWCONS-C` simply add the constraint to the scheduling queue. Similarly, rules `DELCONS-N` and `DELCONS-C` remove the constraint from the scheduling queue and clean up its dependencies from D .

Activating the Solver. Rules `SOLVER-1` and `SOLVER-2` specify the behavior of the scheduler, which is started by rules `ASGN-N2` and `BEGINEND-N2`. Rule `SOLVER-1` defines the termination of the constraint solving phase: this phase ends only when there are no more constraints to be evaluated (i.e., $S = \emptyset$). Rule `SOLVER-2` has an inductive definition. If S is not empty, function `pick` selects from S a new active constraint c_{self} , which is evaluated in constraint mode after removing from D its old dependencies. The final state (σ'') and dependencies (D'') are those obtained by applying the scheduler on the store σ' obtained after the execution of c_{self} and on a new set S'' of constraints. S'' is derived from S by adding any new constraints (S') resulting from the execution of c_{self} along with the constraints depending on reactive memory locations whose content has been changed by c_{self} . The definition of S'' guarantees that constraints can trigger other constraints (even themselves), even if each constraint execution is regarded as an atomic operation and is never interrupted by the scheduler.

2.5 Convergence Properties

The approach we follow in our work consists of modeling dataflow constraint solving as an iterative process that aims at finding a common fixpoint for the current set of constraints. In our context, a fixpoint is a store that satisfies simultaneously all the relations between reactive memory locations specified by the constraints. As we will see in Section 2.6, this provides a unifying framework for solving dataflow constraint systems with both acyclic and cyclic dependencies. So far, we have assumed that the scheduling order of constraint executions is specified by a function `pick` given as a parameter of the solver. A natural question is whether there are any general properties of a set of constraints that let our solver terminate and converge to a common fixpoint independently of the scheduling strategy used by function `pick`.

We adapt the theory in [7, 17] to our framework, modeling constraint executions as the application of functions on stores: let $f_c : \Sigma \rightarrow \Sigma$ be the function computed by a constraint $c \in Cons$, where $f_c(\sigma) = \sigma'$ if $\langle \sigma, c \rangle \Rightarrow_c \sigma'$. A store

$\sigma \in \Sigma$ is a **FIXPOINT** for f_c if $f_c(\sigma) = \sigma$. To simplify the discussion, in this section we assume that constraints only operate on reactive cells and focus our attention on stores where all locations are reactive. The definition of inflationary functions assumes that a partial ordering is defined on the set of stores Σ :

DEFINITION 1 (INFLATIONARY FUNCTIONS). Let (Σ, \preceq) be any partial ordering over the set of stores Σ and let $f : \Sigma \rightarrow \Sigma$ be a function on Σ . We say that f is inflationary if $\sigma \preceq f(\sigma)$ for all $\sigma \in \Sigma$.

A relevant property of partial orderings in our context is the *finite chain property*:

DEFINITION 2 (FINITE CHAIN PROPERTY). A partial ordering (Σ, \preceq) over Σ satisfies the **FINITE CHAIN PROPERTY** if every non-decreasing sequence of elements $\sigma_0 \preceq \sigma_1 \preceq \sigma_2 \preceq \dots$ from Σ eventually stabilizes at some element σ in Σ , i.e., if there exists $j \geq 0$ such that $\sigma_i = \sigma$ for all $i \geq j$.

To describe the store modifications due to the execution of the solver, we use the notion of *iteration of functions* on stores. Let $F = \{f_1, \dots, f_n\}$, $\langle a_1, \dots, a_k \rangle \in [1, n]^k$, and $\sigma \in \Sigma$ be a finite set of functions on Σ , a sequence of indices in $[1, n]$, and an initial store, respectively. An iteration of functions of F starting at σ is a sequence of stores $\langle \sigma_0, \sigma_1, \sigma_2, \dots \rangle$ where $\sigma_0 = \sigma$ and $\sigma_i = f_{a_i}(\sigma_{i-1})$ for $i > 0$. We say that function f_{a_i} is *activated* at step i . Iterations of functions that lead to a fixed point are called *regular*:

DEFINITION 3 (REGULAR FUNCTION ITERATION). A function iteration $\langle \sigma_0, \sigma_1, \sigma_2, \dots \rangle$ is **REGULAR** if it satisfies the following property: for all $f \in F$ and $i \geq 0$, if σ_i is not a fixpoint for f , then f is activated at some step $j > i$.

Using arguments from Chapter 7 of [7], it can be proved that any regular iteration of inflationary functions starting at some initial store stabilizes in a finite number of steps to a common fixpoint. A relevant convergence property of our solver is the following:

THEOREM 1. Let $C = \{c_1, \dots, c_h\}$ be any set of constraints, let $F = \{f_{c_1}, \dots, f_{c_h}\}$ be the functions computed by constraints in C , and let (Σ, \preceq) be any partial ordering over Σ satisfying the finite chain property. If functions in F are inflationary on Σ and $\{f \in F \mid f(\sigma) \neq \sigma\} \subseteq S \subseteq F$, then $\langle \rho, \sigma, D, S \rangle \Rightarrow_s \langle \sigma', D' \rangle$ and σ' is a common fixpoint of the functions in F such that $\sigma \preceq \sigma'$.

If functions in Theorem 1 are also *monotonic* (i.e., $\sigma \preceq \sigma'$ implies $f(\sigma) \preceq f(\sigma')$ for all $\sigma, \sigma' \in \Sigma$), then the solver always converges to the *least* common fixpoint [7], yielding deterministic results independently of the scheduling order.

2.6 Cyclic vs. Acyclic Systems of Constraints

It is possible to prove that the class of inflationary constraints discussed in Theorem 1 includes any program that can be de-

```

insert( $u, v, w$ ):
   $E := E \cup \{(u, v)\}$ 
   $w(u, v) := w$ 
  newcons( $\underbrace{\text{if } d[u] + w(u, v) < d[v] \text{ then } d[v] := d[u] + w(u, v)}_{c_{uv}}$ )
decrease( $u, v, \delta$ ):
   $w(u, v) := w(u, v) - \delta$ 

```

Figure 5. Incremental shortest path updates in a mixed imperative/dataflow style.

scribed in terms of an acyclic dataflow graph, such as computational circuits [6], non-circular attribute grammars [37], and spreadsheets [36]. Hence, if a system of constraints is acyclic, then our solver always converges to the correct result, without the need for programmers to prove any stabilization properties of their constraints. This is the most common case in many applications [6, 19, 34], and several efficient techniques can be adopted by constraint solvers to automatically detect cycles introduced by programming errors [46].

In addition to the acyclic case, our abstract machine can also handle the most general case of cyclic constraints embedded within an imperative program. This opens up the possibility to address problems that would not be solvable using acyclic dataflow graphs. We exemplify this concept by considering the well known algorithmic problem of maintaining distances in a graph subject to local changes to its nodes or edges. In the remainder of this section we show how to specify an incremental variant of the classical Bellman-Ford’s single-source shortest path algorithm [8] in terms of a (possibly cyclic) system of one-way constraints. Compared to purely imperative specifications [20], the formulation of the incremental algorithm in our mixed imperative/dataflow framework is surprisingly simple and requires just a few lines of code.

Example. Let $G = (V, E, w)$ be a directed graph with real edge weights $w(u, v)$, and let s be a source node in V . The *incremental shortest path problem* consists of updating the distances $d[u]$ of all nodes $u \in V$ from the source s after inserting any new edge in the graph, or decreasing the weight of any existing edge. To solve this problem in our framework, we keep edge weights and distances in reactive memory. Assuming to start from a graph with no edges, we initialize $d[s] := 0$ and $d[u] := +\infty$ for all $u \neq s$. The pseudocode of update operations that insert a new edge and decrease the weight of an existing edge by a positive amount δ are shown in Figure 5. Operation $\text{insert}(u, v, w)$ adds edge (u, v) to the graph with weight w and creates a new constraint c_{uv} for the edge: c_{uv} simply *relaxes* the edge if Bellman’s inequality $d[u] + w(u, v) \geq d[v]$ is violated [8]. The constraint is immediately executed after creation (see rule NEWCONS-N1 in Figure 4) and the three pairs $(d[u], c_{uv})$, $(d[v], c_{uv})$, and $(w(u, v), c_{uv})$ are added to the set of dependencies D . Any later change to $d[u]$, $d[v]$ or $w(u, v)$, which may violate the

inequality $d[u] + w(u, v) \geq d[v]$, will cause the re-execution of c_{uv} . Decreasing the weight of an existing edge (u, v) by any positive constant δ with $\text{decrease}(u, v, \delta)$ can be done by just updating $w(u, v)$. In view of rule ASGN-N2 of Figure 4, the system reacts to the change and automatically re-executes c_{uv} and any other affected constraints.

Using the machinery developed in Section 2.5 and suitably defining a partial order on Σ and an appropriate *pick* function, it can be proved that our solver finds a correct solution within the best known worst-case time bounds for the problem. In Section 6.3 we will analyze experimentally our constraint-based approach showing that in practice it can be orders of magnitude faster than recomputing from scratch, even when all weights are non-negative.

