

An Order-Aware Dataflow Model for Parallel Unix Pipelines

SHIVAM HANDA*, CSAIL, MIT, USA

KONSTANTINOS KALLAS*, University of Pennsylvania, USA

NIKOS VASILAKIS*, CSAIL, MIT, USA

MARTIN C. RINARD, CSAIL, MIT, USA

We present a dataflow model for modelling parallel UNIX shell pipelines. To accurately capture the semantics of complex UNIX pipelines, the dataflow model is order-aware, *i.e.*, the order in which a node in the dataflow graph consumes inputs from different edges plays a central role in the semantics of the computation and therefore in the resulting parallelization. We use this model to capture the semantics of transformations that exploit data parallelism available in UNIX shell computations and prove their correctness. We additionally formalize the translations from the UNIX shell to the dataflow model and from the dataflow model back to a parallel shell script. We implement our model and transformations as the compiler and optimization passes of a system parallelizing shell pipelines, and use it to evaluate the speedup achieved on 47 pipelines.

CCS Concepts: • **Software and its engineering** → **Compilers; Massively parallel systems; Scripting languages.**

Additional Key Words and Phrases: UNIX, POSIX, Shell, Parallelism, Dataflow, Order-awareness

ACM Reference Format:

Shivam Handa, Konstantinos Kallas, Nikos Vasilakis, and Martin C. Rinard. 2021. An Order-Aware Dataflow Model for Parallel Unix Pipelines. *Proc. ACM Program. Lang.* 5, ICFP, Article 65 (August 2021), 28 pages. <https://doi.org/10.1145/3473570>

1 INTRODUCTION

UNIX pipelines are an attractive choice for specifying succinct and simple programs for data processing, system orchestration, and other automation tasks [McIlroy et al. 1978]. Consider, for example, the following program based on the original `spell` written by Johnson [Bentley 1985], lightly modified for modern environments:¹

```
cat f1.md f2.md | tr A-Z a-z | tr -cs A-Za-z '\n' | sort | uniq | # (Spell)
grep -vx -f dict.txt - > out ; cat out | wc -l | sed 's/$/ misspelled words!/'
```

The first command streams two markdown files into a pipeline that converts characters in the stream into lower case, removes punctuation, sorts the stream in alphabetical order, removes duplicate words, and filters out words from a dictionary file (lines 1 and 2, up to “;”). A second

*Equal contribution.

¹Johnson’s program additionally used `tr off`, `prepare`, and `col -bx` to clean up now-legacy formatting metadata that does not exist in markdown. Moreover, `comm -13` was replaced with `grep -xvf` to highlight crucial features of the model.

Authors’ addresses: Shivam Handa, MIT Computer Science & Artificial Intelligence Laboratory, 32 Vassar St, Cambridge, MA 02139, United States; Konstantinos Kallas, University of Pennsylvania, Philadelphia, PA 19104, USA; Nikos Vasilakis, MIT Computer Science & Artificial Intelligence Laboratory, 32 Vassar St, Cambridge, MA 02139, United States; and Martin C. Rinard, MIT Computer Science & Artificial Intelligence Laboratory, 32 Vassar St, Cambridge, MA 02139, United States.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2021 Copyright held by the owner/author(s).

2475-1421/2021/8-ART65

<https://doi.org/10.1145/3473570>

pipeline (line 2, after “;”) counts the resulting lines to report the number of misspelled words to the user.

As this example illustrates, the UNIX shell offers a programming model that facilitates the composition of commands using unidirectional communication channels that feed the output of one command as an input to another. These channels are either ephemeral, unnamed pipes expressed using the `|` character and lasting for the duration of the producer and consumer, or persistent, named pipes (UNIX FIFOs) created with `mkfifo` and lasting until explicitly deleted. Each command executes sequentially, with pipelined parallelism available between commands executing in the same pipeline. Unfortunately, this model leaves substantial data parallelism, *i.e.*, parallelism achieved by splitting inputs into pieces and feeding the pieces to parallel instances of the script, unexploited.

This fact is known in the UNIX community and has motivated the development of a variety of tools that attempt to exploit latent data parallelism in UNIX scripts [Raghavan et al. 2020; Tange 2011; Vasilakis et al. 2021]. On the one hand, tools such as GNU Parallel [Tange 2011] can be used by experienced users to achieve parallelism, but could also easily lead to incorrect results. On the other hand, two recent systems, PaSh [Vasilakis et al. 2021] and POSH [Raghavan et al. 2020], focus respectively on extracting data parallelism latent in UNIX pipelines to improve the execution time of (i) CPU-intensive shell scripts (PaSh), and (ii) networked IO-intensive shell scripts (POSH). These systems achieve order-of-magnitude performance improvements on sequential pipelines, but their semantics and associated transformations are not clearly defined, making it difficult to ensure that the optimized parallel scripts are sound with respect to the sequential ones.

To support the ability to reason about and correctly transform UNIX shell pipelines, we present a new *dataflow model*. In contrast to standard dataflow models [Kahn 1974; Kahn and MacQueen 1977; Karp and Miller 1966; Lee and Messerschmitt 1987a,b], our dataflow model is *order-aware*—*i.e.*, the order in which a node in the dataflow graph consumes inputs from different edges plays a central role in the semantics of the computation and therefore in the resulting parallelization. This model is different from models that allow multiplexing different chunks of data in a single channel, such as sharding or tagging, or ones that are oblivious to ordering, such as shuffling—and is a direct byproduct of the ordered semantics of the shell and the opacity of UNIX commands. In the *Spell* script shown earlier, for example, while all commands consume elements from an input stream in order—a property of UNIX streams *e.g.*, pipes and FIFOs—they differ in how they consume across streams: `cat` reads input streams in the order of its arguments, `sort -m` reads input streams in interleaved fashion, and `grep -vx -f` first reads `dict.txt` before reading from its standard input.

We use this order-aware dataflow model (ODFM) to express the semantics of transformations that exploit data parallelism available in UNIX shell computations. These transformations capture the parallelizing optimizations performed by both PaSh [Vasilakis et al. 2021] and POSH [Raghavan et al. 2020]. We also use our model to prove that these transformations are correct, *i.e.*, that they do not affect the program behavior with respect to sequential output. Finally, we formalize bidirectional translations between the shell and ODFM, namely from the shell language to ODFM and *vice versa*, closing the loop for a complete shell-to-shell parallelizing compiler.

To illustrate the applicability of our model, we extend PaSh by reimplementing its compilation and optimization phases with ODFM as the centerpiece. The new implementation translates fragments of a shell script to our dataflow model, applies a series of parallelizing transformations, and then translates the resulting parallel dataflow graph back to a parallel shell script that is executed instead of the original fragment. The implementation improves modularity, facilitates the development

of different transformations independently on top of ODFM, has been incorporated into the PaSh mainline, and is available as part of the PaSh open-source repository.²

We use the new implementation to evaluate the benefit of a specific transformation by parallelizing 47 unmodified shell scripts with and without this transformation and measuring their execution times. Finally, we present a case study in which we parallelize two scripts using GNU Parallel [Tange 2011]; our experience indicates that while it is easy to parallelize shell scripts using such a tool, it is also easy to introduce bugs, leading to incorrect results.

In summary, this paper makes the following contributions:

- **Order-Aware Dataflow Model:** It introduces the *order-aware dataflow model* (ODFM), a dataflow model tailored to the UNIX shell that captures information about the order in which nodes consume inputs from different input edges (§4).
- **Translations:** It formalizes the bidirectional translations between shell and ODFM required to translate shell scripts into dataflow graphs and dataflow graphs back to parallel scripts (§5).
- **Transformations and Proofs of Correctness:** It presents a series of ODFM transformations for extracting data parallelism. It also presents proofs of correctness for these transformations (§6).
- **Results:** It reimplements the optimizing compiler of PaSh and presents experimental results that evaluate the speedup afforded by a specific dataflow transformation (§7).

The paper starts with an informal development building the necessary background (§2) and expounding on *Spell* (§3). It then presents the three main contributions outlined above (§4–7), compares with prior work (§8), and offers a discussion (§9), before closing the paper (§10).

2 BACKGROUND

This section reviews background on commands and abstractions in the UNIX shell.

UNIX Streams: A key UNIX abstraction is the data *stream*, operated upon by executing commands or *processes*. Streams are sequences of bytes, but most commands process them as higher-level sequences of line elements, with the newline character delimiting each element and the EOF condition representing the end of a stream. Streams are often referenced using a filename, that is an identifier in a global name-space made available by the UNIX file-system such as `/home/user/x`. Streams and files in UNIX are isomorphic: some streams can persist as files beyond the execution of the process, whereas other streams are ephemeral in that they only exist to connect the output of one process to the input of another process during their execution. The sequence order is maintained when changing between persistent files and ephemeral streams.

Commands: Each command is an independent computation unit that reads one or more input streams, performs a computation, and produces one or more output streams. Contrary to languages with a closed set of primitives, there is an unlimited number of UNIX commands, each one of which may have arbitrary behaviors—with the command’s side-effects potentially affecting the entire environment on which it is executing. These commands may be written in any language or exist only in binary form, and thus UNIX is not easily amenable to a single parallelizability analysis. Parallelization tools such as GNU `parallel` leave such analysis to developers that have to ensure that the script behavior will not be affected by parallelization, whereas transformation-based tools such as PaSh and POSH identify key invariants that hold for entire classes of commands and then resort to annotation libraries to infer whether each invariant is satisfied by each command. For example, an invariant that is used in both PaSh and POSH is whether a command is stateless, *i.e.*, whether it maintains state when processing its inputs, or whether it processes each input line

²The PaSh repository can be found at github.com/binpash/pash. The results of this paper are maintained in a frozen `icfp-ae` branch, and replicated as an image on Zenodo [Handa et al. 2021].

independently. Commands that satisfy this invariant can be parallelized by splitting their inputs in lines and then combining their outputs.

Command flags: UNIX commands are often configurable, customizing their behavior based on the task at hand. This is usually achieved via environment variables and command flags. Tools like PaSh and POSH address command behavior variability due to flags by including flags and command arguments in their annotation frameworks. Given command annotations, PaSh and POSH can abstract specific command invocations in a pipeline to black boxes for which some assumptions hold. This makes them applicable in the context of the shell where the space of possible command and flag combinations is exceedingly large.

Order of input consumption: In UNIX, all streams are ordered and all commands can safely assume that they can consume elements from their streams in the order they were produced. Additionally, most commands have the ability to operate on multiple files or streams. The order in which commands access these streams is important. In some cases, they read streams in the order of the stream identifiers provided. In other cases, the order is different—for example, an input stream may configure a command, and thus must be read before all the others. Consider for example `grep -f words.txt input.txt`, which first reads `words.txt` to determine the keywords for which it needs to search, and then reads `input.txt` line by line, emitting all lines that contain one of the words in `words.txt`. In other cases, reads from multiple streams are interleaved according to some command-specific semantics.

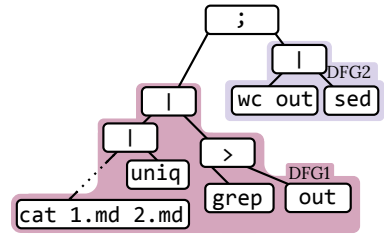
Composition: UNIX Operators: UNIX provides several primitives for program composition, each of which imposes different scheduling constraints on the program execution. Central among them is the *pipe* (`|`), a primitive that passes the output of one process as input to the next. The two processes form a pipeline, producing output and consuming input concurrently and possibly at different rates. The UNIX kernel facilitates program scheduling, communication, and synchronization behind the scenes. For example, *Spell*'s first `tr` transforms each character in the input stream to lower case, passing the stream to the second `tr`: the two `tr`s form a parallel producer-consumer pair of processes.

Apart from pipes, the language of the UNIX shell provides several other forms of program composition—e.g., the sequential composition operator (`;`) for executing one process after another has completed, and control structures such as `if` and `while`. All of these constructs enforce execution ordering between their components. To preserve such ordering and thus ensure correctness, systems such as PaSh and POSH do not “push” parallelization beyond these constructs. Instead, they focus on exploiting parallelism in script regions that do not face ordering constraints—which, as they demonstrate, is enough to significantly improve the performance of scripts found out in the wild [Raghavan et al. 2020; Vasilakis et al. 2021].

3 EXAMPLE AND OVERVIEW

This section provides intuition of the order-aware dataflow model proposed in this paper by following the different phases of a shell-to-shell parallelizing compiler (inspired by PaSh and POSH), formalized in the later sections. Given a script such as *Spell* (§1), the compiler identifies its dataflow regions, translates them to DFGs (Shell→ODFM), applies graph transformations that expose data parallelism on these DFGs, and replaces the original dataflow regions with the now-parallel regions (ODFM→Shell).

Shell→ODFM: Provided a shell script, the compiler starts by identifying subexpressions that are potentially parallelizable. The first step is to parse the script, creating an abstract syntax tree like the one presented on the right. Here we omit any non-stream flags and refer to all the stages between (and including) `tr` and `sort` as a dotted edge ending with `cat`.



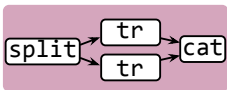
The compiler then identifies parallelism barriers within the shell script: these barriers are operators that enforce synchronization constraints such as the sequential composition operator (“;”). We call any set of commands that does not include a dataflow barrier a *dataflow region*. Dataflow regions are then transformed to dataflow graphs (DFGs), i.e., instances of our order-aware dataflow model. In our example, there are two dataflow regions corresponding to the following dataflow graphs:



As mentioned earlier (§2), the compiler exposes parallelism in each DFG separately to preserve the ordering requirements imposed to ensure correctness. For the rest of this section we focus on the parallelization of DFG1.

Parallelizable Commands: Individual nodes of the dataflow graphs are shell commands. Systems like PaSh and POSH assume key information for individual commands, e.g., whether they are amenable to divide-and-conquer data parallelism. Such data parallelism is achieved by splitting the input into pieces (at stream element boundaries), processing partial inputs in parallel, and finally applying an aggregation function to partial outputs to produce the final output. This decomposition breaks a command into two components—a data-parallel function, which is often the command itself, and an aggregation function. The table on the right presents aggregation functions for the shell commands in our example (all of which are parallelizable).

