

# A live coding environment for a data exploration language

ANONYMOUS AUTHOR(S)

A growing amount of code is written to explore and analyze data, but the way data analysts work with code differs from the way software engineers do. First, many data scientists write code that introduces few or no abstractions. Second, data exploration is an interactive process that blurs the traditional distinction between development-time and run-time. This poses an interesting challenge for programming language research, inviting us to revisit a number of traditional assumptions about programming languages and design a new kind of programming tools that are suited for data analyst needs.

We present the *data exploration calculus*, a formal language that captures the structure of scripts written to explore data. We focus on the simplest kind of data analyses such as those written by journalists to analyse open government data. We implement a programming environment, based on this calculus, that evaluates code instantly during editing and shows previews of the results, while allowing the user to modify code in an unrestricted way in a text editor. Supporting interactive editing is tricky. Any edit can change the structure of code and fully recomputing the output would be too expensive. We present a technique that allows evaluating and type-checking code while it is being written and that reuses results of previous runs. We formalize the technique using the data exploration calculus, prove that it is correct and specify when previous results are reused. Finally, we illustrate the practicality of our approach using empirical evaluation and a case study.

As data analysis becomes an ever more important use of programming, research on programming languages and tools needs to consider new kinds of programming workflows appropriate for those domains and conceive new kinds of tools that can support them. The present paper is one step in this important direction.

## 1 INTRODUCTION

One of the aspects that make spreadsheets easier to use than other programming tools is their liveness. When you change a value in a cell in Excel, the whole spreadsheet updates instantly and you immediately see new results, without having to explicitly trigger re-computation and without having to wait for an extensive period of time.

Increasing numbers of programming environments aim to provide a live development experience for standard programming languages, but doing this is not easy. Fully recomputing the whole program after every keystroke is inefficient. Calculating how a change in the source code changes the result is extremely hard when the text editor allows arbitrary changes. Consider the following snippet that gets the release years of the 10 most expensive movies from a data set `movies`:

```
let top = movies.sortBy( $\lambda x \rightarrow x.getBudget()$ )
    .take(10).map( $\lambda x \rightarrow x.getReleased().format("yyyy")$ )
```

A live coding environment computes and shows the list of years. Suppose that the programmer then modifies the code by making `10` a variable and changing the date format to see the full date:

```
let count = 10
let top = movies.sortBy( $\lambda x \rightarrow x.getBudget()$ )
    .take(count).map( $\lambda x \rightarrow x.getReleased().format("dd-mm-yyyy")$ )
```

Ideally, the live coding environment should understand the change, reuse a cached result of the first two transformations (sorting and taking first 10 elements) and only evaluate the map operation to differently format the release dates of already computed top 10 movies. This can already be done in a number of systems based on structured code editors, but our approach allows this behaviour using a standard text editor. We discuss related work in more detail in Section 8, but we review the most important directions here, in order to situate our contributions.

## 1.1 Related work

Simple data exploration performed, for example, by journalists [Gray et al. 2012] is done either programmatically or using spreadsheets. The former is reproducible and transparent, but requires expert programming skills. The latter is easy, but it limits reproducibility and is error-prone. We aim to bring some aspects of spreadsheets, especially liveness, to programmatic data exploration. Recomputation in spreadsheets is well studied, e.g. by Sestoft [2012], but extending it to code rewritten and edited in an ordinary text editor poses interesting challenges.

Many visual data exploration tools support interactivity [Crotty et al. 2015; Hellerstein et al. 1999; Wesley et al. 2011] and some can even export the workflow as a script [Kandel et al. 2011]. Data analysts who prefer working with code typically resort to notebook systems such as Jupyter or R Markdown [Allaire et al. 2016; Kluyver et al. 2016]. Those are text-based, but have a limited model of recomputation. Users structure code in cells and manually reevaluate cells. Many notebook systems do not track dependencies between cells, which can lead to inconsistencies, an issue that has been addressed by e.g. Koop and Patel [2017]; Pimentel et al. [2015] and Petricek et al. [2018].

Existing text-based editors that provide instant feedback, such as Lighttable [Granger 2012] and Chrome DevTools, often work only in certain situations or require full recomputation. A more principled approach can be used by systems based on structured editing [Lotem and Chuchem 2018; Omar et al. 2019], which benefit from the fact that code can be modified via known operations with known effect on the computation graph (e.g. “extract variable” has no effect on the result; “change constant value” forces recomputation of subsequent code). However, we aim to cater for the many users who prefer to edit programs as free-form text. The technical aspects of our technique for providing instant feedback are related to the work on incremental computation [Acar 2005; Hammer et al. 2015]. This focuses on cases where *data changes* but program stays the same, while our focus is on cases where *program changes* but data stays. Yet, the technical aspects are related and, as noted in Section 8, a language that supports incremental computation could be used to implement the kind of system presented in this paper.

## 1.2 Contributions

We present the design and implementation of a live coding environment for a data exploration language that provides correct and efficient instant feedback, yet is integrated into an ordinary text editor. Our most important contributions are:

- We introduce the *data exploration calculus* (Section 3), a small formally tractable language for data exploration. The calculus is motivated by our review of how data analysts work (Section 2) and makes it possible to write simple, transparent and reproducible data analyses.
- Implementing a live coding environment requires a different way of thinking than the one presented in classic compiler literature, because it does not operate in a batch mode. We capture the essence of the new perspective (Section 4) and use it to build our *live preview* mechanism (Section 5) that evaluates code instantly during editing.
- We prove that our live preview mechanism is correct (Section 6.1) and that it reuses previously evaluated values when possible (Section 6.2). We illustrate the practicality of the mechanism using an empirical evaluation (Section 6.3) and a case study (Section 6.4).
- We describe two extensions of our core environment. First, we add instant type checker that relies on the same core infrastructure as live previews and enables richer user experience through *type providers* (Section 7.1). Second, we add a code reuse mechanism that supports the live coding experience (Section 7.2).

## 2 UNDERSTANDING HOW DATA SCIENTISTS WORK

Data scientists often use general-purpose programming languages such as Python, but the kind of code they write and the way they interact with the system is very different from how software engineers work. The different nature of data science workflow has been documented by Guo [2013]. This paper focuses on simple data wrangling and data exploration as done, for example, by journalists analysing government datasets. This section illustrates how such data analyses typically look and provides a justification for the design of our data exploration calculus.

### 2.1 Data wrangling in Jupyter

Data analysts increasingly use notebook systems such as Jupyter, which make it possible to combine text, equations and code with results of running the code, such as tables or visualizations. Notebooks blur the distinction between development and execution. Data analysts write small snippets of code, run them to see results immediately and then revise them.

Jupyter notebooks are used by a variety of users ranging from scientists who implement complex models of physical systems to journalists who load data, perform simple aggregations and create visualizations. Our initial focus is on the simplest use cases. We believe that making programmatic data exploration more interactive and spreadsheet-like can encourage users to choose programming tools over spreadsheets, allowing them to produce reproducible and transparent data analyses.

As an example, the analysis done by the Financial Times for an article on plastic waste [Blood 2018; Hook and Reed 2018] joins datasets from Eurostat, UN Comtrade and more, aggregates the data and builds a visualization comparing waste flows in 2017 and 2018. Figure 1 shows an extract from one notebook of the data analysis. The code has a number of important properties:

#### UN Comtrade exports data

```
import pandas as pd
material = 'plastics' # 'plastics', 'paper'
```

#### Loading exports data

```
df_mat = pd.read_csv('{material}-2017.csv').fillna(0).sort_values(['country_name', 'period'])
df_mat.head()
```

	period	country_name	kg	country_code
0	2017-01-01	Algeria	43346.0	12
1	2017-03-01	Algeria	32800.0	12
2	2017-03-01	Antigua and Barbuda	17000.0	28

#### Join to country codes

```
# Set keep_default_na because the Namibia has ISO code NA
df_isos = pd.read_excel('iso.xlsx', keep_default_na=False).drop_duplicates('country_code')
df = df_mat.copy().merge(df_isos, 'left', 'country_code').rename({'iso2': 'country_code' }, axis=1)
df.head()
```

	period	country_name	kg	country_code
0	2017-01-01	Algeria	43346.0	DZ
1	2017-03-01	Algeria	32800.0	DZ
2	2017-03-01	Antigua and Barbuda	17000.0	AG

Fig. 1. An excerpt from a Jupyter notebook by Financial Times reporters, which loads trade data from the UN trade database (stored in a local file) and adds ISO country codes (from an Excel spreadsheet).

- There is no abstraction. The analysis uses lambda functions as arguments to library calls, but it does not define any top-level functions. Parameterization is done by defining a list of inputs and then having a for loop, or by defining a top-level variable and setting its value. In our example, material is set to "plastics" and a comment lists other options. This lets the analyst easily see and check the results of intermediate steps of the analysis.
- The code relies on external libraries. Our example uses Pandas [McKinney 2012], which provides operations for data wrangling such as merge to join datasets or drop\_duplicates to delete rows with duplicate column values. Those libraries are external to the data analysis and may even be implemented in other languages – many Python libraries rely on C++.
- The code is structured as a sequence of commands. Some commands define a variable, either by loading data, or by transforming data loaded previously. Even in Python, data is often treated as immutable. Other commands produce an output that is displayed in the notebook.
- There are many corner cases, such as the fact that the keep\_default\_na parameter needs to be set to handle Namibia correctly. These are discovered interactively by writing and running a code snippet and examining the output, so providing an instant feedback is essential.

Many Jupyter notebooks are more complex than the example discussed here. They might define helper functions or even structure code using object-oriented programming. However, this paper is motivated by our conviction that simple data analyses such as the one discussed here are frequent enough to pose an interesting and important domain for programming tools research.

## 2.2 Dot-driven data exploration in The Gamma

Data exploration of the kind discussed in the previous section has been the motivation for recent work by Petricek [2017b] that develops a small scripting language called The Gamma. Scripts in The Gamma are sequences of commands that can either define a variable or produce an output. The language does not support top-level function declarations and lambda functions can be used only as method arguments.

```
olympics
  .filter data'.Games is'.Rio (2016)'.then
  .group data'.by Athlete'
  .count distinct Event'
  .su
```

```
count distinct Gender
count distinct Gold
count distinct Medal
count distinct Silver
count distinct Sport
count distinct Team
count distinct Year
sum Bronze
sum Gold
sum Silver
sum Year
then
```

```
olympics
  .filter data'.Games is'.Rio (2016)'.then
  .group data'.by Athlete'
  .count distinct Event'.sum Gold'.then
  .sort data'.by Gold descending'.then
  .paging.take(5)
```

Athlete	Event	Gold
Michael Phelps	6	5
Katie Ledecky	5	4
Simone Biles	5	4
Katinka Hosszu	4	3
Usain Bolt	3	3

(a) The Gamma script to aggregate Olympic medals. The analyst selects Rio 2016 games and counts the number of distinct events per athlete. After typing '.' a type provider offers further aggregation operations.

(b) Our live coding environment for The Gamma language based on the theory presented in this paper. The table is produced as the data analyst types and updates on-the-fly to show result at the current cursor position.

Fig. 2. Previous work on The Gamma by Petricek [2017b] (left) and our new editor with live preview (right).

Given the limited expressiveness of The Gamma, libraries are implemented in other languages, such as JavaScript. The Gamma supports type providers [Syme et al. 2013], which generate types on the fly, for accessing and exploring external data sources. Type providers provide object types with members and The Gamma makes using those convenient by providing auto-complete when the user types the dot symbol (‘.’) to access a member.

As illustrated by Petricek [2017b], the combination of type providers and auto-complete on ‘.’ makes it possible to complete a large number of data exploration tasks through the very simple interaction of selecting operations from a list. An example in Figure 2a summarizes data on Olympic medals using a type provider. Identifiers such as ‘sum Bronze’ are merely names of members generated by the type provider, based on information about the data source. The type provider used in this example generates an object with members for different data transformations, such as ‘group data’, which return further objects with members for specifying transformation parameters, such as selecting the grouping key using ‘by Athlete’.

The Gamma language is richer, but the example in Figure 2a shows that non-trivial data exploration can be done using a very simple language. The assumptions about structure of code that are explicit in The Gamma are implicitly present in Python and R data analyses produced by journalists, economists and other users with other than programming background. When we refer to The Gamma in this paper, readers not familiar with it can consider a subset of Python.

### 2.3 Live coding for data exploration

The implementation that accompanies this paper builds a live coding environment for The Gamma language. The implementation is discussed in Section 6.4 and it replaces the text editor with auto-complete, originally used by The Gamma, with a live coding environment that provides *live previews*. We choose to extend The Gamma over using a subset of Python for its simplicity as it lets us focus on the principles of live coding systems.

An example of a live preview is shown in Figure 2b. As noted earlier, The Gamma programs consist of lists of commands which are either expressions or let bindings. Our editor displays a live preview below the command that the user is currently editing. The preview shows the result of evaluating the expression or the value assigned to a bound variable. When the user changes the code, the preview is updated automatically, without any additional interaction with the user.

There are a number of guiding principles that inform our design. First, we allow the analyst to edit code in an unrestricted form in a text editor. Although structured editors provide an appealing alternative and make recomputation easier, we prefer the flexibility of plain text. Second, we focus on the scenario when code changes, but input does not. Rapid feedback allows the analyst to quickly adapt code to correctly handle corner cases that typical analysis involves. In contrast to work on incremental computation, we do not consider the case when data changes, although supporting interactive data exploration of streaming data is an interesting future direction.

