

# Reversing Parallel Programs with Blocks and Procedures

James Hoey

Department of Informatics  
University of Leicester, UK  
jbh11@leicester.ac.uk

Irek Ulidowski

Department of Informatics  
University of Leicester, UK  
iu30@leicester.ac.uk

Shoji Yuen

Graduate School of Informatics  
Nagoya University, Japan  
yuen@i.nagoya-u.ac.jp

We show how to reverse a while language extended with blocks, local variables, procedures and the interleaving parallel composition. Annotation is defined along with a set of operational semantics capable of storing necessary reversal information, and identifiers are introduced to capture the interleaving order of an execution. Inversion is defined with a set of operational semantics that use saved information to undo an execution. We prove that annotation does not alter the behaviour of the original program, and that inversion correctly restores the initial program state.

## 1 Introduction

Reverse execution of programs is the ability to take the final program state of an execution, undo all effects that were the result of that execution, and restore the exact initial program state. This is a desirable capability as it has applications to many active research areas, including debugging [3] and Parallel Discrete Event Simulation [2]. When combined with parallelism, reverse execution removes issues relating to non-deterministic execution orders, allowing specific execution interleavings to be analysed easily.

In our previous work [10], we described a state-saving approach to reversible execution of an imperative *while language*. Similarly to RCC [14], we generated two versions of a program, the *augmented forwards version* to save all necessary *reversal information* alongside its execution, and the *inverted version* that uses this saved data to undo all changes. We proved that augmentation did not alter the behaviour of our program, and that inversion correctly restores the initial program state. We then experimented with reversing a tiny language containing assignments and interleaving parallel composition.

In this paper, we extend the while language with blocks, local variables and procedures, as well as the parallel composition operator. Local variables mean we must recognise scope, for example different versions of a shared name used in parallel. Issues arise with the traditional approach, specifically with recursion, and calls to the same procedure executing in parallel. Annotation and inversion are defined, allowing this extended language to be executed forwards with state-saving, as well as in reverse using this saved information. The process of assigning identifiers to statements as we execute them is described, focusing on *backtracking order*, where statements are undone in the inverted order of the forwards execution. We mention future work on causal-consistent reversibility [1, 13, 15] in the conclusion.

Consider the example shown in Figure 1, where  $w1.0$  and  $\lambda$  can be ignored. This is a simple model of a restaurant with two entrances. One where a single person is continually allowed to enter, increasing the number of current single guests ( $c$ ), until the total capacity ( $c + r$ ) reaches the maximum ( $m$ ). The other allows a reserved group of two to enter, increasing the number of reserved guests ( $r$ ). Let the initial state be that  $m = 4$ ,  $c = 0$  and  $r = 0$ . The execution begins with two full iterations of the while loop, allowing two people to enter meaning  $c = 2$ . Next, the condition of the loop is evaluated, but the body is not yet executed. Interleaving now occurs, setting  $r$  to 2. Finally, the body of the loop is now executed, before the condition evaluates to false and the loop finishes. The final state is  $m = 4$ ,  $c = 3$  and  $r = 2$ , which should be invalid as the total number of guests ( $c + r$ )  $> m$ . This executed version of the annotated program is

```

1 par {
2   while w1.0 ((m - c - r - 1) >= 0) do
3     c = c + 1 λ;
4   end λ; }
5 { r = 2 λ; }

```

Figure 1: Original program

```

1 par {
2   while w1.0 ((m - c - r - 1) >= 0) do
3     c = c + 1 (λ, [2,4,7]);
4   end (λ, [1,3,5,8,9]); }
5 { r = 2 (λ, [6]); }

```

Figure 2: Executed annotated program

shown in Figure 2, where each statement now has a stack populated with identifiers in the order in which the statement occurred (starting at 1). One solution to finding this bug is reverse execution. The inverted version we generate (which, coincidentally is identical to Figure 2) allows step-by-step reversal, using identifiers to remove non-determinism. Backtracking through this execution removes the difficulties of cyclic debugging where different interleavings can occur. Using this, we can see that we wrongly commit to allowing the third single person to enter, meaning the condition gave true when expected to give false. Examining this further, we can see that the reserved guests are not considered until they have arrived, meaning this condition is not aware that the maximum capacity is actually  $m - 2$ . This is an example of a *race* between the writing and reading of  $r$ . This is fixed using the condition  $((m - c - 2 - 1) >= 0)$ .

Our main contributions are

1. The definition of three sets of operational semantics for our language, namely for *traditional* forwards only execution, annotated forwards execution, and for reverse execution.
2. Annotation allows all necessary state-saving, and the use of identifiers to record the interleaving order of execution. Inversion then uses the saved information to reverse via backtracking order.
3. Results showing that annotation does not alter the behaviour of the original program, and that inversion correctly restores to the initial program state.

