# Reactive Traversal of Recursive Data Types

Francisco Sant'Anna
Departamento de Informática
— PUC-Rio, Brazil
fsantanna@inf.puc-rio.br

Hisham Muhammad
Departamento de Informática
— PUC-Rio, Brazil
hisham@inf.puc-rio.br

Johnicholas Hines
IDEXX Laboratories
johnicholas.hines@gmail.com

## ABSTRACT

We propose a structured mechanism to traverse recursive data types incrementally, in successive reactions to external input events. `traverse` is an iterator-like anonymous block that can be invoked recursively and suspended at any point, retaining the full state and stack frames alive. `traverse` is designed for the synchronous language CÉU, inheriting all of its concurrency functionality and safety properties, such as parallel compositions with orthogonal abortion, static memory management, and bounded reaction time and memory usage. We discuss three applications in the domain of control-oriented DSLs that contain reactive and recursive behavior at the same time.

## Categories and Subject Descriptors

D.3.3 [**Programming Languages**]: Language Constructs and Features

## General Terms

Design, Languages

## Keywords

Reactive Programming, Recursive Data Types, Structured Programming, Behavior Trees, Domain Specific Languages, Incremental Computation, Logo

## 1. INTRODUCTION

The facilities a given language offers for constructing data types have a direct impact on the nature of algorithms that programmers will write on that language. As an example, the aim for referential transparency in functional languages enforces data structures to be immutable. Under these constraints, one must avoid excessive memory copying through specialized algorithms [14].

In this paper, we discuss the design of recursive data types and an associated control facility for a language developed under a different set of constraints. CÉU [17, 18]

```
1   input void RESET;   // declares an external event
2   var int v = 0;      // variable shared by the trails
3   par do
4      loop do          // 1st trail
5         await 1s;
6         v = v + 1;
7         _printf("v = %d\n", v);
8      end
9   with
10     loop do          // 2nd trail
11        await RESET;
12        v = 0;
13     end
14  end
```

**Figure 1: Introductory example in Céu.**

is an imperative, concurrent and reactive language in which lines of execution, known as *trails*, react together continuously and in synchronous steps to external stimuli. CÉU supports mutable data, with static memory and safe pointer manipulation. However, these features are incompatible with garbage-collected immutable data structures, as well as with general records with arbitrary pointers such as *structs* in C.

The solution to this problem is twofold, with data and control aspects. For data management, we introduce a restricted form of recursive data types that can describe trees (but not general graphs). To control reactive behavior, we propose a structured mechanism that can traverse data types safely and incrementally, in successive reactions to events. After we present the design of these constructs, we discuss three applications in the domain of control-oriented DSLs.

## 2. CÉU CONSTRUCTS

The introductory example[1] in Figure 1 gives a general flavor of CÉU. It first defines an input event RESET (line 1), a shared variable v (ln. 2), and starts two trails with the `par` construct (ln. 3-14): the first (ln. 4-8) increments variable v on every second and prints its value on screen; the second (ln. 10-13) resets v on every external request to RESET.

In the synchronous model of CÉU, a program reacts to an occurring event completely before handling the next. A reaction represents a logical instant in which all trails awaiting the occurring event awake and execute *atomically*, one after the other, until they await again or terminate. As a consequence, all consecutive operations to shared variable v in Figure 1 are atomic because reactions to events 1s and RESET can never interrupt each other. If multiple trails awake

---

[1]A screencast of all examples in the paper is available at https://vimeo.com/135297440.

to the same event, the scheduler employs lexical order, i.e., the trail that appears first in the source code executes first. For this reason, programs are deterministic even in the presence of side effects in concurrent lines of execution. To avoid infinite execution for reactions, Céu ensures that all loops contain `await` statements [17].

Céu relies on a source-to-source compiler that generates single-threaded code in *C*. The generated code has a negligible overhead in terms of memory and CPU usage in comparison to handcrafted event-driven code written directly in *C* [17].

## 2.1 Recursive Data Types

The `data` construct in Céu provides a safer alternative to C's `struct`, `union`, and `enum` definitions. All allocations are made in terms of a memory pool, and instances must form a tree structure with respect to the pool's *root element*.

Figure 2 illustrates the recursive `List` data type, declared as a tagged union (ln. 1–5). The first tag of a recursive type has a special meaning and represents the union's *null tag*. In the example, the tag `NIL` also represents an empty list (ln. 2). The second tag, `CONS` (ln. 4), represents list nodes.