## 3 DATA EXPLORATION CALCULUS

The *data exploration calculus* is a small formal language for data exploration. The calculus is not, in itself, Turing-complete and it can only be used together with external libraries that define what objects are available and what the behaviour of their members is. This is sufficient to capture the simple data analyses discussed in Section 2. We define the calculus in this section and then use it as the basis for discussing our live preview mechanism in Section 4.

We initially present the data exploration calculus without a static type system as the live preview mechanism does not rely on types. We consider type checking later in Section 7.1. Interestingly, the mechanism that is used for live previews can also be used for instant type checking.

Programs, commands, terms, expressions and values

$$\begin{array}{lll} p ::= c_1; \dots; c_n & t ::= o \mid x & e ::= t \mid \lambda x \rightarrow e \\ c ::= \text{let } x = t \mid t & & v ::= o \mid \lambda x \rightarrow e \end{array}$$

Evaluation contexts of expressions

$$\begin{array}{ll} C_e[-] &= C_e[-].m(e_1, \dots, e_n) \mid o.m(v_1, \dots, v_m, C_e[-], e_1, \dots, e_n) \mid - \\ C_c[-] &= \text{let } x = C_e[-] \mid C_e[-] \\ C_p[-] &= o_1; \dots; o_k; C_c[-]; c_1; \dots; c_n \end{array}$$

Let elimination and member reduction

$$\begin{array}{ll} o_1; \dots; o_k; \text{let } x = o; c_1; \dots; c_n \rightsquigarrow & (let) \\ o_1; \dots; o_k; o; c_1[x \leftarrow o]; \dots; c_n[x \leftarrow o] & \\ o.m(v_1, \dots, v_n) \rightsquigarrow_{\epsilon} o' \implies C_p[o.m(v_1, \dots, v_n)] \rightsquigarrow C_p[o'] & (external) \end{array}$$

Fig. 3. Syntax, contexts and reduction rules of the data exploration calculus

### 3.1 Language syntax

The calculus combines object-oriented features such as member access with functional features including lambda functions. The syntax is defined in Figure 3. Object values  $o$  are defined by external libraries that are used in conjunction with the core calculus.

A program  $p$  in the data exploration calculus consists of a sequence of commands  $c$ . A command can be either a let binding or a term. Let bindings define variables  $x$  that can be used in subsequent commands. The calculus supports lambda functions, but they can only appear as arguments in method calls. A term  $t$  can be a value, variable or a member access, while an expression  $e$ , which can appear as an argument in member access, can be a lambda function or a term.<sup>1</sup>

### 3.2 Operational semantics

The data exploration calculus is a call-by-value language. We model evaluation as a small-step reduction  $\rightsquigarrow$ . Fully evaluating a program results in an irreducible sequence of objects  $o_1; \dots; o_n$  (one object for each command, including let bindings) which can be displayed as intermediate results of the data analysis. The operational semantics is parameterized by a relation  $\rightsquigarrow_{\epsilon}$  that models the functionality of the external libraries used with the calculus and defines the reduction behaviour for member accesses. The relation has the following form:

$$o_1.m(v_1, \dots, v_n) \rightsquigarrow_{\epsilon} o_2$$

Here, the operation  $m$  is invoked on an object and takes values (objects or function values) as arguments. The reduction always results in an object. Figure 3 defines the reduction rules in terms of  $\rightsquigarrow_{\epsilon}$  and evaluation contexts;  $C_e$  specifies left-to-right evaluation of arguments of a method call,  $C_c$  specifies evaluation of a command and  $C_p$  defines left-to-right evaluation of a program. The rule (external) calls a method provided by an external library in a call-by-value fashion and (let) substitutes a value of an evaluated variable in all subsequent commands and leaves the result in the list of commands. Note that our semantics does not define how  $\lambda$  functions are reduced. This is done by external libraries, which will typically supply functions with arguments using standard  $\beta$ -reduction. The behaviour is subject to constraints discussed next.

<sup>1</sup>Similar but weaker restrictions on the use of lambda functions exist in other languages. For example, in order to guide type inference, lambda functions in C# can appear as either method arguments or assigned to an explicitly typed variable, but they cannot be assigned to an implicitly typed variable.



### 3.3 Properties

The data exploration calculus has a number of desirable properties. Some of those require that the relation  $\rightsquigarrow_\epsilon$ , which defines evaluation for external libraries, satisfies a number of conditions. We discuss *normalization* and *let elimination* in this section. Those two are particularly important as they will allow us to prove correctness of our method of evaluating live previews in Section 6.1.

**Definition 1** (Further reductions). We define two additional reduction relations:

- We write  $\rightsquigarrow^*$  for the reflexive, transitive closure of  $\rightsquigarrow$
- We write  $\rightsquigarrow_{\text{let}}$  for a call-by-name let binding elimination
 
$$c_1; \dots; c_{k-1}; \text{let } x = t; c_{k+1}; \dots; c_n \rightsquigarrow_{\text{let}} c_1; \dots; c_{k-1}; t; c_{k+1}[x \leftarrow t]; \dots; c_n[x \leftarrow t]$$

We say that two expressions  $e$  and  $e'$  are *observationally equivalent* if, for any context  $C$ , the expressions  $C[e]$  and  $C[e']$  reduce to the same value. Lambda functions  $\lambda x \rightarrow 2$  and  $\lambda x \rightarrow 1+1$  are not equal, but they are observationally equivalent. We require that external libraries satisfy two conditions. First, when a method is called with observationally equivalent values as arguments, it should return the same value. Second, the evaluation of  $o.m(v_1, \dots, v_n)$  should be defined for all  $o, i$  and  $v_i$ . The external library will typically satisfy this by defining an error object and reducing all invalid calls to the error object, following the standard method of Milner [1978].

**Definition 2** (External library). An external library consists of a set of objects  $O$  and a reduction relation  $\rightsquigarrow_\epsilon$  that satisfies the following two properties:

- *Totality*. For all  $o, m, i$  and all  $v_1, \dots, v_i$ , there exists  $o'$  such that  $o.m(v_1, \dots, v_i) \rightsquigarrow_\epsilon o'$ .
- *Compositionality*. For observationally equivalent arguments, the reduction returns the same object, i.e. given  $e_0, e_1, \dots, e_n$  and  $e'_0, e'_1, \dots, e'_n$  and  $m$  such that  $e_0.m(e_1, \dots, e_n) \rightsquigarrow^* o$  and  $e'_0.m(e'_1, \dots, e'_n) \rightsquigarrow^* o'$  then if for any contexts  $C_0, C_1, \dots, C_n$  it holds that if  $C_i[e_i] \rightsquigarrow^* o_i$  and  $C_i[e'_i] \rightsquigarrow^* o_i$  for some  $o_i$  then  $o = o'$ .

The compositionality condition is essential for proving the correctness of our live preview mechanism. Totality allows us to prove normalization, i.e. all programs reduce to a value – although the resulting value may be an error value provided by the external library.

**Theorem 1** (Normalization). *For all  $p$ , there exists  $n$  and  $o_1, \dots, o_n$  such that  $p \rightsquigarrow^* o_1; \dots; o_n$ .*

PROOF. A program that is not a sequence of values can be reduced and reduction decreases the size of the program. See Appendix A.1 for more detail.  $\square$

Although the reduction rules (let) and (external) of the data exploration calculus define an evaluation in a call-by-value order, eliminating let bindings in a call-by-name way using the  $\rightsquigarrow_{\text{let}}$  reduction does not affect the result. This simplifies our later proof of live preview correctness in Section 6.1.

**Lemma 2** (Let elimination for a program). *Given any program  $p$  such that  $p \rightsquigarrow^* o_1; \dots; o_n$  for some  $n$  and  $o_1, \dots, o_n$  then if  $p \rightsquigarrow_{\text{let}} p'$  for some  $p'$  then also  $p' \rightsquigarrow^* o_1; \dots; o_n$ .*

PROOF. The elimination of let binding transforms a program  $c_1; \dots; c_{k-1}; \text{let } x = t; c_{k+1}; \dots; c_n$  to a program  $c_1; \dots; c_{k-1}; t; c_{k+1}[x \leftarrow t]; \dots; c_n[x \leftarrow t]$ . The reduction steps for the new program can be constructed using the steps of  $p \rightsquigarrow^* o_1; \dots; o_n$ . The new command  $t$  reduces to an object  $o$  using the same steps as the original term  $t$  in  $\text{let } x = t$  but with context  $C_c = -$  rather than  $C_c = \text{let } x = -$ ; the terms  $t$  introduced by substitution also reduce using the same steps as before, but using contexts in which the variable  $x$  originally appeared.  $\square$

## 4 FORMALISING A LIVE CODING ENVIRONMENT

A naive way of providing live previews during code editing would be to re-evaluate the code after each change. This would be wasteful – when writing code to explore data, most changes are additive and the preview can be updated by evaluating just the newly added code. In this section, we develop foundations for an efficient way of displaying live previews for the data exploration calculus.

### 4.1 Maintaining dependency graph

The key idea behind our method is to maintain a dependency graph [Kuck et al. 1981] with nodes representing individual operations of the computation that can be evaluated to obtain a preview. Each time the program text is modified, we parse it afresh (using an error-recovering parser) and bind the abstract syntax tree to the dependency graph. When binding a new expression to the graph, we reuse previously created nodes as long as they have the same dependencies. For expressions that have a new structure, we create new nodes.

The nodes of the graph serve as unique keys into a lookup table containing previously evaluated parts of the program. When a preview is requested for an expression, we use the graph node bound to the expression to find a preview. If a preview has not been evaluated, we force the evaluation of all dependencies in the graph and then evaluate the operation represented by the current node.

**4.1.1 Elements of the graph.** The nodes of the graph represent individual operations of the computation. In our design, the nodes are used as cache keys, so we attach a unique symbol  $s$  to some of the nodes. That way, we can create two unique nodes representing, for example, access to a member named `take` which differ in their dependencies.

Furthermore, the graph edges are labelled with labels indicating the kind of dependency. For a method call, the labels are “first argument”, “second argument” and so on. Writing  $s$  for symbols and  $i$  for integers, nodes (vertices)  $v$  and edge labels  $l$  are defined as:

$$\begin{aligned} v &= \text{val}(o) \mid \text{var}(x) \mid \text{mem}(m, s) \mid \text{fun}(x, s) & (\text{Vertices}) \\ l &= \text{body} \mid \text{arg}(i) & (\text{Edge labels}) \end{aligned}$$

The `val` node represents a primitive value and contains the object itself. Two occurrences of `10` in the source code will be represented by the same node. Member access `mem` contains the member name, together with a unique symbol  $s$  – two member access nodes with different dependencies will contain a different symbol. Dependencies of member access are labelled with `arg` indicating the index of the argument (the instance has index 0 and arguments start with 1). Finally, nodes `fun` and

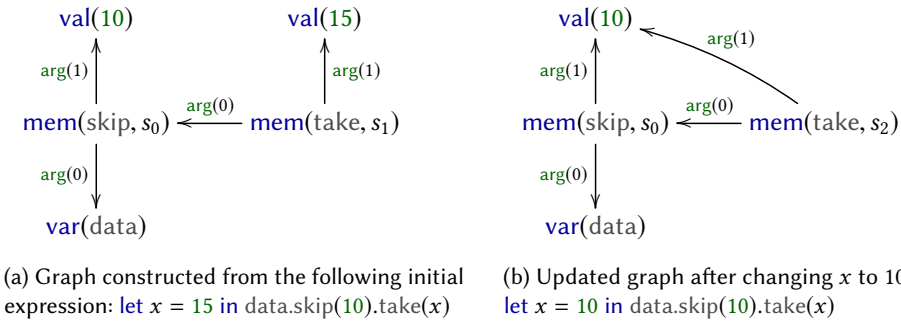


Fig. 4. Dependency graphs formed by two steps of the live programming process.



`var` represent function values and variables bound by  $\lambda$  abstraction. For simplicity, we use variable names rather than de Bruijn indices and so renaming a bound variable forces recomputation.

**4.1.2 Example graph.** Figure 4 illustrates how we construct and update the dependency graph. Node representing `take(x)` depends on the argument – the number 15 – and the instance, which is a node representing `skip(10)`. This, in turn, depends on the instance data and the number 10. Note that variables bound via `let` binding such as `x` do not appear as `var` nodes. The node using it depends directly on the node representing the expression that is assigned to `x`.

After changing the value of `x`, we create a new graph. The dependencies of the node `mem(skip, s0)` are unchanged and so the symbol `s0` attached to the node remains the same and previously computed previews can be reused. This part of the program is not recomputed. The `arg(1)` dependency of the `take` call changed and so we create a new node `mem(skip, s2)` with a fresh symbol `s2`. The preview for this node is then computed as needed using the already known values of its dependencies.