We also have a prototype simulator under development. This will be capable of implementing both the forwards and reverse execution, and used for both performance evaluation and validation of our results.

## 1.1 Related Work

Program inversion has been the focus of many works for many years, including Jefferson [12], Gries [8] and Glück and Kawabe [4, 5]. The Reverse C Compiler (RCC) by Perumalla [14] describes a state-saving approach to reversibility of C programs. The Backstroke framework [19] and extensions of it by Schordan et al [17] describe an approach to reversing C++ in the setting of Parallel Discrete Event Simulation [2]. The reversible programming language Janus, worked on in [20, 21], adds additional information into the source code, making all programs reversible. More recent work on reversible imperative programs by Glück and Yokoyama [6, 7] introduce the languages R-WHILE and R-CORE. Reversibility of algebraic process calculi is the focus of work by Phillips and Ulidowski [15, 16], where the notion of identifiers was introduced. There has been work on reversible object oriented programming languages, including that of Schultz [18] and the language ROOPL [9]. The application of reverse computation to debugging of message passing concurrent programs is considered by Giachino et al [3].

## 2 Programming Language, Environments and Scope

Let  $\mathbf{P}$  be the set of all programs and  $\mathbf{S}$  be the set of all statements. Each program  $P$  will be either a statement  $S$ , the sequential composition of programs  $P; Q$  or the parallel composition of programs  $P \text{ par } Q$

(sometimes written as  $\text{par } \{P\}\{Q\}$ ). Each statement will either be a skip operation (empty statement), an assignment, a conditional, a loop, a block, a variable or procedure declaration, a variable or procedure removal or a call. A block consists of the declaration of both local variables DV and procedures DP, a body that uses these, and then the removal of local procedures RP and variables RV. Procedures do not have arguments, static scope is assumed and recursion is permitted. The syntax of this language is shown below, including arithmetic and boolean expressions. Note that the constructs `runC` and `runB` are reserved words that appear in our syntax, but not in original programs, and will be explained in Section 3. These allow static operational semantics to be defined, needed to aid state-saving and in our results.

Many statements contain a path `pa` that is explained in Section 2.2. Each conditional, loop, block, procedure and procedure call statement has a unique identifier named `In`, `Wn`, `Bn`, `Pn` and `Cn` respectively, each of which is an element of the sets **In**, **Wn**, **Bn**, **Pn** and **Cn** respectively. The set union of these gives us the set of *construct identifiers* **CI**. Note a procedure has a name from the set **n** (appears in code, and is potentially duplicated), as well as a unique identifier `Pn`.

$$\begin{aligned}
P &::= \varepsilon \mid S \mid P; P \mid P \text{ par } P \\
S &::= \text{skip} \mid X = E \text{ pa} \mid \text{if } In \ B \ \text{then } P \ \text{else } Q \ \text{end pa} \mid \\
&\quad \text{while } Wn \ B \ \text{do } P \ \text{end pa} \mid \text{begin } Bn \ DV \ DP \ P \ RP \ RV \ \text{end} \mid \\
&\quad \text{call } Cn \ n \ \text{pa} \mid \text{runC } Cn \ P \ \text{end} \mid \text{runB } P \ \text{end} \\
\\
DV &::= \varepsilon \mid \text{var } X = v \ \text{pa}; \ DV & \quad DP &::= \varepsilon \mid \text{proc } Pn \ n \ \text{is } P \ \text{pa}; \ DP \\
RV &::= \varepsilon \mid \text{remove } X = v \ \text{pa}; \ RV & \quad RP &::= \varepsilon \mid \text{remove } Pn \ n \ \text{is } P \ \text{pa}; \ RP \\
E &::= \text{Var} \mid n \mid (E) \mid E \ \text{Op} \ E & \quad B &::= T \mid F \mid \neg B \mid (B) \mid E == E \mid E > E \mid B \wedge B
\end{aligned}$$

## 2.1 Environments

We complete our setting with the definition of several environments. Let **V** be the set of all program variables, **Loc** be the set of all memory locations, and **Num** be the set of integers.

As in [11], we first have a *variable environment*  $\gamma$ , responsible for mapping a variable name and the block to which it is local ( $\lambda$  in the case of global variables) to its bound memory location. This is defined as  $\gamma: (\mathbf{V} \times \mathbf{Bn}) \mapsto \mathbf{Loc}$ . The notation  $\gamma[(X, Bn) \Rightarrow l]$  indicates that the pair  $(X, Bn)$  maps to the memory location  $l$ , while  $\gamma[(X, Bn)]$  represents an update to  $\gamma$  with the mapping for the pair  $(X, Bn)$  removed.