Command	Aggreg. Function
<code>cat</code>	<code>cat \$*</code>
<code>tr A-Z a-z</code>	<code>cat \$*</code>
<code>tr -d a</code>	<code>cat \$*</code>
<code>sort</code>	<code>sort -m \$*</code>
<code>uniq</code>	<code>uniq \$*</code>
<code>grep -f a -</code>	<code>cat \$*</code>
<code>wc -l</code>	<code>paste -d+ \$* bc</code>
<code>sed 's/a/b/'</code>	<code>cat \$*</code>

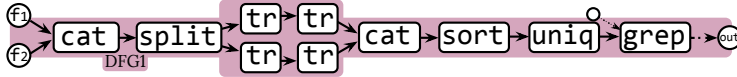


For example, consider the decomposition of the `tr` command. Applying `tr` over the entire input produces the same result as `split`ting the input into two, applying `tr` to the two partial inputs, and then merging the partial results with a `cat` aggregation function. Note that both `split` and `cat` are order-aware, i.e., `split` sends the first half of its input to the first `tr` and the rest to the second, while `cat` concatenates its inputs in order. This guarantees that the output of the DFG is the same as the one before the transformation.

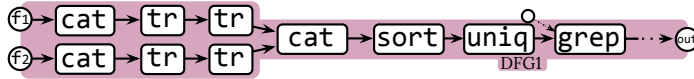
Parallelization Transformations: Given the decomposition of individual commands, the compiler’s next step is to apply graph transformations to exploit parallelism present in the computation represented by the DFG. As each parallelizable UNIX command comes with a corresponding aggregation function, the compiler’s transformations first convert the DFG into one that exploits parallelism at each stage. After applying the transformation to the two `tr` stages, the DFG looks as follows:



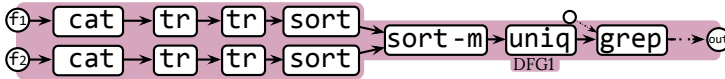
After these transformations are applied to all DFG nodes, the next transformation pass is applied to pairs of `cat` and `split` nodes: whenever a `cat` is followed by a `split` of the same width, the transformation removes the pair and connects the parallel streams directly to each other. The goal is to *push* data parallelism transformations as far down the pipeline as possible to expose the maximal amount of parallelism. Here is the resulting DFG for the transformation applied to the two `tr` stages:



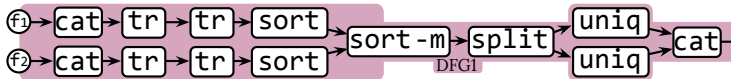
Applying this transformation to the first three stages—*i.e.*, `cat`, `tr`, and `tr`—of DFG1 produces the following transformed DFG.



The next node to parallelize is `sort`. To merge the partial output of parallel sorts, we need to apply a sorted merge. (In GNU systems, this is available as `sort -m` so we use this as the label of the merging node.) The transformation then removes `cat`, replicates `sort`, and merges their outputs with `sort -m`:



It then continues pushing parallelism down the pipeline, after applying a `split` function to split `sort -m`'s outputs.



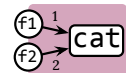
As mentioned earlier, a similar pass of iterative transformations is applied to DFG2, but the two DFGs are not merged to preserve the synchronization constraint of the dataflow barrier “;”.

Order Awareness: Data-parallel systems [Dean and Ghemawat 2008; Zaharia et al. 2012] often achieve parallelism using sharding, *i.e.*, partitioning input based on some key, or using shuffling, *i.e.*, arbitrary partitioning of inputs to parallel instances of an operator.

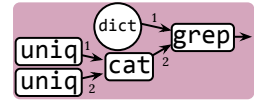
However, these techniques cannot be directly applied to the context of the shell, since (1) UNIX commands and pipelines assume strict ordering of their input elements, (2) most commands are not independent on the basis of some key (to enable sharding), and (3) many commands are not commutative (*e.g.*, `uniq`, `cat -n`). Since our goal is to define a model that applies directly to existing shell scripts, we cannot simply introduce new primitives that support sharding or shuffling, as is done in the case of systems that design an abstraction that fits their needs (*e.g.*, MapReduce, Spark). Thus, data parallelism in the shell requires a careful treatment of input and output ordering.

To further explain the need for order-awareness in a model for data parallel UNIX pipelines, let's look at the following examples. Consider *Spell*'s `cat f1.md f2.md` command that starts reading from `f2.md` only after it has completed reading `f1.md`;

note that any or both input streams may be pipes waiting for results from other processes. This order can be visualized as a label over each input edge. Correctly parallelizing this command requires ensuring that parallel `cat` (and possibly followup stages) maintains this order.

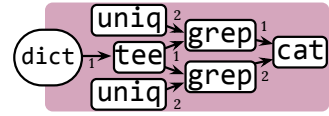


As a more interesting example, consider *Spell*'s `grep`, whose DFG is shown on the right. Parallelizing `grep` without taking order into account is not trivial, because `grep -vx -f`'s set difference is not commutative: we cannot simply split its input streams into two pairs of partial inputs



fed into two copies of `grep`. Taking input ordering into account, however, highlights an important dependency between `grep`'s inputs. The `dict` stream can be viewed as configuring `grep`, and thus `grep` can be modeled as consuming the entire `dict` stream before consuming partial inputs.

Armed with this insight, the compiler parallelizes `grep` by passing the same `dict.txt` stream to both `grep` copies. This requires an intermediary tee for duplicating the `dict.txt` stream to both copies of `grep`, each of which consumes the stream in its entirety before consuming the results of the preceding `uniq`.



Order-awareness is also important for the DFG translation back to a shell script. In the specific example we need to know how to instantiate the arguments of each `grep` of all possible options—*e.g.*, `grep -vx -f p1 p2, cat p1 | grep -vx -f - p2, etc.` Aggregators are UNIX commands with their own ordering characteristics that need to be accounted for.

The order of input consumption in the examples of this section is statically known and can be represented for each node as a set of configuration inputs, plus a sequence of the rest of its inputs. To accurately capture the behavior of shell programs, however, ODFM is more expressive, allowing any order of input consumption. The correctness of our parallelization transformations is predicated upon static but configurable orderings: a command reads a set of *configuration* streams to setup the consumption order of its input streams which are then consumed in-order, one after the other.

ODFM→Shell: The transformed graph is finally compiled back to a script that uses POSIX shell primitives to drive parallelism explicitly. A benefit of the dataflow model is that it can be directly implemented on top of the shell, simply translating each node to a command, and each edge to a stream. The generated parallel script for *Spell* can be seen below.

```

mkfifo t{0..14}          # DFG1: start
tr A-Z a-z < f1.md > t0 &
tr A-Z a-z < f2.md > t1 &
tr -d[:punct:] < t0 > t2 &
tr -d[:punct:] < t1 > t3 &
sort < t2 > t4 &
sort < t3 > t5 &
sort -m t4 t5 > t6 &
split t7 t8 < t6 &
# ...
tee t9 > t10 < dict.txt &
grep -vx -f t9 - < t11 > t13 &
grep -vx -f t10 - < t12 > t14 &
cat t13 t14 > out &
wait
rm t{0..14}              # DFG1: end

mkfifo t{0..8}          # DFG2: start
split t0 t1 < out &
wc -l < t0 > t2 &
wc -l < t1 > t3 &
paste -d+ t2 t3 | bc > t4 &
split t5 t6 < t4 &
sed 's/$/ misspelled words!/' < t5 > t7 &
  
```

$$\begin{aligned}
P &:= I; O; \bar{E} \\
I &:= \text{input } \bar{x} \\
O &:= \text{output } \bar{x} \\
\mathcal{E} &:= \bar{x}_o \leftarrow f(\bar{x}_i)
\end{aligned}$$

Fig. 1. Dataflow Description Language (DDL).

```

sed 's/$/ misspelled words!/' < t6 > t8 &
cat t7 t8 &
wait
rm t{0..8}           # DFG2: end

```

The two DFGs are compiled into the two fragments that start with `mkfifo` and end with `rm`. Each fragment uses a series of named pipes (FIFOs) to explicitly manipulate the input and output streams of each data-parallel instance, effectively laying out the structure of the DFG using explicit channel naming (UNIX FIFOs are named in the filesystem similar to normal files.) Aggregation functions are used to merge partial outputs from previous commands coming in through multiple FIFOs—for example, `sort -m t4 t6` and `cat t11 t12` for the first fragment, and `paste -d+ t2 t3 | bc` and `cat t7 t8` for the second. A `wait` blocks until all commands executing in parallel complete.

The parallel script is simplified for clarity of exposition: it does not show the details of input splitting, handling of SIGPIPE deadlocks, and other technical details that are handled by the current implementation.

Readers might be wondering about the correctness of having two `sed` commands in the parallel script: won't the string “misspelled words” appear twice in the output? Note, however, that the output of the `wc` stage (`fifo t4`) contains a single line. As a result, the second `sed` will not be given any input line and thus will not produce any output.

4 AN ORDER-AWARE DATAFLOW MODEL

In this section we describe the order-aware dataflow model (ODFM) and its semantics.

4.1 Preliminaries

As discussed earlier (§2), the two main shell abstractions are (i) data streams, and (ii) commands communicating via streams. We represent streams as named variables and commands as functions that read from and write to streams.

We first introduce some basic notation formalizing data streams on which our dataflow description language works. For a set D , we write D^* to denote the set of all finite words over D . For words $x, y \in D^*$, we write $x \cdot y$ or xy to denote their concatenation. We write ϵ for the empty word and \perp for the End-of-File condition. We say that x is a *prefix* of y , and we write $x \leq y$, if there is a word z such that $y = xz$. The \leq order is reflexive, antisymmetric, and transitive (*i.e.*, it is a partial order), and is often called the *prefix order*. We use the notation $D^* \cdot \perp$ to denote a *closed stream*, abstractly representing a file/pipe stream that has been closed, *i.e.*, one which no process will open for writing. The notation D^* is used to denote an *open stream*, abstractly representing an open pipe. Later, other process may add new elements at the end of this value. In the rest of our formalization we focus on terminating streams, and therefore terminating programs, since all of the data processing scripts that we have encountered are terminating. We discuss extensions for infinite streams in §9.

$$\begin{array}{c}
\langle x'_1, \dots, x'_p \rangle \leftarrow f(x_1, \dots, x_k \dots x_n) \in \mathcal{E} \\
|v_x| = 1 \vee v_x = \perp \\
k \in \text{choice}_f(v_1, \dots, v_n) \\
v_k \cdot v_x \leq \Gamma(x_k) \\
i \in [1, k-1] \cup [k+1, n]. v_i = \sigma(x_i) \\
\langle v_1^m, \dots, v_p^m \rangle = \llbracket f(v_1, \dots, v_k \circ v_x \dots v_n) \rrbracket_s
\end{array}
\quad \text{STEP}$$

$$\begin{array}{c}
\mathcal{I}, \mathcal{O}, \mathcal{E} \vdash \Gamma[x'_1 \rightarrow v'_1, \dots, x'_p \rightarrow v'_p], \sigma[x_k \rightarrow v_k] \rightsquigarrow \\
\Gamma[x'_1 \rightarrow v'_1 \cdot v_1^m, \dots, x'_p \rightarrow v'_p \cdot v_p^m], \sigma[x_k \rightarrow v_k \cdot v_x]
\end{array}$$

Fig. 2. Small Step Execution Semantics

4.2 Dataflow Description Language

Figure 1 presents the Dataflow Description Language (DDL) for defining dataflow graphs (DFG). A program $p \in P$ in DDL is of the form $\mathcal{I}; \mathcal{O}; \mathcal{E}$. \mathcal{I} and \mathcal{O} represent sets of edges, vectors of the form $\bar{x} = \langle x_1, x_2, \dots, x_n \rangle$. Variables x_1, x_2, \dots represent DFG edges, *i.e.*, streams used as a communication channel between DFG nodes and as the input and output of the entire DFG.

\mathcal{I} is of the form input \bar{x} , where \bar{x} is the set of the input variables. Each variable $x \in \mathcal{I}$ represents a file `file(x)` that is read from the UNIX filesystem. Note that multiple input variables can refer to the same file.

\mathcal{O} is of the form output \bar{x} , where \bar{x} is the set of output variables. Each variable $x \in \mathcal{O}$ represents a file `file(x)` that is written to the UNIX filesystem.

\mathcal{E} represents the nodes of the DFG. A node $\bar{x}_o \leftarrow f(\bar{x}_i)$ represents a function from list of input variables (edges) \bar{x}_i to output variables (edges) \bar{x}_o . We require that f is monotone with respect to a lifting of the prefix order for a sequence of inputs; that is, $\forall v, v', \bar{v}_i$, if $v \leq v'$, $\langle v_1, \dots, v_n \rangle = f(v, \bar{v}_i)$ and $\langle v'_1, \dots, v'_n \rangle = f(v', \bar{v}_i)$, then $\forall k \in [1, n]$. $v_k \leq v'_k$. This captures the idea that a node cannot retract output that it has already produced.

We wrap all functions f with an execution wrapper $\llbracket \cdot \rrbracket$ that ensures that all outputs of f are closed when its inputs are closed:

$$\llbracket f(v_1 \cdot \perp, v_2 \cdot \perp, \dots, v_n \cdot \perp) \rrbracket = \langle v'_1 \cdot \perp, v'_2 \cdot \perp, \dots, v'_k \cdot \perp \rangle$$

This is helpful to ensure termination. From now on, we only refer to the wrapped function semantics.