**4.1.3 Reusing graph nodes.** The binding process takes an expression and constructs a dependency graph reusing existing nodes when possible. For this, we use a lookup table of member access and function value nodes. The key in the lookup table is formed by a node kind (for disambiguation) together with a list of dependencies. A node kind is a member access containing the member name or a function containing the bound variable name; a lookup table  $\Delta$  then maps a node kind with a list of dependencies to a cached node:

$$\begin{aligned} k &::= \text{fun}(x) \mid \text{mem}(m) && (\text{Node kinds}) \\ \Delta(k, [(v_1, l_1), \dots, (v_n, l_n)]) && (\text{Lookup for a node}) \end{aligned}$$

The example on the second line looks for a node of a kind `k` that has dependencies `v1, ..., vn` labelled with labels `l1, ..., ln`. We write  $\Delta(k, l) \downarrow$  when a value for a given key is not defined. For example, when creating the graph in Figure 4b, we perform the following lookup for the `skip` member access:

$$\Delta(\text{mem}(\text{skip}), [(\text{var}(\text{data}), \text{arg}(0)), (\text{val}(10), \text{arg}(1))])$$

The lookup returns the node `mem(skip, s0)` known from the previous step. We then perform the following lookup for the `take` member access:

$$\Delta(\text{mem}(\text{take}), [(\text{mem}(\text{skip}, s_0), \text{arg}(0)), (\text{val}(10), \text{arg}(1))])$$

In the previous graph, the argument of `take` was 15 rather than 10 and so this lookup fails. We then construct a new node `mem(take, s2)` using a fresh symbol `s2` and later insert it into the cache.

## 4.2 Binding expressions to a graph

After parsing updated code, we update the dependency graph and link each node of the abstract syntax tree to a node of the dependency graph. This process is called binding and is defined by the `bind-expr` function (Figure 5) and `bind-prog` function (Figure 6). Both functions are annotated with a lookup table  $\Delta$  discussed in Section 4.1 and a variable context  $\Gamma$ . The variable context is a map from variable names to dependency graph nodes and is used for variables bound using `let` binding.

When applied on an expression `e`, binding `bind-exprΓ,Δ(e)` returns a node `v` corresponding to the expression `e` paired with a dependency graph  $(V, E)$ . In the graph,  $V$  is a set of nodes `v` and  $E$  is a set of labelled edges  $(v_1, v_2, l)$ . We attach the label directly to the edge rather than keeping a separate colouring function as this makes the formalisation simpler. The `bind-progΓ,Δ` function works similarly, but takes a sequence of commands and returns a sequence of nodes.

$$\begin{aligned} \text{bind-expr}_{\Gamma, \Delta}(e_0.m(e_1, \dots, e_n)) &= v, (\{v\} \cup V_0 \cup \dots \cup V_n, E \cup E_0 \cup \dots \cup E_n) \\ &\text{when } v_i, (V_i, E_i) = \text{bind-expr}_{\Gamma, \Delta}(e_i) \text{ and } v = \Delta(\text{mem}(m), [(v_0, \text{arg}(0)), \dots, (v_n, \text{arg}(n))]) \\ &\text{let } E = \{(v, v_0, \text{arg}(0)), \dots, (v, v_n, \text{arg}(n))\} \end{aligned} \quad (1)$$

$$\begin{aligned} \text{bind-expr}_{\Gamma, \Delta}(e_0.m(e_1, \dots, e_n)) &= v, (\{v\} \cup V_0 \cup \dots \cup V_n, E \cup E_0 \cup \dots \cup E_n) \\ &\text{when } v_i, (V_i, E_i) = \text{bind-expr}_{\Gamma, \Delta}(e_i) \text{ and } \Delta(\text{mem}(m), [(v_0, \text{arg}(0)), \dots, (v_n, \text{arg}(n))]) \downarrow \\ &\text{let } v = \text{mem}(m, s), s \text{ fresh and } E = \{(v, v_0, \text{arg}(0)), \dots, (v, v_n, \text{arg}(n))\} \end{aligned} \quad (2)$$

$$\begin{aligned} \text{bind-expr}_{\Gamma, \Delta}(\lambda x \rightarrow e) &= v, (\{v\} \cup V, \{e\} \cup E) \\ &\text{when } \Gamma_1 = \Gamma \cup \{x, \text{var}(x)\} \text{ and } v_0, (V, E) = \text{bind-expr}_{\Gamma_1, \Delta}(e) \text{ and } v = \Delta(\text{fun}(x), [(v_0, \text{body})]) \\ &\text{let } e = (v, v_0, \text{body}) \end{aligned} \quad (3)$$

$$\begin{aligned} \text{bind-expr}_{\Gamma, \Delta}(\lambda x \rightarrow e) &= v, (\{v\} \cup V, \{e\} \cup E) \\ &\text{when } \Gamma_1 = \Gamma \cup \{x, \text{var}(x)\} \text{ and } v_0, (V, E) = \text{bind-expr}_{\Gamma_1, \Delta}(e) \text{ and } \Delta(\text{fun}(x), [(v_0, \text{body})]) \downarrow \\ &\text{let } v = \text{fun}(x, s), s \text{ fresh and } e = (v, v_0, \text{body}) \end{aligned} \quad (4)$$

$$\text{bind-expr}_{\Gamma, \Delta}(o) = \text{val}(o), (\{\text{val}(o)\}, \emptyset) \quad (5)$$

$$\text{bind-expr}_{\Gamma, \Delta}(x) = v, (\{v\}, \emptyset) \text{ when } v = \Gamma(x) \quad (6)$$

Fig. 5. Binding rules that define a construction of a dependency graph for an expression.

$$\begin{aligned} \text{bind-prog}_{\Gamma, \Delta}(\text{let } x = e; c_2; \dots; c_n) &= v_1; \dots; v_n, (\{v_1\} \cup V \cup V_1, E \cup E_1) \\ &\text{let } v_1, (V_1, E_1) = \text{bind-expr}_{\Gamma, \Delta}(e) \text{ and } \Gamma_1 = \Gamma \cup \{(x, v_1)\} \\ &\text{and } v_2; \dots; v_n, (V, E) = \text{bind-prog}_{\Gamma_1, \Delta}(c_2; \dots; c_n) \end{aligned} \quad (7)$$

$$\begin{aligned} \text{bind-prog}_{\Gamma, \Delta}(e; c_2; \dots; c_n) &= v_1; \dots; v_n, (\{v_1\} \cup V \cup V_1, E \cup E_1) \\ &\text{let } v_1, (V_1, E_1) = \text{bind-expr}_{\Gamma, \Delta}(e) \text{ and } v_2; \dots; v_n, (V, E) = \text{bind-prog}_{\Gamma_1, \Delta}(c_2; \dots; c_n) \end{aligned} \quad (8)$$

$$\text{bind-prog}_{\Gamma, \Delta}([]) = [], (\emptyset, \emptyset) \quad (9)$$

Fig. 6. Binding rules that define a construction of a dependency graph for a program.

**4.2.1 Binding expressions.** When binding member access, we use `bind-expr` recursively to get a node and a dependency graph for each sub-expression. The nodes representing sub-expressions are then used as dependencies for lookup into  $\Delta$ , together with their labels. If a node already exists in  $\Delta$  it is reused (1). Alternatively, we create a new node containing a fresh symbol (2).

If a lambda function uses its argument, we will not be able to evaluate its body. In this case, the graph node bound to a function will depend on a synthetic node `var(x)` that represents the variable with unknown value. When binding a function, we create the synthetic variable and add it to the variable context  $\Gamma_1$  before binding the body. As with member access, the node representing a function may (3) or may not (4) be already present in the lookup table.

**4.2.2 Binding programs.** When binding a program, we bind the first command and then recursively process remaining commands until we reach an empty list of commands (9). For `let` binding (7), we bind the expression  $e$  assigned to the variable to obtain a graph node  $v_1$ . We then bind the remaining commands using a variable context  $\Gamma_1$  that maps the value of the variable to the graph

$\text{update}_{V,E}(\Delta_{i-1}) = \Delta_i$  such that:

$$\begin{aligned} \Delta_i(\text{mem}(m), [(v_0, \text{arg}(0)), \dots, (v_n, \text{arg}(n))]) &= \text{mem}(m, s) \\ &\text{when } \text{mem}(m, s) \in V \text{ and } (\text{mem}(m, s), v_i, \text{arg}(i)) \in E \text{ for } i = 0, \dots, n \\ \Delta_i(\text{fun}(x), [(v_1, \text{body})]) &= \text{fun}(x, s) \\ &\text{when } \text{fun}(x, s) \in V \text{ and } (\text{fun}(x, s), v_1, \text{body}) \in E \\ \Delta_i(v) &= \Delta_{i-1}(v) \quad (\text{otherwise}) \end{aligned}$$

Fig. 7. Updating the node cache after binding a new graph

node  $v_1$ . The variable context is used when binding a variable (6) and so all variables declared using `let` binding will be bound to a graph node representing the value assigned to the variable. When the command is just an expression (8), we bind the expression using `bind-expr`.

### 4.3 Edit and rebind loop

The binding process formalised in Section 4.2 specifies how to update the dependency graph after program is updated and parsed. During live coding, this is done repeatedly as the programmer edits code. Throughout the process, we maintain a series of lookup table states  $\Delta_0, \Delta_1, \Delta_2, \dots$ . The initial lookup table is empty, i.e.  $\Delta_0 = \emptyset$ . At a step  $i$ , we parse a program  $p_i$  (consisting of several commands) and obtain a new dependency graph using the previous  $\Delta$ . The result is a sequence of nodes corresponding to commands of the program and a graph  $(V, E)$ :

$$v_1; \dots; v_n, (V, E) = \text{bind-prog}_{\emptyset, \Delta_{i-1}}(p_i)$$

The new state of the cache is computed by calling the  $\text{update}_{V,E}(\Delta_{i-1})$  function defined in Figure 7. The function adds newly created nodes from the graph  $(V, E)$  to the previous cache  $\Delta_{i-1}$  and returns a new cache  $\Delta_i$ . We only cache nodes for function and member accesses – nodes for variables and primitive values will remain the same thanks to the way they are constructed.

## 5 COMPUTING LIVE PREVIEWS

The binding process described in the previous section constructs a dependency graph after code changes. The nodes in the dependency graph correspond to individual operations that will be performed when running the program. When evaluating a preview, we attach (partial) results to nodes of the graph. Since the binding process reuses nodes when their dependencies do not change, previews for expressions (or sub-expressions) of a program are reused when updating a preview.

In this section, we describe how previews are evaluated. The evaluation is done over the dependency graph, rather than over the structure of program expressions as in the operational semantics given in Section 3.2. We analyse the preview evaluation formally in Section 6 and show that the resulting previews are the same as the result we would get by directly evaluating the code and, equally importantly, we show that no recomputation occurs when code is edited in certain ways.

### 5.1 Previews and delayed previews

Programs in the data exploration calculus consist of sequence of commands. Those are evaluated to a value with a preview that can be displayed to the user. However, we also support previews for sub-expressions. This can be problematic if the current sub-expression is inside the body of a function. For example:

```
let top = movies.take(10).map( $\lambda x \rightarrow x.getReleased().format("dd-mm-yyyy")$ )
```

Here, we can directly evaluate sub-expressions `movies` and `movies.take(10)`, but not `x.getReleased()` and `x.getReleased().format("dd-mm-yyyy")` because they contain a free variable `x`. Our preview evaluation algorithm addresses this by producing two kinds of previews. A *fully evaluated preview* is just a value, while a *delayed preview* is a partially evaluated expression with free variables:

$$\begin{aligned} p &= o \mid \lambda x \rightarrow e && \text{(Fully evaluated previews)} \\ d &= p \mid \llbracket e \rrbracket_\Gamma && \text{(Evaluated and delayed previews)} \end{aligned}$$

A fully evaluated preview  $p$  can be either a primitive object or a function value with no free variables. A possibly delayed preview  $d$  can be either an evaluated preview  $p$  or an expression  $e$  that requires variables  $\Gamma$ . For simplicity, we use an untyped language and so  $\Gamma$  is just a list of variables  $x_1, \dots, x_n$ .

A delayed preview is not necessarily the body of a lambda function as it appears in the source code. We partially evaluate sub-expressions of the body that do not have free variables or that have free variables bound by an earlier let binding.

Our implementation does not currently display delayed previews to the user, but there is a number of possible approaches for doing that. Most interestingly, since lambda functions always appear as arguments of a member access, we could force the evaluation of the surrounding expression, capture (a number of) values passed as inputs to the lambda function and display a preview based on those. In Section 7.2 we also consider a more speculative design for an abstraction mechanism that supports live previews, which could, in some cases, replace lambda function.