2.7 Glitch Avoidance

The performances of our solver depend on the ability to minimize *glitches* [15], i.e., redundant evaluations of constraints. In the case of acyclic systems, glitch freedom can be achieved by creating constraints in topological order and by letting the *pick* function return the least recently created/executed constraint. For cyclic systems, glitch avoidance is problem-specific, and relies on the definition of appropriate *pick* functions. In some cases, glitches may be entailed by algorithmic properties of the program at hand, and therefore may be unavoidable. For instance, glitches may arise in the shortest paths example discussed in Section 2.6 if multiple edge insertions are performed as a batch inside an atomic block. This is a consequence of the fact that each edge may be relaxed up to $|V|$ times by Bellman-Ford’s algorithm. Defining a *pick* function that guarantees the glitch freedom property in this situation would improve the long-standing $O(|V| \cdot |E|)$ time bound of the shortest paths problem with arbitrary edge weights.

3. Embodiment into C/C++

In this section we show how to apply the concepts developed in Section 2 to the C and C++ languages, deriving an extension that we call DC. DC has exactly the same syntax as C/C++, but operations that read or modify objects have a different semantics. All other primitives, including creating/deleting constraints, allocating/deallocating reactive objects, and opening/closing atomic blocks, are provided as runtime library functions¹ (see Figure 6).

Reactive Memory Allocation. Similarly to other automatic change propagation approaches (e.g., [5, 41]), in DC all objects allocated statically or dynamically are non-reactive by default. Reactive locations are allocated dynamically using library functions `rmalloc` and `rfree`, which work just like `malloc` and `free`, but on a separate heap.

¹A detailed documentation of the DC application programming interface, including stricter library naming conventions and several additional features not covered in this paper, is available at the URL: <http://www.dis.uniroma1.it/~demetres/dc/>

```

typedef void (*cons_t)(void*);
int newcons(cons_t cons, void* param);
void delcons(int cons_id);
void* rmalloc(size_t size);
void rfree(void* ptr);
void begin_at();
void end_at();
void arm_final(int cons_id, cons_t final);
void set_comp(int (*comp)(void*, void*));

```

Figure 6. Main functions of the DC language extension.

Opening and Closing Atomic Blocks. Atomic blocks are supported in DC using two library functions `begin_at` and `end_at`. Calling `begin_at` opens an atomic block, which should be closed with a matching call to `end_at`. Nested atomic blocks are allowed, and are handled using a counter of nesting levels so that the solver is only resumed at the end of the outer block, processing any pending constraints that need to be first executed or brought up to date as a result of the block’s execution.

Creating and Deleting Constraints. For the sake of simplicity, in Section 2 constraints have been modeled as ordinary statements. DC takes a more flexible approach: constraints are specified as closures formed by a function that carries out the computation and a user-defined parameter to be passed to the function. Different constraints may therefore share the same function code, but have different user-defined parameters. New constraint instances can be created by calling `newcons`, which takes as parameters a pointer `cons` to a function and a user-defined parameter `param`. When invoked in non-atomic normal execution mode, `newcons` executes immediately function `cons` with parameter `param`, and logs all dependencies between the created constraint and the reactive locations read during the execution. If a constraint is created inside an atomic block (or inside another constraint), its first evaluation is deferred until the end of the execution of the current block (or constraint). All subsequent re-executions of the constraint triggered by modifications of the reactive cells it depends on will be performed with the same value of `param` specified at the creation time. `newcons` returns a unique id for the created constraint, which can be passed to `delcons`.

Reading and Modifying Objects. Reading and modifying objects in reactive memory can be done in DC by evaluating ordinary C/C++ expressions. We remark that no syntax extensions or explicit macro/function invocations are required.

Customizing the Scheduler. Differently from other approaches [41], DC allows programmers to customize the execution order of scheduled constraints. While the default `pick` function of DC (which gives higher priority to least recently executed constraints) works just fine in practice for a large class of problems (see Section 6), the ability to replace it can play an important role for some specific problems. DC provides a function `set_comp` that installs a

```

struct robject {
    void* operator new(size_t size) { return rmalloc(size); }
    void operator delete(void* ptr) { rfree(ptr); }
};

static void con_h(void*), fin_h(void*);
class rcons {
    int id;
public:
    virtual void cons() = 0;
    virtual void final() {}
    rcons() { id = -1; }
    ~rcons() { disable(); }
    void enable() { if (id == -1) id = newcons(con_h, this); }
    void disable() { if (id != -1) { delcons(id); id = -1; } }
    void arm_final() { if (id != -1) arm_final(id, fin_h); }
    void unarm_final() { if (id != -1) arm_final(id, NULL); }
};

void con_h(void* p) { ((rcons*)p)->cons(); }
void fin_h(void* p) { ((rcons*)p)->final(); }

```

Figure 7. C++ wrapping of DC primitives.

user-defined comparator to determine the relative priority of two constraints. The comparator receives as arguments the user-defined parameters associated with the constraints to be compared.

Final Handlers. An additional feature of DC, built on top of the core constraint handling mechanisms described in Section 4, is the ability to perform some finalization operations only when the results of constraint evaluations are stable, i.e., when the solver has found a common fixpoint. For instance, a constraint computing the attribute of a widget in a graphic user interface may also update the screen by calling drawing primitives of the GUI toolkit: if a redrawing occurs at each constraint execution, this may cause unnecessary screen updates and flickering effects. Another usage example of this feature will be given in Section 4.3.

DC allows users to specify portions of code for a constraint to be executed as final actions just before resuming the underlying imperative program interrupted by the solver activation. This can be done by calling function `arm_final` during constraint solving: the operation schedules a *final handler* to be executed at the end of the current solving session. The function takes as parameters a constraint id and a pointer to a final handler, or `NULL` to cancel a previous request. A final handler receives the same parameter as the constraint it is associated to, but no dependencies from reactive locations are logged during its execution. All final handlers are executed in normal execution mode inside an atomic block.

C++ Wrapping of DC Primitives. The examples in the remainder of this paper are based on a simple C++ wrapping of the DC primitives, shown in Figure 7. We abstract the concepts of reactive object and constraint using two classes: `robject` and `rcons`. The former is a base class for objects stored in reactive memory. This is achieved by overloading the `new` and `delete` operators in terms of the corresponding DC primitives `rmalloc` and `rfree`, so that all member variables of the object are reactive. Class `rcons` is a virtual

base class for objects representing dataflow constraints. The class provides a pure virtual function called `cons`, to be defined in subclasses, which provides the user code for a constraint. An additional empty `final` function can be optionally overridden in subclasses to define the finalization code for a constraint. The class also provides functions `enable` and `disable` to activate/deactivate the constraint associated with the object, and functions `arm_final` and `unarm_final` to schedule/unschedule the execution of final handlers.

4. Applications and Programming Examples

In this section, we discuss how DC can improve C/C++ programmability in three relevant application scenarios. To the best of our knowledge, these applications have not been explored before in the context of dataflow programming. All the code we show is real.

4.1 Incremental Computation

In many applications, input data is subject to continuous updates that need to be processed efficiently. For instance, in a networking scenario, routers must react quickly to link failures by updating routing tables in order to minimize communication delays. When the input is subject to small changes, a program may incrementally fix only the portion of the output affected by the update, without having to recompute the entire solution from scratch. For many problems, efficient *ad hoc* algorithms are known that can update the output asymptotically faster than recomputing from scratch, delivering in practice speedups of several orders of magnitude [19, 21]. Such dynamic algorithms, however, are typically difficult to design and implement, even for problems that are easy to be solved from-scratch. A language-centric approach, which was extensively explored in both functional and imperative programming languages, consists of automatically turning a conventional static algorithm into an incremental one, by selectively recomputing the portions of a computation affected by an update of the input. This powerful technique, known as self-adjusting computation [3, 5], provides a principled way of deriving efficient incremental code for several problems. We now show that dataflow constraints can provide an effective alternative for specifying incremental programs. Later in this section we discuss differences and similarities of the two approaches.