We have a *data store*  $\sigma$ , responsible for mapping each memory location to the value it currently holds, defined as  $\sigma: (\mathbf{Loc} \mapsto \mathbf{Num})$ . The notion  $\sigma[l \mapsto v]$  indicates location  $l$  now holds the value  $v$ .

The *procedure environment*  $\mu$  is responsible for mapping either a procedure or call identifier to both the actual procedure name (used in code) and (a copy of) the body. This environment is defined as  $\mu: (\mathbf{Pn} \cup \mathbf{Cn}) \mapsto (\mathbf{n} \times \mathbf{P})$ . The notation  $\mu[Pn \Rightarrow (n, P)]$  represents that  $Pn$  maps to the pair  $(n, P)$ ,  $\mu[\text{refC}(Pn, P)]$  represents the updating of the mapping for  $Pn$  with changes retrieved from  $P$ , and  $\mu[Pn]$  indicates the removal of the mapping for  $Pn$  (in each case,  $Pn$  could also be  $Cn$ ).

Finally, the *while environment*  $\beta$  is responsible for mapping a unique loop identifier to a copy of that loop. This serves the purpose of storing both the original condition and program, allowing our semantics to be static. This is defined as  $\beta: \mathbf{Wn} \mapsto \mathbf{P}$ . The notation  $\beta[Wn \Rightarrow P]$  indicates that  $Wn$  now maps to the program  $P$ ,  $\beta[\text{refW}(Wn, P)]$  represents the updating of the mapping for  $Wn$  with changes retrieved from  $P$ , and the notation  $\beta[Wn]$  shows the removal of the mapping for  $Wn$ .

We now combine all environments, and use the notation  $\square$  to represent the set  $\{\sigma, \gamma, \mu, \beta\}$ . Each environment has a prime version, indicating a potential and arbitrary change.

## 2.2 Scope

Local variables can share their name with a global variable, as well as local variables declared in different blocks. The traditional method of handling this, as described for example in [11], is to implement a stack of environments, storing a copy for each scope. This is not suitable when we use parallel composition as there can be several active scopes in an execution at once. We therefore implement a single environment that will store all versions of variables. Variables will either be global, or local to a specific block. Using  $\lambda$  to represent the empty block name (a global variable), associating a variable with the identifier of the block in which it is declared is sufficient. Now all versions of a variable name are stored distinctly.

We must be able to access the correct version of a given variable name. Under the traditional approach, each environment will have only one mapping of any variable, something we do not have. We must be able to determine the block identifier in which the variable was defined, which will not necessarily be the current block. We achieve this by assigning a *path* to each statement. Each path will be the sequence of the block identifiers  $B_n$ , for blocks in which this statement resides, separated using ‘\*’. Consider statement  $F = S(b_2 * b_1, A)$  (from Figure 4) that has a path  $b_2 * b_1$ , meaning it occurs within a block  $b_2$ , which is nested within  $b_1$ . Therefore we have the function  $evalV()$ , that takes a variable name and a path, traverses the sequence of block names until the first is found that has a local variable of this name, and uses this block name (or  $\lambda$  if no match) to return the desired memory location. A similar reasoning, and function  $evalP()$ , exists to evaluate potentially shared procedure names, returning the correct unique procedure identifier.

One complication is *code reuse*, where the same program code is executed multiple times, with the two cases being procedure and loop bodies. Consider two calls to the same procedure in parallel. Both may create a local variable of a block, where on each side the block has the same name, meaning both will incorrectly use the same version of the local variable. A similar case exists for recursive calls, as shown in Example 1. Therefore, we *rename* any reused code prior to its execution. This must make all constructs unique, a task achieved using the unique call identifier  $C_n$ . All construct names are modified to now start with the unique call identifier, with paths also updated to reflect changes made to block identifiers. Consider again the statement for  $F$  from Figure 4. When block  $b_2$  is renamed to  $c_1 : c_2 : b_2$ , the path becomes  $c_1 : c_2 : b_2 * b_1$  (Figure 7). This removes the issue described above, as each version of the variable will have a different block identifier. This process must modify the call identifier of any recursive call statement similarly. Therefore we have the function  $reP()$  that implements this renaming of procedure bodies, and as renaming occurs in reverse, we have  $IreP()$ .

The reuse of loop bodies is different, as it is not possible for the same code (with same construct names) to be executed in parallel when not in a procedure body (handled above). However, in order to keep all identifiers unique, and to aid future extensions to causal-consistent reversibility, we *version* each construct name, incrementing it by 1 for each loop iteration. For example, a conditional statement  $i1.0$  will become  $i1.1$ . These versions are maintained via the function  $nextID()$ , used in  $reL()$  that performs this renaming, and the function  $previousID()$ , used in the reverse renaming function  $IreL()$ .