## 5.2 Evaluation of previews

The evaluation of previews is defined in Figure 8. Given a dependency graph  $(V, E)$ , we define a relation  $v \Downarrow d$  that evaluates a sub-expression corresponding to the node  $v$  to a possibly delayed preview  $d$ . The nodes  $V$  and edges  $E$  of the dependency graph are parameters of the  $\Downarrow$  relation, but they do not change during the evaluation and so we do not explicitly write them.

$$\begin{array}{l} \text{(lift-expr)} \quad \frac{v \Downarrow \llbracket e \rrbracket_\Gamma}{v \Downarrow_{\text{lift}} \llbracket e \rrbracket_\Gamma} \qquad \text{(fun-val)} \quad \frac{(\text{fun}(x, s), v, \text{body}) \in E \quad v \Downarrow p}{\text{fun}(x, s) \Downarrow \lambda x \rightarrow p} \\ \text{(lift-prev)} \quad \frac{v \Downarrow p}{v \Downarrow_{\text{lift}} \llbracket p \rrbracket_\emptyset} \qquad \text{(fun-bind)} \quad \frac{(\text{fun}(x, s), v, \text{body}) \in E \quad v \Downarrow \llbracket e \rrbracket_x}{\text{fun}(x, s) \Downarrow \lambda x \rightarrow e} \\ \text{(val)} \quad \frac{}{\text{val}(o) \Downarrow o} \qquad \text{(fun-expr)} \quad \frac{(\text{fun}(x, s), v, \text{body}) \in E \quad v \Downarrow \llbracket e \rrbracket_{x, \Gamma}}{\text{fun}(x, s) \Downarrow \llbracket \lambda x \rightarrow e \rrbracket_\Gamma} \\ \text{(var)} \quad \frac{}{\text{var}(x) \Downarrow \llbracket x \rrbracket_x} \\ \text{(mem-val)} \quad \frac{\forall i \in \{0 \dots k\}. (\text{mem}(m, s), v_i, \text{arg}(i)) \in E \quad v_i \Downarrow p_i \quad p_0.m(p_1, \dots, p_k) \rightsquigarrow_\epsilon p}{\text{mem}(m, s) \Downarrow p} \\ \text{(mem-expr)} \quad \frac{\forall i \in \{0 \dots k\}. (\text{mem}(m, s), v_i, \text{arg}(i)) \in E \quad \exists j \in \{0 \dots k\}. v_j \not\Downarrow p_j \quad v_i \Downarrow_{\text{lift}} \llbracket e_i \rrbracket_{\Gamma_i}}{\text{mem}(m, s) \Downarrow \llbracket e_0.m(e_1, \dots, e_k) \rrbracket_{\Gamma_0, \dots, \Gamma_k}} \end{array}$$

Fig. 8. Rules that define evaluation of previews over a dependency graph for a program

The auxiliary relation  $v \Downarrow_{\text{lift}} d$  always evaluates to a delayed preview. If the ordinary evaluation returns a delayed preview, so does the auxiliary relation (lift-expr). If the ordinary evaluation returns a value, the value is wrapped into a delayed preview requiring no variables (lift-prev). A node representing a value is evaluated to a value (val) and a node representing an unbound variable is reduced to a delayed preview that requires the variable and returns its value (var).

For member access, we distinguish two cases. If all arguments evaluate to values (member-val), then we use the evaluation relation defined by external libraries  $\rightsquigarrow_e$  to immediately evaluate the member access and produce a value. If some of the arguments are delayed (member-expr), because the member access is inside the body of a lambda function, we produce a delayed member access expression that requires the union of the variables required by the individual arguments.

The evaluation of function values is similar, but requires three cases. If the body can be reduced to a value with no unbound variables (fun-val), we return a lambda function that returns the value. If the body requires only the bound variable (fun-bind), we return a lambda function with the delayed preview as the body. If the body requires further variables, the result is a delayed preview.

### 5.3 Caching of evaluated previews

For simplicity, the relation  $\Downarrow$  in Figure 8 does not specify how previews are cached and linked to graph nodes. In practice, this is done by maintaining a lookup table from graph nodes  $v$  to (possibly delayed) previews  $p$ . Whenever  $\Downarrow$  is used to obtain a preview for a graph node, we first attempt to find an already evaluated preview using the lookup table. If the preview has not been previously evaluated, we evaluate it and add it to the lookup table.

The cached evaluated previews can be reused in two ways. First, multiple nodes can depend on one sub-graph in a single dependency graph (if the same sub-expression appears twice in the program). Second, the keys of the lookup table are graph nodes and nodes are reused when a new dependency graph is constructed after the user edits the source code.

### 5.4 Theories of delayed previews

The operational semantics presented in this paper serves two purposes. It gives a simple guide for implementing text-based live coding environments for data science and we use it to prove that our optimized way of producing live previews is correct. However, some aspects of our mechanism are related to important work in semantics of programming languages and deserve to be mentioned.

The construction of delayed previews is related to meta-programming. Assuming we have delayed previews  $\llbracket e_0 \rrbracket_x$  and  $\llbracket e_1 \rrbracket_y$  and we invoke a member  $m$  on  $e_0$  using  $e_1$  as an argument. To do this, we construct a new delayed preview  $\llbracket e_0.m(e_1) \rrbracket_{x,y}$ . This operation is akin to expression splicing from meta-programming [Syme 2006; Taha and Sheard 2000].

The semantics of delayed previews can be more formally captured by Contextual Modal Type Theory (CMTT) [Nanevski et al. 2008] and comonads [Gabbay and Nanevski 2013]. In CMTT,  $[\Psi]A$  denotes that a proposition  $A$  is valid in context  $\Psi$ , which is similar to our delayed previews written as  $\llbracket A \rrbracket_\Psi$ . CMTT defines rules for composing context-dependent propositions that would allow us to express the splicing operation used in (mem-expr). In categorical terms, the context-dependent proposition can be modelled as a graded comonad [Gaborardi et al. 2016; Mycroft et al. 2016]. The evaluation of a preview with no context dependencies (built implicitly into our evaluation rules) corresponds to the counit operation of a comonad and would be explicitly written as  $\llbracket A \rrbracket_\emptyset \rightarrow A$ .

## 6 EVALUATING LIVE CODING ENVIRONMENT

Evaluating previews using a dependency graph makes it possible to cache partial results. The mechanism satisfies two properties. First, if we evaluate a preview using dependency graph with caching, it is the same as the value we would obtain by evaluating the expression directly. Second,

the evaluation of previews using dependency graphs reuses – in some cases – previously evaluated partial results. In other words, we show that the mechanism is correct and implements an effective optimization. We discuss correctness first and then evaluate effectiveness, both theoretically (Section 6.2) and empirically (Section 6.3). Finally, we describe a case study where we used the methods outlined in the paper to develop a larger online service for data exploration (Section 6.4).

### 6.1 Correctness of previews

To show that the previews are correct, we prove two properties. Correctness (Theorem 6) guarantees that, no matter how a dependency graph is constructed, when we use it to evaluate previews for a program, the previews are the same as the values we would obtain by evaluating the program commands directly. Determinacy (Theorem 7) guarantees that if we cache a preview for a graph node and update the graph, the preview we would evaluate using the updated graph would be the same as the cached preview.

To simplify the proofs, we consider programs without let bindings. This is possible, because eliminating let bindings does not change the result of evaluation, as shown earlier in Lemma 2, and it also does not change the constructed dependency graph as shown below in Lemma 3.

**Lemma 3** (Let elimination for a dependency graph). *Given programs  $p_1, p_2$  such that  $p_1 \rightsquigarrow_{\text{let}} p_2$  and a lookup table  $\Delta_0$  then if  $v_1; \dots; v_n, (V, E) = \text{bind-prog}_{\emptyset, \Delta_0}(p_1)$  and  $v'_1; \dots; v'_n, (V', E') = \text{bind-prog}_{\emptyset, \Delta_1}(p_2)$  such that  $\Delta_1 = \text{update}_{V, E}(\Delta_0)$  then for all  $i$ ,  $v_i = v'_i$  and also  $(V, E) = (V', E')$ .*

**PROOF.** Assume  $p_1 = c_1; \dots; c_{k-1}; \text{let } x = e; c_{k+1}; \dots; c_n$  and the let binding is eliminated resulting in  $p_2 = c_1; \dots; c_{k-1}; e; c_{k+1}[x \leftarrow e]; \dots; c_n[x \leftarrow e]$ . When binding  $p_1$ , the case  $\text{bind-prog}_{\Gamma, \Delta}(\text{let } x = e)$  is handled using (7) and the node resulting from binding  $e$  is added to the graph  $V, E$ . It is then referenced each time  $x$  appears in subsequent commands  $c_{k+1}; \dots; c_n$ . When binding  $p_2$ , the node resulting from binding  $e$  is a primitive value or a node already present in  $\Delta_1$  (added by  $\text{update}_{V, E}$ ) and is reused each time  $\text{bind-expr}_{\Gamma, \Delta_1}(e)$  is called.  $\square$

The Lemma 3 provides a way of removing let bindings from a program, such that the resulting dependency graph remains the same. Here, we bind the original program first, which adds the node for  $e$  to  $\Delta$ . In our implementation, this is not needed because  $\Delta$  is updated while the graph is being constructed using  $\text{bind-expr}$ . To keep the formalisation simpler, we separate the process of building the dependency graph and updating  $\Delta$  and thus Lemma 3 requires an extra binding step.

Now, we can show that, given a let-free expression, the preview obtained using a correctly constructed dependency graph is the same as the one we would obtain by directly evaluating the expression. This requires a simple auxiliary lemma.

**Lemma 4** (Lookup inversion). *Given  $\Delta$  obtained using update as defined in Figure 7 then:*

- If  $v = \Delta(\text{fun}(x), [(v_0, l_0)])$  then  $v = \text{fun}(x, s)$  for some  $s$ .
- If  $v = \Delta(\text{mem}(m), [(v_0, l_0), \dots, (v_n, l_n)])$  then  $v = \text{mem}(m, s)$  for some  $s$ .

**PROOF.** By construction of  $\Delta$  in Figure 7.  $\square$

**Theorem 5** (Term preview correctness). *Given a term  $t$  that has no free variables, together with a lookup table  $\Delta$  obtained from any sequence of programs using  $\text{bind-prog}$  (Figure 6) and  $\text{update}$  (Figure 7), then let  $v, (V, E) = \text{bind-expr}_{\emptyset, \Delta}(t)$ .*

*If  $v \Downarrow p$  over a graph  $(V, E)$  then  $p = o$  for some value  $o$  and  $t \rightsquigarrow^* o$ .*

**PROOF.** First note that, when combining recursively constructed sub-graphs, the  $\text{bind-expr}$  function adds new nodes and edges leading from those new nodes. Therefore, an evaluation using  $\Downarrow$



over a sub-graph will also be valid over the new graph – the newly added nodes and edges do not introduce non-determinism to the rules given in Figure 8.

We prove a more general property showing that for any  $e$ , its binding  $\mathbf{v}, (V, E) = \text{bind-expr}_{\emptyset, \Delta}(e)$  and any evaluation context  $C$  such that  $C[e] \rightsquigarrow o$  for some  $o$ , one of the following holds:

- a. If  $FV(e) = \emptyset$  then  $\mathbf{v} \Downarrow p$  for some  $p$  and  $C[p] \rightsquigarrow o$
- b. If  $FV(e) \neq \emptyset$  then  $\mathbf{v} \Downarrow \llbracket e_p \rrbracket_{FV(e)}$  for some  $e_p$  and  $C[e_p] \rightsquigarrow o$

In the first case,  $p$  is a value, but it is not always the case that  $e \rightsquigarrow^* p$ , because  $p$  may be lambda function and preview evaluation may reduce sub-expression in the body of the function. Using a context  $C$  in which the value reduces to an object avoids this problem.

The proof of the theorem follows from the more general property. Using a context  $C[-] = -$ , the term  $t$  reduces  $t \rightsquigarrow^* t' \rightsquigarrow_{\epsilon} o$  for some  $o$  and the preview  $p$  is a value  $o$  because  $C[p] = p = o$ . The proof is by induction over the binding process, which follows the structure of the expression. The full proof is shown in Appendix A.2.  $\square$

The correctness theorem combines the previous two results.

**Theorem 6** (Program preview correctness). *Consider a program  $p = c_1; \dots; c_n$  that has no free variables, together with a lookup table  $\Delta_0$  obtained from any sequence of programs using  $\text{bind-prog}$  (Figure 6) and update (Figure 7). Assume a let-free program  $p' = t_1; \dots; t_n$  such that  $p \rightsquigarrow_{\text{let}}^* p'$ .*

*Let  $\mathbf{v}_1; \dots; \mathbf{v}_n, (V, E) = \text{bind-prog}_{\emptyset, \Delta_0}(p)$  and define updated lookup table  $\Delta_1 = \text{update}_{V, E}(\Delta_0)$  and let  $\mathbf{v}'_1; \dots; \mathbf{v}'_n, (V', E') = \text{bind-prog}_{\emptyset, \Delta_1}(p')$ .*

*If  $\mathbf{v}'_i \Downarrow p_i$  over a graph  $(V', E')$  then  $p_i = o_i$  for some value  $o_i$  and  $t_i \rightsquigarrow o_i$ .*

PROOF. Direct consequence of Lemma 3 and Theorem 5.  $\square$

Our implementation updates  $\Delta$  during the recursive binding process and so a stronger version of the property holds: previews calculated over a graph obtained directly for the original program  $p$  are the same as the values of the fully evaluated program. Our formalisation omits this for simplicity.

The second important property is determinacy, which makes it possible to cache the previews evaluated using  $\Downarrow$  using the corresponding graph node as a lookup key.