In the first block of the example (ln. 7–16), we declare a `pool` of `List` objects of size 1 (ln. 8). All pools have static memory management based on their enclosing lexical scope (ln. 7–16). The pool variable represents the root element, and is implicitly initialized to the *null* tag of its data type, i.e., `lst1` receives `List.NIL` (ln. 8). Then, we use the `=new` construct which performs allocation and assignment at the same time (ln. 9–12): it attempts to allocate a list of three elements (10, 20, and 30), and assigns the result to `lst1`, which is also the destination memory pool for the allocation (inferred from the *l-value* in the assignment). Since the pool has size 1, only the allocation of first element succeeds, with the failed subtree allocation returning the *null* tag (i.e., `List.NIL`). The print command (ln. 13–14) outputs "10, 1": the `head` of the first element, and *true* in the `NIL` check for the second element. Finally, the end of the block (ln. 16) deallocates the pool along with all elements inside it.

In the second block (ln. 18–24), we declare the `lst2` pool with an unbounded memory limit (i.e., `List[]` in line 19). Now, all allocations succeed (ln. 20).[2] Then, we mutate the tail of the first element to point to a newly allocated element in the same pool, which also succeeds (ln. 21). At the moment of the mutation, the old subtree (containing values "20" and "30") is completely removed from memory. The print command (ln. 22) outputs "50", displaying the `head` of the new second element. Again, the end of the block (ln. 24) deallocates the pool along with all of its remaining elements.

In Céu, recursive data types impose some restrictions. Elements in different pools cannot be mixed without copying. Types may contain arbitrary weak pointers, but weak pointers to Céu structures must be observed via the `watching` construct, as they can be invalidated at any time (to be discussed in Section 2.2).

## 2.2 Traversing Data Types

Céu introduces the `traverse` structured mechanism for traversing recursive data types incrementally. The code in

---

[2]To save space, in the next examples we omit the data type prefix in tags (e.g., `List.CONS` becomes `CONS`).

```
1   data List with
2       tag NIL ();
3   or
4       tag CONS (int head, List tail);
5   end
6
7   do
8       pool List[1] lst1;
9       lst1 = new List.CONS(10,
10                  List.CONS(20,
11                   List.CONS(30,
12                    List.NIL())));
13      _printf("%d, %d\n", lst1.CONS.head,
14                          lst1.CONS.tail.NIL);
15          // prints 10, 1
16  end
17
18  do
19      pool List[] lst2;
20      lst2 = new CONS(10, CONS(20, CONS(30, NIL())));
21      lst2.CONS.tail = new CONS(50, NIL());
22      _printf("%d\n", lst2.CONS.tail.CONS.head);
23          // prints 50 (20 and 30 have been freed)
24  end
```

**Figure 2: A recursive `List` data type definition (ln. 1–5) with uses (ln. 7–16 and 18–24).**

```
1   pool List[3] lst = <...>;  // [10, 20, 30]
2
3   var int sum =
4       traverse e in lst do
5           if e:NIL then
6               escape 0;
7           else
8               var int sum_tail = traverse e:CONS.tail;
9               escape sum_tail + e:CONS.head;
10          end
11      end;
12
13  _printf("sum = %d\n", sum); // prints 60
```

**Figure 3: Calculating the *sum* of a list.**

Figure 3 creates a list (ln. 1) and traverses it to calculate the sum of elements (ln. 3–11). The `traverse` block (ln. 4–11) starts with the element `e` pointing to the root of the list `lst`. The `escape` statements (ln. 6,9) return a value to the enclosing assignment to `sum` (ln. 3). The sum of a `NIL` list is `0` (ln. 5–6).[3] A `CONS` list needs to calculate the `sum` of its tail recursively, invoking `traverse` again (ln. 8), which creates a nested instance of the enclosing `traverse` block (ln. 4–11), now with `e` pointing to `e:CONS.tail`. Only after the recursive traversal of its subtree terminates the `CONS` clause adds its `head` and returns the sum (ln. 9).

When used without event-based control mechanisms, as in Figure 3, a `traverse` block is equivalent to an anonymous closure called recursively. However, `traverse` complies with the event system and memory management discipline of Céu. In fact, it is an abstraction defined in terms of a more fundamental concept, *organisms* [18], which are objects with concurrent trails of execution (akin to Simula [3]). Figure 4 depicts the expansion of the `traverse` construct along with a non-trivial example.