We also assume that commands do not produce output if they have not consumed any input, *i.e.*, the following is true:

$$\langle \epsilon, \dots, \epsilon \rangle = \llbracket f(\epsilon, \dots, \epsilon) \rrbracket$$

A variable in DDL is assigned *only once* and consumed by *only one* node. DDL does not allow the dataflow graph to contain *any cycles*. This also holds for variables in \mathcal{I} and \mathcal{O} , which cannot refer to the same variables in \mathcal{I} and never assigned a different value in \mathcal{E} . Similarly, variables in \mathcal{O} are not read by any node in \mathcal{E} . All variables which are not included in \mathcal{I} and \mathcal{O} abstractly represent temporary files/pipes which are created during the execution of a shell script. We assume that within a dataflow program, all variables are reachable from some input variables.

Execution Semantics: Figure 2 presents the small step execution semantics for DDL. Maps Γ associates variable names to the data contained in the stream it represents. Map σ associates variable names to the data in the stream that has already been processed—representing the read-once semantics of UNIX pipes. Let $\langle x'_1, \dots, x'_p \rangle \leftarrow f(x_1, \dots, x_n)$ be a node in our DFG program. The function choice_f represents the order in which a commands consumes its inputs by returning a set of input indexes on which the function blocks on waiting to read. For example, the choice_{cat} function for the command `cat` always returns the next non-closed index—as `cat` reads its inputs in sequence, each one until depletion.

$$\{k+1\} = \text{choice}_{cat}(v_1 \cdot \perp, \dots, v_k \cdot \perp, v_{k+1}, \dots, v_n)$$

$$\frac{\langle x'_1, \dots, x'_p \rangle \leftarrow f(x_1, \dots, x_n) \in \mathcal{E} \quad \langle v'_1 \cdot \perp, \dots, v'_p \cdot \perp \rangle = \llbracket f(v_1 \cdot \perp, \dots, v_n \cdot \perp) \rrbracket}{\mathcal{I}, \mathcal{O}, \mathcal{E} \vdash \Gamma[x'_1 \rightarrow v'_1 \cdot \perp, \dots, x'_p \rightarrow v'_p \cdot \perp, x_1 \rightarrow v_1 \cdot \perp, \dots, x_n \rightarrow v_n \cdot \perp]} \text{COMPLETION}$$

Fig. 3. Execution Constraints

For a choice_f function to be valid, it has to return an input index that has not been closed yet. Formally,

$$\mathcal{S} = \text{choice}_f(v_1, \dots, v_k \cdot \perp, \dots, v_n) \implies k \notin \mathcal{S}$$

We assume that the set returned by choice_f cannot be *empty* unless all input indexes are closed, meaning that all nodes consume all of their inputs until depletion even if they do not need the rest of it for processing.

The small step semantics nondeterministically picks a variable x_k , such that $k \in \text{choice}_f(v_1, \dots, v_n)$, *i.e.*, f is waiting to read some input from x_k , and $\sigma(x_k) < \Gamma(x_k)$, *i.e.*, there is data on the stream represented by variable x_k that has to be processed. The execution then retrieves the next message v_x to process, and computes new messages v_1^m, \dots, v_p^m to pass on to the output streams x'_1, \dots, x'_p . Note that any of these messages (input or output) might be \perp . We pass $v_k \circ v_x$, which denotes that the previous data v_k is now being combined with the new message v_x , to function f . For all functions f and new messages v , given $\langle v'_1, v'_2, \dots, v'_p \rangle = \llbracket f(v_1, \dots, v_k, \dots, v_n) \rrbracket$ we assume the following constraint holds:

$$\begin{aligned} \langle v_1^m, \dots, v_p^m \rangle &= \llbracket f(v_1, \dots, v_k \circ v_x, \dots, v_n) \rrbracket_s \\ \iff \langle v'_1 \cdot v_1^m, \dots, v'_p \cdot v_p^m \rangle &= \llbracket f(v_1 \dots, v_k \cdot v_x, \dots, v_n) \rrbracket \end{aligned}$$

This constraint ensures that first processing arguments $v_1, \dots, v_k, \dots, v_n$ and then message v_x append to argument k^{th} stream is equivalent to processing messages $v_1, \dots, v_k \cdot v_x, \dots, v_n$ at once. Having this property, allows our system to process messages as they arrive or wait for all the messages to arrive, without changing the semantics of the execution.

The messages v_1^m, \dots, v_p^m are passed on to their respective output streams (by updating Γ). Note that the size of the output messages could vary, and they could even be empty. Finally, σ is updated to denote that v_x has been processed.

Execution: Let $\langle \mathcal{I}, \mathcal{O}, \mathcal{E} \rangle$ be a dataflow program, where \mathcal{I} = input \bar{x}_i are the input variables, and output \bar{x}_o are the output variables. Let σ_i be the initial mapping from all variable names in the dataflow program $\langle \mathcal{I}, \mathcal{O}, \mathcal{E} \rangle$ to empty string ϵ . Let Γ_i be the initial mapping for variables in the dataflow program, such that all non-input variables $x \notin \bar{x}_i$, map to the empty string $\Gamma_i(x) = \epsilon$. In contrast, all input variables $x \in \bar{x}_i$, *i.e.*, files already present in the file system, are mapped to the contents of the respective input file $\Gamma_i(x) = v \cdot \perp$.

When no more small step transitions can take place (*i.e.*, all commands have finished processing), the dataflow execution terminates and the contents of output variables in \mathcal{O} can be written to their respective output files. Figure 3 represents the constraint that has to be satisfied by Γ at the end of execution, *i.e.*, when all variables are processed. We now prove some auxiliary theorems and lemmas to show that dataflow programs always terminate and that when they terminate, the constraint in Figure 3 holds.

THEOREM 4.1. *Let $\langle x'_1, x'_2, \dots, x'_p \rangle \leftarrow f(x_1, \dots, x_n) \in \mathcal{E}$. During any point within the execution of the DFG, the following statement is true:*

$$\langle \Gamma(x'_1), \dots, \Gamma(x'_p) \rangle = \llbracket f(\sigma(x_1), \dots, \sigma(x_n)) \rrbracket$$

PROOF. Proof by induction on the number of execution steps.

Base Case: Let Γ_i and σ_i be the initial mappings. For $x \in \langle x_1, \dots, x_n \rangle$, $\sigma_i(x) = \epsilon$. For $x \in \langle x'_1, \dots, x'_p \rangle$, $\Gamma_i(x) = \epsilon$ (since x'_1, \dots, x'_p are not input variables to the DFG, they will be initialized to ϵ). The following property is true for all functions f :

$$\langle \epsilon, \dots, \epsilon \rangle = \llbracket f(\epsilon, \dots, \epsilon) \rrbracket$$

Therefore, for the initial mappings, Γ_i and σ_i ,

$$\langle \Gamma_i(x'_1), \dots, \Gamma_i(x'_p) \rangle = \llbracket f(\sigma_i(x_1), \dots, \sigma_i(x_n)) \rrbracket$$

Induction Hypothesis: Let Γ and σ be a snapshot of stream mappings during the execution of the DFG such that the following statement is true:

$$\langle \Gamma(x'_1), \dots, \Gamma(x'_p) \rangle = \llbracket f(\sigma(x_1), \dots, \sigma(x_n)) \rrbracket$$

Induction Case: Let Γ and σ be a snapshot of stream mappings such that the induction hypothesis is true. Let Γ' and σ' be the snapshot after a single step of the execution takes place, given the snapshots Γ and σ .

If for all $i \in [1, n]$, $\sigma(x_i) = \sigma'(x_i)$, then a message updating x'_1, \dots, x'_p was not processed (they can only be written by this node). Therefore, for all $k \in [1, p]$, $\Gamma(x'_k) = \Gamma'(x'_k)$ and the following statement is true (assuming Induction Hypothesis)

$$\langle \Gamma'(x'_1), \dots, \Gamma'(x'_p) \rangle = \llbracket f(\sigma'(x_1), \dots, \sigma'(x_n)) \rrbracket$$

If there exists an $i \in [1, n]$ such that $\sigma'(x_i) = \sigma(x_i) \cdot v_x$, $v_x \neq \epsilon$, then a message v_x was processed. Note that the above statement can only be true for a single i . For all $k \neq i$, $\sigma(x_k) = \sigma'(x_k)$.

The following statement is true from Induction Hypothesis:

$$\langle \Gamma(x'_1), \dots, \Gamma(x'_p) \rangle = \llbracket f(\sigma(x_1), \dots, \sigma(x_k), \dots, \sigma(x_n)) \rrbracket$$

and from small step semantics, for all $k \in [1, p]$:

$$\Gamma'(x_k) = \Gamma(x_k) \cdot v_k^m$$

where:

$$\langle v_1^m, \dots, v_p^m \rangle = \llbracket f(\sigma(x_1), \dots, \sigma(x_k) \circ v_x, \dots, \sigma(x_n)) \rrbracket_s$$

Using the definition of $\llbracket \cdot \rrbracket_s$, the following statement is true:

$$\langle \Gamma(x'_1) \cdot v_1^m, \dots, \Gamma(x'_p) \cdot v_p^m \rangle = \llbracket f(\sigma(x_1), \dots, \sigma(x_k) \cdot v_x, \dots, \sigma(x_n)) \rrbracket$$

Therefore, the following is true:

$$\langle \Gamma'(x'_1), \dots, \Gamma'(x'_p) \rangle = \llbracket f(\sigma'(x_1), \dots, \sigma'(x_k), \dots, \sigma'(x_n)) \rrbracket$$

The small step semantics will preserve this property. Therefore, by induction, for all $\langle x'_1, x'_2, \dots, x'_p \rangle \leftarrow f(x_1, \dots, x_n) \in \mathcal{E}$, the following is always true about Γ and σ , during any point within the execution:

$$\langle \Gamma(x'_1), \dots, \Gamma(x'_p) \rangle = \llbracket f(\sigma(x_1), \dots, \sigma(x_n)) \rrbracket$$

□

LEMMA 4.2. Let $\langle x'_1, x'_2, \dots, x'_p \rangle \leftarrow f(x_1, \dots, x_n) \in \mathcal{E}$. If $\forall i \in [1, n]$ $\Gamma(x_i) = v_i \cdot \perp$, then eventually $\forall i \in [1, n]$, $\sigma(x_i) = v_i \cdot \perp$ and eventually $\forall i \in [1, p]$, $\Gamma(x'_i) = v'_i \cdot \perp$.

PROOF. If for all $i \in [1, n]$, $\Gamma(x_i)$ is closed and choice_f is non empty unless $\sigma(x_i)$ is closed for all i , then eventually the execution will take a step to update σ till, for all $i \in [1, n]$, $\sigma(x_i)$ is closed. When all inputs are closed, $\llbracket \cdot \rrbracket$ dictates that all outputs will be closed as well. Using theorem 4.1, $\Gamma(x'_i)$ will be closed. □

Commands	c	$:=$	$(s = w)^* wr^* \mid \text{pipe} \mid c^+ \&^? \mid c r^+$
			$\mid c \& \mid (c) \mid c_1; c_2 \mid c_1 \&\& c_2$
			$\mid c_1 \mid \mid c_2 \mid ! c \mid \text{while } c_1 c_2$
			$\mid \text{for } s w c \mid \text{if } c_1 c_2 c_3 \mid \text{case } w cb^*$
			$\mid s() c$
Redirections	r	$:=$	\dots
Words	w	$:=$	$(s \mid _ \mid k)^*$
Control Codes	k	$:=$	\dots
Strings	s	\in	Σ^+

Fig. 4. A relevant subset of shell syntax presented in Smoosh [Greenberg and Blatt 2020].

THEOREM 4.3. *Eventually for all variables x , $\exists v. \Gamma(x) = v \cdot \perp$, i.e., all variables will eventually be closed.*

PROOF. Let C be the set of variables which will be closed eventually. Note that $I \subseteq C$ (all input variables to the DFG will be eventually closed). Using Lemma 4.2, for any node $\langle x'_1, x'_2, \dots, x'_p \rangle \leftarrow f(x_1, \dots, x_n) \in \mathcal{E}$, if $x_1, \dots, x_n \in C$, then $x'_1, \dots, x'_p \in C$. Since the dataflow program contains no cycles, eventually all variables reachable from the input variables are in C . \square

THEOREM 4.4. *The Dataflow program will always terminate. Let Γ and σ be the stream mappings when the DFG terminates. The for Γ and σ , the constraint 3 will be true.*

PROOF. The DFG graph terminates when all variables are closed. From Theorem 4.3, all variables will eventually be closed. Constraint 3 follows from Theorem 4.1, all variables being closed when DFG terminates, and the properties of $\llbracket \cdot \rrbracket$. \square

5 FROM SHELL SCRIPTS TO DFGS AND BACK AGAIN

This section formalizes the translations between the shell and our order-aware dataflow model.

5.1 Shell to ODFM

Given a shell script the compiler starts by recursing on the AST, replacing subtrees in a bottom-up fashion with dataflow programs. Fig. 4 shows a relevant subset of shell syntax, adapted from Smoosh [Greenberg and Blatt 2020]. Intuitively, some shell constructs (such as pipes `|`) allow for the composition of the dataflow programs of their components, while others (such as `;`) prevent it. Figure 5 shows the translation rules for some interesting constructs, and Figure 6 shows several auxiliary relations that are part of this translation. We denote compilation from a shell AST to a shell AST as $c \uparrow c'$, and compilation to a dataflow program as $c \uparrow \langle p, b \rangle$ where p is a dataflow program and $b \in \{\text{bg}, \text{fg}\}$ denotes whether the program is to be executed in the foreground or background.