**Theorem 7** (Program preview determinacy). *For some  $\Delta$  and for any programs  $p, p'$ , assume that the first program is bound, i.e.  $\mathbf{v}_1; \dots; \mathbf{v}_n, (V, E) = \text{bind-prog}_{\emptyset, \Delta}(p)$ , the graph node cache is updated  $\Delta' = \text{update}_{V, E}(\Delta)$  and the second program is bound, i.e.  $\mathbf{v}'_1; \dots; \mathbf{v}'_m, (V', E') = \text{bind-prog}_{\emptyset, \Delta'}(p')$ . Now, for any  $\mathbf{v}$ , if  $\mathbf{v} \Downarrow p$  over  $(V, E)$  then also  $\mathbf{v} \Downarrow p$  over  $(V', E')$ .*

PROOF. By induction over  $\Downarrow$  over  $(V, E)$ , we show that the same evaluation rules also apply over  $(V', E')$ . This is the case, because graph nodes added to  $\Delta'$  by  $\text{update}_{V, E}$  are added as new nodes in  $\text{bind-prog}_{\emptyset, \Delta'}$  and nodes and edges of  $(V, E)$  used during the evaluation are unaffected.  $\square$

The mechanism used for caching previews (Section 5.3), keeps a preview  $d$  in a lookup table indexed by nodes  $\mathbf{v}$ . Theorem 7 guarantees that this is a valid – as we update dependency graph during code editing, previous nodes will continue representing the same sub-expressions.

## 6.2 Reuse of previews

In the motivating example given in Section 1, the data analyst first extracted a constant value into a let binding and then modified a parameter of the last method call in a call chain. We argued that a live coding environment should reuse partially evaluated previews for these two cases. We now show that this is, indeed, the case in our system.

In this section, we identify a number of code edit operations where the previously evaluated values for a sub-expression can be reused. The list of operations is shown in Figure 9. To express



Edit contexts of expressions

$$\begin{aligned} K_e[-] &= K_e[-].m(e_1, \dots, e_n) \mid e.m(e_1, \dots, e_{l-1}, K_e[-], e_{l+1}, \dots, e_n) \mid - \\ K_c[-] &= \text{let } x = K_e[-] \mid K_e[-] \end{aligned}$$

Code edit operations preserving preview for a sub-expression

- (let-intro-var)  $\bar{c}_1; \langle e \rangle; \bar{c}_2$  changes to  $\bar{c}_1; \text{let } x = e; \langle x \rangle; \bar{c}_2$  where  $x$  is a fresh name.
- (let-intro-ins)  $\bar{c}_1; \bar{c}_2; \langle K_c[e] \rangle; \bar{c}_3$  is changed to  $\bar{c}_1; \text{let } x = e; \bar{c}_2; \langle K_c[x] \rangle; \bar{c}_3$  via a semantically non-equivalent expression  $\bar{c}_1; \bar{c}_2; K_c[x]; \bar{c}_3$  where  $x$  is free.
- (let-intro-del)  $\bar{c}_1; \bar{c}_2; \langle K_c[e] \rangle; \bar{c}_3$  is changed to  $\bar{c}_1; \text{let } x = e; \bar{c}_2; \langle K_c[x] \rangle; \bar{c}_3$  via an expression  $\bar{c}_1; \text{let } x = e; \bar{c}_2; K_c[e]; \bar{c}_3$  with unused variable  $x$ .
- (let-elim-del)  $\bar{c}_1; \text{let } x = e; \bar{c}_2; \langle K_c[x] \rangle; \bar{c}_3$  is changed to  $\bar{c}_1; \bar{c}_2; \langle K_c[e] \rangle; \bar{c}_3$  via a semantically non-equivalent expression  $\bar{c}_1; \bar{c}_2; K_c[x]; \bar{c}_3$  where  $x$  is free.
- (let-elim-ins)  $\bar{c}_1; \text{let } x = e; \bar{c}_2; \langle K_c[x] \rangle; \bar{c}_3$  is changed to  $\bar{c}_1; \bar{c}_2; \langle K_c[e] \rangle; \bar{c}_3$  via an expression  $\bar{c}_1; \text{let } x = e; \bar{c}_2; K_c[e]; \bar{c}_3$  with unused variable  $x$ .
- (edit-mem)  $\bar{c}_1; K_c[\langle e_0 \rangle.m(\bar{e})]; \bar{c}_2$  is changed to  $\bar{c}_1; K_c[\langle e_0 \rangle.m'(\bar{e}')]; \bar{c}_2$
- (edit-let)  $\bar{c}_1; \text{let } x = e_1; \bar{c}_2; K_c[\langle e_2 \rangle]; \bar{c}_3$  is changed to  $\bar{c}_1; \text{let } x = e'_1; \bar{c}_2; K_c[\langle e_2 \rangle]; \bar{c}_3$  when  $x \notin FV(e_2)$ .

Fig. 9. Code edit operations that preserve previously evaluated preview

the operations we define an editing context  $K$  which is defined similarly to evaluation context  $C$  from Figure 3, but captures sub-expressions appearing anywhere in the program.

We use the notation  $\langle e \rangle$  to mark the part of the expressions before and after the edit that have the same value and will not need to be recomputed after the change; we write  $\bar{c}$  and  $\bar{e}$  for a list of commands and expressions, respectively. In some of the edit operations, we also specify an intermediate program that may be semantically different and only has a partial preview. This illustrates a typical way of working with code in a text editor using cut and paste. For example, in (let-intro-ins), we the analyst cuts a sub-expression  $e$ , replaces it with a variable  $x$  and then adds a let binding for a variable  $x$  and inserts the expression  $e$  from the clipboard. The (let-intro-del) operation captures the same edit, but done in a different order.

The of preview-preserving code edit operations includes several ways of introducing a let binding, several ways of eliminating a let binding and two edit operations where the data analyst modifies an unrelated part of the program. It is worth noting that the list is not exhaustive. Rather, it aims to illustrate some of the typical operations that the data analyst might perform when writing code.

Theorem 9 proves that the operations given in Figure 9 preserve the preview for a marked sub-expression. It relies on the following lemma that characterizes one of the common kinds of edits more generally. Given two versions of a program that both contain the same sub-expression  $e$ , if the let bindings that define the values of variables used in  $e$  do not change, then the graph node assigned to  $e$  will be the same when binding the original and the updated program.

**Lemma 8** (Binding sub-expressions). *Assume we have programs  $p_1 = c_1; \dots; c_k; K_c[e]; c_{k+1}; \dots; c_n$  and  $p_2 = c'_1; \dots; c'_k; K'_c[e]; c'_{k+1}; \dots; c'_n$  and  $I \subseteq \{1 \dots k\}$  such that  $\forall i \in I. c_i = c'_i$  and for each  $x \in \bigcup_{i \in I} FV(c_i) \cup FV(e)$  there exists  $j \in I$  such that  $c_j = \text{let } x = e$  for some  $e$ . Given any  $\Delta$ , assume that the first program is bound, i.e.  $v_1; \dots; v_n, (V, E) = \text{bind-prog}_{0, \Delta}(p_1)$ , the cache is updated  $\Delta' = \text{update}_{V, E}(\Delta)$  and the second program is bound, i.e.  $v'_1; \dots; v'_n, (V', E') = \text{bind-prog}_{0, \Delta'}(p_2)$ .*

Now, assume  $v, G = \text{bind-expr}_{\Gamma, \Delta}(e)$  and  $v', G' = \text{bind-expr}_{\Gamma', \Delta'}(e)$  are the recursive calls to bind  $e$  during the first and the second binding, respectively. Then, the graph nodes assigned to the sub-expression  $e$  are the same, i.e.  $v = v'$ .

PROOF. First, assuming that  $\forall x \in FV(e). \Gamma(x) = \Gamma'(x)$ , we show by induction over the binding process of  $e$  for the first program that the result is the same. In cases (1) and (3), the updated  $\Delta'$  contains the required key and so the second binding proceeds using the same case. In cases (2) and (4), the second binding reuses the node created by the first binding using case (1) and (3), respectively. Cases (5) and (6) are the same.

Second, when binding let bindings in  $c_1; \dots; c_k$ , the initial  $\Gamma = \emptyset$  during both bindings. Nodes added to  $\Gamma$  and  $\Gamma'$  for commands  $c_j$  such that  $j \in I$  are the same and nodes added for remaining commands do not add any new nodes referenced from  $e$  and so  $v = v'$  using the above.  $\square$

**Theorem 9** (Preview reuse). *Given the sequence of expressions as specified in Figure 9, if the expressions are bound in sequence and graph node cache updated as specified in Figure 7, then the graph nodes assigned to the specified sub-expressions are the same.*

PROOF. Cases (edit-let) and (edit-mem) are direct consequences of Lemma 8; for (let-intro-var), the node assigned to  $x$  is the node assigned to  $e$  which is the same as before the edit from Lemma thm:sub-expr. Cases (let-intro-ins) and (let-intro-del) are similar to (let-intro-var), but also require using induction over the binding of  $K_c[e]$ . Finally, cases (let-elim-ins) and (let-elim-del) are similar and also use Lemma 8 together with induction over the binding of  $K_c[x]$ .  $\square$

### 6.3 Empirical evaluation of efficiency

The key performance claim about our method of providing instant feedback is that it is more efficient than recomputing values for the whole program (or the current command) after every keystroke. In the previous section, we formally proved that this is true and gave examples of code edit operations that do not cause recomputation. In this section, we further support this claim with an empirical evaluation. The purpose of this section is not to precisely evaluate overheads of our implementation, but to compare how much recomputation different evaluation strategies perform.

For the purpose of the evaluation, we use a simple image manipulation library that provides operations for loading, greyscaling, blurring and combining images. We compare delays in updating the preview for three different evaluation strategies, while performing the same sequence of code edit operations. Using image processing as an example gives us a way to visualize the reuse of previously computed values. As in a typical data exploration scenario, the individual operations are relatively expensive compared to the overheads of building the dependency graph.

**6.3.1 Code edit operations and methodology.** To avoid clutter when visualizing the performance, we update the preview after complete tokens are added (such as identifier, number, dot or left parenthesis) rather than after individual keystrokes. Figure 10 shows the sequence of code edit operations that we use to measure the delays in updating a live preview. We first enter an expression to load, greyscale and blur an image (1) then introduce let binding (2) and add more operations (3). Finally, we extract one of the parameters into a variable (4). Most of the operations are simply adding code, but there are two cases where we modify existing code and change value of a parameter for blur and combine immediately after (1) and (3), respectively.

We implement the algorithm described in Sections 4 and 5 in a simple web-based environment that allows the user to modify code and explicitly trigger recomputation. It then measures time needed to recompute values for the whole program and displays the resulting image. If the parsing fails, e.g. because of an unclosed parenthesis, previews are not updated and we record only the time taken by parsing attempt. We compare the delays of three different evaluation strategies:

- *Call-by-value*. Following the semantics in Section 3.2, all sub-expressions are fully evaluated before an expression is evaluated. In this case, work may be done even if the evaluation results in an error. For example, we parse `image.load("shadow.png").blur` as member access with no arguments. The evaluation loads the image, but then fails because `blur` requires one argument.
- *Lazy*. To address the wastefulness of call-by-value strategy discussed above, we simulate lazy evaluation by implementing a version of the image processing library where operations build a delayed computation and only evaluate it when rendering an image. Using this strategy, failing computations do not perform unnecessary work.
- *Live*. Finally, the live strategy uses the algorithm described in Sections 4 and 5. The cache is empty at the beginning of the experiment and we update it after each token is added. This is the only strategy that does not start afresh after reparsing code.

6.3.2 *Efficiency evaluation*. The experimental environment is implemented in F# and compiled to JavaScript using the Fable compiler. We run the experiments in Firefox (version 64.0.2, 32bit) on Windows 10 (build 1809, 64bit) on a computer with Intel Core i7-7660U CPU and 16Gb RAM.

Figure 11 shows times needed to recompute previews after individual tokens of the program are added, deleted or modified, according to the script in Figure 10, resulting in 38 measurements for each of the 3 evaluation strategies. We mark a number of notable points in the chart:

- Loading image for the first time incurs small extra overhead in the live strategy.
- Greyscaling using the live strategy does not need to re-load the image.
- Accessing the `blur` member without arguments causes delay for call-by-value strategy.
- When varying the parameter of `blur`, the live strategy reuses previously greyscaled image.
- Introducing `let` binding does not cause recomputation when using live strategy.
- As in (c), accessing `combine` member without valid argument only affects call-by-value.
- Varying parameter of `combine` in live strategy does not recompute `blur` and is much faster.
- Introducing `let` binding, again, causes full recomputation for lazy can call-by-value.