The example in *CODE-1* of Figure 4 (ⓐ) extends the body of the previous example in Figure 3 with reactive behavior. Now, for each recursive iteration, we print the current `head` and `await` 1 second before traversing the `tail` (ln. 10-12). Note that while nested iterations of `traverse` await

---

[3]The operator ':', as in `e:NIL`, is equivalent to C's '->'.

```
1   pool List[3] lst = <...>;  // [10, 20, 30]
2
3   par do
4     var int sum =
5       traverse e in lst do
6         if e:NIL then
7           escape 0;
8         else
9           watching e do
10            _printf("me = %d\n", e:CONS.head);
11            await 1s;
12            var int sum_tail = traverse e:CONS.tail;
13            escape sum_tail + e:CONS.head;
14          end
15          escape 0;
16        end
17      end;
18    _printf("sum = %d\n", sum);
19        // prints 60 (with no mutations)
20  with
21    <...>
22      lst.CONS.tail = NIL();
23        // possible concurrent mutation
24    <...>
25  end
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40  .
```

Ⓒ₁ **CODE-1: Original code (with `traverse`)**

```
1   pool List[3] lst = <...>;  // [10, 20, 30]
2
3   par do
4     class Frame (Frame[3]& frames,
5                  _Dummy*    parent,
6                  List[3]*   e)
7     do
8       watching *this.parent do
9         if e:NIL then
10          escape 0;
11        else
12          watching e do
13            _printf("me = %d\n", e:CONS.head);
14            await 1s;
15            do
16              var _Dummy scope;
17              var Frame* frame = spawn Frame(this.frames,
18                                      &scope, e:CONS.tail)
19                                  in this.frames;
20              var int sum_tail = await *frame;
21            end
22            escape sum_tail + e:CONS.head;
23          end
24          escape 0;
25        end
26      end
27      escape 0;
28    end
29    pool Frame[3] frames;
30    var _Dummy scope;
31    var Frame* frame = spawn Frame(frames, &scope, lst)
32                        in frames;
33    var int sum = await *frame;
34    _printf("sum = %d\n", sum); // prints 60
35  with
36    <...>
37      lst.CONS.tail = NIL();
38        // possible concurrent mutation
39    <...>
40  end
```

Ⓒ₂ **CODE-2: Expanded code (without `traverse`)**

**Figure 4: Calculating the *sum* of a list, one element each second. The `traverse` construct is a syntactic sugar that can be "desugared" with explicit organisms.**

1 second, all previous iterations are blocked, retaining their full state of execution. Note also that a parallel trail (ln. 21–24) remains reactive in the meantime, and can mutate the list at any time. We handle the issue of parallel mutations using the `watching` construct. The `watching` block aborts if the object being watched is released from memory [18]. Using `watching e` (ln. 9–14) we ensure that if element `e` is released from memory during `await` (ln. 11) due to a mutation in the list (ln. 22), we simply ignore the whole subtree and return 0 (ln. 15). CÉU enforces at compile time that all accesses to pointers that cross `await` statements are protected with an enclosing `watching` block. This ensures that if concurrent side effects affect the pointed object, no code uses the stale pointer, because the whole block is aborted.

*CODE-2* in Figure 4 (Ⓒ₂) is the equivalent expansion of Ⓒ₁ without the `traverse` construct. Because it contains concurrency constructs (i.e., `await` and `watching`), the body of the `traverse` (Ⓒ₁: 6–16) is abstracted in an organism of the `Frame` class (Ⓒ₂: 4–26), which is analogous to "stack frames" for subroutines in standard programming languages. In CÉU, organisms also reside in lexically-scoped memory pools [18], hence, the `pool` of frames (Ⓒ₂: 29) is analogous to a runtime "call stack". We limit the number of stack frames to match the exact maximum number of elements to traverse (Ⓒ₁: 1 and Ⓒ₂: 4, 29). To "call" the first `traverse` iteration,

we dynamically `spawn` a `Frame` instance into the `frames` pool (Ⓒ₂: 31–32), which runs implicitly in parallel with the rest of the application (e.g., Ⓒ₂: 36–39). Then, we immediately `await` the termination of this `frame` (Ⓒ₂: 33). Only after the whole traversal finishes that we acquire the `sum` and print it (Ⓒ₂: 33–34).

A `Frame` receives three arguments in the constructor (Ⓒ₂: 4–6): a reference to a `pool` (to recursively `spawn` new frames); a pointer to its parent scope (to handle abortion); and a pointer to the subtree of the data type (to be able to manipulate it). The `Frame` constructor for the first call (Ⓒ₂: 31–32) receives the static pool of frames, a dummy `scope` organism attached to the current scope, and the original tree to traverse (Ⓒ₁: 5). The `Frame` constructor for recursive calls (Ⓒ₂: 17–19) receives the same pool of frames, another dummy `scope` as parent, and the subtree for the recursive invocation (Ⓒ₂: 12). The `Frame` body (Ⓒ₂: 8–27) watches the dummy organism attached to its `parent` scope and aborts if the parent terminates. Watching an object in the parent scope guarantees that the hierarchy in the call stack is preserved, i.e., that there are no orphan frames executing. The remaining code is almost the same in the original `traverse` body and in the `Frame` body (Ⓒ₁: 6–16 and Ⓒ₂: 9–25), with the exception of the recursive invocation explained above (Ⓒ₂: 16–19).