Example. To put our approach into the perspective of previous work on self-adjusting computation, we revisit the problem of incremental re-evaluation of binary expression trees discussed in [30]. This problem is a special case of the circuit evaluation: input values are stored at the leaves and the value of each internal node is determined by applying a binary operator (e.g., sum or product) on the values of its children. The final result of the evaluation is stored at the root. We start from the conventional node structure that a programmer would use for a binary expression tree, containing the type of the operation computed at the node (only rel-

```
template<typename T> struct node : robject, rcons {
    enum op_t { SUM, PROD };
    T val;
    op_t op;
    node *left, *right;
    node(T v): val(v), left(NULL), right(NULL) { enable(); }
    node(op_t o): op(o), left(NULL), right(NULL) { enable(); }
    void cons() {
        if (left == NULL || right == NULL) return;
        switch (op) {
            case SUM: val = left->val + right->val; break;
            case PROD: val = left->val * right->val; break;
        }
    }
};
```

Figure 8. Incremental evaluation of expression trees.

evant for internal nodes), the node’s value, and the pointers to the subtrees. Our DC-based solution (see Figure 8) simply extends the node declaration by letting it inherit from classes `robject` and `rcons`, and by providing the code of a constraint that computes the value of the node in terms of the values stored at its children. Everything else is exactly what the programmer would have done anyway to build the input data structure. An expression tree can be constructed by just creating nodes and connecting them in the usual way:

```
node<int> *root = new node<int>(node<int>::SUM);
root->left = new node<int>(10);
root->right = new node<int>(node<int>::PROD);
root->right->left = new node<int>(2);
root->right->right = new node<int>(6);
```

The example above creates the tree shown in Figure 9 (left). Since all fields of the node are reactive and each node is equipped with a constraint that computes its value, at any time during the tree construction, `root->value` contains the correct result of the expression evaluation. We remark that this value not only is given for free without the need to compute it explicitly by traversing the tree, but is also updated automatically after any change of the tree. For instance, changing the value of the rightmost leaf with `root->right->right->val = 3` triggers the propagation chain shown in Figure 9 (right). Other possible updates that would be automatically propagated include changing the operation type of a node or even adding/removing entire subtrees. Notice that a single change to a node may trigger the re-execution of the constraints attached to all its ancestors, so the total worst-case time per update is $O(h)$, where h is the height of the tree. For a balanced expression tree, this is exponentially faster than recomputing from scratch. If a batch of changes are to be performed and only the final value of the tree is of interest, performance can be improved by grouping updates with `begin_at()` and `end_at()` so that the re-execution of constraints is deferred until the end of the batch, e.g.:

```
begin_at(); // put the solver to sleep
root->op = node<int>::SUM; // change node operation type
delete root->right->left // delete leaf
... // etc...
end_at(); // wake up the solver
```

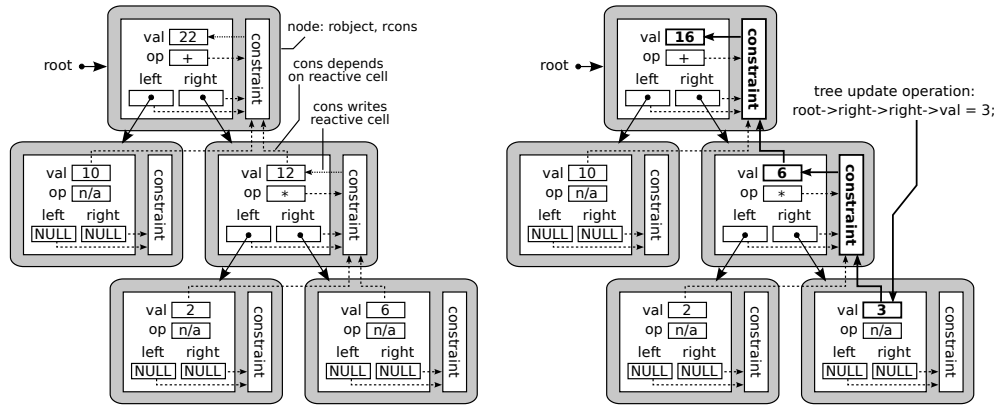


Figure 9. Reactive expression tree (left) and change propagation chain after a leaf value update (right).

Discussion. DC and imperative self-adjusting computation languages such as CEAL [30] share the basic idea of change propagation, and reactive memory is very similar to CEAL’s modifiables. However, the two approaches differ in a number of important aspects. In CEAL, the solution is initially computed by a core component and later updated by a *mutator*, which performs changes to the input. In DC there is no explicit distinction between an initial run and a sequence of updates, and in particular there is no static algorithm that is automatically dynamized. Instead, programmers explicitly break down the solution of a complex problem into a collection of reactive code fragments that locally update small portions of the program state as a function of other portions. This implies a paradigm shift that may be less straightforward for the average programmer than writing static algorithms, but it can make it easier to exploit specific properties of the problem at hand, which can be crucial for coding algorithms provably faster than recomputing from scratch.

4.2 Implementing the Observer Design Pattern

As a second example, we show how the reactive nature of our framework can be naturally exploited to implement the observer software design pattern. A common issue arising from partitioning a system into a collection of cooperating software modules is the need to maintain consistency between related objects. In general, a tight coupling of the involved software components is not desirable, as it would reduce their reusability. For example, graphical user interface toolkits almost invariably separate presentational aspects from the underlying application data management, allowing data processing and data presentation modules to be reused independently. The *observer software design pattern* [14] answers the above concerns by defining one-to-many dependencies between objects so that when one object (the *subject*) changes state, all its dependents (the *observers*) are automatically notified. A key aspect is that subjects send out notifications of their change of state, without having to know who their observers are, while any number

of observers can be subscribed to receive these notifications (subjects and observers are therefore not tightly coupled). A widely deployed embodiment of this pattern is provided by the Qt application development framework [26].

Qt is based on a signal-slot communication mechanism: a signal is emitted when a particular event occurs, whereas a slot is a function that is called in response to a particular signal. An object acting as a subject emits signals in response to changes of its state by explicitly calling a special member function designated as a signal. Observers and subjects can be explicitly connected so that any signal emitted by a subject triggers the invocation of one or more observer slots. Programmers can connect as many signals as they want to a single slot, and a signal can be connected to as many slots as they need. Since the connection is set up externally after creating the objects, this approach allows objects to be unaware of the existence of each other, enhancing information encapsulation and reuse of software components. Subjects and observers can be created in Qt as instances of the `QObject` base class. Qt’s signal-slot infrastructure hinges upon an extension of the C++ language with three new keywords: `signal` and `slot`, to designate functions as signals or slots, and `emit`, to generate signals.

A Minimal Example: Qt vs. DC. To illustrate the concepts discussed above and compare Qt and DC as tools for implementing the observer pattern, we consider a minimal example excerpted from the Qt 4.6 reference documentation. The goal is to set up a program in which two counter variables `a` and `b` are connected together so that the value of `b` is automatically kept consistent with the value of `a`. The example starts with the simple declaration shown in Figure 10(a) (all except the framed box), which encapsulates the counter into an object with member functions `value/setValue` for accessing/modifying it. Figure 10(b) shows how the `Counter` class can be modified in Qt so that counter modifications can be automatically propagated to other objects as prescribed by the observer pattern. First of all, the class inherits from Qt’s `QObject` base class and starts with the `QObject`

```

class Counter : public robject {
public:
    Counter() { m_value = 0; }
    int value() const { return m_value; }
    void setValue(int value) { m_value = value; }
private:
    int m_value;
};
(a) A counter class and its DC observer pattern version (framed box).

class Counter : public QObject {
    Q_OBJECT
public:
    Counter() { m_value = 0; }
    int value() const { return m_value; }
public slots:
    void setValue(int value);
signals:
    void valueChanged(int newValue);
private:
    int m_value;
};

void Counter::setValue(int value) {
    if (value != m_value) {
        m_value = value;
        emit valueChanged(value);
    }
}
(b) Qt observer pattern version of the counter class.

```

Figure 10. Observer pattern example excerpted from the Qt 4.6 reference documentation: DC vs. Qt implementation.

macro. Function `setValue` is declared as a slot and it is augmented by calling explicitly the `valueChanged` signal with the `emit` keyword every time an actual change occurs. Since Qt `Counter` objects contain both signal and slot functions they can act both as subjects and as observers. The following code snippet shows how two counters can be created and connected so that each change to the former triggers a change of the latter:

```

Counter *a = new Counter, *b = new Counter;

QObject::connect(a, SIGNAL(valueChanged(int)),
                b, SLOT(setValue(int)));

a->setValue(12); // a->value() == 12, b->value() == 12
b->setValue(48); // a->value() == 12, b->value() == 48

```

The `QObject::connect` call installs a connection between counters `a` and `b`: every time `emit valueChanged(value)` is issued by `a` with a given actual parameter, `setValue(int value)` is automatically invoked on `b` with the same parameter. Therefore, the call `a->setValue(12)` has as a side-effect that the value of `b` is also set to 12. Conversely, the call `b->setValue(48)` entails no change of `a` as no connection exists from `b` to `a`.

The same result can be achieved in DC by just letting the `Counter` class of Figure 10(a) inherit from the `robject` base class of Figure 7. As a result, the `m_value` member variable is stored in reactive memory. The prescribed connection

between reactive counters can be enforced with a one-way dataflow constraint that simply assigns the value of `b` equal to the value of `a`:

```

Counter *a = new Counter, *b = new Counter;

struct C : rcons {
    Counter *a, *b;
    C(Counter *a, Counter *b) : a(a), b(b) { enable(); }
    void cons() { b->setValue(a->value()); }
} c(a,b);

a->setValue(12); // a->value() == 12, b->value() == 12
b->setValue(48); // a->value() == 12, b->value() == 48

```

We notice that the role of the `QObject::connect` of the Qt implementation is now played by a dataflow constraint, yielding exactly the same program behavior.

Discussion. The example above shows that DC's runtime system handles automatically a number of aspects that would have to be set up explicitly by the programmers using Qt's mechanism:

- there is no need to define slots and signals, relieving programmers from the burden of extending the definition of subject and observer classes with extra machinery (see Figure 10);
- only actual changes of an object's state trigger propagation events, so programmers do not have to make explicit checks such as in `Counter::setValue`'s definition to prevent infinite looping in the case of cyclic connections (see Figure 10(b));
- DC does not require extensions of the language, and thus the code does not have to be preprocessed before being compiled.