A summarized view of the delays is provided in Figure 12, which shows histograms illustrating the distribution of delays for each of the three evaluation methods. A large proportion of delays is very

- (1) Enter the following code and then change parameter of `blur` from 4 to 8:

```
image.load("shadow.png").greyScale().blur(4)
```

- (2) Assign the result to a variable and start writing code for further operations:

```
let shadow = image.load("shadow.png").greyScale().blur(8)
shadow.combine
```

- (3) Finish code to combine two images and then change parameter of `combine` from 20 to 80:

```
let shadow = image.load("shadow.png").greyScale().blur(8)
shadow.combine(image.load("pope.png"), 20)
```

- (4) Extract the parameter of `combine` to a `let` bound variable:

```
let ratio = 80
let shadow = image.load("shadow.png").greyScale().blur(8)
shadow.combine(image.load("pope.png"), ratio)
```

Fig. 10. Code edit operations that are used in the experimental evaluation

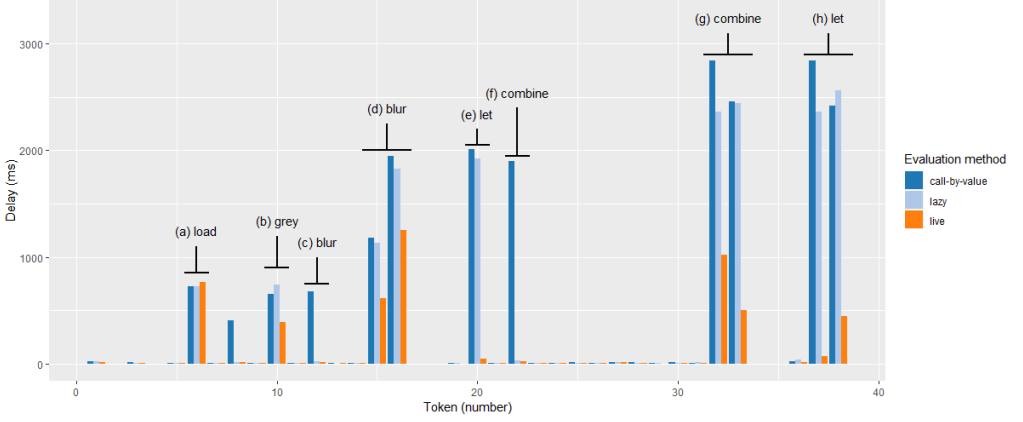


Fig. 11. Time required to recompute the results of a sample program after individual tokens of the program are added or modified for three different evaluations strategies.

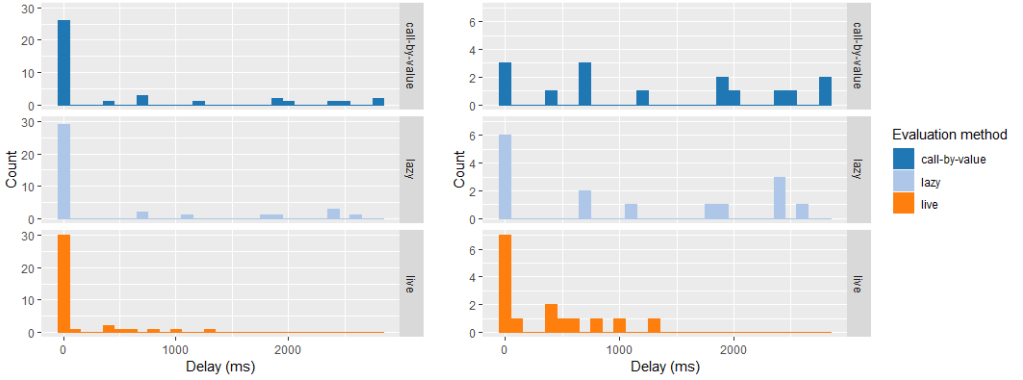


Fig. 12. Histogram showing the distribution of delays incurred when updating previews. We show a histogram computed from all delays (left) and only from delays larger than 15ms (right).

small (less than 15ms) because the parser used in our experimental environment often fails (e.g. for unclosed parentheses). The histogram on the right summarizes only delays for edit operations where the delay for at least one of the strategies was over 15ms. A more robust parser would be able to recover and the delays for call-by-value and lazy strategies would be even more significant.

**6.3.3 Summary of results.** The purpose of our experiment is not to exactly assess the overhead of our implementation. Our goal is to illustrate how often can previously evaluated results be reused and the impact this has when writing code. The experiment presented in this section is small-scale, but it is sufficient for this purpose. When recomputing results after every edit using the *call-by-value* strategy, the time needed to update results grows continually. The *lazy* strategy removes the overhead for programs that fail, but keeps the same trend. Our *live* strategy reuses values computed previously. Consequently, expensive operations (such as (d) and (g) in Figure 11) are significantly faster, because they do not need to recompute operations done previously when writing the code. As shown in Figure 12, there are almost no very expensive operations (taking over 1 second) in the *live* strategy in contrast to several in the other two strategies.

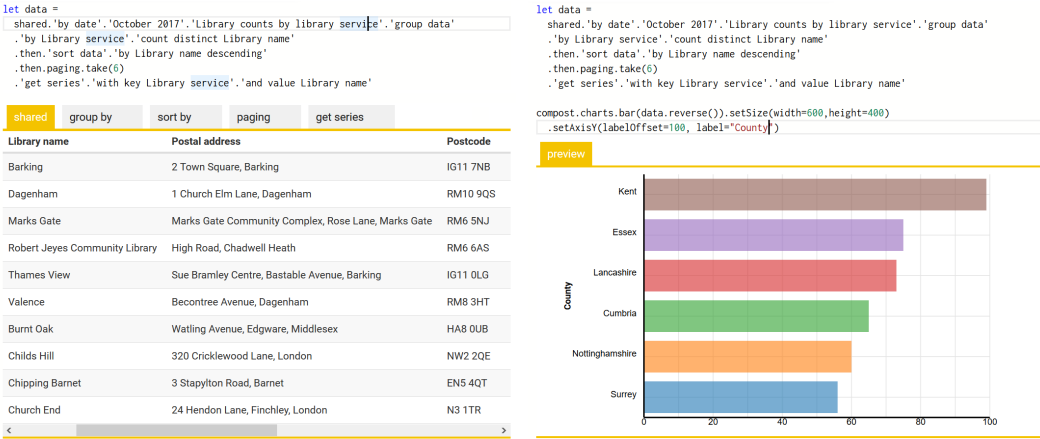


Fig. 13. Data analysis counting the number of libraries per county in the UK. We load and aggregate data (left) and then create a chart with a label (right). The example is adapted from <http://gallery.thegamma.net/73/>.

#### 6.4 Transparent tools for data journalism

In Section 2.1, we motivated our work by considering how journalists explore open data. In addition to the theoretical and experimental work presented in this paper, we also implemented an online data exploration environment, equipped with live editor for The Gamma language that provides instant feedback during coding. The environment uses the principles presented in this paper to build a more comprehensive system that allows users, such as journalists, to analyse, summarize and visualize open data. In this section, we briefly report on our experience with the system.

The environment is available at <http://gallery.thegamma.net>. Two screenshots shown in Figure 13 illustrate a number of interesting features:

- The left screenshot uses a type provider for data aggregation [Petricek 2017b] (Section 2.2). Our system views the type provider as an external library (with objects, members and reduction relation). Type providers rely on type information to provide editor auto-complete, which we support by implementing type checking over a dependency graph as discussed in Section 7.1.
- When displaying live preview for code written using the data aggregation type provider (left), our environment also generates tabs that show individual steps of the data transformation. A tab is selected based on the cursor position and shows preview for the current sub-expression (e.g. raw data before grouping in the screenshot).
- Another external library provides support for charting (right). The environment does not handle charts in a special way and displays the result of evaluating the whole command. The screenshot shows a case where the user modifies parameters of the chart. Thanks to our live evaluation strategy, this is done efficiently without reevaluating the data transformation.

A user study to evaluate the usability of the system from a human-computer interaction perspective is left for future work. As an anecdotal evidence of the usability, the code in Figure 13 was developed by an attendee of a tutorial at Mozilla Festival 2017 who had no prior programming experience.

### 7 EXTENDING DATA EXPLORATION CALCULUS

The data exploration calculus presented in Section 3 allowed us to formalise the live preview mechanism, but it lacks many important features. We outline two extensions. First, we discuss how

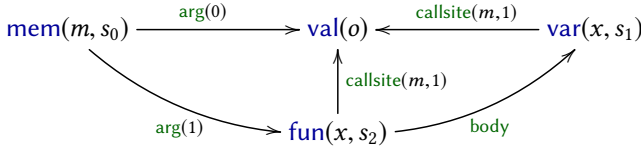


Fig. 14. Dependency graph for  $o.m(\lambda x \rightarrow x)$  with a newly added `callsite` edges.

to add instant type checking by using the dependency graph construction described in Section 4. Second, we consider an abstraction mechanism that preserves the live coding experience.

### 7.1 Type checking

Instant live previews give analysts quick feedback when they write incorrect code, but having type information is still valuable. First, it can help give better error messages. Second, types can be used to provide auto-complete – when the user types ‘.’ we can offer available members without having to evaluate the wait until the value of the object is available.

**7.1.1 Revised dependency graph.** Type checking of small programs is typically fast enough that no caching is necessary. However, The Gamma supports *type providers* [Christiansen 2013; Syme et al. 2013], which can generate types based on an external file or a REST service call, e.g. [Petricek et al. 2016]. For this reason, type checking can be relatively time consuming and can benefit from the same caching facilities as those available for live previews.

Adding type checking requires revising the way we construct the dependency graph introduced in Section 4. Previously, a variable bound by a lambda function had no dependencies. However, the type of the variable depends on the context in which it appears. Given an expression  $o.m(\lambda x \rightarrow x)$ , we infer the type of  $x$  from the type of the first argument of the member  $m$ . A variable node for  $x$  thus needs to depend on the call site of  $m$ . We capture that by adding an edge `callsite`( $m, i$ ) from  $x$  to  $o$  which indicates that  $x$  is the input variable of a function passes as the  $i^{\text{th}}$  argument to the  $m$  member of the expression represented by the target node. We also add `callsite`( $m, i$ ) as an edge from the node of the function. Figure 14 shows the revised dependency graph for  $o.m(\lambda x \rightarrow x)$ .

**7.1.2 Type checking.** The structure of typing rules is similar to the evaluation relation  $v \Downarrow d$  defined earlier. Given a dependency graph  $(V, E)$ , we define typing judgements in the form  $v \vdash \tau$ . The type  $\tau$  can be a primitive type, a function  $\tau \rightarrow \tau$  or an object type  $\{m_1 : \sigma_1, \dots, m_n : \sigma_n\}$  with member types  $\sigma = (\tau_1, \dots, \tau_n) \rightarrow \tau$ .

The typing rules for variables, functions and member access are shown in Figure 15. When type checking a member access (`mem`), we find its dependencies  $v_i$  and check that the instance is an object with the required member  $m$ . The types of arguments of the member then need to match the types of the remaining (non-instance) nodes. Type checking a function (`fun`) and a variable (`var`) is similar. In both cases, we follow the `callsite` edge to find the member that accepts the function as an argument and use the type of the argument to check the type of the function or infer the type of the variable.

The results of type checking can be cached and reused in the same way as live previews, although we leave out the details. A property akin to correctness (Theorem 6) requires defining standard type checking over the structure of expressions, which we also omit for space reasons.

## 7.2 Feedback-friendly abstraction

The data analysis by Financial Times in Section 2.1 illustrates why notebook users often avoid abstraction. Wrapping code in a function makes it impossible to split code into cells and see results of intermediate steps. Instead, the analysis used a global variable with possible values in a comment.

Providing live previews inside ordinary functions is problematic, because we do not have readily available values for input parameters and our mechanism for lambda functions only provides delayed previews inside body of a function. We believe that extending the data exploration calculus with an abstraction mechanism that would support development with instant feedback is an interesting design problem and we briefly outline a possible solution here.

Data scientists often write code interactively using a sample data set and, when it works well, wrap it into a function that they then call on other data. Similarly, spreadsheet users often write equation in the first row of a table and then use the “drag down” operation to apply it to other rows. One way of adding similar functionality to the data exploration calculus is to label a sequence of commands such that the sequence can be reused later with different inputs:

$$\begin{aligned} p &= c_1; \dots; c_n \\ c &= \text{let } x = t \mid t \mid \text{lbl}: p \mid \text{lbl} \end{aligned}$$

We introduce two new kinds of commands: a labelled sequence of commands and a reference to a label. When evaluating, the command *lbl* is replaced with the associated sequence of commands *p* before any other reductions. Consequently, variables used in the labelled block are dynamically scoped and we can use let binding to redefine a value of a variable before invoking the block repeatedly. The correct use of dynamic scoping can be checked using coeffs [Petricek 2017a].

This minimalistic abstraction mechanism supports code reuse without affecting how live previews are computed. Commands in a labelled block require variables to be defined before the block. Those define sample data for development and can be redefined before reusing the block. We intend to implement this mechanism in a future version of our data exploration environment (Section 6.4).

## 8 RELATED AND FUTURE WORK

We aim to make creating transparent data analyses, such as those done by journalists, easier and we present a tool that provides instant feedback when writing code in a text editor. Many of our design decisions have alternatives that have been explored in a rich body of literature.