The first two rules **COMMANDTRANS** and **COMMANDID** describe compilation of commands. The bulk of the work is done in **cmd2node**, which, when possible, defines a correspondence between a command and a dataflow node. Predicate **pure** indicates whether the command is pure, i.e., whether it only interacts with its environment by reading and writing to files. All commands that we have seen until now (`grep`, `sort`, `uniq`) satisfy this predicate. The relations **ins** and **outs** define a correspondence between a commands arguments and the nodes inputs and outputs. We assume that a variable is uniquely identified from the file that it refers too, therefore if two variables have the same name, then they also refer to the same files. Finally, relation **func** extracts information

$$\begin{array}{c}
\frac{\text{cmd2node}(w, \bar{x}_o \leftarrow f(\bar{x}_i)) \quad \text{add_metadata}(f, \bar{a}s, \bar{r}) = f' \\
\text{redir}(\bar{x}_o, \bar{x}_i, \bar{r}, \bar{x}'_o, \bar{x}'_i)}{\bar{a}s\bar{w}\bar{r} \uparrow \langle \text{input}\bar{x}'_i; \text{output}\bar{x}'_o; \bar{x}'_o \leftarrow f'(\bar{x}'_i), \text{fg} \rangle} \text{COMMANDTRANS} \\
\\
\frac{\text{cmd2node}(w, \perp)}{\bar{a}s\bar{w}\bar{r} \uparrow \bar{a}s\bar{w}\bar{r}} \text{COMMANDID} \qquad \frac{c \uparrow \langle p, b \rangle}{c \& \uparrow \langle p, \text{bg} \rangle} \text{BACKGROUNDDFG} \\
\\
\frac{c \uparrow c'}{c \& \uparrow c' \&} \text{BACKGROUNDCMD} \qquad \frac{c_1 \uparrow \langle p_1, \text{bg} \rangle \quad c_2 \uparrow \langle p_2, b \rangle}{c_1 ; c_2 \uparrow \langle \text{compose}(p_1, p_2), b \rangle} \text{SEQBOTHBG} \\
\\
\frac{c_1 \uparrow c'_1 \quad c_2 \uparrow \langle p_2, \text{bg} \rangle}{\text{opt}(p_2) \Downarrow c'_2} \text{SEQRIGHTBG} \qquad \frac{c_1 \uparrow \langle p_1, \text{fg} \rangle \quad c_2 \uparrow \langle p_2, \text{bg} \rangle}{\text{opt}(p_1) \Downarrow c'_1 \quad \text{opt}(p_2) \Downarrow c'_2} \text{SEQBOTHFGBG} \\
\frac{c_1 ; c_2 \uparrow c'_1; (c'_2 \&)}{c_1 ; c_2 \uparrow c'_1; (c'_2 \&)} \\
\\
\frac{c_1 \uparrow c'_1 \quad c_2 \uparrow \langle p_2, \text{fg} \rangle}{\text{opt}(p_2) \Downarrow c'_2} \text{SEQRIGHTFG} \qquad \frac{c_1 \uparrow c'_1 \quad c_2 \uparrow c'_2}{c_1 ; c_2 \uparrow c'_1 ; c'_2} \text{SEQNONE} \\
\\
\frac{c_1 \uparrow \langle p_1, b_1 \rangle, \dots, c_n \uparrow \langle p_n, b_n \rangle, \quad p'_1 \dots p'_{n-1} = \text{map}(\text{connectpipe}, p_1 \dots p_{n-1}) \\
p = \text{fold_left}(\text{compose}, p'_1 p'_2 \dots p'_{n-1} p_n)}{\text{pipe}|c_1 c_2 \dots c_n \& \uparrow \langle p, \text{bg} \rangle} \text{PIPEBG} \\
\\
\frac{c_1 \uparrow \langle p_1, b_1 \rangle, \dots, c_n \uparrow \langle p_n, b_n \rangle, \quad p'_1 \dots p'_{n-1} = \text{map}(\text{connectpipe}, p_1 \dots p_{n-1}) \\
p = \text{fold_left}(\text{compose}, p'_1 p'_2 \dots p'_{n-1} p_n)}{\text{pipe}|c_1 c_2 \dots c_n \uparrow \langle p, \text{fg} \rangle} \text{PIPEFG}
\end{array}$$

Fig. 5. A subset of the compilation rules.

about the execution of the command (such as its choice function and w) to be able to reconstruct it later on. Note that the four relations **pure**, **ins**, **outs**, and **func** act as axioms, and the soundness of our model and translations depends on their correctness. Prior work [Raghavan et al. 2020; Vasilakis et al. 2021] has shown how to construct such relations for specific commands using annotation frameworks, with PaSh providing annotations for more than 50 commands in POSIX and GNU Coreutils—two large and widely used sets of commands.

The rule **BACKGROUNDDFG** sets the background flag for the underlying dataflow program; if the operand of a $\&$ is not compiled to a dataflow program then it is simply left as is. The last part holds for all shell constructs, we currently only create dataflow nodes from a single command.

The next set of rules refer to the sequential composition operator “;”. This operator acts as a dataflow barrier since it enforces an execution ordering between its two operands. Because of that, it forces the dataflow programs that are generated from its operands to be optimized (with **opt**) and then compiled back to shell scripts (with **ll**). However, there is one case (**SEQBOTHBG**) where a dataflow region can propagate through a “;” and that is if the first component is to be executed in the background. In this case “;” does not enforce an execution order constraint between its two

$$\begin{array}{c}
(\text{vars}(\mathcal{E}_1) \setminus \text{vars}(\mathcal{I}_1)) \cap (\text{vars}(\mathcal{E}_2) \setminus \text{vars}(\mathcal{I}_2)) = \emptyset \quad \mathcal{I}' = \mathcal{I}_1 \setminus \mathcal{O}_2 \cup \mathcal{I}_2 \setminus \mathcal{O}_1 \\
\mathcal{O}' = \mathcal{O}_1 \setminus \mathcal{I}_2 \cup \mathcal{O}_2 \setminus \mathcal{I}_1 \quad p' = \mathcal{I}'; \mathcal{O}'; \mathcal{E}_1 \cup \mathcal{E}_2 \\
\text{vars}(\mathcal{E}_1) \setminus (\text{vars}(\mathcal{I}_1) \cup \text{vars}(\mathcal{O}_1) \cap \text{vars}(\mathcal{I}_2)) = \emptyset \quad \text{valid}(p') \\
\text{vars}(\mathcal{E}_2) \setminus (\text{vars}(\mathcal{I}_2) \cup \text{vars}(\mathcal{O}_2) \cap \text{vars}(\mathcal{I}_1)) = \emptyset \\
\vec{x}_1 = \text{vars}(\mathcal{E}_1) \cap \text{vars}(\mathcal{E}_2) \setminus (\text{vars}(\mathcal{I}_1) \cup \text{vars}(\mathcal{I}_2) \cup \text{vars}(\mathcal{O}_3) \cup \text{vars}(\mathcal{O}_2)) \\
\vec{x}_2 \notin \text{vars}(\mathcal{E}_1) \cup \text{vars}(\mathcal{E}_2) \quad |\vec{x}_1| = |\vec{x}_2| \quad \mathcal{E}'_2 = \mathcal{E}_2[\vec{x}_2/\vec{x}_1] \\
\forall x \in (\mathcal{I}_1 \cup \mathcal{I}_2 \setminus \mathcal{I}') \cup (\mathcal{O}_1 \cup \mathcal{O}_2 \setminus \mathcal{O}') . \text{pipe}(x) \\
\hline
\text{compose}(\mathcal{I}_1; \mathcal{O}_1; \mathcal{E}_1, \mathcal{I}_2; \mathcal{O}_2; \mathcal{E}_2) = p' \\
\hline
\frac{\text{pure}(w) \quad \text{ins}(w, \bar{x}_i, f) \quad \text{outs}(w, \bar{x}_o, f) \quad \text{func}(w, f)}{\text{cmd2node}(w, \bar{x}_o \leftarrow f(\bar{x}_i))} \quad \frac{\neg \text{pure}(w)}{\text{cmd2node}(w, \perp)} \\
\frac{\text{ins}_f(\bar{x}_i, w, x_{in}?) \quad \text{outs}_f(\bar{x}_o, w, x_{out}?) \quad \text{func}(w, f)}{\text{node2cmd}(\bar{x}_o \leftarrow f(\bar{x}_i), w, x_{in}?, x_{out}?)} \quad \frac{\mathcal{O}' = \mathcal{O}[\bar{x}_{stdIn}/\bar{x}_{stdOut}]}{\text{connectpipe}(\mathcal{I}; \mathcal{O}; \mathcal{E}) = \mathcal{I}; \mathcal{O}'; \mathcal{E}} \\
\frac{\text{node2cmd}(\bar{x}'_o \leftarrow f(\bar{x}'_i), w, x_{in}?, x_{out}?) \quad \text{get_metadata}(f) = \langle \bar{a}s, \bar{r} \rangle}{\text{redir}(\bar{x}'_o, \bar{x}'_i, \bar{r}, \bar{x}_o, \bar{x}_i) \quad \bar{r}' = \text{in_out}(\bar{r}, x_{in}?, x_{out}?)}}{\text{instantiate}(\bar{x}_o \leftarrow f(\bar{x}_i)) = \bar{a}s w \bar{r}}
\end{array}$$

Fig. 6. Auxiliary relations for translating commands to nodes and back.

operands and the generated dataflow programs can be safely composed into a bigger one. The rules for “&&” and “||” are similar (omitted).

The relation **compose** unifies two dataflow programs by combining the inputs of one with the outputs of the other and vice versa. Before doing that, it ensures that the composed dataflow graph will be valid by checking that there is at most one reader and one writer for each internal and output variable, as well as all the rest of the dataflow program invariants, *e.g.*, the absence of cycles (§4).

The remaining rules (not shown) introduce synchronization constraints and are not part of our parallelization effort—for example, we consider all branching operators as strict dataflow barriers.

5.2 ODFM to Shell

Figure 7 presents the compilation \Downarrow of a dataflow program $p = \mathcal{I}; \mathcal{O}; \bar{\mathcal{E}}$ to a shell program. The compilation can be separated in a prologue, the main body, and an epilogue.

The prologue creates a named pipe (*i.e.*, UNIX FIFO) for every variable in the program. Named pipes are created in a temporary directory using the `mkfifo` command, and are similar in behavior to ephemeral pipes except that they are explicitly associated to a file-system identifier—*i.e.*, they are a special file in the file-system. Named pipes are used in place of ephemeral pipes (`|`) in the original script.

The epilogue inserts a `wait` to ensure that all the nodes in the dataflow have completed execution, and then removes all named pipes from the temporary directory using `rm`. The design of the prologue-epilogue pair mimics how UNIX treats ephemeral pipes, which correspond to temporary identifiers in a hidden file-system.

The main body expresses the parallel computation and can also be separated into three components. For each of the input variables $x_i \in \mathcal{I}$, we add a command that copies the file $f = \text{file}(x_i)$ to its designated pipe. Similarly, for all output variables $x_i \in \mathcal{O}$ we add a command that copies the

$$\begin{array}{c}
I = \text{input } x_{i_1}, \dots, x_{i_n} \quad O = \text{output } x_{o_1}, \dots, x_{o_m} \quad \mathcal{E} = n_1, \dots, n_k \\
\text{cin} = \text{cat } \text{file}(x_{i_1}) > \text{pipe}(x_{i_1}) \& ; \dots ; \text{cat } \text{file}(x_k) > \text{pipe}(x_{i_n}) \& \\
\text{cin} = \text{cat } \text{pipe}(x_{o_1}) > \text{file}(x_{o_1}) \& ; \dots ; \text{cat } \text{pipe}(x_{o_m}) > \text{file}(x_{o_m}) \& \\
\text{cnodes} = \text{instantiate}(n_1) \& ; \dots ; \text{instantiate}(n_k) \& \\
\hline
\text{body}(I, O, \mathcal{E}) = \text{cin} ; \text{cout} ; \text{cnodes} \\
\hline
\text{vars}(p) = \langle v_1, \dots, v_n \rangle \qquad \text{prologue}(p) = pr \quad \text{epilogue}(p) = ep \\
\text{prologue}(p) = \text{mkfifo } / \text{tmp}/p\{1..n\} \qquad p \Downarrow pr ; \text{body}(p) ; ep \\
\hline
\text{vars}(p) = \langle v_1, \dots, v_n \rangle \\
\text{epilogue}(p) = \text{wait} ; \text{rm } / \text{tmp}/p\{1..n\}
\end{array}$$

Fig. 7. DFG to shell transformations.

designated pipe to the output file in the filesystem $f = \text{file}(x_i)$. Finally, we translate each node in \mathcal{E} to a shell command that reads from the pipes corresponding to its input variables and writes to the pipes corresponding to its output variables. In order to correctly translate a node back to a command, we use the node-command correspondence functions (similar to the ones for \uparrow) that were used for the translation of the command to a node. Since a translated command might get its input from (or send its output to) a named pipe, we need to also add those as new redirections with in_out . For example, for a node $x_3 \leftarrow f(x_1, x_2)$ that first reads x_1 and then reads x_2 , where $f = \text{grep } -f$, the following command would be produced:

```
grep -f p1 p2 > p3 &
```

6 PARALLELIZATION TRANSFORMATIONS

In this section we define a set of transformations that expose data parallelism on a dataflow graph. We start by defining a set of helper DFG nodes and a set of auxiliary transformations to simplify the graph and enable the parallelization transformations. Then we identify a property on dataflow nodes that indicates whether the node can be executed in a data parallel fashion. We then define the parallelization transformations and we conclude with a proof that applying all of the transformations preserves the semantics of the original DFG.