We list below further points that make dataflow constraints a flexible framework for supporting some aspects of component programming, putting it into the perspective of mainstream embodiments of the observer pattern such as Qt:

- in DC, only subjects need to be reactive, while observers can be of any C++ class, even of third-party libraries distributed in binary code form. In Qt, third-party observers must be wrapped using classes equipped with slots that act as stubs;
- relations between Qt objects are specified by creating explicitly one-to-one signal-slot connections one at a time; a single DC constraint can enforce simultaneously any arbitrary set of many-to-many relations. Furthermore, as the input variables of a dataflow constraint are detected automatically, relations may change dynamically depending on the state of some objects;
- Qt signal-slot connections let subjects communicate copies of values to their observers; in contrast, DC constraints can compute the values received by the observers as an arbitrary function of the state of multiple subjects, encapsulating complex update semantics.

```

0 template<class T, class N> class snode : public rcons {
1   map<N**, snode<T,N>*> *m;
2   N   *head, **tail;
3   snode *next;
4   int  refc;
5 public:
6   snode(N *h, N **t, map<N**, snode<T,N>*> *m) :
7     m(m), head(h), tail(t), next(NULL), refc(0) {
8     (*m)[tail] = this;
9     enable();
10  }
11  ~snode() {
12    m->erase(tail);
13    if (next != NULL && --next->refc == 0) delete next;
14  }
15  void cons() {
16    snode<T,N>* cur_next;
17
18    if (*tail != NULL) {
19      typename map<N**, snode<T,N>*>::iterator it =
20        m->find( &(*tail)->next );
21      if (it != m->end())
22        cur_next = it->second;
23      else cur_next = new snode<T,N>(*tail,
24                                   &(*tail)->next, m);
25    } else cur_next = NULL;
26
27    if (next != cur_next) {
28      if (next != NULL && --next->refc == 0)
29        next->arm_final();
30      if (cur_next != NULL && cur_next->refc++ == 0)
31        cur_next->unarm_final();
32      next = cur_next;
33    }
34    if (head != NULL) T::watch(head);
35  }
36  void final() { delete this; }
37 };
38
39 template<class T, class N> class watcher {
40   snode<T,N> *gen;
41   map<N**, snode<T,N>*> m;
42 public:
43   watcher(N** h) { gen = new snode<T,N>(NULL, h, &m); }
44   ~watcher()    { delete gen; }
45 };

```

Figure 11. Data structure checking and repair: *list watcher*.

4.3 Data Structure Checking and Repair

Long-living applications inevitably experience various forms of damage, often due to bugs in the program, which could lead to system crashes or wrong computational results. The ability of a program to perform automatic consistency checks and self-healing operations can greatly improve reliability in software development. One of the most common causes of faults is connected with different kinds of data structure corruptions, which can be mitigated using data structure repair techniques [24].

In this section, we show how dataflow constraints can be used to check and repair reactive data structures. We exemplify this concept by considering the simple problem of repairing a corrupt doubly-linked list [38]. We first show how to build a generic *list watcher*, which is able to detect any changes to a list and perform actions when modifications occur. This provides an advanced example of DC programming, where constraints are created and destroyed by other constraints. Differently from the expression trees of Section 4.1, where constraints are attributes of nodes, the

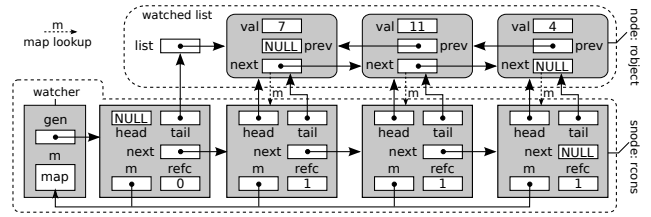


Figure 12. A reactive doubly-linked list, monitored by a *watcher*.

main challenge here is how to let the watched list be completely unaware of the watcher, while still maintaining automatically a constraint for each node. The complete code of the watcher is shown in Figure 11. The only assumption our watcher makes on list nodes to be monitored (of generic type N) is that they are reactive and contain a *next* field pointing to the successor. The main idea is to maintain a *shadow list* of constraints that mirrors the watched list (Figure 12). Shadow nodes are *snode* objects containing pointers to the monitored nodes (*head*) and to their next fields (*tail*). A special shadow generator node (*gen*) is associated to the reactive variable (*list*) holding the pointer to the first node of the input list. A lookup table (*m*) maintains a mapping from list nodes to the corresponding shadow nodes. The heart of the watcher is the constraint associated with shadow nodes (lines 15–35). It first checks if the successor of the monitored node, if any, is already mapped to a shadow node (lines 18–21). If not, it creates a new shadow node (line 23). Lines 27–33 handle the case where the successor of the shadow node has changed and its *next* field has to be updated. Line 34 calls a user-defined *watch* function (provided by template parameter T), which performs any desired checks and repairs for an input list node. To dispose of shadow nodes when the corresponding nodes are disconnected from the list, we use a simple reference counting technique, deferring to a final handler the task of deallocating dead shadow nodes (line 36).

The following code snippet shows how to create a simple repairer for a doubly-linked list based on the watcher of Figure 11:

```

struct node : robject { int val; node *next, *prev; };

struct myrepairer {
  static void watch(node* x) {
    // check
    if (x->next != NULL && x != x->next->prev)
      // repair
      x->next->prev = x;
  }
};

// create reactive list head and repairer
node** list = ...;
watcher<myrepairer,node> rep(list);

// manipulate the list
...

```

The repairer object `rep` checks if the invariant property $x == x \rightarrow \text{next} \rightarrow \text{prev}$ is satisfied for all nodes in the list, and recovers it to a consistent state if any violation is detected during the execution of the program. We notice that several different watchers may be created to monitor the same list.

5. Implementation

In this section we discuss how DC can be implemented via a combination of runtime libraries, hardware/operating system support, and dynamic code patching, without requiring any source code preprocessing. The overall architecture of our DC implementation, which was developed on a Linux IA-32 platform, is shown in Figure 13. At a very high level, the DC runtime library is stratified into two modules: 1) a *reactive memory manager*, which defines the `rmalloc` and `rfree` primitives and provides support for tracing accesses to reactive memory locations; 2) a *constraint solver*, which schedules and dispatches the execution of constraints, keeping track of dependencies between reactive memory locations and constraints. We start our description by discussing how to support reactive memory, which is the backbone of the whole architecture.

5.1 Reactive Memory

Taking inspiration from transactional memories [1], we implemented reactive memory using off-the-shelf memory protection hardware. Our key technique uses access violations (AV) combined with dynamic binary code patching as a basic mechanism to trace read/write operations to reactive memory locations.

Access Violations and Dynamic Code Patching. Reactive memory is kept in a *protected* region of the address space so that any read/write access to a reactive object raises an AV. Since access violation handling is very inefficient, we use it just to incrementally detect instructions that access reactive memory. When an instruction x first tries to access a reactive location, a segmentation fault with offending instruction x is raised. In the SIGSEGV handler, we patch the trace t containing x by overwriting its initial 5 bytes with a jump to a dynamically recompiled trace t' derived from t , which is placed in a code cache. In trace t' , x is instrumented with additional inline code that accesses reactive locations without generating AVs, and possibly activates the constraint solver. Trace t' ends with a jump that leads back to the end of t so that control flow can continue normally in the original code. Since t' may contain several memory access instructions, it is re-generated every time a new instruction that accesses reactive memory is discovered. To identify traces in the code, we analyze statically the binary code when it is loaded and we construct a lookup table that maps the address of each memory access instruction to the trace containing it. To handle the cases where a trace in a function f is smaller than 5 bytes and thus cannot be patched, we overwrite the beginning of f with a jump to a new version

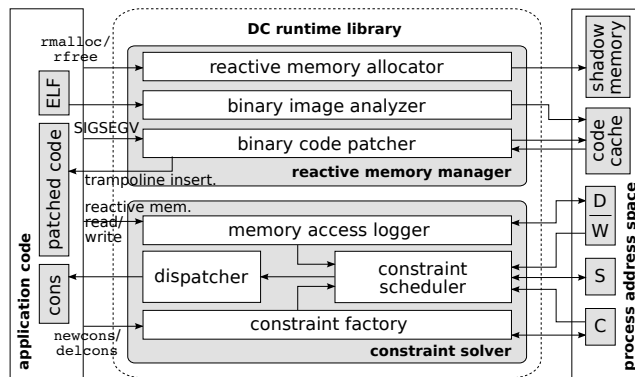
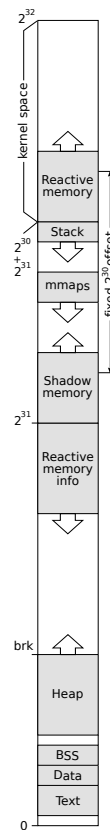


Figure 13. DC’s software architecture.

f' of f where traces are padded with trailing `nop` instructions so that the smallest trace is at least 5-bytes long.