As the expansion illustrates, three aspects make `traverse`

fundamentally different from recursive function calls:

1. Each `traverse` invocation spawns a new organism for the frame which executes concurrently with other parts of the application. Also, each frame itself can execute arbitrary code in Céu with nested concurrent trails (to be illustrated in Section 3).

2. A `traverse` is attached to a lexically-scoped memory pool for a specific data structure. Therefore, we can infer at compile time the maximum traversal depth if the data is bounded (e.g., `List[3] lst`). Enforcing bounded limits is an important requirement for constrained and real-time embedded systems, which is the original application domain of Céu [17]. In addition, when a recursive data structured goes out of scope, all associated stack frames are automatically aborted, given that the pool of frames expands to a narrower lexical scope (e.g., ⓒ: 1 *vs.* 29) [18].

3. The execution body of a `traverse` block is implicitly wrapped by a concurrency construct that watches for mutations of the current node. In practice, this means that it reacts consistently if another trail of execution modifies the data structure being traversed (e.g., ⓒ₁: 22).

We believe that the `traverse` construct, more than a simple convenience, reduces the complexity of programs considerably, handling hierarchy of behaviors associated with recursive data types automatically.

## 3. APPLICATIONS

In this section, we present three applications that explore the reactive nature of the `traverse` construct. We start with *Behavior Trees* used in video games for AI modeling. Then, we show a *Logo Turtle* that can execute commands in parallel. Finally, we extend the Turtle example with a dynamic and concurrent queue of commands that can affect the running program.

### 3.1 Behavior Trees

*Behavior Trees* are a family of DSLs used for game AI [10, 7]. The DSLs vary between languages, but they usually include sequence (`SEQ`) and selection (`SEL`) combinators to model concurrent creature behavior. The `SEQ` can be understood as short-circuit evaluation of an 'and', while the `SEL` corresponds to an 'or'. This skeleton is extensible with leaves to test and set properties, perform animations and sounds, etc., and is an effective alternative to finite state machines for authoring game AI.

However, because the evaluation of trees extends across multiple game frames, specifying node behaviors in generic languages via event-driven programming becomes a challenge due to "stack ripping" [11]. By lowering the barrier to writing custom nodes and leaves, Céu lightweight event control mechanisms make behavior trees more usable.

Figure 5 describes a generic grammar for behavior trees (ln. 1–9). The `SEQ` and `SEL` tags (ln. 4,6) are recursive and behave as described above. The `LEAF` tag (ln. 8) receives a reference to an opaque `Leaf` data type, which is defined externally and is specific to the application domain. The interpreter for behavior trees is abstracted in a class definition (ln. 11–35) and receives the tree to traverse as the single argument (ln. 11). The body acquires the return status of the traversal (ln. 12) and returns it as the final result (ln. 34). For the `SEQ` tag (ln. 14–20), we traverse the

```
1   data BTree with
2      tag NIL ();
3   or
4      tag SEQ (BTree first, BTree second);
5   or
6      tag SEL (BTree first, BTree second);
7   or
8      tag LEAF (Leaf& leaf);
9   end
10
11  class BTreeInterpreter (BTree[]& btree) do
12     var int ret =
13        traverse t in btree do
14           if t:SEQ then
15              var int ok = traverse t:SEQ.first;
16              if ok == 0 then
17                 escape ok;
18              end
19              ok = traverse t:SEQ.second;
20              escape ok;
21           else/if t:SEL then
22              var int ok = traverse t:SEL.first;
23              if ok != 0 then
24                 escape ok;
25              end
26              ok = traverse t:SEL.second;
27              escape ok;
28           else/if t:LEAF then
29              var int ret =
30                 do LeafHandler(t:LEAF.leaf);
31              escape ret;
32           end
33        end;
34     escape ret;
35  end
```

**Figure 5: A simple grammar of behavior trees with `SEQ` and `SEL` nodes and a straightforward interpreter.**

`first` subtree (ln. 15) and only if it succeeds, we traverse the `second` subtree (ln. 19). For the `SEL` tag (ln. 21–27), we traverse the `first` subtree (ln. 22) and only if it fails, we traverse the `second` subtree (ln. 26). Finally, the `LEAF` tag (ln. 28–31) delegates the behavior to another class, which does real work and is domain specific. The `do Class` syntax (ln. 29–30) creates an anonymous and lexically scoped organism and awaits its termination to return the final status (ln. 31). The organism itself can contain any valid code in Céu (including parallel compositions) and executes for an arbitrary amount of time [18].

```
1   pool BTree[] btree =
2      new SEQ(SEQ,
3         SEL(
4            LEAF(SENSE_ON_TABLE(C)),
5            LEAF(MOVE_BLOCK_TO_TABLE(C))),
6         LEAF(MOVE_BLOCK_TO_BLOCK(B, A)),
7         LEAF(MOVE_BLOCK_TO_BLOCK(C, B)));
```
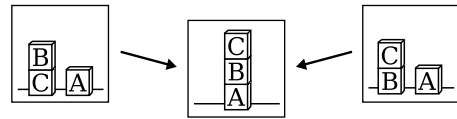