6.1 Helper Nodes and Auxiliary Transformations

Before we define the parallelization transformations, we introduce several helper functions that can be used as dataflow nodes. The first function is `split`. `split` takes a single input variable (file or pipe) and sequentially splits into multiple output variables. The exact size of the data written in each output variable is left abstract since it does not affect correctness but only performance.

$$\begin{array}{c}
\bar{x} \leftarrow \text{split}(x_i) \\
\bar{v} = \langle v_1 \cdot \perp, v_2 \cdot \perp, \dots, v_{k-1} \cdot \perp, v_k, \epsilon, \dots, \epsilon \rangle, \\
v_c = v_1 \cdot v_2 \cdot \dots \cdot v_k \\
\hline
\forall v_c. \llbracket \text{split}(v_c) \rrbracket = \bar{v} \\
\hline
\bar{v} = \langle v_1 \cdot \perp, v_2 \cdot \perp, \dots, v_m \cdot \perp \rangle, v_c = v_1 \cdot v_2 \cdot \dots \cdot v_m \cdot \perp \\
\hline
\forall v_c. \llbracket \text{split}(v_c) \rrbracket = \bar{v}
\end{array}$$

The second function is `cat`, which coincidentally behaves the same as the UNIX command `cat`. `cat`, given a list of input variables, combines their values and assigns it to a single output variable. Formally `cat` is defined below:

$$\begin{array}{c}
 x_i \leftarrow \text{cat}(\bar{x}) \\
 \bar{v} = \langle v_1 \cdot \perp, v_2 \cdot \perp, \dots, v_{k-1} \cdot \perp, v_k, \dots, v_m \rangle, \\
 v_c = v_1 \cdot v_2 \dots \cdot v_k \\
 \hline
 \forall \bar{v}. \llbracket \text{cat}(\bar{v}) \rrbracket = v_c \\
 \bar{v} = \langle v_1 \cdot \perp, v_2 \cdot \perp, \dots, v_m \cdot \perp \rangle, v_c = v_1 \cdot v_2 \dots \cdot v_m \cdot \perp \\
 \hline
 \forall \bar{v}. \llbracket \text{cat}(\bar{v}) \rrbracket = v_c
 \end{array}$$

The third function is `tee`, which behaves the same way as the UNIX command `tee`, i.e. copying its input variable to several output variables. Formally `tee` is defined below:

$$\begin{array}{c}
 \bar{x} \leftarrow \text{tee}(x_i) \\
 \bar{v} = \langle v_c, v_c, \dots, v_c \rangle, \\
 \hline
 \forall v_c. \llbracket \text{tee}(v_c) \rrbracket = \bar{v} \\
 \bar{v} = \langle v_c \cdot \perp, v_c \cdot \perp, \dots, v_c \cdot \perp \rangle, \\
 \hline
 \forall v_c. \llbracket \text{tee}(v_c \cdot \perp) \rrbracket = \bar{v}
 \end{array}$$

The final function is `relay`. `relay` works as an identity function. Formally `relay` is defined below:

$$x_i \leftarrow \text{relay}(x_j), \forall v. \llbracket \text{relay}(v) \rrbracket = v, \forall v. \llbracket \text{relay}(v \cdot \perp) \rrbracket = v \cdot \perp$$

Using these helper nodes our compiler performs a set of auxiliary transformations that are depicted in Figure 8. `relay` acts an identity function, therefore any edge can be transformed to include a `relay` between them. Splitting in multiple stages to get n edges is the same as splitting in one step into n edges. Similarly, combining n edges in multiple stages is the same as combining n edges in a single stage. If we split an edge into n edges, and then combine the n edges back, this behaves as an identity. A `cat` can be pushed following a `tee` by creating n copies of the `tee` function. If a `cat` has single incoming edge, we can convert it into a `relay`. If a split has a single outgoing edge, we can convert it into a `relay`.

The first seven transformations can be performed both ways. The last transformations is one way. A split after a `cat` can be converted into `relays`, if the input arity of `cat` is the same as output arity of split. The reverse transformation in this case is not allowed as, using (`Relay`) rule, we can `cat` and split any two or more streams in the dataflow graph. This will allow us to pass the output of any function in our graph to any other function as an input. This will break the semantics of our Dataflow graph.

6.2 Data Parallelism and Transformations

The dataflow model exposes task parallelism as each different node can execute independently—only communicating with the other nodes through their communication channels. In addition to that, it is possible to achieve data parallelism by executing some nodes in parallel by partitioning part of their input.

Sequential Consumption Nodes: We are interested in nodes that produce a single output and consume their inputs in sequence (one after the other when they are depleted), after having consumed the rest of their inputs as an initialization and configuration phase. Note that there are several examples of shell commands that correspond to such nodes, e.g. `grep`, `sort`, `grep -f`, and `sha1sum`. Let such a node $x' = f(x_1, \dots, x_{n+m})$, where w.l.o.g. x_1, x_2, \dots, x_n represent the configuration

$$\begin{array}{c}
\frac{x_j \leftarrow \text{Unused}(I, O, \mathcal{E}), \mathcal{E}' = \mathcal{E}[x_j/x_i]}{I, O, \mathcal{E} \iff I, O, \mathcal{E}' \cup \{x_j \leftarrow \text{relay}(x_i)\}} \text{RELAY} \\
\\
\frac{E = \{\langle x_s, x'_s \rangle \leftarrow \text{split}(x), \langle x_1, \dots, x_k \rangle \leftarrow \text{split}(x_s), \langle x_{k+1}, \dots, x_m \rangle \leftarrow \text{split}(x'_s)\}}{I, O, \mathcal{E} \cup \{\langle x_1, \dots, x_m \rangle \leftarrow \text{split}(x)\} \iff I, O, \mathcal{E} \cup E} \text{SPLIT-SPLIT} \\
\\
\frac{E = \{x_c \leftarrow \text{cat}(\langle x_1, \dots, x_k \rangle), x'_c \leftarrow \text{cat}(\langle x_{k+1}, \dots, x_m \rangle), x \leftarrow \text{cat}(\langle x_c, x'_c \rangle)\}}{I, O, \mathcal{E} \cup \{x \leftarrow \text{cat}(\langle x_1, \dots, x_m \rangle)\} \iff I, O, \mathcal{E} \cup E} \text{CONCAT-CONCAT} \\
\\
\frac{\bar{x} \leftarrow \text{Unused}(I, O, \mathcal{E}), E = \{\bar{x} \leftarrow \text{split}(x_j), x_i \leftarrow \text{cat}(\bar{x})\}}{I, O, \mathcal{E} \cup \{x_i \leftarrow \text{relay}(x_j)\} \iff I, O, \mathcal{E} \cup E} \text{SPLIT-CONCAT} \\
\\
\frac{E = \{\langle x_1^u, x_1^d \rangle \leftarrow \text{tee}(x_1), \langle x_2^u, x_2^d \rangle \leftarrow \text{tee}(x_2), \dots, \langle x_n^u, x_n^d \rangle \leftarrow \text{tee}(x_n), \\ x_o \leftarrow \text{cat}(x_1^u, x_2^u, \dots, x_n^u), x'_o \leftarrow \text{cat}(x_1^d, x_2^d, \dots, x_n^d)\}}{I, O, \mathcal{E} \cup \{x \leftarrow \text{cat}(x_1, x_2, \dots, x_n), \langle x_o, x'_o \rangle \leftarrow \text{tee}(x)\} \iff I, O, \mathcal{E} \cup E} \text{TEE-CONCAT} \\
\\
\frac{}{I, O, \mathcal{E} \cup \{x_j \leftarrow \text{cat}(x_i)\} \iff I', O, \mathcal{E}' \cup \{x_j \leftarrow \text{relay}(x_i)\}} \text{ONE-CONCAT} \\
\\
\frac{}{I, O, \mathcal{E} \cup \{x_j \leftarrow \text{split}(x_i)\} \iff I', O, \mathcal{E}' \cup \{x_j \leftarrow \text{relay}(x_i)\}} \text{ONE-SPLIT} \\
\\
\frac{E = \{x \leftarrow \text{Unused}(I, O, \mathcal{E}), \\ \bar{x}_i = \langle x_1, x_2, \dots, x_n \rangle, \bar{x}_j = \langle x'_1, x'_2, \dots, x'_n \rangle, \\ \langle x'_1 \leftarrow \text{relay}(x_1), x'_2 \leftarrow \text{relay}(x_2), \dots, x'_n \leftarrow \text{relay}(x_n) \rangle\}}{I, O, \mathcal{E} \cup \{x \leftarrow \text{cat}(\bar{x}_i), \bar{x}_j \leftarrow \text{split}(x)\} \implies I, O, \mathcal{E} \cup E} \text{CONCAT-SPLIT}
\end{array}$$

Fig. 8. Auxiliary Transformation

$$\frac{x_r, x_m, x_1^m, \dots, x_n^m, x_1^r, \dots, x_n^r, x_1^t, \dots, x_n^t, x_1^p, \dots, x_n^p \leftarrow \text{Unused}(I, O, \mathcal{E}) \quad \text{dp}(f, f_m, f_r)}{E = \{\langle x_1^c, \dots, x_n^c, x_{n+1}^c \rangle \leftarrow \text{tee}(x^c) : \forall x^c \in \bar{x}\} \cup \\ \{x_i^m \leftarrow f_m(x_i, \bar{x}_i) : \forall i \in \{1 \dots n\}\} \cup \{x_r \leftarrow f_r(x_1^m, \dots, x_n^m, \bar{x}_{n+1})\}}{I, O, \mathcal{E} \cup \{x \leftarrow \text{cat}(x_1, \dots, x_n), x_i \leftarrow f(x, \bar{x})\} \implies I, O, \mathcal{E} \cup E} \text{PARALLEL}$$

Fig. 9. Parallelization Transformation

inputs and x_{n+1}, \dots, x_{n+m} represent the sequential consumption inputs. The consumption order of such a command is shown below:

$$\text{choice}_f(\bar{v}) = \begin{cases} \{i : i \leq n \wedge \neg \text{closed}(v_i)\} & \text{if } \neg \forall i \leq n, \text{closed}(v_i) \\ \{i : \forall j < i, \text{closed}(v_j)\} & \text{otherwise} \end{cases}$$

If we know that a command f satisfies the above property we can safely transform it to a $x^i = \text{cat}(x_{n+1}, \dots, x_{n+m})$ followed by a command $x' = f'(x^i, x_1, \dots, x_n)$, without altering the semantics of the graph.

Data Parallel Nodes: We now shift our focus to a subset of the sequential consumption nodes, namely those that can be executed in a data parallel fashion by splitting their inputs. These are nodes that can be broken down in a parallel *map* f_m and an associative *aggregate* f_r . Formally, these nodes have to satisfy the following equation:

$$\forall n, \llbracket f(v_1^i \dots v_n^i, \bar{v}) \rrbracket = \llbracket f_r(\bar{v}_1, \dots, \bar{v}_n, \bar{v}) \rrbracket$$

$$\text{where } \forall i \in \{1 \dots n\} \bar{v}_i = \llbracket f_m(v_i^i, \bar{v}) \rrbracket$$

We denote data parallel nodes as $\text{dp}(f, f_m, f_r)$. Example of such a node that satisfies this property is the sort command, where $f_m = \text{sort}$ and $f_r = \text{sort} - \text{m}$.

In addition to the above equation, a *map* function f_m should not output anything when its input closes.

$$\llbracket f_m(v, \bar{v}) \rrbracket = \llbracket f_m(v \cdot \perp, \bar{v}) \rrbracket$$

Note that f_m could have multiple outputs and be different than the original function f . As has been noted in prior research [Farzan and Nicolet 2017] this is important as some functions require auxiliary information in the map phase in order to be parallelized. An important observation is that a subset of all data parallel nodes are completely stateless, meaning that $f_m = f$ and $f_r = \text{cat}$, and therefore are embarrassingly parallel.

We can now define a transformation on any data parallel node f , that replaces it with a map followed by an aggregate. This transformation is formally shown in Figure 9. Essentially, all the sequential consumption inputs (that are concatenated using cat) are given to different f_m nodes the outputs of which are then aggregated using f_r while preserving the input order. Note that the configuration inputs have to be duplicated using tee to ensure that all parallel f_m s and f_r s will be able to read them in case they are pipes and not files on disk.

Using the auxiliary transformations—by adding a `split` followed by `cat` before a data parallel node, we can always parallelize them using the parallelization transformation.

Correctness of Transformations: We now proceed to prove a series of statements regarding the semantics-preservation properties of dataflow programs.

Program Equivalence: Let $p = \langle I, O, \mathcal{E} \rangle$ and $p' = \langle I', O', \mathcal{E}' \rangle$ be two dataflow programs, where $I = \langle x_1^i, \dots, x_n^i \rangle$, $I' = \langle y_1^i, \dots, y_n^i \rangle$, $O = \langle x_1^o, \dots, x_m^o \rangle$, and $O' = \langle y_1^o, \dots, y_m^o \rangle$. These programs are equivalent if and only if, assuming initial value of input values are equal (for all $k \in [1, n]$. x_k^i is equal to the value of y_k^i), for the value of the output variables is the same, when both of these DFGs terminate. Formally, for all values v_1, \dots, v_n , if

$$\forall j \in [1, n]. \Gamma_i(x_j^i) = \Gamma_i'(y_j^i) = v_j$$

$$\forall j \in [1, m]. \Gamma_o(x_j^o) = \Gamma_o'(y_j^o)$$

where Γ_i, Γ_i' are the initial mappings for p and p' respectively, and Γ_o, Γ_o' are the mappings when p and p' have completed their execution.

THEOREM 6.1. *Let $p = \langle I, O, \mathcal{E} \cup E \rangle$ and $p' = \langle I, O, \mathcal{E} \cup E' \rangle$ be two dataflow programs. Let S_i be the set of input variables in node set E (variables read in E but not assigned inside E). Let S_o be the set of output variables in the node set E (variables assigned in E but not read inside E). Let S_i', S_o' be the input variables and output variables of E' . We assume $S_i = S_i'$ and $S_o = S_o'$. If $\langle S_i, S_o, E \rangle$ is equivalent to $\langle S_i, S_o, E' \rangle$, then program $\langle I, O, \mathcal{E} \cup E \rangle$ is equivalent to $\langle I, O, \mathcal{E} \cup E' \rangle$.*

PROOF. Given any initial mapping Γ_i , let Γ_o, Γ'_o be the mappings when p and p' complete their execution. For all $x \in \mathcal{S}_i$, $\Gamma_o(x) = \Gamma'_o(x)$ as there are no cycles in the dataflow graph, and the subgraph which computes \mathcal{S}_i is same in both p and p' .

Since $\langle \mathcal{S}_i, \mathcal{S}_o, E \rangle$ is equivalent to $\langle \mathcal{S}_i, \mathcal{S}_o, E' \rangle$, and for all $x \in \mathcal{S}_i$, $\Gamma_o(x) = \Gamma'_o(x)$, for all $x \in \mathcal{S}_o$, $\Gamma_o(x) = \Gamma'_o(x)$.