## 3 Forwards Only Operational Semantics

We now define the traditional, forwards only semantics of our language. We give a set of transition rules for each construct of our language and a set of environments. These rules specify how *configurations* (namely, pairs containing a program and a set of environments) compute by performing single transition steps. The transition rules define our small step transition relation *configuration*  $\hookrightarrow$  *configuration*. The transitive closure  $\hookrightarrow^*$  represents executions of programs. The forwards only semantics do not perform

any state-saving, thus making them irreversible. Transitions labelled with  $a$  or  $b$  are steps of arithmetic or boolean expression evaluation respectively, while  $\hookrightarrow_a^*$  and  $\hookrightarrow_b^*$  are the transitive closure of each. Semantics of both are omitted as they are as expected, see [11]. By abuse of notation, we now use  $\square$  to represent all environments of the set  $\{\sigma, \gamma, \mu, \beta\}$  that are not modified via the specific rule. For example, if a rule only changes the procedure environment  $\mu$ , then  $\square$  will be the set  $\{\sigma, \gamma, \beta\}$ . The semantics listed below are static, necessary for later sections including state-saving and our results.

**Sequential and Parallel Composition** Programs can be of the form  $S;P$  or  $P \text{ par } Q$ . As such, programs either execute sequentially, or allow each side of a parallel statement to interleave their execution.

$$\begin{array}{ll}
\text{[S1]} \quad \frac{(S \mid \square) \hookrightarrow (S' \mid \square')}{(S; P \mid \square) \hookrightarrow (S'; P \mid \square')} & \text{[S2]} \quad \frac{}{(\text{skip}; P \mid \square) \hookrightarrow (P \mid \square)} \\
\text{[P1]} \quad \frac{(P \mid \square) \hookrightarrow (P' \mid \square')}{(P \text{ par } Q \mid \square) \hookrightarrow (P' \text{ par } Q \mid \square')} & \text{[P2]} \quad \frac{(Q \mid \square) \hookrightarrow (Q' \mid \square')}{(P \text{ par } Q \mid \square) \hookrightarrow (P \text{ par } Q' \mid \square')} \\
\text{[P3]} \quad \frac{}{(P \text{ par skip} \mid \square) \hookrightarrow (P \mid \square)} & \text{[P4]} \quad \frac{}{(\text{skip par } Q \mid \square) \hookrightarrow (Q \mid \square)}
\end{array}$$

**Assignment** All assignments are considered destructive, with the overwritten value being lost. A single atomic rule both evaluates the expression and assigns the new value to the appropriate memory location. Arithmetic expressions do not contain side effects, meaning evaluation of these does not change the environments, as shown in rule [D1]. Similarly, this is also the case for boolean expressions, as shown in rule [I1] and there after.

$$\text{[D1]} \quad \frac{(e \text{ pa} \mid \sigma, \gamma, \square) \hookrightarrow_a^* (v \mid \sigma, \gamma, \square) \quad \text{evalV}(\gamma, \text{pa}, X) = 1}{(X = e \text{ pa} \mid \sigma, \gamma, \square) \hookrightarrow (\text{skip} \mid \sigma[1 \mapsto v], \gamma, \square)}$$

**Conditional** Condition evaluation is atomic via [I1], before the appropriate branch is executed completely to skip (potentially interleaved).

$$\begin{array}{l}
\text{[I1]} \quad \frac{(b \text{ pa} \mid \square) \hookrightarrow_b^* (V \mid \square)}{(\text{if In } b \text{ then } P \text{ else } Q \text{ end pa} \mid \square) \hookrightarrow (\text{if In } V \text{ then } P \text{ else } Q \text{ end pa} \mid \square)} \\
\text{[I2]} \quad \frac{(P \mid \square) \hookrightarrow (P' \mid \square')}{(\text{if In } T \text{ then } P \text{ else } Q \text{ end pa} \mid \square) \hookrightarrow (\text{if In } T \text{ then } P' \text{ else } Q \text{ end pa} \mid \square')} \\
\text{[I3]} \quad \frac{(Q \mid \square) \hookrightarrow (Q' \mid \square')}{(\text{if In } F \text{ then } P \text{ else } Q \text{ end pa} \mid \square) \hookrightarrow (\text{if In } F \text{ then } P \text{ else } Q' \text{ end pa} \mid \square')} \\
\text{[I4]} \quad \frac{}{(\text{if In } T \text{ then skip else } Q \text{ end pa} \mid \square) \hookrightarrow (\text{skip} \mid \square)} \\
\text{[I5]} \quad \frac{}{(\text{if In } F \text{ then } P \text{ else skip end pa} \mid \square) \hookrightarrow (\text{skip} \mid \square)}
\end{array}$$