Shadow Memory and Address Redirecting. To avoid expensive un-protect and re-protect page operations at each access to reactive memory, we mirror reactive memory pages with unprotected *shadow pages* that contain the actual data. The shadow memory region is kept under the control of our reactive memory allocator, which maps it onto the swap space with the `mmap` system call. Any access to a reactive object is transparently redirected to the corresponding object in the shadow memory. As a result, memory locations at addresses within the reactive memory region are never actually read or written by the program. To avoid wasting memory without actually accessing it, reactive memory can be placed within the Kernel space, located in the upper 1GB of the address space on 32-bit Linux machines with the classical 3/1 virtual address split. Kernel space is flagged in the page tables as exclusive to privileged code (ring 2 or lower), thus an AV is triggered if a user-mode instruction tries to touch it. More recent 64-bit platforms offer even more flexibility to accommodate reactive memory in protected regions of the address space. We let the reactive memory region start at address $2^{30} + 2^{31} = 0xC000000$ and grow upward as more space is needed (see the figure on the right). The shadow memory region starts at address $2^{31} = 0x8000000$ and grows upward, eventually hitting the memory mapping segment used by Linux to keep dynamic libraries, anonymous mappings, etc. Any reactive object at address x is mirrored by a shadow object at address $x - \delta$, where $\delta = 2^{30} = 0x4000000$ is a fixed offset. This makes address redirecting very efficient.



5.2 Constraint Solver

Our implementation aggregates reactive locations in 4-byte *words* aligned at 32 bit boundaries. The solver is activated every time such a word is read in constraint execution mode, or its value is modified by a write operation. The main involved units are (see Figure 13):

1. A *dispatcher* that executes constraints, maintaining a global *timestamp* that grows by one at each constraint execution. For each constraint, we keep the timestamp of its latest execution.
2. A *memory access logger* that maintains the set of dependencies D and a list W of all reactive memory words written by the execution of the current constraint c_{self} , along with their initial values before the execution. To avoid logging information about the same word multiple times during the execution of a constraint, the logger stamps each word with the time of the latest constraint execution that accessed it. Information is logged only if the accessed word has a timestamp older than the current global timestamp, which can only happen once for any constraint execution. To represent D , the logger keeps for each word v the address of the head node of a linked list containing the id's of constraints depending upon v .
3. A *constraint scheduler* that maintains the set of scheduled constraints S . By default S is a priority queue, where the priority of a constraint is given by the timestamp of its latest execution: the scheduler repeatedly picks and lets the dispatcher execute the constraint with the highest priority, until S gets empty. Upon completion of a constraint's execution, words are scanned and removed from W : for each $v \in W$ whose value has changed since the beginning of the execution, the constraint id's in the list of nodes associated with v are added to S , if not already there.

Nodes of the linked lists that represent D and data structures S and W are kept in contiguous chunks allocated with `malloc`. To support direct lookup, timestamps and dependency list heads for reactive memory words are stored in a contiguous *reactive memory info* region that starts at address $2^{31} = 0x8000000$ and grows downward, eventually hitting the heap's `brk`.

A critical aspect is how to clean up old dependencies in D when a constraint is re-evaluated. To solve the problem efficiently in constant amortized time per list operation, we keep for each node its insertion time into the linked list. We say that a node is *stale* if its timestamp is older than the timestamp of the constraint it refers to, and *up to date* otherwise. Our solver uses a lazy approach and disposes of stale nodes only when the word they refer to is modified and the linked list is traversed to add constraints to S . To prevent the number of stale nodes from growing too large, we use an incremental garbage collection technique.

6. Experimental Evaluation

In this section we present an experimental analysis of the performances of DC in a variety of different settings, showing that our implementation is effective in practice.

6.1 Benchmark Suite

We have evaluated DC on a set of benchmarks that includes a variety of problems on lists, grids, trees, and graphs, as well as full and event-intensive interactive applications.

Linked Lists. We considered several fundamental primitives on linear linked data structures, which provide a variety of data manipulation patterns. Our benchmarks include data structures for: computing the sum of the elements in a list (`adder`), filtering the items of a list according to a given function (`filter`), randomly assigning each element of a list to one of two output lists (`halver`), mapping the items of a list onto new values according to a given mapping function (`mapper`), merging two sorted lists into a single sorted output list (`merger`), producing a sorted version of an input list (`msorter`), producing a reversed version of an input list (`reverser`), and splitting a list into two output lists, each containing only elements smaller or, respectively, greater than a given pivot (`splitter`). All benchmarks are subject to operations that add or remove nodes from the input lists.

Graphs and Trees. Benchmarks in this class include classical algorithmic problems for routing in networks and tree computations:

- `sp`: given a weighted directed graph and a source node s , computes the distances of all graph nodes from s . Graph edges are subject to edge weight decreases.
- `exptrees`: computes the value of an expression tree subject to operations that change leaf values or operators computed by internal nodes (see Section 4.1).

Interactive Applications. We considered both full real applications and synthetic worst-case scenarios, including:

- `othello`: full application that implements the well-known board game in which two players in turn place colored pieces on a square board, with the goal of reversing as many of their opponent's pieces as possible;
- `buttongrid`: event-intensive graphic user interface application with a window containing $n \times n$ push buttons embedded in a grid layout. This is an extreme artificial scenario in which many events are generated, since a quadratic number of buttons need to be resized and repositioned to maintain the prescribed layout at each interactive resize event.

Some benchmarks, such as `sp`, are very computationally demanding. For all these benchmarks we have considered an implementation based on DC, obtained by making the base data structures (e.g., the input list) reactive, and a conven-

Benchmark	From-scratch time (secs)						Propagation time (msecs)			Mem peak usage (Mbytes)			DC statistics		
	conv	dc	ceal	dc conv	ceal conv	ceal dc	dc	ceal	ceal dc	dc	ceal	ceal dc	avg cons per update	instr time	patched instr
adder	0.10	1.44	1.40	14.40	14.00	0.97	0.68	85.80	126.17	211.54	232.87	1.10	1.5	0.030	26
exptrees	0.14	1.02	1.07	7.28	7.64	1.04	4.11	5.46	1.32	143.30	225.32	1.57	15.6	0.028	72
filter	0.19	2.08	1.11	10.94	5.84	0.53	0.63	2.49	3.95	265.78	189.47	0.71	0.5	0.032	39
halver	0.20	2.08	1.33	10.40	6.65	0.63	0.61	3.95	6.47	269.10	218.22	0.81	0.5	0.030	38
mapper	0.19	2.04	1.30	10.73	6.84	0.63	0.61	2.63	4.31	261.53	214.34	0.81	0.5	0.032	39
merger	0.19	2.12	1.37	11.15	7.21	0.64	0.66	4.43	6.71	284.41	218.21	0.81	0.5	0.031	57
msorter	0.91	5.18	3.91	5.69	4.29	0.75	5.55	15.91	2.86	689.59	820.14	1.18	37.6	0.031	75
reverser	0.18	2.04	1.30	11.33	7.22	0.63	0.62	2.63	4.24	267.45	214.34	0.80	0.5	0.030	37
splitter	0.18	2.27	1.31	12.61	7.27	0.57	1.54	3.92	2.54	344.60	222.34	0.64	1.5	0.031	56

Table 1. Performance evaluation of DC versus CEAL, for a common set of benchmarks. Input size is $n = 1,000,000$ for all tests except `msorter`, for which $n = 100,000$.

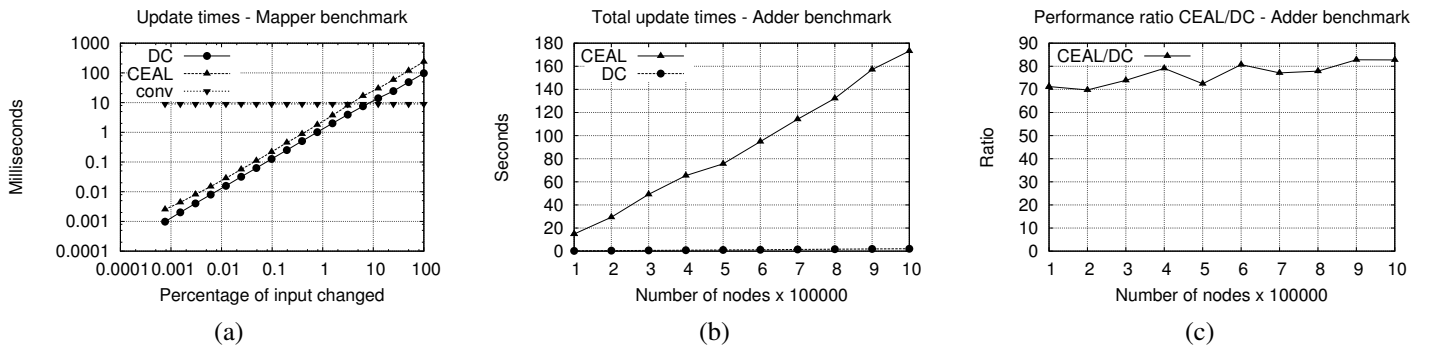


Figure 14. (a) Change propagation times on the `mapper` benchmark for complex updates with input size $n = 100,000$; (b-c) performance comparison of the change propagation times of DC and CEAL on the `adder` benchmark.

tional implementation in C based on non-reactive data structures. Interactive applications (`othello` and `buttongrid`) are written in the Qt-4 framework: change propagation throughout the GUI is implemented either using constraints (DC versions), or using the standard signal-slot mechanism provided by Qt (conventional versions). To assess the performances of DC against competitors that can quickly respond to input changes, we have also considered highly tuned *ad-hoc* dynamic algorithms [20, 44] and incremental solutions realized in CEAL [30], a state-of-the-art C-based language for self-adjusting computation. Benchmarks in common with CEAL are `adder`, `exptrees`, `filter`, `halver`, `mapper`, `merger`, `msorter`, `reverser`, and `splitter`. For these benchmarks, we have used the optimized implementations provided by Hammer *et al.* [30].