**Figure 6: A *blocks world* behavior tree.**

As an example of a domain, the *blocks world* is a classical planning domain in AI [19]. The tree in Figure 6 is based on the output from a Contingent-FF benchmark that extends the blocks domain with sensor actions [8]. We want to achieve an `ABC` stack and assume two decision possibili-

ties, as illustrated in the figure. We use a `SEL` node (ln. 3) with a sensor leaf (ln. 4) to decide which strategy is appropriate: If `C` is not sensed on top of the table, we first move it to the table (ln. 5). Then, in both situations, we stack `B` on top of `A`, and `C` on top of `B` (ln. 6–7). The example illustrates how the behavior tree can exhibit goal-directed behavior specified directly by domain designers.

## 3.2 Logo Turtle

Our second example is an interpreter for a simple variant of the classic Logo turtle-graphics interpreter [15]. The aim of this example is to demonstrate parallel traversal: we can instruct the turtle to move and rotate at the same time to trace curves.

Figure 7 presents the `Command` data type (ln. 1–15), which specifies the abstract syntax of our Logo variant. As in traditional Logo, commands can execute in sequence through the `SEQ` tag (ln. 4), and can also repeat a number of times through the `REPEAT` tag (ln. 6). We extend the `MOVE` and `ROTATE` commands to take as arguments the speed at which they should affect the turtle (ln. 8,10). For example, a `Command.MOVE(300)` node directs the turtle to move at the speed of 300 pixels per second, indefinitely. Therefore, the only way to make the turtle stop moving or rotating is through two CÉU-like extensions: The `AWAIT` tag (ln. 12) simply awaits a given number of milliseconds. The `PAROR` tag (ln. 14) launches two commands in parallel, and aborts both as soon as one of them terminates. As an example, the program in lines 54–56 makes the turtle to move along a semicircle.

The interpreter for the commands is also abstracted in a class definition (ln. 17–51). It holds as attributes a reference to a `Turtle` object (which implements the UI) and a reference to the commands (ln. 17–18). The execution body of the class uses the `traverse` construct to interpret the commands (ln. 20–50). The `SEQ` tag (ln. 22–24) traverses each of its subcommands in sequence (in contrast with `BTreeInterpreter`, it does not handle failures). The `REPEAT` tag (ln. 26–29) traverses its subcommand the specified number of times. The `MOVE` and `ROTATE` tags (ln. 31–37) rely on predefined classes of organisms to update the position and orientation of the received `turtle`. The `AWAIT` tag (ln. 39–40) simply causes the current iteration of the `traverse` to await the given amount of time. Finally, the `PAROR` tag (ln. 42–47) uses the `par/or` construct to traverse both subcommands at the same time: as soon as one of the subtrees terminates its execution, the other one is safely aborted [17].

Note that the entire interpreter block is surrounded by a `watching` construct (ln. 22). As discussed in Section 2.2, the CÉU compiler enforces the presence of the guard, due to the use of the `cmd` pointer in code that spans multiple reactions. This ensures clean abortion in case the AST is mutated by code running in other trails.

## 3.3 Enqueuing Commands

All examples so far create a fixed tree that does not vary during traversal. Figure 8 extends the Turtle application with a queue of pending commands to execute.