**While Loop** Evaluation of the condition is always atomic. [W1] handles the first iteration of a loop, where no mapping for  $Wn$  is present in  $\beta$ . The mapping  $Wn \Rightarrow R$  is inserted, and the condition is evaluated. [W2] handles any other iteration, evaluating the condition and updating the mapping within  $\beta$  to  $Wn \Rightarrow R'$ . Both rules rename the body, making constructs unique. [W3] executes the body. [W4] continues the loop using the program  $P$ , retrieved from  $\beta$  for  $Wn$ , until the condition is false, when [W5] will conclude the statement. The premise  $\beta(Wn) = R$  indicates an arbitrary mapping exists. This is necessary as the rule [W5] requires the removal of some mapping via  $\beta[Wn]$ . Note that these semantics (and the semantics defined in Sections 4 and 5) are correct for all while loops with conditions that require evaluation. In the case of  $b$  initially being  $T$  or  $F$ , there may be ambiguity in our rules.

$$\begin{aligned}
\text{[W1]} \quad & \frac{\beta(\text{Wn}) = \text{und} \quad (\text{b pa} \mid \beta, \square) \hookrightarrow_b^* (V \mid \beta, \square)}{(\text{while Wn b do P end pa} \mid \beta, \square) \leftrightarrow (\text{while Wn V do } \text{reL}(P) \text{ end pa} \mid \beta[\text{Wn} \Rightarrow R], \square)} \\
& \text{where } R = \text{while Wn b do } \text{reL}(P) \text{ end pa} \\
\text{[W2]} \quad & \frac{\beta(\text{Wn}) = \text{while Wn b do P end pa} \quad (\text{b pa} \mid \beta, \square) \hookrightarrow_b^* (V \mid \beta, \square)}{(\text{while Wn b do P end pa} \mid \beta, \square) \leftrightarrow (\text{while Wn V do } \text{reL}(P) \text{ end pa} \mid \beta[\text{Wn} \Rightarrow R'], \square)} \\
& \text{where } R' = \text{while Wn b do } \text{reL}(P) \text{ end pa} \\
\text{[W3]} \quad & \frac{(R \mid \square) \leftrightarrow (R' \mid \square')}{(\text{while Wn T do R end pa} \mid \square) \leftrightarrow (\text{while Wn T do R' end pa} \mid \square')} \\
\text{[W4]} \quad & \frac{\beta(\text{Wn}) = P}{(\text{while Wn T do skip end pa} \mid \beta, \square) \leftrightarrow (P \mid \beta, \square)} \\
\text{[W5]} \quad & \frac{\beta(\text{Wn}) = R}{(\text{while Wn F do P end pa} \mid \beta, \square) \leftrightarrow (\text{skip} \mid \beta[\text{Wn}], \square)}
\end{aligned}$$

**Block** Blocks begin with [B1] that creates the runB construct. This then executes the block body via [B2], beginning with the declaration of local variables and procedures. The program then executes using these local definitions, before all such information is removed. Finally [B3] concludes the statement.

$$\begin{aligned}
\text{[B1]} \quad & \frac{}{(\text{begin Bn P end} \mid \square) \leftrightarrow (\text{runB P end} \mid \square)} \quad \text{where } P = DV; DP; Q; RP; RV \\
\text{[B2]} \quad & \frac{(P \mid \square) \leftrightarrow (P' \mid \square')}{(\text{runB P end} \mid \square) \leftrightarrow (\text{runB P' end} \mid \square')} \quad \text{[B3]} \quad \frac{}{(\text{runB skip end} \mid \square) \leftrightarrow (\text{skip} \mid \square)}
\end{aligned}$$

**Variable and Procedure Declaration** A variable declaration [L1] associates the given variable name and current block name Bn (first element of the sequence pa, written Bn\*pa') to the next available memory location l (via nextLoc()) in  $\gamma$ , while also mapping this location to the value v in  $\sigma$  (via the notation  $l \mapsto v$ ). A procedure declaration [L2] inserts the basis mapping between the unique procedure identifier Pn and a pair containing both the procedure name and the procedure body (n, P). A call statement uses this mapping to create a renamed version.

$$\begin{aligned}
\text{[L1]} \quad & \frac{\text{nextLoc}() = l \quad \text{pa} = \text{Bn*pa}'}{(\text{var X} = v \text{ pa} \mid \sigma, \gamma, \square) \leftrightarrow (\text{skip} \mid \sigma[l \mapsto v], \gamma[(X, \text{Bn}) \Rightarrow l], \square)} \\
\text{[L2]} \quad & \frac{}{(\text{proc Pn n is P pa} \mid \mu, \square) \leftrightarrow (\text{skip} \mid \mu[\text{Pn} \Rightarrow (n, P)], \square)}
\end{aligned}$$