*Notebooks and data science tools.* Basic feedback mechanism, the REPL (read-eval-print-loop), originated in Lisp [McCarthy 1978] and is now common in many languages [Findler et al. 2002]. Notebook systems such as Jupyter and R markdown [Allaire et al. 2016; Kluyver et al. 2016; Pérez and Granger 2007] structure code in cells, but still rely on explicitly executing individual cells using a stateful REPL environment. Integrating provenance tracking with notebooks [Koop and Patel

$$\begin{aligned} (\text{mem}) \quad & \frac{\forall i \in \{0 \dots k\}. (\text{mem}(m, s), v_i, \text{arg}(i)) \in E \quad v_0 \vdash \{.., m : (\tau_1, \dots, \tau_k) \rightarrow \tau, ..\} \quad v_i \vdash \tau_i}{\text{mem}(m, s) \vdash \tau} \\ (\text{var}) \quad & \frac{(\text{var}(x, s), v, \text{callsite}(m, i)) \in E \quad v \vdash \{.., m : (\tau_1, \dots, \tau_k) \rightarrow \tau, ..\} \quad \tau_i = \tau' \rightarrow \tau''}{\text{var}(x, s) \vdash \tau'} \\ (\text{fun}) \quad & \frac{\{(\text{fun}(x, s), v_b, \text{body}), (\text{var}(x, s), v_c, \text{callsite}(m, i))\} \subseteq E \quad v_c \vdash \{.., m : (\tau_1, \dots, \tau_k) \rightarrow \tau, ..\} \quad \tau_i = \tau' \rightarrow \tau'' \quad v_b \vdash \tau''}{\text{fun}(x, s) \vdash \tau' \rightarrow \tau''} \end{aligned}$$

Fig. 15. Rules that define type checking of terms and expressions over a dependency graph  $(V, E)$



2017; Petricek et al. 2018; Pimentel et al. 2015] addresses the reproducibility issues of this model, but keeps cells as units of work. Ideas such as dependency tracking and efficient recomputation exist in many visual data exploration tools [Crotty et al. 2015; Hellerstein et al. 1999; Wesley et al. 2011] and scientific workflow systems [Altintas et al. 2004; Oinn et al. 2004]. Some of the implementation techniques are related to our work, but we focus on text-based scripts. Tempe [DeLine et al. 2015] focuses on streaming data, but provides a text-based scripting environment with automatic code update and DeLine and Fisher [2015] empirically evaluate its usability in contrast to REPLs. The challenges faced by Tempe are more akin to dynamic software update [Hicks et al. 2001].

*Live coding and editing.* Live coding based on textual programs has been popularised by Victor [2012a,b] and is actively developed in domains such as live coded music [Aaron and Blackwell 2013; Roberts et al. 2015]. Kubelka et al. [2018] reviews the use of live coding in a Smalltalk derived environment. Lighttable by Granger [2012] and Chrome DevTools provide limited live previews akin to those presented in this paper, but without well specified recomputation model. Finally, work on keeping state during code edits [Burckhardt et al. 2013; McDirmid 2007] would be relevant for supporting streaming data. In the future, we also plan to explore how direct manipulation [Shneiderman 1981] could be integrated with our preview mechanism. Kubelka et al. [2018]

Using structured editors [Szwilius and Neal 1996] rather than text editors is a promising alternative direction as editors can recompute program based on understanding of edit operations [Lotem and Chuchem 2018; Omar et al. 2019; Perera 2013]. This can be elegantly implemented using bi-directional lambda calculus [Omar et al. 2017], but it also makes us consider more human-centric abstractions [Edwards 2005, 2018] along the lines of the one outlined in Section 7.2.

*Incremental computation and dependency analysis.* Work on self-adjusting and incremental computation [Acar 2005; Hammer et al. 2014] handles recomputation when the program stays the same, but input changes. It relies on dependency tracking similar to ours. We rely on static code analysis whereas most incremental systems, e.g. [Acar et al. 2006; Acar and Ley-Wild 2009; Hammer et al. 2015] evaluate the program and use programmer-supplied information to build a dependency graph. Hammer et al. [2014] implement a small adaptive interpreter that treats code as changing input data, suggesting an interesting implementation technique for live coding systems. Our use of dependency graphs [Kuck et al. 1981] is static and first-order and can be seen as a form of program slicing [Weiser 1981], although our binding process is more directly inspired by Roslyn [Ng et al. 2011], which uses it for efficient background type-checking.

*Semantics and partial evaluation.* The evaluation of previews is a form of partial evaluation [Consel and Danvy 1993], done in a way that allows caching. This can be done implicitly or explicitly in the form of multi-stage programming [Taha and Sheard 2000]. Semantically, the evaluation of previews can be seen as a modality [Fairtlough et al. 2001] and delayed previews are linked to contextual modal type theory [Nanevski et al. 2008], formally modelled using comonads [Gabbay and Nanevski 2013]. This suggests a direction for rigorous analysis of the presented system.

## 9 SUMMARY

One of the reasons that make spreadsheets easier to use than programming tools is that they provide instant feedback. We aim to make programming tools as instant as spreadsheets. We formalized a number of key aspects of simple data analyses such as those don by journalists and then used our observations to build both theory and simple practical data analytics tools.

Our *data exploration calculus* is a simple formally tractable language for data exploration. The calculus captures key observations about simple data analyses. They rely on logic from external

libraries, implement few abstractions and are written as lists of commands. We later revised the calculus to add type checking and simple abstraction mechanism.

Our main technical contribution is a *live preview* mechanism that efficiently evaluates code during editing and instantly provides a preview of the result. We allow users to edit code in unconstrained way in a text editor, which makes this particularly challenging. The key trick is to separate the process into fast *binding phase*, which constructs a dependency graph and slower *evaluation phase* that can cache results. This makes it possible to quickly parse updated code, reconstruct dependency graph and compute preview using previous, partially evaluated, results.

We evaluated our approach in three ways. First, we proved that our mechanism is correct and that it reuses evaluated values for many common code edit operations. Second, we conducted an experimental study that illustrates how often are previously evaluated results reused during typical programming scenario. Thirdly, we used our research as a basis for online data exploration environment, which shows the practical usability of our work.

## REFERENCES

- Samuel Aaron and Alan F Blackwell. 2013. From sonic Pi to overtone: creative musical experiences with domain-specific and functional languages. In *Proceedings of the first ACM SIGPLAN workshop on Functional art, music, modeling & design*. ACM, 35–46.
- Umut A. Acar. 2005. *Self-adjusting Computation*. Ph.D. Dissertation. Pittsburgh, PA, USA. AAI3166271.
- Umut A. Acar, Guy E. Blelloch, and Robert Harper. 2006. Adaptive Functional Programming. *ACM Trans. Program. Lang. Syst.* 28, 6 (Nov. 2006), 990–1034. <https://doi.org/10.1145/1186632.1186634>
- Umut A. Acar and Ruy Ley-Wild. 2009. *Self-adjusting Computation with Delta ML*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1–38. [https://doi.org/10.1007/978-3-642-04652-0\\_1](https://doi.org/10.1007/978-3-642-04652-0_1)
- J Allaire, Joe Cheng, Yihui Xie, Jonathan McPherson, Winston Chang, Jeff Allen, Hadley Wickham, Aron Atkins, Rob Hyndman, and R Arslan. 2016. rmarkdown: Dynamic Documents for R. *R package version 1* (2016), 9010.
- Ilkay Altintas, Chad Berkley, Efrat Jaeger, Matthew Jones, Bertram Ludascher, and Steve Mock. 2004. Kepler: an extensible system for design and execution of scientific workflows. In *Scientific and Statistical Database Management*. IEEE, 423–424.
- David Blood. 2018. Recycling is broken – notebooks. Retrieved February 1, 2019 from <https://github.com/ft-interactive/recycling-is-broken-notebooks>
- Sebastian Burckhardt, Manuel Fahndrich, Peli de Halleux, Sean McDirmid, Michal Moskal, and Nikolai Tillmann. 2013. It’s Alive! Continuous Feedback in UI Programming. In *PLDI*. ACM SIGPLAN.
- David Raymond Christiansen. 2013. Dependent type providers. In *Proceedings of the 9th ACM SIGPLAN workshop on Generic programming*. ACM, 25–34.
- Charles Consel and Olivier Danvy. 1993. Tutorial notes on partial evaluation. In *Proceedings of the 20th ACM SIGPLAN-SIGACT symposium on Principles of programming languages*. ACM, 493–501.
- Andrew Crotty, Alex Galakatos, Emanuel Zraggen, Carsten Binnig, and Tim Kraska. 2015. Vizdom: Interactive Analytics Through Pen and Touch. *Proc. VLDB Endow.* 8, 12 (Aug. 2015), 2024–2027. <https://doi.org/10.14778/2824032.2824127>
- R. DeLine and D. Fisher. 2015. Supporting exploratory data analysis with live programming. In *2015 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. 111–119. <https://doi.org/10.1109/VLHCC.2015.7357205>
- R. DeLine, D. Fisher, B. Chandramouli, J. Goldstein, M. Barnett, J. Terwilliger, and J. Wernsing. 2015. Tempe: Live scripting for live data. In *2015 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. 137–141. <https://doi.org/10.1109/VLHCC.2015.7357208>
- Jonathan Edwards. 2005. Subtext: uncovering the simplicity of programming. *ACM SIGPLAN Notices* 40, 10 (2005), 505–518.
- Jonathan Edwards. 2018. Direct Programming. Retrieved February 9, 2019 from <https://vimeo.com/274771188>
- Matt Fairtlough, Michael Mendler, and Eugenio Moggi. 2001. Special issue: Modalities in type theory. *Mathematical Structures in Computer Science* 11, 4 (2001), 507–509.
- Robert Bruce Findler, John Clements, Cormac Flanagan, Matthew Flatt, Shriram Krishnamurthi, Paul Steckler, and Matthias Felleisen. 2002. DrScheme: A programming environment for Scheme. *Journal of functional programming* 12, 2 (2002), 159–182.
- Murdoch J Gabbay and Aleksandar Nanevski. 2013. Denotation of contextual modal type theory (CMTT): Syntax and meta-programming. *Journal of Applied Logic* 11, 1 (2013), 1–29.
- Marco Gaboardi, Shin-ya Katsumata, Dominic A Orchard, Flavien Breuvar, and Tarmo Uustalu. 2016. Combining effects and coeffects via grading. In *ICFP*. 476–489.

- Chris Granger. 2012. LightTable: A new IDE concept. <http://www.chris-granger.com/2012/04/12/light-table-a-new-ide-concept/> (2012).
- Jonathan Gray, Lucy Chambers, and Liliana Bounegru. 2012. *The data journalism handbook: how journalists can use data to improve the news*. "O'Reilly Media, Inc".
- P Guo. 2013. Data science workflow: Overview and challenges. *blog CACM, Communications of the ACM* (2013).
- Matthew A. Hammer, Joshua Dunfield, Kyle Headley, Nicholas Labich, Jeffrey S. Foster, Michael Hicks, and David Van Horn. 2015. Incremental Computation with Names. In *Proceedings of the 2015 ACM SIGPLAN International Conference on Object-Oriented Programming, Systems, Languages, and Applications (OOPSLA 2015)*. ACM, New York, NY, USA, 748–766. <https://doi.org/10.1145/2814270.2814305>
- Matthew A. Hammer, Khoo Yit Phang, Michael Hicks, and Jeffrey S. Foster. 2014. Adapton: Composable, Demand-driven Incremental Computation. In *Proceedings of the 35th ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI '14)*. ACM, New York, NY, USA, 156–166. <https://doi.org/10.1145/2594291.2594324>
- J. M. Hellerstein, R. Avnur, A. Chou, C. Hidber, C. Olston, V. Raman, T. Roth, and P. J. Haas. 1999. Interactive data analysis: the Control project. *Computer* 32, 8 (Aug 1999), 51–59. <https://doi.org/10.1109/2.781635>
- Michael Hicks, Jonathan T. Moore, and Scott Nettles. 2001. Dynamic Software Updating. In *Proceedings of the ACM SIGPLAN 2001 Conference on Programming Language Design and Implementation (PLDI '01)*. ACM, New York, NY, USA, 13–23. <https://doi.org/10.1145/378795.378798>
- Leslie Hook and John Reed. 2018. Why the world's recycling system stopped working. Retrieved February 1, 2019 from <https://www.ft.com/content/360e2524-d71a-11e8-a854-33d6f82e62f8>
- Sean Kandel, Andreas Paepcke, Joseph Hellerstein, and Jeffrey Heer. 2011. Wrangler: Interactive Visual Specification of Data Transformation Scripts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, 3363–3372. <https://doi.org/10.1145/1978942.1979444>
- Ken Kennedy and John R Allen. 2001. *Optimizing compilers for modern architectures: a dependence-based approach*. Morgan Kaufmann Publishers Inc.
- Thomas Kluyver, Benjamin Ragan-Kelley, Fernando Pérez, Brian E Granger, Matthias Bussonnier, Jonathan Frederic, Kyle Kelley, Jessica B Hamrick, Jason Grout, Sylvain Corlay, et al. 2016. Jupyter Notebooks—a publishing format for reproducible computational workflows.. In *ELPUB*. 87–90.
- David Koop and Jay Patel. 2017. Dataflow Notebooks: Encoding and Tracking Dependencies of Cells. In *9th USENIX Workshop on the Theory and Practice of Provenance, TaPP 2017, Seattle, WA, USA, June 23, 2017*.
- Juraj Kubelka, Romain Robbes, and Alexandre Bergel. 2018. The Road to Live Programming: Insights from the Practice. In *Proceedings of the 40th International Conference on Software Engineering (ICSE '18)*. ACM, New York, NY, USA, 1090–1101. <https://doi.org/10.1145/3180155.3180200>
- David J Kuck, Robert H Kuhn, David A Padua, Bruce Leasure, and Michael Wolfe. 1981. Dependence graphs and compiler optimizations. In *Proceedings of the 8th ACM SIGPLAN-SIGACT symposium on Principles of programming languages*. ACM, 207–218.
- Eyal Lotem and Yair Chuchem. 2018. Lamdu Project. <https://github.com/lamdu/lamdu> (2018).
- John McCarthy. 1978. History of LISP. *SIGPLAN Not.* 13, 8 (Aug. 1978), 217–223. <https://doi.org/10.1145/960118.808387>
- Sean McDirmid. 2007. Living it up with a live programming language. In *ACM SIGPLAN Notices*, Vol. 42. ACM, 623–638.
- Wes McKinney. 2012. *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*. O'Reilly Media, Inc.
- Robin Milner. 1978. A theory of type polymorphism in programming. *J. Comput. System Sci.* 17, 3 (1978), 348 – 375. [https://doi.org/10.1016/0022-0000\(78\)90014-4](https://doi.org/10.1016/0022-0000(78)90014-4)
- Alan Mycroft, Dominic Orchard, and Tomas Petricek. 2016. Effect systems revisited—control-flow algebra and semantics. In *Semantics, Logics, and Calculi*. Springer, 1–32.
- Aleksandar Nanevski, Frank Pfenning, and Brigitte Pientka. 2008. Contextual modal type theory. *ACM Transactions on Computational Logic (TOCL)* 9, 3 (2008), 23.
- Karen Ng, Matt Warren, Peter Golde, and Anders Hejlsberg. 2011. The Roslyn Project, Exposing the C# and VB compiler's code analysis. *White paper, Microsoft* (2011).
- Tom Oinn, Matthew Addis, Justin Ferris, Darren Marvin, Martin Senger, Mark Greenwood, Tim Carver, Kevin Glover, Matthew R Pocock, Anil Wipat, et al. 2004. Taverna: a tool for the composition and enactment of bioinformatics workflows. *Bioinformatics* 20, 17 (2004), 3045–3054.
- Cyrus Omar, Ian Voysey, Ravi Chugh, and Matthew A. Hammer. 2019. Live Functional Programming with Typed Holes. *PACMPL* 3, POPL (2019). <https://doi.org/10.1145/3290327>
- Cyrus Omar, Ian Voysey, Michael Hilton, Jonathan Aldrich, and Matthew A Hammer. 2017. Hazelnut: a bidirectionally typed structure editor calculus. In *Proceedings of the 44th ACM SIGPLAN Symposium on Principles of Programming Languages*. ACM, 86–99.
- Roland Perera. 2013. *Interactive functional programming*. Ph.D. Dissertation. University of Birmingham.