We define a new `Queue` data type in *CODE-3*: The `ROOT` tag (ln. 4–6) has a `running` subtree of commands, a `waiting` queue of pending commands, and a `tmp` node to allow in-place manipulation of the tree (to be discussed further). The `ITEM` tag (ln. 8–9) represents a queue item and contains a `cmds` subtree with the actual commands to execute (as de-

```
1   data Command with
2       tag NOTHING ();
3   or
4       tag SEQ (Command first, Command second);
5   or
6       tag REPEAT (int times, Command command);
7   or
8       tag MOVE (int pixels);
9   or
10      tag ROTATE (int angle);
11  or
12      tag AWAIT (int ms);
13  or
14      tag PAROR (Command first, Command second);
15  end
16
17  class CommandInterpreter (Turtle&    turtle,
18                            Command[]& cmds)
19  do
20      traverse cmd in cmds do
21          watching cmd do
22              if cmd:SEQ then
23                  traverse cmd:SEQ.first;
24                  traverse cmd:SEQ.second;
25
26              else/if cmd:REPEAT then
27                  loop i in cmd:REPEAT.times do
28                      traverse cmd:REPEAT.command;
29                  end
30
31              else/if cmd:MOVE then
32                  do TurtleMove(turtle,
33                                cmd:MOVE.pixels);
34
35              else/if cmd:ROTATE then
36                  do TurtleRotate(turtle,
37                                  cmd:ROTATE.angle);
38
39              else/if cmd:AWAIT then
40                  await (cmd:AWAIT.ms) ms;
41
42              else/if cmd:PAROR then
43                  par/or do
44                      traverse cmd:PAROR.first;
45                  with
46                      traverse cmd:PAROR.second;
47                  end
48              end
49          end
50      end
51  end
52
53  pool Command[] cmds =
54      new PAROR(
55          AWAIT(1000),
56          PAROR(MOVE(300), ROTATE(180)));
57
58  var Turtle turtle;
59  do CommandInterpreter(turtle, cmds);
```

**Figure 7: Grammar, interpreter, and sample program for a Logo Turtle DSL.**

scribed in Figure 7: 1–15), and a `prv` queue item pointing to an older item (i.e., the queue is in reverse order to facilitate insertion). As Figure 9 illustrates in box `0`, a queue instance has a single `ROOT` node with linked lists of `ITEM` nodes in the `running` and `waiting` fields. Except when creating a new command, the `tmp` field is always `NIL`.

We define the queue traversal in *CODE-4*. The `ROOT` traversal (ln. 3–14) is a continuous `loop` that executes the `running` subtree and swaps it with the `waiting` queue on termination. The `par/and` (ln. 5–9) ensures that that the swap only occurs after the current `running` commands terminate (ln. 6) *and* something (in parallel) mutates the `waiting` subtree (ln. 8), meaning that the queue is no longer empty. The swapping process (ln. 10–13) is illustrated in Figure 9:

```
data Queue with          1     traverse qu in queue do      1     input (char*,int,int) ENQUEUE;       1
  tag NIL ();            2       watching qu do            2     every (cmd,vel,time) in ENQUEUE do   2
or                       3         if qu:ROOT then          3       if _strcmp(cmd,"move")==0 then     3
  tag ROOT (Queue running, 4         loop do               4         queue.ROOT.tmp =                 4
            Queue waiting, 5           par/and do          5           new ITEM(                      5
            Queue tmp);   6             traverse qu:ROOT.running; 6             NOTHING(),              6
or                       7           with                  7             ITEM(                        7
  tag ITEM (Command cmds, 8             await qu:ROOT.waiting; 8               PAROR(                   8
            Queue prv);   9           end                  9                 MOVE(vel),               9
end                     10           qu:ROOT.running =    10                 AWAIT(time)),           10
                        11             qu:ROOT.waiting;   11               NIL()));                 11
                        12           qu:ROOT.waiting =    12       else/if _strcmp(cmd,"rotate")==0 then 12
                        13             new ITEM(NOTHING(), NIL()); 13         <...> // analogous to the MOVE above 13
                        14         end                    14       end                              14
                        15       else/if qu:ITEM then     15       queue.ROOT.tmp.ITEM.prv.ITEM.prv = 15
                        16         traverse qu:ITEM.prv;  16         queue.ROOT.waiting.ITEM.prv;    16
                        17         do CommandInterpreter( 17       queue.ROOT.waiting = queue.ROOT.tmp; 17
                        18             turtle, qu:ITEM.cmds); 18     end                              18
                        19       end                      19                                        19
                        20     end                        20                                        20
.                       21     end                        21     .                                  21
```

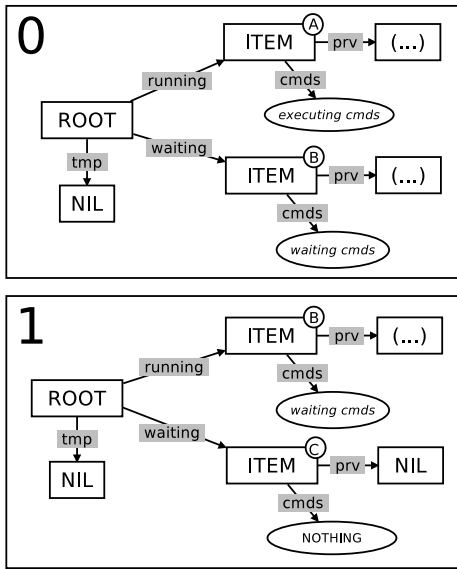| CODE-3: Queue type | CODE-4: Queue traversal | CODE-5: Enqueuing commands |
| --- | --- | --- |

Figure 8: Queue extension for the Turtle DSL of Figure 7.