**Procedure Call** [G1] evaluates the procedure name to Pn, and retrieves the basis entry (n, P) from  $\mu$ . The renamed version of P, written P', is then inserted into  $\mu$  via the mapping Cn  $\Rightarrow$  (n, P'), and the runC construct is formed. [G2] executes the body of the call statement, before [G3] concludes the statement by removing the mapping for Cn within  $\mu$ , written  $\mu[\text{Cn}]$ .

$$\begin{aligned}
\text{[G1]} \quad & \frac{\text{evalP}(n, \text{pa}) = \text{Pn} \quad \mu(\text{Pn}) = (n, P) \quad \text{reP}(P, \text{Cn}) = P'}{(\text{call Cn n pa} \mid \mu, \square) \leftrightarrow (\text{runC Cn P' end} \mid \mu[\text{Cn} \Rightarrow (n, P')], \square)} \\
\text{[G2]} \quad & \frac{(P \mid \square) \leftrightarrow (P' \mid \square')}{(\text{runC Cn P end} \mid \square) \leftrightarrow (\text{runC Cn P' end} \mid \square')} \\
\text{[G3]} \quad & \frac{}{(\text{runC Cn skip end} \mid \mu, \square) \leftrightarrow (\text{skip} \mid \mu[\text{Cn}], \square)}
\end{aligned}$$

**Variable and Procedure Removal** A local variable  $X$  is local to the inner most block  $B_n$  of its path, written as  $B_n * pa'$ . Removal of a variable [H1] removes the mapping of  $(X, B_n)$  from  $\gamma$ , written  $\gamma[(X, B_n)]$ . The location associated with this mapping is set to 0 ( $1 \mapsto 0$ ) and marked free for future use. A procedure removal [H2] removes the mapping for the given procedure identifier  $P_n$ , written as  $\mu[P_n]$ .

$$[H1] \quad \frac{pa = B_n * pa' \quad \gamma(X, B_n) = 1}{(\text{remove } X = v \text{ pa} \mid \sigma, \gamma, \square) \leftrightarrow (\text{skip} \mid \sigma[1 \mapsto 0], \gamma[(X, B_n)], \square)}$$

$$[H2] \quad \frac{}{(\text{remove } P_n \text{ n is } P \text{ pa} \mid \mu, \square) \leftrightarrow (\text{skip} \mid \mu[P_n], \square)}$$

## 4 Forwards Semantics of Annotated Programs

Annotation is the process of generating our annotated version of a program. This will make small changes to the syntax of our program, adding the capability of storing necessary reversal information. Before defining this in detail, we must have an environment for storing this information, keeping it separate from the program state. To do this, we use an updated version of our auxiliary store  $\delta$  [10]. There is a stack for each program variable name, storing any overwritten values the variable holds throughout the execution, whether the variable is global, local or both. Using one stack for all versions of a variable helps us to devise a technique to handle *races* on that variable. There is a single stack  $B$  for *all* conditional statements. After completion of the conditional, a pair containing an identifier and a boolean value indicating which branch was executed will be saved. Identifiers allow us to resolve any races, meaning all pairs for all conditionals can be pushed to a single stack. There is a single stack  $W$  for *all* while loops. As explained and illustrated in our previous work [10], this stack will contain pairs of an identifier and a boolean value. These pairs produce a sequence of boolean values necessary for inverse execution. Stack  $WI$  stores all *annotation information* (order of identifiers) of a while loop, before it's removed from  $\beta$ . Stack  $Pr$  performs a similar task for procedure bodies. Let  $\mathbf{V}$  be the set of all variables,  $\mathbb{S}(\mathbf{V})$  be a set of stacks, one for each element of  $\mathbf{V}$ ,  $\mathbb{B}$  be the set of boolean values,  $\mathbb{C}$  be the annotation information and  $\mathbf{K}$  be the set of identifiers. Then  $\delta : (\mathbb{S}(\mathbf{V}) \mapsto (\mathbf{K} \times \mathbf{Num})) \cup (\mathbb{S}(B) \mapsto (\mathbf{K} \times \mathbb{B})) \cup (\mathbb{S}(W) \mapsto (\mathbf{K} \times \mathbb{B})) \cup (\mathbb{S}(WI) \mapsto (\mathbf{K} \times \mathbb{C})) \cup (\mathbb{S}(Pr) \mapsto (\mathbf{K} \times \mathbb{C}))$ . The notation  $\delta[e1 \mapsto st]$  pushes  $e1$  to the stack  $st$ , while  $\delta[st/st']$  pops the top element of stack  $st$ , leaving the remaining stack  $st'$ .