6.2 Performance Metrics and Experimental Setup

We tested our benchmarks both on synthetic and on real test sets, considering a variety of performance metrics:

Running times: we measured the time required to initialize the data structures with the input data (from-scratch execution), the time required by change propagation, and binary code instrumentation time. All reported times are wall-clock times, averaged over three independent trials. Times were

measured with `gettimeofday()`, turning off any other processes running in the background.

Memory usage: we computed the memory peak usage as well as a detailed breakdown to assess which components of our implementation take up most memory (constraints, shadow memory, reactive memory, stale and non-stale dependencies, etc.).

DC-related statistics: we collected detailed profiling information including counts of patched instructions, stale dependencies cleanups, allocated/deallocated reactive blocks, created/deleted constraints, constraints executed per update, and distinct constraints executed per update.

All DC programs considered in this section, except for `sp` that will be discussed separately, use the default timestamp-based comparator for constraint scheduling.

The experiments were performed on a PC equipped with a 2.10 GHz Intel Core 2 Duo with 3 GB of RAM, running Linux Mandriva 2010.1 with Qt 4.6. All programs were compiled with `gcc 4.4.3` and optimization flag `-O3`.

6.3 Incremental Computation

The reactive nature of our mixed imperative/dataflow framework makes it a natural ground for incremental computation. In this section, we present experimental evidence that

Road network			From-scratch time (msec)		Propagation time (msec)		Speedup		Mem peak usage (Mbytes)			Statistics	
Graph	$n \cdot 10^3$	$m \cdot 10^3$	sq	sp	rr	$\frac{sq}{sp}$	$\frac{sq}{rr}$	sp	rr	sq	sp cons per update	rr node scans per update	
NY	264	733	50.99	0.16	0.07	318.6	728.4	76.75	26.62	26.19	143.9	143.9	
BAY	321	800	59.99	0.15	0.07	399.9	857.0	84.84	30.21	29.82	170.6	170.5	
COL	435	1,057	79.98	0.28	0.17	285.6	470.4	108.61	39.09	38.97	378.3	378.2	
FLA	1,070	2,712	192.97	0.63	0.35	306.3	551.3	251.26	93.42	93.29	687.5	687.3	
NW	1,207	2,840	236.96	0.87	0.54	272.3	438.8	270.66	102.15	101.53	1002.4	1002.3	
NE	1,524	3,897	354.94	0.27	0.16	1314.5	2218.3	350.86	132.85	132.15	320.2	320.1	

Table 2. Performance evaluation of DC for incremental routing in US road networks using up to 1.5 million constraints.

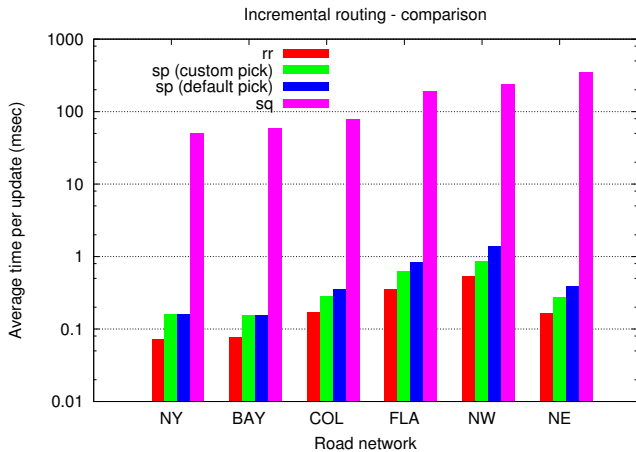


Figure 15. Analysis of different pick function definitions on the incremental routing problem.

a constraint-based solution in our framework can respond to input updates very efficiently. We first show that the propagation times are comparable to state of the art automatic change propagation frameworks, such as CEAL [30], and for some problems can be orders of magnitude faster than recomputing from scratch. We then consider a routing problem on real road networks, and compare our DC-based solution both to a conventional implementation and to a highly optimized *ad hoc* dynamic algorithm supporting a class of specific update operations.

Comparison to CEAL. Table 1 summarizes the outcome of our experimental comparison with the conventional version and with CEAL for all common benchmarks. Input size is $n = 1,000,000$ for all tests (with the exception of *msorter*, for which $n = 100,000$), where n is the length of the input list for the list-based benchmarks, and the number of nodes in the (balanced) input tree for *exptrees*. Table 1 reports from-scratch execution times of both DC and CEAL (compared to the corresponding conventional implementations), average propagation times in response to small changes of the input, memory usage and some DC stats (average number of executed constraints per update, executable instrumentation time, and total number of patched instruc-

tions). The experiments show that our DC implementation performs remarkably well. From-scratch times are on average a factor of 1.4 higher than those of CEAL, while propagation times are smaller by a factor of 4 on average for all tests considered except the adder, yielding large speed-ups over complete recalculation. In the case of the adder benchmark, DC leads by a huge margin in terms of propagation time (see Figure 14a and Figure 14b), which can be attributed to the different asymptotic performance of the algorithms handling the change propagation (constant for DC, and logarithmic in the input size for the list reduction approach used by CEAL). We remark that the logarithmic bound of self-adjusting computation could be reduced to constant by using a traceable accumulator [5]; however, support for traceable data structures is not yet integrated in CEAL.

We also investigated how DC and CEAL scale in the case of batches of updates that change multiple input items simultaneously. The results are reported in Figure 14a for the representative mapper benchmark, showing that the selective recalculations performed by DC and CEAL are faster than recomputing from scratch for changes up to significant percentages of the input.

Comparison to ad hoc Incremental Shortest Paths. We now consider an application of the shortest path algorithm discussed in Section 2.6 to incremental routing in road networks. We assess the empirical performance of a constraint-based solution implemented in DC (sp) by comparing it with Goldberg’s *smart queue* implementation of Dijkstra’s algorithm (sq), a highly-optimized C++ code used as the reference benchmark in the 9th DIMACS Implementation Challenge [22], and with an engineered version of the *ad hoc* incremental algorithm by Ramalingam and Reps (rr) [20, 44]. Our code supports update operations following the high-level description given in Figure 5, except that we create one constraint per node, rather than one constraint per edge. We used as input data a suite of US road networks of size up to $n = 1.5$ million nodes and $m = 3.8$ million edges derived from the UA Census 2000 TIGER/Line Files [45]. Edge weights are large and represent integer positive travel times. We performed on each graph a sequence of $m/10$ random edge weight decreases, obtained by picking edges uni-

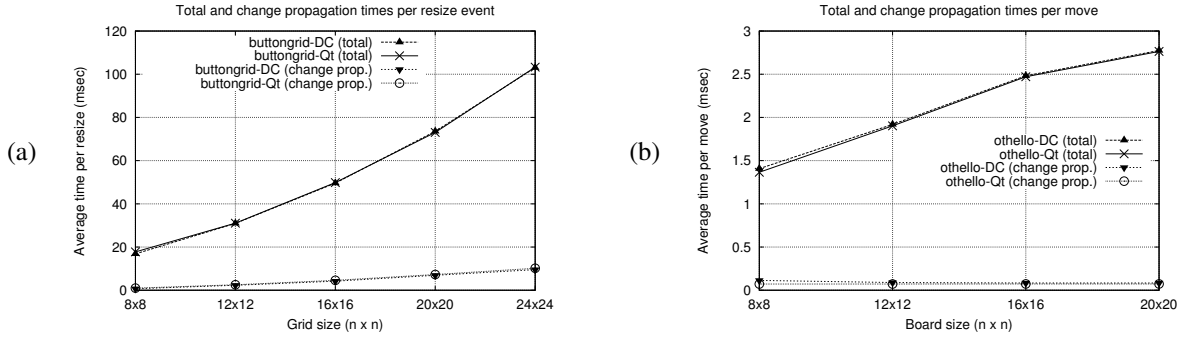


Figure 16. Comparison with signal-slot mechanism in Qt: (a) buttongrid; (b) othello.

formly at random and reducing their weights by a factor of 2. Updates that did not change any distances were not counted.