Figure 9: Swapping waiting and running commands.

- The initial state (box 0) assumes pre-existing running and waiting items.
- Lines 10–11 assign the waiting subtree (marked Ⓑ) to the running field, releasing the old subtree Ⓐ. Recursive data types in CÉU have *move semantics*, hence,

the waiting field is automatically set to the union null type (i.e., Queue.NIL).
- Lines 12–13 assign a new neutral ITEM with a dummy command NOTHING (box 1: Ⓒ) to the waiting field, completing the swapping operation.

After the swapping process, the loop restarts and traverses the new running commands (ln. 4). The ITEM traversal (ln. 15–18) is straightforward: first we traverse the previous item (ln. 16), and then we reuse the CommandInterpreter class of Figure 7 to traverse the commands (ln. 17–18).

Even though this example mutates the running field only *after* its traversal terminates (ln. 10–11), it is safe to do an arbitrary mutation at any time. Note that the compiler enforces the use of watching (ln. 2) which encloses the running turtle interpreter (ln. 17–18). Hence, if its enclosing ITEM (ln. 15) is mutated, the watching will awake and abort the interpreter running inside the lexical scope.

The enqueuing of new commands is depicted in *CODE-5*. The external input event ENQUEUE (ln. 1) accepts *"move"* and *"rotate"* strings with an associated velocity and time (i.e., "char*,int,int" arguments). The every loop (ln. 2–18) reacts to each occurrence of ENQUEUE, creating and enqueuing the requested command, as illustrated in Figure 10:

- The initial state (box 0) assumes a pre-existing neutral ITEM in the root of the waiting field Ⓓ.
- Lines 4–11 assign a new subtree to the tmp field (box 1) with a new neutral ITEM (Ⓔ, l. 5–6) linked to the set of commands to MOVE the turtle (Ⓕ, l. 7–11).
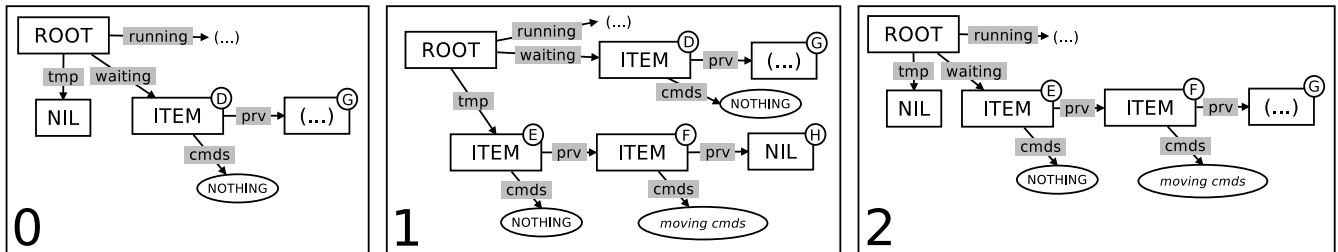


Figure 10: Enqueuing new commands.

- Lines 15–16 move the already `waiting` commands (ⓖ) to the tail of `tmp`, in the place of `NIL` (ⓗ). The old location is automatically set to `NIL`. Note that we skip the neutral `ITEM` nodes of both `waiting` and `tmp` fields. This prevents the `waiting` root from becoming `NIL` and awaking the `ROOT` node (*CODE-4*: 8) before we finish the enqueuing operation.
- Line 17 moves the `tmp` subtree (ⓔ) back to the `waiting` field, releasing the abandoned neutral `ITEM` (ⓓ), and notifying the `ROOT` node that the queue is no longer empty. The `tmp` field is automatically set to `NIL`. Note that the new `waiting` subtree preserves the neutral `ITEM` ⓔ for subsequent enqueuing operations.

## 4. RELATED WORK

Traversing data reactively in an imperative language requires dealing with concurrent updates. Some attempts make this process more transparent, but performance concerns ultimately require the programmer to specify behavior explicitly. One-way dataflow constraints were used to track updates in data structures in a reactive imperative language [4]. In it, classes are annotated with constraint handler functions which are recursively called on value updates. Another approach focuses on incremental computation and uses a combination of dynamic dependency graphs and memoization [1]. Keeping track of dependencies incurs significant overhead: traceable data types [2] mitigates this issue, letting the programmer control the granularity of tracked data. In CÉU, tracking pointer updates is explicit (and enforced), and the move semantics for recursive data fields propagates changes to tracked pointers automatically.