We now define annotation. Each statement, excluding blocks and parallel, receives a stack  $A$  for identifiers. The association of a statement to its stack persists throughout the execution. Each time a statement executes, the (global and atomic) function  $next()$  retrieves the next available identifier, which is pushed to that statements stack (and  $\delta$  if necessary). The functions  $ann()$  and  $a()$  are defined below.

$$ann(\varepsilon) = \varepsilon \quad ann(S;P) = a(S);ann(P) \quad ann(P \text{ par } Q) = ann(P) \text{ par } ann(Q)$$

$$a(\text{skip}) = \text{skip } I$$

$$a(X = e \text{ pa}) = X = e \text{ (pa, A)}$$

$$a(\text{if } In \text{ b then } P \text{ else } Q \text{ end pa}) = \text{if } In \text{ b then } ann(P) \text{ else } ann(Q) \text{ end (pa, A)}$$

$$a(\text{while } Wn \text{ b do } P \text{ end pa}) = \text{while } Wn \text{ b do } ann(P) \text{ end (pa, A)}$$

$$a(\text{begin } Bn \text{ DV DP P RP RV end}) = \text{begin } Bn \text{ ann(DV) ann(DP) ann(P) ann(RP) ann(RV) end}$$

$$a(\text{var } X = v \text{ pa}) = \text{var } X = v \text{ (pa, A)}$$

$$a(\text{proc } Pn \text{ n is } P \text{ pa}) = \text{proc } Pn \text{ n is } ann(P) \text{ (pa, A)}$$

$$a(\text{call } Cn \text{ n pa}) = \text{call } Cn \text{ n (pa, A)}$$

$$a(\text{remove } X = v \text{ pa}) = \text{remove } X = v \text{ (pa, A)}$$

$$a(\text{remove } Pn \text{ n is } P \text{ pa}) = \text{remove } Pn \text{ n is } ann(P) \text{ (pa, A)}$$

```

1 begin b1
2  proc p1 fib is
3    begin b2
4      var T = 0 b2;
5      if i1 (N - 2 > 0) then
6        T = F + S b2;
7        F = S b2;
8        S = T b2;
9        N = N - 1 b2;
10     call c2 fib b2;
11   end b2
12   remove T = 0 b2;
13 end
14 end b1
15 call c1 fib b1;
16 remove p1 fib is P b1;
17 end

```

Figure 3: Original program

```

1 begin b1
2  proc p1 fib is
3    begin b2
4      var T = 0 (b2*b1,A);
5      if i1 (N - 2 > 0) then
6        T = F + S (b2*b1,A);
7        F = S (b2*b1,A);
8        S = T (b2*b1,A);
9        N = N - 1 (b2*b1,A);
10     call c2 fib (b2*b1,A);
11   end (b2*b1,A)
12   remove T = 0 (b2,A);
13 end
14 end (b1,A)
15 call c1 fib (b1,A);
16 remove p1 fib is AP (b1,A);
17 end

```

Figure 4: Renamed and annotated program

We use  $I$  to represent either nothing, a path, an identifier stack, or a pair consisting of both a path and an identifier stack. After application of these functions, the resulting annotated version is of the following, modified syntax (with expressions omitted as they match Section 2). We note that  $a(\text{runB } P \text{ end})$  is  $\text{runB } \text{ann}(P) \text{ end}$ , and  $a(\text{runC } C_n P \text{ end})$  is  $\text{runC } C_n \text{ ann}(P) \text{ end } A$ . Execution of the annotated version produces the *executed annotated version*, an identical copy but with populated identifier stacks.

$$\begin{aligned}
 AP &::= \varepsilon \mid AS \mid AP; \quad AP \mid AP \text{ par } AP \\
 AS &::= \text{skip } I \mid X = E \text{ (pa,A)} \mid \text{if } I_n B \text{ then } AP \text{ else } AQ \text{ end (pa,A)} \mid \\
 &\quad \text{while } W_n B \text{ do } AP \text{ end (pa,A)} \mid \text{begin } B_n \text{ ADV ADP AP ARP ARV end} \mid \\
 &\quad \text{call } C_n n \text{ (pa,A)} \mid \text{runC } C_n AP \text{ end } A \mid \text{runB } AP \text{ end} \\
 ADV &::= \varepsilon \mid \text{var } X = v \text{ (pa,A)}; \quad ADV \quad ADP ::= \varepsilon \mid \text{proc } P_n n \text{ is } AP \text{ (pa,A)}; \quad ADP \\
 ARV &::= \varepsilon \mid \text{remove } X = v \text{ (pa,A)}; \quad ARV \quad ARP ::= \varepsilon \mid \text{remove } P_n n \text{ is } AP \text{ (pa,A)}; \quad ARP
 \end{aligned}$$

As all statements within an annotated version will be of this syntax, this must be reflected in our environments, specifically  $\mu$  and  $\beta$  that store programs. As a result, we now use  $\square$  to represent the set of annotated environments, unless explicitly stated otherwise.