The results of our experiments are shown in Table 2 and Figure 15. Both `sp` and `rr` were initialized with distances computed using `sq`, hence we report from-scratch time only for this algorithm. Due to the nature of the problem, the average number of node distances affected by an update is rather small and almost independent of the size of the graph. Analogously to the incremental algorithm of Ramalingam and Reps, the automatic change propagation strategy used by our solver takes full advantage of this strong locality, re-evaluating only affected constraints and delivering substantial speedups over static solutions in typical scenarios. Our DC-based implementation yields propagation times that are, on average, a factor of 1.85 higher than the conventional *ad hoc* incremental algorithm, but it is less complex, requires fewer lines of code, is fully composable, and is able to respond seamlessly to multiple data changes, relieving the programmer from the task of implementing explicitly change propagation. We also tested `sp` with different types of schedulers. By customizing the `pick` function of the default priority queue scheduler (giving highest priority to nodes closest to the source), a noticeable performance improvement has been achieved (see Figure 15). We also tried a simple stack scheduler, which, however, incurred a slowdown of a factor of 4 over the default scheduler.

6.4 Comparison to Qt’s Signal-slot Mechanism

Maintaining relations between widgets in a graphic user interface is one of the most classical applications of dataflow constraints [46]. We assess the performance of DC in event-intensive interactive applications by comparing the DC implementations of `buttongrid` and `othello` with the conventional versions built atop Qt’s signal-slot mechanism.

In `buttongrid`, each constraint computes the size and position of a button in terms of the size and position of adjacent buttons. We considered user interaction sessions with continuous resizing, which induce intensive scheduling activity along several propagation chains in the acyclic dataflow graph. In `othello`, constraints are attached to cells

of the game board (stored in reactive memory) and maintain a mapping between the board and its graphical representation: in this way, the game logic can be completely unaware of the GUI backend, as prescribed by the observer pattern (see Section 4.2). For both benchmarks, we experimented with different grid/board sizes. Figure 16 plots the average time per resize event (`buttongrid`) and per game move (`othello`), measured over 3 independent runs. Both the total time and the change propagation time are reported. For all values of n , the performance differences of the DC and Qt conventional implementations are negligible and the curves are almost overlapped. Furthermore, the time spent in change propagation is only a small fraction of the total time, showing that the overhead introduced by access violations handling, instrumentation, and scheduling in DC can be largely amortized over the general cost of widget management and event propagation in Qt and in its underlying layers.

7. Related Work

The ability of a program to respond to modifications of its environment is a feature that has been widely explored in a large variety of settings and along rather different research lines. While this section is far from being exhaustive, we discuss some previous works that appear to be more closely related to ours.

GUI and Animation Toolkits. Although dataflow programming is a general paradigm, dataflow constraints have gained popularity in the 90’s especially in the creation of interactive user interfaces. Amulet [41] is a graphic user interface toolkit based on the dataflow paradigm. It integrates a constraint solver with a prototype-instance object model implemented on top of C++, and is closely related to our work. Each object, created by making an instance of a prototype object, consists of a set of properties (e.g., appearance or position) that are stored in reactive variables, called slots. Constraints are created by assigning formulas to slots. Values of slots are accessed through a `Get` method that, when invoked from inside of a formula, sets up a dependency be-

tween slots. A variety of approaches have been tested by the developers to solve constraints [46].

FRAN (Functional Reactive Animation) provides a reactive environment for composing multimedia animations through temporal modeling [25]: graphical objects in FRAN use time-varying, reactive variables to automatically change their properties, achieving an animation that is function of both events and time.

The data-driven Alpha language provided by the Leonardo software visualization system [18] allows programmers to specify declarative mappings between the state of a C program and a graphical representation of its data structures.

Reactive Languages. The dataflow model of computation can also be supported directly by programming languages. Most of them are visual languages, often used in industrial settings [10], and allow the programmer to directly manage the dataflow graph by visually putting links between the various entities. Only a few non-visual languages provide a dataflow environment, mostly for specific domains. Among them, Signal [29] and Lustre [13] are dedicated to programming real-time systems found in embedded software, and SystemC [28] is a system-level specification and design language based on C++.

Functional Reactive Programming (FRP) is a declarative programming model for constructing interactive applications [25, 42, 47]. FRP offers two kinds of reactive inputs: *behaviors* (e.g., time-continuous variables whose value changes are automatically propagated by the language), and *events* (e.g., potentially infinite streams of discrete events, each of which triggers additional computations). Early implementations of FRP have been embedded in the programming language Haskell [33]. FrTime [15] extends a purely functional subset of PLT Scheme with an instantiation of the FRP paradigm, supporting eager evaluation and *benign impurities* (e.g., imperative commands for drawing, and for creating and varying mutable references). Additionally, in FrTime behaviors and events can be turned into one another through the use of primitives such as *hold* and *changes*. The former takes as input an initial value and an event stream, and returns a behavior that starts with the initial value and changes to the last event value every time a new event occurs. The latter consumes a behavior and returns an event stream that emits the value of the behavior whenever it changes. The problem of integrating FrTime and object-oriented graphics toolkits has also been the object of research [35]. Recently, Meyerovich *et al.* [40] have introduced Flapjax, a reactive extension to the JavaScript language targeted at Web applications, whose approach is mainly informed by FrTime. Frappé [16] integrates the FRP model with the Java Beans technology, allowing reactive programming in Java.

SugarCubes [12] and ReactiveML [39] allow reactive programming (in Java and OCAML, respectively) by relying not on operating system and runtime support, as our approach does, but rather on causality analysis and a custom

interpreter/compiler. Both systems, however, track dependencies between functional units, through the use of specific language constructs, such as *events*, and explicit commands for generating and waiting for events.

Several other systems with reactive capabilities have been proposed in recent years, especially in the field of Web programming. For a more comprehensive discussion on these systems, we refer the interested reader to the paper by Meyerovich *et al.* [40] and the references therein.

Constraint Programming. Dataflow constraints fit within the more general field of constraint programming [7]. Terms such as “constraint propagation” and “constraint solving” have often been used in papers related to dataflow since the early developments of the area [11, 41, 46]. However, the techniques developed so far in dataflow programming are quite distant from those appearing in the constraint programming literature [9]. In constraint programming, relations between variables can be stated in the form of multi-way constraints, typically specified over restricted domains such as real numbers, integers, or Booleans. Domain-specific solvers use knowledge of the domain in order to forbid explicitly values or combinations of values for some variables [9], while dataflow constraint solvers are domain-independent.

Kaleidoscope [27] integrates imperative object-oriented programming with constraint programming, by allowing the specification of multi-way constraints between user defined objects. Constraints are asserted by statements containing a duration keyword (e.g., *once*, *always*), an optional constraint strength parameter (e.g., *weak*, *strong*, *required*), effectively supporting constraint hierarchies, and an arbitrary object-oriented expression. Differently from our approach, constraint satisfaction is guaranteed by a specialized compiler/interpreter pair.

In the field of Constraint Logic Programming, *attributed variables*, a new data type that associates variables with arbitrary attributes, have proven to be a powerful mechanism for extending logic programming systems with the ability of constraint solving [31, 32]. Attributes may represent user-defined constraints, and extensible unifications are supported: when an attributed variable is to be unified with a term (possibly another attributed variable), a user-defined unification handler is invoked to process the objects involved and change the variables’ attributes. While being a relatively low-level construct for constraint programming, attributed variables have also served as a basis for Constraint Handling Rules [32], a high level declarative language to write constraint solvers. Recently, Action Rules [48] has extended logic programming languages, such as Prolog, with the concept of *agents* that can be used for both event-handling and constraint propagation. Agents carry out specific actions and are activated when certain events are posted.

Algorithms for Constraint Satisfaction. Moving from early work on attribute grammars [19, 37], a variety of incremental algorithms for performing efficient dataflow con-

straint satisfaction have been proposed in the literature and integrated in dataflow systems such as Amulet. These algorithms are based either on a mark-sweep approach [19, 34], or on a topological ordering [6]. In contrast, DC uses a priority-based approach, which allows users to customize the constraint scheduling order. Mark-sweep algorithms are preferable when the dataflow graph can change dynamically during constraint evaluation: this may happen if constraints use indirection and conditionals, and thus cannot be statically analyzed. With both approaches, if there are cyclic dependencies between constraints, they are arbitrarily broken, paying attention to evaluate each constraint in a cycle at most once. Compared to our iterative approach, this limits the expressive power of constraints.

Self-adjusting Computation. A final related area, that we have extensively discussed throughout the paper, is that of self-adjusting computation, in which programs respond to input changes by updating automatically their output. This is achieved by recording data and control dependencies during the execution of programs so that a change propagation algorithm can update the computation as if the program were run from scratch, but executing only those parts of the computation affected by changes. We refer to [3–5, 30] for recent progress in this field.

8. Future Work

The work presented in this paper paves the road to several further developments. Although conventional platforms offer limited support for implementing reactive memory efficiently, we believe that our approach can greatly benefit from advances in the hot field of transactional memories, which shares with us the same fundamental need for a fine-grained, highly-efficient control over memory accesses. Multi-core platforms suggest another interesting direction. Indeed, exposing parallelism was one of the motivations for dataflow architectures, since the early developments of the area. We regard it as a challenging goal to design effective models and efficient implementations of one-way dataflow constraints in multi-core environments.

Acknowledgments

We wish to thank Umut Acar and Matthew Hammer for many enlightening discussions and for their support with CEAL. We are also indebted to Alessandro Macchioni for his contributions to the implementation of reactive memory, and to Pietro Cenciarelli and Ivano Salvo for providing useful feedback on the formal aspects of our work.