- Fernando Pérez and Brian E Granger. 2007. IPython: a system for interactive scientific computing. *Computing in Science & Engineering* 9, 3 (2007).
- Tomas Petricek. 2017a. *Context-aware programming languages*. Ph.D. Dissertation. University of Cambridge, Computer Laboratory.
- Tomas Petricek. 2017b. Data Exploration through Dot-driven Development. In *LIPICs-Leibniz International Proceedings in Informatics*, Vol. 74. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.
- Tomas Petricek, James Geddes, and Charles A. Sutton. 2018. Wrattler: Reproducible, live and polyglot notebooks. In *10th USENIX Workshop on the Theory and Practice of Provenance, TaPP 2018, London, UK, July 11-12, 2018*.
- Tomas Petricek, Gustavo Guerra, and Don Syme. 2016. Types from data: Making structured data first-class citizens in F. *ACM SIGPLAN Notices* 51, 6 (2016), 477–490.
- João Felipe Nicolaci Pimentel, Vanessa Braganholo, Leonardo Murta, and Juliana Freire. 2015. Collecting and analyzing provenance on interactive notebooks: when IPython meets noWorkflow. In *Workshop on the Theory and Practice of Provenance (TaPP)*. 155–167.
- Charles Roberts, Matthew Wright, and JoAnn Kuchera-Morin. 2015. Beyond editing: extended interaction with textual code fragment. In *NIME*. 126–131.
- Peter Sestoft. 2012. *Spreadsheet technology*. Technical Report. Citeseer.
- Ben Shneiderman. 1981. Direct manipulation: A step beyond programming languages. In *ACM SIGSOC Bulletin*, Vol. 13. ACM, 143.
- Don Syme. 2006. Leveraging .net meta-programming components from F#: integrated queries and interoperable heterogeneous execution. In *Proceedings of the 2006 workshop on ML*. ACM, 43–54.
- Donald Syme, Keith Battocchi, Kenji Takeda, Donna Malayeri, and Tomas Petricek. 2013. Themes in information-rich functional programming for internet-scale data sources. In *Proceedings of the 2013 workshop on Data driven functional programming*. ACM, 1–4.
- Gerd Szwillus and Lisa Neal. 1996. *Structure-based editors and environments*. Academic Press, Inc.
- Walid Taha and Tim Sheard. 2000. MetaML and multi-stage programming with explicit annotations. *Theoretical computer science* 248, 1 (2000), 211–242.
- Bret Victor. 2012a. Inventing on Principle. <http://worrydream.com/InventingOnPrinciple> (2012).
- Bret Victor. 2012b. Learnable programming: Designing a programming system for understanding programs. <http://worrydream.com/LearnableProgramming> (2012).
- Mark Weiser. 1981. Program slicing. In *Proceedings of the 5th international conference on Software engineering*. IEEE Press, 439–449.
- Richard Wesley, Matthew Eldridge, and Pawel T. Terlecki. 2011. An Analytic Data Engine for Visualization in Tableau. In *Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data (SIGMOD '11)*. ACM, 1185–1194. <https://doi.org/10.1145/1989323.1989449>

## A APPENDIX

### A.1 Normalization

**Theorem 10** (Normalization). *For all  $p$ , there exists  $n$  and  $o_1, \dots, o_n$  such that  $p \rightsquigarrow^* o_1; \dots; o_n$ .*

PROOF. We define size of a program in data exploration calculus as follows:

$$\begin{aligned}
 \text{size}(c_1; \dots; c_n) &= 1 + \sum_{i=1}^n \text{size}(c_i) \\
 \text{size}(\text{let } x = t) &= 1 + \text{size}(t) \\
 \text{size}(e_0.m(e_1, \dots, e_n)) &= 1 + \sum_{i=0}^n \text{size}(e_i) \\
 \text{size}(\lambda x \rightarrow e) &= 1 + \text{size}(e) \\
 \text{size}(o) = \text{size}(x) &= 1
 \end{aligned} \tag{1}$$

The property holds because, first, both (let) and (external) decrease the size of the program and, second, a program is either fully evaluated, i.e.  $o_1; \dots; o_n$  for some  $n$  or, it can be reduced using one of the reduction rules.  $\square$

### A.2 Term preview correctness

**Theorem 11** (Term preview correctness). *Given a term  $t$  that has no free variables, together with a lookup table  $\Delta$  obtained from any sequence of programs using bind-prog (Figure 6) and update (Figure 7), then let  $\mathbf{v}, (V, E) = \text{bind-expr}_{\emptyset, \Delta}(t)$ . If  $\mathbf{v} \Downarrow p$  over a graph  $(V, E)$  then  $p = o$  for some value  $o$  and  $t \rightsquigarrow^* o$ .*

PROOF. First note that, when combining recursively constructed sub-graphs, the bind-expr function adds new nodes and edges leading from those new nodes. Therefore, an evaluation using  $\Downarrow$  over a sub-graph will also be valid over the new graph – the newly added nodes and edges do not introduce non-determinism to the rules given in Figure 8.

We prove a more general property showing that for any  $e$ , its binding  $\mathbf{v}, (V, E) = \text{bind-expr}_{\emptyset, \Delta}(e)$  and any evaluation context  $C$  such that  $C[e] \rightsquigarrow o$  for some  $o$ , one of the following holds:

- If  $FV(e) = \emptyset$  then  $\mathbf{v} \Downarrow p$  for some  $p$  and  $C[p] \rightsquigarrow o$
- If  $FV(e) \neq \emptyset$  then  $\mathbf{v} \Downarrow \llbracket e_p \rrbracket_{FV(e)}$  for some  $e_p$  and  $C[e_p] \rightsquigarrow o$

In the first case,  $p$  is a value, but it is not always the case that  $e \rightsquigarrow^* p$ , because  $p$  may be lambda function and preview evaluation may reduce sub-expression in the body of the function. Using a context  $C$  in which the value reduces to an object avoids this problem.

The proof of the theorem follows from the more general property. Using a context  $C[-] = -$ , the term  $t$  reduces  $t \rightsquigarrow^* t' \rightsquigarrow_{\epsilon} o$  for some  $o$  and the preview  $p$  is a value  $o$  because  $C[p] = p = o$ . The proof is by induction over the binding process, which follows the structure of the expression:

(1)  $\text{bind-expr}_{\Gamma, \Delta}(e_0.m(e_1, \dots, e_n))$  – Here  $e = e_0.m(e_1, \dots, e_n)$ ,  $\mathbf{v}_i$  are graph nodes obtained by induction for expressions  $e_i$  and  $\{(\mathbf{v}, \mathbf{v}_0, \text{arg}(0)), \dots, (\mathbf{v}, \mathbf{v}_n, \text{arg}(n))\} \subseteq E$ . From lookup inversion Lemma 4,  $\mathbf{v} = \text{mem}(m, s)$  for some  $s$ .

If  $FV(e) = \emptyset$ , then  $\mathbf{v}_i \Downarrow p_i$  for  $i \in 0 \dots n$  and  $\mathbf{v} \Downarrow p$  using (mem-val) such that  $p_0.m(p_1, \dots, p_n) \rightsquigarrow p$ . From induction hypothesis and *compositionality* of external libraries (Definition 2), it holds that for any  $C$  such that  $C[e_0.m(e_1, \dots, e_n)] \rightsquigarrow o$  for some  $o$  then also  $C[p_0.m(p_1, \dots, p_n)] \rightsquigarrow C[p] \rightsquigarrow o$ .

If  $FV(e) \neq \emptyset$ , then  $\mathbf{v}_i \Downarrow_{\text{lift}} \llbracket e'_i \rrbracket$  for  $i \in 0 \dots n$  and  $\mathbf{v} \Downarrow \llbracket e'_0.m(e'_1, \dots, e'_n) \rrbracket_{FV(e)}$  using (mem-expr). From induction hypothesis and *compositionality* of external libraries (Definition 2), it holds that for any  $C$  such that  $C[e_0.m(e_1, \dots, e_n)] \rightsquigarrow o$  for some  $o$  then also  $C[e'_0.m(e'_1, \dots, e'_n)] \rightsquigarrow o$ .

(2)  $\text{bind-expr}_{\Gamma, \Delta}(e_0.m(e_1, \dots, e_n))$  – This case is similar to (1), except that the fact that  $\mathbf{v} = \text{mem}(m, s)$  holds by construction, rather than using Lemma 4.

(3)  $\text{bind-expr}_{\Gamma, \Delta}(\lambda x \rightarrow e_b)$  – Here  $e = \lambda x \rightarrow e_b$ ,  $v_b$  is the graph node obtained by induction for the expression  $e_b$  and  $(v, v_b, \text{body}) \in E$ . From lookup inversion Lemma 4,  $v = \text{fun}(x, s)$  for some  $s$ . The evaluation can use one of three rules:

- i. If  $FV(e) = \emptyset$  then  $v_b \Downarrow p_b$  for some  $p_b$  and  $v \Downarrow \lambda x \rightarrow p_b$  using (fun-val). Let  $e'_b = p_b$ .
- ii. If  $FV(e_b) = \{x\}$  then  $v_b \Downarrow \llbracket e'_b \rrbracket_x$  for some  $e'_b$  and  $v \Downarrow \lambda x \rightarrow e'_b$  using (fun-bind).
- iii. Otherwise,  $v_b \Downarrow \llbracket e'_b \rrbracket_{x, \Gamma}$  for some  $e'_b$  and  $v \Downarrow \llbracket \lambda x \rightarrow e'_b \rrbracket_{\Gamma}$  using (fun-expr).

For i.) and ii.) we show that a.) is the case; for iii.) we show that b.) is the case; that is for any  $C$ , if  $C[\lambda x \rightarrow e_b] \rightsquigarrow o$  then also  $C[\lambda x \rightarrow e'_b] \rightsquigarrow o$ . For a given  $C$ , let  $C'[-] = C[\lambda x \rightarrow -]$  and use the induction hypothesis, i.e. if  $C'[e_b] \rightsquigarrow o$  for some  $o$  then also  $C'[e'_b] \rightsquigarrow o$ .

(4)  $\text{bind-expr}_{\Gamma, \Delta}(\lambda x \rightarrow e)$  – This case is similar to (3), except that the fact that  $v = \text{fun}(x, s)$  holds by construction, rather than using Lemma 4.

(5)  $\text{bind-expr}_{\Gamma, \Delta}(o)$  – In this case  $e = o$  and  $v = \text{val}(o)$  and  $\text{val}(o) \Downarrow o$  using (val) and so the case a.) trivially holds.

(6)  $\text{bind-expr}_{\Gamma, \Delta}(x)$  – The initial  $\Gamma$  is empty, so  $x$  must have been added to  $\Gamma$  by case (3) or (4). Hence,  $v = \text{var}(x)$ ,  $v \Downarrow \llbracket x \rrbracket_x$  using (var) and so  $e_p = e = x$  and the case b.) trivially holds.

□