Only variables in set \mathcal{S}_o are the variables assigned in E and E' , that are used in computing the value of the output variables \mathcal{O} . Since the value of the variables is same in both these programs, given the same input mapping Γ_i , for all output variables $x \in \mathcal{O}$, $\Gamma_o(x) = \Gamma'_o(x)$.

Therefore, both these programs are equivalent. \square

THEOREM 6.2. *Transformations presented in Figure 8 and Figure 9 preserve program equivalence.*

PROOF. The (Relay) transformation preserves program equivalence as the program terminates, the value of x_i is equal to the value of x_j .

The remaining transformations, transforming an input program $\langle J, \mathcal{O}, \mathcal{E} \cup E \rangle$ to an output program $\langle J', \mathcal{O}', \mathcal{E} \cup E' \rangle$. For all transformations $\mathcal{S}_i = \mathcal{S}'_i$ and $\mathcal{S}_o = \mathcal{S}'_o$ (where $\mathcal{S}_i, \mathcal{S}'_i, \mathcal{S}_o, \mathcal{S}'_o$ are defined above). First seven transformations, equivalence of programs $\langle \mathcal{S}_i, \mathcal{S}_o, E \rangle$ and $\langle \mathcal{S}_i, \mathcal{S}_o, E' \rangle$ follow from the execution semantics for cat, relay, split, tee, the properties of f_m and f_r for data parallel commands f .

The (Concat-Split) transformation relies on the additional property that the program produces the same output independent of how the split breaks the input stream. Choice of a particular way of breaking the stream does not change the value of the program's output variables when it terminates.

Since, $\langle \mathcal{S}_i, \mathcal{S}_o, E \rangle$ is equivalent to $\langle \mathcal{S}_i, \mathcal{S}_o, E' \rangle$, these transformations preserve equivalence (Theorem 6.1). \square

7 EVALUATION

Our evaluation consists of two parts. The first part is a case study of applying GNU Parallel to two scripts, demonstrating the difficulty of manually reasoning about parallel shell pipelines and the challenges that one has to address in order to achieve a parallel implementation. The second part demonstrates the performance benefits of our transformations on 47 unmodified shell scripts. Before discussing our evaluation, we offer a brief outline of the compiler implementation.

Implementation: We reimplement the compilation and optimization phases of PaSh [Vasilakis et al. 2021] according to our model and associated transformations. The new implementation is about 1500 lines of Python code and uses the order-aware dataflow model as the centerpiece intermediate representation. It is also more modular and facilitates the development of additional transformations, closely mirroring the back-and-forth shell-to-ODFM translations described in Section 5 and the paralleling transformations described in Section 6. While we expect that most users would use PaSh by writing shell scripts, completely ignoring the ODFM, it is also possible to manually describe programs in the intermediate representation, enabling other frontend and backend frameworks to interface with it.

By reimplementing PaSh's optimization phase to mirror our transformations we also discovered and solved a bug in PaSh. The old implementation did not tee the configuration inputs of a parallelized command, but rather allowed all parallel copies to read from the same input. While this is correct if the configuration input is a file on disk, the semantics indicated that in the general case it leads to incorrect results—for example, in cases where this input is a stream—because all parallel commands consume items from a single stream, only reading a subset of them.

7.1 Case Study: GNU Parallel

We describe an attempt to achieve data parallelism in two scripts using GNU parallel [Tange 2011], a tool for running shell commands in parallel. We chose GNU parallel because it compares favorably to other alternatives in the literature [Tange 2020], but note that GNU parallel sits somewhere between an automated compiler, like PaSh and POSH, and a fully manual approach—illustrating only some of the issues that one might face while manually trying to parallelize their shell scripts.

Spell: We first apply parallel on *Spell*'s first pipeline (§1):

```
TEMP_C1="/tmp/{/}.out1"
TEMP1=$(seq -w 0 ${JOBS - 1}) | sed 's^+/tmp/in+' | sed 's/$/.out1/' | tr '\n' ' ')
TEMP1=$(echo $TEMP1)
mkfifo $TEMP1
parallel "cat {} | col -bx | tr -cs A-Za-z '\n' | tr A-Z a-z | \
  tr -d '[:punct:]' | sort > $TEMP_C1" ::: $IN &
sort -m $TEMP1 | parallel -k --jobs ${JOBS} --pipe --block "$BLOCK_SIZE" "uniq" |
  uniq | parallel -k --jobs ${JOBS} --pipe --block "$BLOCK_SIZE" "grep -vx -f $dict -"
rm $TEMP1
```

It took us a few iterations to get the parallel version right, leading to a few observations. First, despite its automation benefits, parallel still requires manual placement of the intermediate FIFO pipes and *agg* functions. Additionally, achieving ideal performance requires some tweaking: setting `--block` to 10K, 250K, and 250M yields widely different execution times—27, 4, and 3 minutes respectively.

Most importantly, omitting the `-k` flag in the last two fragments breaks correctness due to re-ordering related to scheduling non-determinism. These fragments are fortunate cases in which the `-k` flag has the desired effect, because their output order follows the same order as the arguments of the commands they parallelize. Other commands face problems, in that the correct output order is not the argument order nor an arbitrary interleaving.

Set-difference: We apply parallel to Set-diff, a script that compares two streams using sort and comm:

```
mkfifo s1 s2
TEMP_C1="/tmp/{/}.out1"
TEMP1=$(seq -w 0 ${JOBS - 1}) | sed 's^+/tmp/in+' | sed 's/$/.out1/' | tr '\n' ' ')
TEMP1=$(echo $TEMP1)
TEMP_C2="/tmp/{/}.out2"
TEMP2=$(seq -w 0 ${JOBS - 1}) | sed 's^+/tmp/in+' | sed 's/$/.out2/' | tr '\n' ' ')
TEMP2=$(echo $TEMP2)
mkfifo ${TEMP1} ${TEMP2}
parallel "cat {} | cut -d ' ' -f 1 | tr [:lower:] [:upper:] | sort > $TEMP_C1" ::: $IN &
sort -m ${TEMP1} > s1 &
parallel "cat {} | cut -d ' ' -f 1 | sort > $TEMP_C2" ::: $IN2 &
sort -m ${TEMP2} > s2 &
cat s1 | parallel -k --pipe --jobs ${JOBS} \
  --block "$BLOCK_SIZE" "grep -vx -f s2 -"
rm ${TEMP1} ${TEMP2}
rm s1 s2
```

In addition to the issues highlighted in Spell, this parallel implementation has a subtle bug. GNU parallel spawns several instances of `grep -vx -f s2 -` that all read FIFO `s2`. When the first parallel instance exits, the kernel sends a SIGPIPE signal to the second `sort -m`. This forces `sort` to exit, in turn leaving the rest of the parallel `grep -vx -f` instances blocked waiting for new input.

The most straightforward way we have found to address this bug is to remove (1) the “&” operator after the second sort `-m`, and (2) `s2` from `mkfifo`. This modification sacrifices pipeline parallelism, as the first stage of the pipeline completes before executing `grep -vx -f`. The parallel pipeline modified for correctness completes in 4m54s. Our compiler does not sacrifice pipeline parallelism by using `tee` to replicate `s2` for all parallel instances of `grep -vx -f` (§2), completing in 4m7s.

7.2 Performance Results

Methodology: We use three sets of benchmark programs from various sources, including GitHub, StackOverflow, and the UNIX literature [Bentley 1985; Bentley et al. 1986; Bhandari 2020; Jurafsky 2017; McIlroy et al. 1978; Taylor and Perry 2016].

- **Expert Pipelines:** The first set contains 9 pipelines: NFA-regex, Sort, Top-N, WF, Spell, Difference, Bi-grams, Set-Difference, and Shortest-Scripts. Pipelines in this set contain 2–7 stages (mean: 5.2), ranging from a scalable CPU-intensive `grep` stage in NFA-regex to a non-parallelizable `diff` stage in Difference. These scripts are written by UNIX experts: a few pipelines are from UNIX legends [Bentley 1985; Bentley et al. 1986; McIlroy et al. 1978], one from a book on UNIX scripting [Taylor and Perry 2016], and a few are from top Stackoverflow answers [Jurafsky 2017].
- **Unix50 Pipelines:** The second set contains 34 pipelines solving the UNIX 50 game [Labs 2019]. This set is from a recent set of challenges celebrating of UNIX’s 50-year legacy, solvable by UNIX pipelines. The problems were designed to highlight UNIX’s modular philosophy [McIlroy et al. 1978]. We found unofficial solutions to all-but-three problems on GitHub [Bhandari 2020], expressed as pipelines with 2–12 stages (mean: 5.58). They make extensive use of standard commands under a variety of flags, and appear to be written by non-experts—contrary to the previous set, they often use sub-optimal or non-UNIX-y constructs. We execute each pipeline as-is, without any modification.
- **COVID-19 Mass-Transit Analysis Pipelines:** The third set contains 4 pipelines that were used to analyze real telemetry data from bus schedules during the COVID-19 response in one of Europe’s largest cities [Tsaliki and Spinellis 2021]. The pipelines compute several statistics on the transit system per day—such as average serving hours per day and average number of vehicles per day. Pipelines range between 9 and 10 stages (mean: 9.2) and use typical UNIX staples such as `sed`, `awk`, `sort`, and `uniq`.

We use our implementation of PaSh to parallelize all of the pipelines in these benchmark sets, working with three configurations:

- **Baseline:** Our compiler simply executes the script using a standard shell (in our case `bash`) without performing any optimizations. This configuration is used as our baseline. Note that it is not completely sequential since the shell already achieves pipeline and task parallelism based on `|` and `&`.
- **No Cat-Split:** Our compiler performs all transformations except `CONCAT-SPLIT`. This configuration achieves parallelism by splitting the input before each command and then merging it back. It is used as a baseline to measure the benefits achieved by the `CONCAT-SPLIT` transformation.
- **Parallel:** Our compiler performs all transformations. The `CONCAT-SPLIT` transformation, which removes a `cat` with n inputs followed by a `split` with n outputs, ensures that data is not merged unnecessarily between parallel stages.

Experiments were run on a 2.1GHz Intel Xeon E5-2683 with 512GB of memory and 64 physical cores, Debian 4.9.144-3.1, GNU Coreutils 8.30-3, GNU Bash 5.0.3(1), OCaml 4.05.0, and Python 3.7.3. All pipelines are set to (initially) read from and (finally) write to the file system. For “Expert

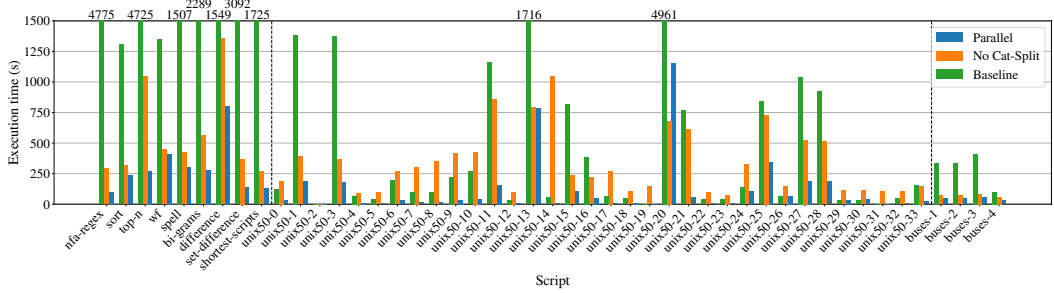
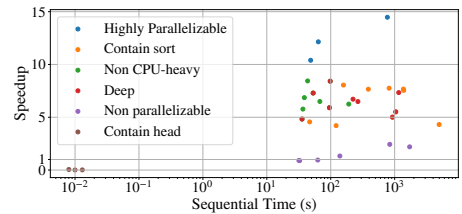


Fig. 10. Execution times with configurations **Baseline**, **Parallel**, and **No Cat-Split**, with 16 \times -parallelism.

Pipelines”, we use 10GB-collections of inputs from Project Gutenberg [Hart 1971]; for “Unix50 Pipelines”, we gather their inputs from each level in the game [Labs 2019] and multiply them up to 10GB. For “Bus Route Analysis Pipelines” we use the real bus telemetry data for the year 2020 (~ 3.4GB). The inputs for all pipelines are split in 16 equal sized chunks, corresponding to the intended parallelism level.

Fig. 10 shows the execution times on all programs with 16 \times parallelism and for all three configurations mentioned in the beginning of the evaluation. It shows that all programs achieve significant improvements with the addition of the **Concat-Split** transformation. The average speedup without **Concat-Split** over the bash baseline is 2.26 \times . The average speedup with the transformation is 6.16 \times .

The figure on the right explains the differences in the effect of the transformation based on the kind of commands involved in the pipelines. It offers a correlation between sequential time and speedup, and shows that different programs that involve commands with similar characteristics (color) see similar speedups (y-axis). Programs containing only parallelizable commands see the highest speedup (10.4–14.5 \times). Programs with limited speedup either (1) contain sort, which does not scale linearly, (2) are not CPU-intensive, resulting in pronounced IO and constant costs, or (3) are deep pipelines, already exploiting significant pipeline-based parallelism. Programs with non-parallelizable commands see no significant change in execution time (0.9–1.3 \times). Finally, programs containing head have a very small sequential execution, typically under 1s, and thus their parallel equivalents see a slowdown due to constant costs—still remaining under 1s.



8 RELATED WORK

Dataflow Graph Models: Graph models of computation where nodes represent units of computation and edges represent FIFO communication channels have been studied extensively [Dennis 1974; Kahn 1974; Kahn and MacQueen 1977; Karp and Miller 1966; Lee and Messerschmitt 1987a,b]. ODFM sits somewhere between Kahn Process Networks [Kahn 1974; Kahn and MacQueen 1977] (KPN), the model of computation adopted by UNIX pipes, and Synchronous Dataflow [Lee and Messerschmitt 1987a,b] (SDF). A key difference between ODFM and SDF is that ODFM does not assume fixed item rates—a property used by SDF for efficient scheduling determined at compile-time. Two differences between ODFM from KPNs is that (i) ODFM does not allow cycles, and (ii)

ODFM exposes information about the input consumption order of each node. This order provides enough information at compile time to perform parallelizing transformations while also enabling translation of the dataflow back to a UNIX shell script.