More universal approaches such as generators first appeared in CLU [12], but with a stackless implementation that did not support recursion. Icon [5] allows yielding through recursive functions, but delegation is explicit via the `suspend` keyword. Python originally introduced CLU-like stackless generators in version 2.2, but those were later internally promoted to coroutines in order to support recursion. Finally, in version 3.3, a form of delegation was introduced [16]. Lua [9] has first-class stackful coroutines: generator functions have to be constructed by wrapping coroutine objects, but explicit delegation in recursive calls is not required. More recently, extensions to support asynchronous concurrent updates during traversal have been proposed for Dart [13] and Scala [6]. Like our work, the Scala extension is translated into the language's core concurrency model. CÉU builds its higher-level `traverse` construct on top of organisms, which also provides cooperative multitasking, but with transparent support for recursion and concurrency. Organisms allow composing parallel traversals more naturally in comparison to coroutines, which requires a specialized scheduler and does not impose safety guarantees.

## 5. CONCLUSION

We presented a new construct for traversing recursive data types incrementally in the context of CÉU, an imperative reactive language with synchronous concurrency. Combining concurrency and safety while traversing recursive data structures is not straightforward, especially with support for mutation.

The `traverse` construct encapsulates an idiom that handles each recursive step in a separate organism (CÉU's abstraction mechanism), allowing concurrent traversal while preserving the language's safety properties. A `traverse` block is tied with its associated data structure, ensuring bounded execution time. By dealing with concurrency in a primitive control mechanism, we make reactive traversal as easy to perform correctly as a recursive function call. We impose some restrictions on the data structures that can be represented. The requirement of tree hierarchies with move semantics demands care when designing algorithms that manipulate them. In trade, these restrictions enable static memory management with deterministic deallocation. Still, we do not feel that the restrictions are prohibitively limiting. For instance, persistent data structures in functional languages operate under similar constraints.

## 6. REFERENCES

[1] U. A. Acar et al. An experimental analysis of self-adjusting computation. *ACM Trans. Program. Lang. Syst.*, 32(1):3:1–3:53, Nov. 2009.

[2] U. A. Acar et al. Traceable data types for self-adjusting computation. In *Proceedings of PLDI '10*, PLDI '10, pages 483–496. ACM, 2010.

[3] O.-J. Dahl and K. Nygaard. SIMULA: an ALGOL-based simulation language. *Communications of the ACM*, 9(9):671–678, 1966.

[4] C. Demetrescu et al. Reactive imperative programming with dataflow constraints. *ACM Trans. Program. Lang. Syst.*, 37(1):3:1–3:53, Nov. 2014.

[5] R. E. Griswold, D. R. Hanson, and J. T. Korb. Generators in Icon. *ACM Trans. Program. Lang. Syst.*, 3(2):144–161, Apr. 1981.

[6] P. Haller and H. Miller. RAY: Integrating Rx and Async for direct-style reactive streams. In *Proceedings of REM '13*, Oct. 2013.

[7] C. Hecker. My liner notes for Spore. `http://chrishecker.com/My_liner_notes_for_spore`, 2009. Accessed: 2015-07-30.

[8] J. Hoffmann and R. Brafman. Contingent planning via heuristic forward search with implicit belief states. In *Proc. ICAPS*, volume 2005, 2005.

[9] R. Ierusalimschy. *Programming in Lua, Third Edition.* Lua.Org, 3rd edition, 2013.

[10] D. Isla. Handling complexity in the Halo 2 AI. In *Game Developers Conference*, 2005.

[11] M. N. Krohn et al. Events can make sense. In *USENIX Annual Technical Conference*, pages 87–100, 2007.

[12] B. Liskov. A history of CLU. *SIGPLAN Not.*, 28(3):133–147, Mar. 1993.

[13] E. Meijer, K. Millikin, and G. Bracha. Spicing up dart with side effects. *Queue*, 13(3):40:40–40:59, Mar. 2015.

[14] C. Okasaki. *Purely Functional Data Structures.* Cambridge University Press, 1998.

[15] S. Papert. *Mindstorms: Children, Computers, and Powerful Ideas.* Basic Books, Inc., 1980.

[16] Python Software Foundation. (PEPs) 255, 342, and 380. `https://www.python.org/dev/peps/`. Accessed: 2015-08-03.

[17] F. Sant'Anna et al. Safe system-level concurrency on resource-constrained nodes. In *SenSys'13*. ACM, 2013.

[18] F. Sant'Anna et al. Structured synchronous reactive programming with Céu. In *Modularity'15*, 2015.

[19] J. Slaney and S. Thiébaux. Blocks world revisited. *Artif. Intell.*, 125(1-2):119–153, Jan. 2001.