This work was supported in part by the Italian Ministry of Education, University, and Research (MIUR) under PRIN 2008TFBWL4 national research project “AlgoDEEP: Algorithmic Challenges for Data-Intensive Processing on Emerging Computing Platforms”.

References

- [1] M. Abadi, T. Harris, and M. Mehrara. Transactional Memory with Strong Atomicity Using Off-the-Shelf Memory Protection Hardware. In *PPoPP*, pages 185–196, 2009.
- [2] R. Abraham, M. M. Burnett, and M. Erwig. Spreadsheet Programming. In *Wiley Encyclopedia of Computer Science and Engineering*. John Wiley & Sons, Inc., 2008.
- [3] U. A. Acar. Self-Adjusting Computation: (an Overview). In *PEPM*, pages 1–6, 2009.
- [4] U. A. Acar, G. E. Blelloch, M. Blume, and K. Tangwongsan. An Experimental Analysis of Self-Adjusting Computation. In *PLDI*, pages 96–107, 2006.
- [5] U. A. Acar, G. E. Blelloch, R. Ley-Wild, K. Tangwongsan, and D. Türkoglu. Traceable Data Types for Self-Adjusting Computation. In *PLDI*, pages 483–496, 2010.
- [6] B. Alpern, R. Hoover, B. K. Rosen, P. F. Sweeney, and F. K. Zadeck. Incremental Evaluation of Computational Circuits. In *SODA*, pages 32–42, 1990.
- [7] K. R. Apt. *Principles of Constraint Programming*. Cambridge University Press, 2003.
- [8] R. Bellmann. On a Routing Problem. *Quarterly of Applied Mathematics*, 16:87–90, 1958.
- [9] C. Bessiere. Constraint Propagation. In F. Rossi, P. van Beek, and T. Walsh, editors, *Handbook of Constraint Programming*. 2006.
- [10] P. A. Blume. *The LabVIEW Style Book*. Prentice Hall, 2007.
- [11] A. Borning. The Programming Language Aspects of ThingLab, a Constraint-Oriented Simulation Laboratory. *ACM Transactions on Programming Languages and Systems*, 3(4):353–387, 1981.
- [12] F. Boussinot and J.-F. Susini. The SugarCubes Tool Box: a Reactive Java Framework. *Software: Practice and Experience*, 28(14):1531–1550, 1998.
- [13] P. Caspi, P. Pilaud, N. Halbwachs, and J. Plaice. Lustre, a Declarative Language for Programming Synchronous Systems. In *POPL*, pages 178–188, 1987.
- [14] C. Chambers, B. Harrison, and J. Vlissides. A Debate on Language and Tool Support for Design Patterns. In *POPL*, pages 277–289, 2000.
- [15] G. H. Cooper and S. Krishnamurthi. Embedding Dynamic Dataflow in a Call-by-Value Language. In *ESOP*, pages 294–308, 2006.
- [16] A. Courtney. Frappé: Functional Reactive Programming in Java. In *PADL*, pages 29–44, 2001.
- [17] P. Cousot and R. Cousot. Abstract Interpretation: A Unified Lattice Model for Static Analysis of Programs by Construction or Approximation of Fixpoints. In *POPL*, pages 238–252, 1977.
- [18] P. Crescenzi, C. Demetrescu, I. Finocchi, and R. Petreschi. Reversible Execution and Visualization of Programs with Leonardo. *Journal of Visual Languages and Computing*, 11(2):125–150, 2000.
- [19] A. J. Demers, T. W. Reps, and T. Teitelbaum. Incremental Evaluation for Attribute Grammars with Application to Syntax-Directed Editors. In *POPL*, pages 105–116, 1981.

- [20] C. Demetrescu. *Fully Dynamic Algorithms for Path Problems on Directed Graphs*. PhD thesis, Sapienza University of Rome, 2001.
- [21] C. Demetrescu, I. Finocchi, and G. Italiano. *Handbook on Data Structures and Applications*, chapter 36: Dynamic Graphs. D. Mehta and S. Sahni (eds.), CRC Press, 2005.
- [22] C. Demetrescu, A. V. Goldberg, and D. S. Johnson, editors. *The Shortest Path Problem: Ninth DIMACS Implementation Challenge*. American Mathematical Society, 2009.
- [23] C. Demetrescu, I. Finocchi, and A. Ribichini. Reactive Imperative Programming with Dataflow Constraints. Technical Report arXiv:1104.2293, April 2011.
- [24] B. Demsky and M. Rinard. Automatic Detection and Repair of Errors in Data Structures. In *OOPSLA*, pages 78–95, 2003.
- [25] C. Elliott and P. Hudak. Functional Reactive Animation. In *ICFP*, pages 263–273, 1997.
- [26] A. Ezust and P. Ezust. *An Introduction to Design Patterns in C++ with Qt 4*. Prentice Hall, 2006.
- [27] B. Freeman-Benson and A. Borning. Integrating Constraints with an Object-Oriented Language. In *ECOOP*, pages 268–286, 1992.
- [28] T. Groetker, S. Liao, G. Martin, and S. Swan. *System Design with SystemC*. Kluwer Academic Publishers, 2002.
- [29] P. L. Guernic, A. Benveniste, P. Bournaï, and T. Gautier. SIGNAL - A Data Flow-Oriented Language for Signal Processing. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 34(2):362–374, 1986.
- [30] M. Hammer, U. A. Acar, and Y. Chen. CEAL: a C-based Language for Self-Adjusting Computation. In *PLDI*, pages 25–37, 2009.
- [31] C. Holzbaur. Metastructures vs. Attributed Variables in the Context of Extensible Unification - Applied for the Implementation of CLP Languages. In *PLILP*, pages 260–268, 1992.
- [32] C. Holzbaur and T. Frühwirth. Compiling Constraint Handling Rules into Prolog with Attributed Variables. In *PPDP*, pages 117–133, 1999.
- [33] P. Hudak, S. Peyton Jones, and P. Wadler (editors). Report on the Programming Language Haskell, A Non-strict Purely Functional Language (Version 1.2). *ACM SIGPLAN Notices*, 27(5), 1992.
- [34] S. E. Hudson. Incremental Attribute Evaluation: A Flexible Algorithm for Lazy Update. *ACM Transactions on Programming Languages and Systems*, 13(3):315–341, 1991.
- [35] D. Ignatoff, G. H. Cooper, and S. Krishnamurthi. Crossing State Lines: Adapting Object-Oriented Frameworks to Functional Reactive Languages. In *FLOPS*, pages 259–276, 2006.
- [36] A. Kay. Computer software. *Scientific American*, 251(3):191–207, 1984.
- [37] D. E. Knuth. Semantics of Context-free Languages. *Theory of Computing Systems*, 2(2):127–145, 1968.
- [38] M. Z. Malik, K. Ghori, B. Elkarablieh, and S. Khurshid. A Case for Automated Debugging Using Data Structure Repair. In *ASE*, pages 620–624, 2009.
- [39] L. Mandel and M. Pouzet. ReactiveML, a Reactive Extension to ML. In *PPDP*, pages 82–93, 2005.
- [40] L. A. Meyerovich, A. Guha, J. Baskin, G. H. Cooper, M. Greenberg, A. Bromfield, and S. Krishnamurthi. Flapjax: a Programming Language for Ajax Applications. In *OOPSLA*, pages 1–20, 2009.
- [41] B. A. Myers, R. G. McDaniel, R. C. Miller, A. S. Ferency, A. Faulring, B. D. Kyle, A. Mickish, A. Klimovitski, and P. Doane. The Amulet Environment: New Models for Effective User Interface Software Development. *IEEE Transactions on Software Engineering*, 23(6):347–365, 1997.
- [42] H. Nilsson, A. Courtney, and J. Peterson. Functional Reactive Programming, Continued. In *HASKELL*, pages 51–64, 2002.
- [43] S. Prasad and S. Arun-Kumar. An Introduction to Operational Semantics. In *Compiler Design Handbook: Optimizations and Machine Code*, pages 841–890. CRC Press, Boca Raton, 2002.
- [44] G. Ramalingam and T. Reps. An Incremental Algorithm for a Generalization of the Shortest-Path Problem. *Journal of Algorithms*, 21(2):267 – 305, 1996.
- [45] U.S. Census Bureau, Washington, DC. UA Census 2000 TIGER/Line Files. <http://www.census.gov/geo/www/tiger/>, 2002.
- [46] B. T. Vander Zanden, R. Halterman, B. A. Myers, R. McDaniel, R. Miller, P. Szekely, D. A. Giuse, and D. Kosbie. Lessons Learned about One-Way, Dataflow Constraints in the Garnet and Amulet Graphical Toolkits. *ACM Transactions on Programming Languages and Systems*, 23(6):776–796, 2001.
- [47] Z. Wan and P. Hudak. Functional Reactive Programming from First Principles. In *PLDI*, pages 242–252, 2000.
- [48] N.-F. Zhou. Programming Finite-Domain Constraint Propagators in Action Rules. *Theory and Practice of Logic Programming*, 6:483–507, 2006.