Systems for batch [Dean and Ghemawat 2008; Murray et al. 2013; Zaharia et al. 2012], stream [Gordon et al. 2006; Mamouras et al. 2017; Thies et al. 2002], and signal processing [Bourke and Pouzet 2013; Lee and Messerschmitt 1987a] provide dataflow-based abstractions. These abstractions are different from ODFM which operates on the UNIX shell, an existing language with its own peculiarities that have guided the design of the model.

One technique for retrofitting order over unordered streaming primitives such as sharding and shuffling is to extend the types of elements using tagging [Arvind et al. 1984; Arvind and Nikhil 1990; Watson and Gurd 1979]. This technique would not work in the UNIX shell, because (1) commands are black boxes operating on stream elements in unconstrained ways (but in known order), and (2) because data streams exchanged between commands contain flat strings, without support for additional metadata extensions, and thus no obvious way to augment elements with tags. ODFM instead captures ordering on the edges of the dataflow graph, and leverages the consumption order of nodes (the choice function) in the graph to orchestrate execution appropriately.

Synchronous languages [Berry and Gonthier 1992; Le Guernic et al. 1986; Maraninchi and Rémond 2001; Ratel et al. 1991] model stream graphs as circuits where nodes are state machines and edges are wires that carry a single value. Lustre [Ratel et al. 1991] is based on a dataflow model that is similar to ours, but its focus is different as it is not intended for exploiting data-parallelism.

Semantics and Transformations: Prior work proposes semantics for streaming extensions to relational query languages based on dataflow [Arasu et al. 2006; Li et al. 2005]. In contrast to our work, it focuses on transformations of time-varying relations.

More recently, there has been significant work on the correct parallelization of distributed streaming applications by proposing sound optimizations and compilation techniques [hir 2014; Schneider et al. 2015], type systems [Mamouras et al. 2019], and differential testing [Kallas et al. 2020]. These efforts aim at producing a parallel implementation of a dataflow streaming computation using techniques that do not require knowledge of the order of consumption of each node—a property that very important in our setting.

Recent work proposes a semantic framework for stream processing that uses monoids to capture the type of data streams [Mamouras 2020]. That work mostly focuses on generality of expression, showing that several already proposed programming models can be expressed on top of it. It also touches upon soundness proofs of optimizations using algebraic reasoning, which is similar to our approach.

Divide and Conquer Decomposition: Prior work has shown the possibility of decomposing programs or program fragments using divide-and-conquer techniques [Farzan and Nicolet 2017, 2019; Rugina and Rinard 1999; Smith and Albarghouthi 2016]. The majority of that work focuses on parallelizing special constructs—*e.g.*, loops, matrices, and arrays—rather than stream-oriented primitives. Techniques for automated synthesis of MapReduce-style distributed programs [Smith and Albarghouthi 2016] can be of significant aid for individual commands. In some cases [Farzan and Nicolet 2017, 2019], the map phase is augmented to maintain additional metadata used by the reducer phase. These techniques complement our work, since they can be used to derive aggregators and the parallelizability properties of yet unknown shell commands, making them possible to capture in our model.

Parallel Shell Scripting: Tools exposing parallelism on modern UNIXes such as qsub [W. Gentzsch (Sun Microsystems) 2001], SLURM [Yoo et al. 2003], AMFS [Zhang et al. 2013] and GNU parallel [Tange 2011] are predicated upon explicit and careful orchestration from their users. Similarly,

several shells [Duff 1990; McDonald and Dix 1988; Spinellis and Fragkoulis 2017; Walker et al. 2009] add primitives for non-linear pipe topologies—some of which target parallelism. Here too, however, users are expected to manually rewrite scripts to exploit these new primitives without jeopardizing correctness.

Our work is inspired by PaSh [Vasilakis et al. 2021] and POSH [Raghavan et al. 2020], two recent systems that use command annotations to parallelize and distribute shell programs by operating on dataflow fragments of the Unix shell. Our work is tied to PaSh [Vasilakis et al. 2021], as it (i) uses its annotation framework for instantiating the correspondence of commands to dataflow nodes (§5.1), and (ii) serves as its formal foundation since it reimplements all of its parallelizing transformations and proves them correct. POSH [Raghavan et al. 2020] too translates shell scripts to dataflow graphs and optimizes them achieving performance benefits, but its goal is to offload commands close to their data in a distributed environment. Thus, POSH only performs limited parallelization transformations, and focusing more on the scheduling problem of determining where to execute each command. It only parallelizes commands that require a concatenation combiner, *i.e.*, a subset of the transformations that we prove correct in this work, and thus replacing its intermediate representation with our ODFM would be possible. POSH also proposes an annotation framework that captures several command characteristics. Some of these characteristics, such as parallelizability, are also captured by PaSh. Others are related to the scheduling problem—for example, whether a command such as `grep` produces output smaller than its input, making it a good candidate for offloading close to the input data.

POSIX Shell Semantics: Our work depends on Smoosh, an effort focused on formalizing the semantics of the POSIX shell [Greenberg and Blatt 2020]. Smoosh focuses on POSIX semantics, whereas our work introduces a novel dataflow model in order to transform programs and prove the correctness of parallelization transformations on them. One of the Smoosh authors has also argued for making concurrency explicit via shell constructs [Greenberg 2018]. This is different from our work, since it focuses on the capabilities of the shell language as an orchestration language, and does not deal with the data parallelism of pipelines.

Parallel Userspace Environments: By focusing on simplifying the development of distributed programs, a plethora of environments inadvertently assist in the construction of parallel software. Such systems [Barak and La’adan 1998; Mullender et al. 1990; Ousterhout et al. 1988], languages [Kilian et al. 2007; Sewell et al. 2005; Viriding et al. 1996], or system-language hybrids [Epstein et al. 2011; Pike et al. 1990; Vasilakis et al. 2015] hide many of the challenges of dealing with concurrency as long as developers leverage the provided abstractions—which are strongly coupled to the underlying operating or runtime system. Even when these efforts are shell-oriented, such as Plan9’s `rc`, they are backward-incompatible with the UNIX shell, and often focus primarily on hiding the existence of a network rather than on modelling data parallelism.

9 DISCUSSION

Command Annotations: The translation of shell scripts to dataflow programs is based on knowledge about command characteristics, such as the predicate `pure` that determines whether a command does not perform any side-effect except for writing to a set of output files (§5.1). In the current implementation, this information is acquired through the annotation language and annotations provided by PaSh [Vasilakis et al. 2021]. An interesting avenue for future research would be to explore analyses for inferring or checking the annotations for commands. Such work could help extend the set of supported commands (which currently requires manual effort). Furthermore, it would be interesting to explore extensions to the annotation language in order to enable additional optimizations; for example, commands that are commutative and associative could be parallelized

more efficiently, by relaxing the requirement for input order and better utilizing the underlying resources.

Directly accessing the IR in the implementation: As described earlier (§7), our implementation currently allows manually developing programs in the ODFM intermediate representation. However, this interface is not that convenient to use as an end-user since it requires manually instantiating each node of the graph with the necessary command metadata, *e.g.*, inputs and outputs. It would be interesting future work to design different frontends that interface with this IR. For example, a frontend compiler from the language proposed by dgsh [Spinellis and Fragkoulis 2017]; a shell that supports extended syntax for creating DAG pipelines. The IR could also act as an interface for different backends, for example one that implements ODFM in a distributed setting.

Parallel Script Debugging: Debugging standard shell pipelines can be hard and it usually requires several iterations of trial and error until the user gets the script right. Our approach does not make the debugging experience any worse, as the system produces as output a parallel shell script, which can be inspected and modified like any standard shell script (as seen in §3). For example, a user could debug a script by removing a few stages of the parallel pipeline, or redirecting some intermediate outputs to permanent files for inspection. This is possible because of the expressiveness of ODFM and the existence of a bidirectional transformation from dataflow programs to shell scripts, which allows the compiler to simply use a standard shell such as bash as its backend.

An approach that is particularly helpful, and which we have used ourselves, is to ask the compiler to add a *relay* node between every two nodes of the graph and then instantiate this *relay* node with an identity command that duplicates its input to its output and also a log file.

```
tee $LOG < $IN > $OUT
```

This allows for stream introspection without affecting the behavior of the pipeline, facilitating debugging since the user can inspect all intermediate outputs at once.

Stream Finiteness and Extensions: In our current model, parallelism is achieved by partitioning the finite stream, processing the partitions in parallel, and merging the results. Like PaSh and POSH, our model is designed to support terminating computations over large but finite data streams. All of the data processing scripts that we have encountered conform to this model and are terminating. One way to extend our work to support and parallelize infinite streams—such as the ones produced by `yes` and `tail -f`—would involve repeated applications of partitioning, processing, and merging.

10 CONCLUSION

We presented an order-aware dataflow model for exploiting data parallelism latent in UNIX shell scripts. The dataflow model is order-aware, accurately capturing the semantics of complex UNIX pipelines: in our model, the order in which a node in the dataflow graph consumes inputs from different edges plays a central role in the semantics of the computation and therefore in the resulting parallelization. The model captures the semantics of transformations that exploit data parallelism available in UNIX shell computations and prove their correctness. We additionally formalized the translations from the UNIX shell to the dataflow model and from the dataflow model back to a parallel shell script. We implemented our model and transformations as the compiler and optimization passes of PaSh, a system parallelizing shell pipelines, and used it to evaluate the speedup achieved on 47 data-processing pipelines.

While the shell has been mostly ignored by the research community for most of its 50-year lifespan, recent work [Greenberg and Blatt 2020; Greenberg et al. 2021a,b; Raghavan et al. 2020; Spinellis and Fragkoulis 2017; Vasilakis et al. 2021] indicates renewed community interest in

shell-related research. We view our work partly as providing the missing correctness piece of the shell-optimization work done by the systems community [Raghavan et al. 2020; Vasilakis et al. 2021], and partly as a stepping stone for further studies on the dataflow fragment of the shell, e.g., the development of more elaborate transformations and optimizations.

ACKNOWLEDGMENTS

We thank Konstantinos Mamouras for preliminary discussions that helped spark an interest for this work, Dimitris Karnikis for help with the artifact, Diomidis Spinellis for benchmarks and discussions, Michael Greenberg and Jiashi Shen for comments on the presentation of our work, the anonymous ICFP reviewers and our shepherd Rishiyur Nikhil for extensive feedback, and the ICFP artifact reviewers for their comments that significantly improved the paper artifact. This research was funded in part by DARPA contracts HR00112020013 and HR001120C0191, and NSF award CCF 1763514. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect those of DARPA or other agencies.

REFERENCES

2014. A Catalog of Stream Processing Optimizations. *ACM Comput. Surv.* 46, 4, Article 46 (March 2014), 34 pages. <https://doi.org/10.1145/2528412>
- Arvind Arasu, Shivnath Babu, and Jennifer Widom. 2006. The CQL Continuous Query Language: Semantic Foundations and Query Execution. *The VLDB Journal* 15, 2 (June 2006), 121–142. <https://doi.org/10.1007/s00778-004-0147-z>
- K. Arvind, E. David Culler, Robert Iannucci, Vinod Kathail, Keshav Pingali, and Robert Thomas. 1984. *The tagged token dataflow architecture*. Technical Report. Technical report, MIT Laboratory for Computer Science.
- K. Arvind and Rishiyur S. Nikhil. 1990. Executing a Program on the MIT Tagged-Token Dataflow Architecture. *IEEE Trans. Comput.* 39, 3 (March 1990), 300–318. <https://doi.org/10.1109/12.48862>
- Amnon Barak and Oren La'adan. 1998. The MOSIX Multicomputer Operating System for High Performance Cluster Computing. *Future Gener. Comput. Syst.* 13, 4–5 (March 1998), 361–372. [https://doi.org/10.1016/S0167-739X\(97\)00037-X](https://doi.org/10.1016/S0167-739X(97)00037-X)
- Jon Bentley. 1985. Programming Pearls: A Spelling Checker. *Commun. ACM* 28, 5 (May 1985), 456–462. <https://doi.org/10.1145/3532.315102>
- Jon Bentley, Don Knuth, and Doug McIlroy. 1986. Programming Pearls: A Literate Program. *Commun. ACM* 29, 6 (June 1986), 471–483. <https://doi.org/10.1145/5948.315654>
- Gérard Berry and Georges Gonthier. 1992. The ESTEREL Synchronous Programming Language: Design, Semantics, Implementation. *Sci. Comput. Program.* 19, 2 (Nov. 1992), 87–152. [https://doi.org/10.1016/0167-6423\(92\)90005-V](https://doi.org/10.1016/0167-6423(92)90005-V)
- Pawan Bhandari. 2020. Solutions to unixgame.io. <https://git.io/Jf2dn> Accessed: 2020-04-14.
- Timothy Bourke and Marc Pouzet. 2013. ZéLus: A Synchronous Language with ODEs. In *Proceedings of the 16th International Conference on Hybrid Systems: Computation and Control (HSCC '13)*. Association for Computing Machinery, New York, NY, USA, 113–118. <https://doi.org/10.1145/2461328.2461348>
- Jeffrey Dean and Sanjay Ghemawat. 2008. MapReduce: Simplified Data Processing on Large Clusters. *Commun. ACM* 51, 1 (Jan. 2008), 107–113. <https://doi.org/10.1145/1327452.1327492>
- Jack B. Dennis. 1974. First Version of a Data Flow Procedure Language. In *Programming Symposium*, B. Robinet (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 362–376. https://doi.org/10.1007/3-540-06859-7_145
- Tom Duff. 1990. Rc—a Shell for Plan 9 and UNIX. (1990), 283–296.
- Jeff Epstein, Andrew P. Black, and Simon Peyton-Jones. 2011. Towards Haskell in the Cloud. In *Proceedings of the 4th ACM Symposium on Haskell (Haskell '11)*. ACM, New York, NY, USA, 118–129. <https://doi.org/10.1145/2034675.2034690>
- Azadeh Farzan and Victor Nicolet. 2017. Synthesis of Divide and Conquer Parallelism for Loops. In *Proceedings of the 38th ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI 2017)*. Association for Computing Machinery, New York, NY, USA, 540–555. <https://doi.org/10.1145/3062341.3062355>
- Azadeh Farzan and Victor Nicolet. 2019. Modular Divide-and-Conquer Parallelization of Nested Loops. In *Proceedings of the 40th ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI 2019)*. Association for Computing Machinery, New York, NY, USA, 610–624. <https://doi.org/10.1145/3314221.3314612>
- Michael I Gordon, William Thies, and Saman Amarasinghe. 2006. Exploiting coarse-grained task, data, and pipeline parallelism in stream programs. *ACM SIGPLAN Notices* 41, 11 (2006), 151–162.
- Michael Greenberg. 2018. The POSIX shell is an interactive DSL for concurrency.
- Michael Greenberg and Austin J. Blatt. 2020. Executable Formal Semantics for the POSIX Shell: Smoosh: the Symbolic, Mechanized, Observable, Operational Shell. *Proc. ACM Program. Lang.* 4, POPL, Article 43 (Jan. 2020), 31 pages. <https://doi.org/10.1145/3314221.3314612>

[//doi.org/10.1145/3371111](https://doi.org/10.1145/3371111)

- Michael Greenberg, Konstantinos Kallas, and Nikos Vasilakis. 2021a. The Future of the Shell: Unix and Beyond. In *Proceedings of the Workshop on Hot Topics in Operating Systems (HotOS '21)*. Association for Computing Machinery, New York, NY, USA, 240–241. <https://doi.org/10.1145/3458336.3465296>
- Michael Greenberg, Konstantinos Kallas, and Nikos Vasilakis. 2021b. Unix Shell Programming: The Next 50 Years. In *Proceedings of the Workshop on Hot Topics in Operating Systems (HotOS '21)*. Association for Computing Machinery, New York, NY, USA, 104–111. <https://doi.org/10.1145/3458336.3465294>
- Shivam Handa, Konstantinos Kallas, Nikos Vasilakis, and Martin Rinard. 2021. An Order-aware Dataflow Model for Parallel Unix Pipelines (Artifact). <https://doi.org/10.5281/zenodo.4776838>
- Michael Hart. 1971. *Project Gutenberg*. <https://www.gutenberg.org/>
- Dan Jurafsky. 2017. Unix for Poets. <https://web.stanford.edu/class/cs124/lec/124-2018-UnixForPoets.pdf>
- Gilles Kahn. 1974. The Semantics of a Simple Language for Parallel Programming. *Information Processing* 74 (1974), 471–475.
- Gilles Kahn and David B. MacQueen. 1977. Coroutines and Networks of Parallel Processes. *Information Processing* 77 (1977), 993–998.
- Konstantinos Kallas, Filip Niksic, Caleb Stanford, and Rajeev Alur. 2020. DiffStream: Differential Output Testing for Stream Processing Programs. *Proc. ACM Program. Lang.* 4, OOPSLA, Article 153 (Nov. 2020), 29 pages. <https://doi.org/10.1145/3428221>
- Richard M. Karp and Raymond E. Miller. 1966. Properties of a Model for Parallel Computations: Determinacy, Termination, Queueing. *SIAM J. Appl. Math.* 14, 6 (1966), 1390–1411. <https://doi.org/10.1137/0114108>
- Charles Edwin Killian, James W. Anderson, Ryan Braud, Ranjit Jhala, and Amin M. Vahdat. 2007. Mace: Language Support for Building Distributed Systems. In *Proceedings of the 28th ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI '07)*. ACM, New York, NY, USA, 179–188. <https://doi.org/10.1145/1250734.1250755>
- Nokia Bell Labs. 2019. The Unix Game—Solve puzzles using Unix pipes. <https://unixgame.io/unix50> Accessed: 2020-03-05.
- Paul Le Guernic, Albert Benveniste, Patricia Bournai, and Thierry Gautier. 1986. Signal—A data flow-oriented language for signal processing. *IEEE transactions on acoustics, speech, and signal processing* 34, 2 (1986), 362–374.
- Edward Ashford Lee and David G. Messerschmitt. 1987a. Static Scheduling of Synchronous Data Flow Programs for Digital Signal Processing. *IEEE Trans. Comput.* 36, 1 (Jan. 1987), 24–35. <https://doi.org/10.1109/TC.1987.5009446>
- Edward A. Lee and David G. Messerschmitt. 1987b. Synchronous data flow. *Proc. IEEE* 75, 9 (1987), 1235–1245.
- Jin Li, David Maier, Kristin Tuft, Vassilis Papadimos, and Peter A. Tucker. 2005. Semantics and Evaluation Techniques for Window Aggregates in Data Streams. In *Proceedings of the 2005 ACM SIGMOD International Conference on Management of Data (SIGMOD '05)*. Association for Computing Machinery, New York, NY, USA, 311–322. <https://doi.org/10.1145/1066157.1066193>
- Konstantinos Mamouras. 2020. Semantic Foundations for Deterministic Dataflow and Stream Processing. In *European Symposium on Programming*. Springer, Cham, 394–427.
- Konstantinos Mamouras, Mukund Raghothaman, Rajeev Alur, Zachary G. Ives, and Sanjeev Khanna. 2017. StreamQRE: Modular Specification and Efficient Evaluation of Quantitative Queries over Streaming Data. In *Proceedings of the 38th ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI 2017)*. Association for Computing Machinery, New York, NY, USA, 693–708. <https://doi.org/10.1145/3062341.3062369>
- Konstantinos Mamouras, Caleb Stanford, Rajeev Alur, Zachary G. Ives, and Val Tannen. 2019. Data-Trace Types for Distributed Stream Processing Systems. In *Proceedings of the 40th ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI 2019)*. ACM, New York, NY, USA, 670–685. <https://doi.org/10.1145/3314221.3314580>
- Florence Maraninchi and Yann Rémond. 2001. Argos: An Automaton-Based Synchronous Language. *Comput. Lang.* 27, 1–3 (April 2001), 61–92. [https://doi.org/10.1016/S0096-0551\(01\)00016-9](https://doi.org/10.1016/S0096-0551(01)00016-9)
- Chris McDonald and Trevor I. Dix. 1988. Support for Graphs of Processes in a Command Interpreter. *Softw. Pract. Exper.* 18, 10 (Oct. 1988), 1011–1016. <https://doi.org/10.1002/spe.4380181007>
- Malcolm D McIlroy, Elliot N Pinson, and Berkley A Tague. 1978. UNIX Time-Sharing System: Foreword. *Bell System Technical Journal* 57, 6 (1978), 1899–1904.
- Sape J. Mullender, Guido van Rossum, Andrew S. Tanenbaum, Robbert van Renesse, and Hans van Staveren. 1990. Amoeba: A Distributed Operating System for the 1990s. *Computer* 23, 5 (May 1990), 44–53. <https://doi.org/10.1109/2.53354>
- Derek G. Murray, Frank McSherry, Rebecca Isaacs, Michael Isard, Paul Barham, and Martín Abadi. 2013. Naiad: A Timely Dataflow System. In *Proceedings of the Twenty-Fourth ACM Symposium on Operating Systems Principles (SOSP '13)*. ACM, New York, NY, USA, 439–455. <https://doi.org/10.1145/2517349.2522738>
- John K Ousterhout, Andrew R. Cherenon, Fred Douglass, Michael N. Nelson, and Brent B. Welch. 1988. The Sprite network operating system. *Computer* 21, 2 (1988), 23–36. <http://www.research.ibm.com/people/f/douglass/papers/sprite.pdf>
- Rob Pike, Dave Presotto, Ken Thompson, Howard Trickey, et al. 1990. Plan 9 from Bell Labs. In *Proceedings of the summer 1990 UKUUG Conference*. 1–9. <http://css.csail.mit.edu/6.824/2014/papers/plan9.pdf>

- Deepti Raghavan, Sadjad Fouladi, Philip Levis, and Matei Zaharia. 2020. POSH: A Data-Aware Shell. In *2020 USENIX Annual Technical Conference (USENIX ATC 20)*. USENIX Association, 617–631. <https://www.usenix.org/conference/atc20/presentation/raghavan>
- Christophe Ratel, Nicolas Halbwichs, and Pascal Raymond. 1991. Programming and Verifying Critical Systems by Means of the Synchronous Data-Flow Language LUSTRE. (1991), 112–119. <https://doi.org/10.1145/125083.123062>
- Radu Rugina and Martin Rinard. 1999. Automatic Parallelization of Divide and Conquer Algorithms. In *Proceedings of the Seventh ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming (PPoPP '99)*. Association for Computing Machinery, New York, NY, USA, 72–83. <https://doi.org/10.1145/301104.301111>
- Scott Schneider, Martin Hirzel, Bugra Gedik, and Kun-Lung Wu. 2015. Safe Data Parallelism for General Streaming. *IEEE Trans. Comput.* 64, 2 (Feb. 2015), 504–517. <https://doi.org/10.1109/TC.2013.221>
- Peter Sewell, James J. Leifer, Keith Wansbrough, Francesco Zappa Nardelli, Mair Allen-Williams, Pierre Habouzit, and Viktor Vafeiadis. 2005. Acute: High-level Programming Language Design for Distributed Computation. In *Proceedings of the Tenth ACM SIGPLAN International Conference on Functional Programming (ICFP '05)*. ACM, New York, NY, USA, 15–26. <https://doi.org/10.1145/1086365.1086370>
- Calvin Smith and Aws Albarghouthi. 2016. MapReduce Program Synthesis. In *Proceedings of the 37th ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI '16)*. Association for Computing Machinery, New York, NY, USA, 326–340. <https://doi.org/10.1145/2908080.2908102>
- Diomidis Spinellis and Marios Fragkoulis. 2017. Extending Unix Pipelines to DAGs. *IEEE Trans. Comput.* 66, 9 (2017), 1547–1561.
- Ole Tange. 2011. GNU Parallel—The Command-Line Power Tool. *login: The USENIX Magazine* 36, 1 (Feb 2011), 42–47. <https://doi.org/10.5281/zenodo.16303>
- Ole Tange. 2020. Differences Between GNU parallel and Alternatives. https://www.gnu.org/software/parallel/parallel_alternatives.html
- Dave Taylor and Brandon Perry. 2016. *Wicked Cool Shell Scripts: 101 Scripts for Linux, OS X, and UNIX Systems* (2nd ed.). No Starch Press, USA.
- William Thies, Michal Karczmarek, and Saman P. Amarasinghe. 2002. StreamIt: A Language for Streaming Applications. In *Proceedings of the 11th International Conference on Compiler Construction (CC '02)*. Springer-Verlag, Berlin, Heidelberg, 179–196.
- Eleftheria Tsaliki and Diomidis Spinellis. 2021. The real statistics of buses in Athens. <https://insidestory.gr/article/noymera-leoforeia-athinas?token=0MFVISB8N6>. Published by the Inside Story news organization.
- Nikos Vasilakis, Konstantinos Kallas, Konstantinos Mamouras, Achilles Benetopoulos, and Lazar Cvetković. 2021. *PaSh: Light-Touch Data-Parallel Shell Processing*. Association for Computing Machinery, New York, NY, USA, 49–66. <https://doi.org/10.1145/3447786.3456228>
- Nikos Vasilakis, Ben Karel, and Jonathan M. Smith. 2015. From Lone Dwarfs to Giant Superclusters: Rethinking Operating System Abstractions for the Cloud. In *Proceedings of the 15th USENIX Conference on Hot Topics in Operating Systems (HOTOS'15)*. USENIX Association, Berkeley, CA, USA, 15–15. <http://dl.acm.org/citation.cfm?id=2831090.2831105>
- Robert Virding, Claes Wikström, Mike Williams, and Joe Armstrong. 1996. *Concurrent Programming in ERLANG (2nd Ed.)*. Prentice Hall International (UK) Ltd., GBR.
- W. Gentsch (Sun Microsystems). 2001. Sun Grid Engine: Towards Creating a Compute Power Grid. In *Proceedings of the 1st International Symposium on Cluster Computing and the Grid (CCGRID '01)*. IEEE Computer Society, USA, 35.
- Edward Walker, Weijia Xu, and Vinoth Chandar. 2009. Composing and Executing Parallel Data-Flow Graphs with Shell Pipes. In *Proceedings of the 4th Workshop on Workflows in Support of Large-Scale Science (WORKS '09)*. Association for Computing Machinery, New York, NY, USA, Article 11, 10 pages. <https://doi.org/10.1145/1645164.1645175>
- Ian Watson and John Gurd. 1979. A prototype data flow computer with token labelling. In *1979 International Workshop on Managing Requirements Knowledge (MARK)*. IEEE, 623–628.
- Andy B Yoo, Morris A Jette, and Mark Grondona. 2003. Slurm: Simple linux utility for resource management. In *Workshop on Job Scheduling Strategies for Parallel Processing*. Springer, 44–60.
- Matei Zaharia, Mosharaf Chowdhury, Tathagata Das, Ankur Dave, Justin Ma, Murphy McCauley, Michael J. Franklin, Scott Shenker, and Ion Stoica. 2012. Resilient Distributed Datasets: A Fault-tolerant Abstraction for In-memory Cluster Computing. In *Proceedings of the 9th USENIX Conference on Networked Systems Design and Implementation (NSDI'12)*. USENIX Association, Berkeley, CA, USA, 2–2. <http://dl.acm.org/citation.cfm?id=2228298.2228301>
- Zhao Zhang, Daniel S. Katz, Timothy G. Armstrong, Justin M. Wozniak, and Ian Foster. 2013. Parallelizing the Execution of Sequential Scripts. In *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis (SC '13)*. Association for Computing Machinery, New York, NY, USA, Article 31, 12 pages. <https://doi.org/10.1145/2503210.2503222>