**Example 1.** We now consider the example implementation of a Fibonacci sequence using our programming language above, shown in Figure 3. Note that version numbers are omitted due to the absence of while loops, and all paths are initially just the most direct block name. Let  $P$  denote the procedure body shown in the declaration statement, namely lines 3-13 of Figure 3. The procedure removal statement on line 16 then uses this. Assume global variables  $F=3$ ,  $S=4$  and  $N=4$ . This program calculates the  $N$ th element of the Fibonacci sequence beginning with the first and second elements  $F$  and  $S$ . After execution, the  $N$ th element will be the value held by  $S$ . Before we execute this program forwards, it must first be both annotated and renamed, shown in Figure 4. All paths have now been updated to include all block identifiers necessary for execution, and all appropriate statements now have a stack  $A$  for storing identifiers. Line 7 of Figure 3 has become line 7 of Figure 4, which now has the path  $b2*b1$ , meaning this statement appears directly within  $b2$ , and indirectly within  $b1$ .



Prior to defining the operational semantics of this, we must first introduce three functions. The first,  $getAI()$ , returns the order and application of identifiers to a given program. This allows us to extract the annotation information that would otherwise be lost when a while or procedure environment mapping is removed. The second,  $refW()$ , reflects a given annotation change to the copy of the program mapped to the given while identifier. The third,  $refC()$ , is identical, but will reflect a change made to a procedure body using a given call identifier. Recall functions  $reL()$  and  $reP()$  from Section 2.2.

We are now ready to give the operational semantics. The transition rules are those in Section 3, but with  $\hookrightarrow$  replaced with  $\rightarrow$ , and with all state-saving performed. We introduce *m-rules*, the name given to each transition rule that assigns an identifier. All other rules that do not use identifiers are now named *non m-rules*. Transitions  $\xrightarrow{m}$  are also called *identifier transitions*.

**Sequential and Parallel Composition** These are identical to Section 3, but with the annotated syntax.

$$\begin{array}{ll}
\text{[S1a]} \quad \frac{(AS \mid \square) \overset{\circ}{\rightarrow} (AS' \mid \square')}{(AS; AP \mid \square) \overset{\circ}{\rightarrow} (AS'; AP \mid \square')} & \text{[S2a]} \quad \frac{}{(\text{skip } I; AP \mid \square) \rightarrow (AP \mid \square)} \\
\text{[P1a]} \quad \frac{(AP \mid \square) \overset{\circ}{\rightarrow} (AP' \mid \square')}{(AP \text{ par } AQ \mid \square) \overset{\circ}{\rightarrow} (AP' \text{ par } AQ \mid \square')} & \text{[P2a]} \quad \frac{(AQ \mid \square) \overset{\circ}{\rightarrow} (AQ' \mid \square')}{(AP \text{ par } AQ \mid \square) \overset{\circ}{\rightarrow} (AP \text{ par } AQ' \mid \square')} \\
\text{[P3a]} \quad \frac{}{(AP \text{ par skip } I \mid \square) \rightarrow (AP \mid \square)} & \text{[P4a]} \quad \frac{}{(\text{skip } I \text{ par } AQ \mid \square) \rightarrow (AQ \mid \square)}
\end{array}$$

**Assignment** This is an m-rule, saving the old value and the next available identifier  $m$ , retrieved via the function  $next()$ , onto this variables stack on  $\delta$ .

$$\text{[D1a]} \quad \frac{(e \text{ pa} \mid \delta, \sigma, \gamma, \square) \hookrightarrow_a^* (v \mid \delta, \sigma, \gamma, \square) \quad m = next() \quad evalV(\gamma, pa, X) = 1}{(X = e \text{ (pa, A)} \mid \delta, \sigma, \gamma, \square) \xrightarrow{m} (\text{skip } m:A \mid \delta[(m, \sigma(1)) \rightarrow X], \sigma[1 \mapsto v], \gamma, \square)}$$

**Conditional** All rules follow as in Section 3, but with annotated programs, and [I4a] and [I5a] both being m-rules. These save the next available identifier  $m$  (via  $next()$ ) and a boolean value indicating which branch was executed (after execution of the branch) onto stack  $B$  on  $\delta$ .

$$\begin{array}{l}
\text{[I1a]} \quad \frac{(b \text{ pa} \mid \square) \hookrightarrow_b^* (v \mid \square)}{(\text{if In } b \text{ then } AP \text{ else } AQ \text{ end (pa, A)} \mid \square) \rightarrow (\text{if In } v \text{ then } AP \text{ else } AQ \text{ end (pa, A)} \mid \square)} \\
\text{[I2a]} \quad \frac{(AP \mid \square) \overset{\circ}{\rightarrow} (AP' \mid \square')}{(\text{if In } T \text{ then } AP \text{ else } AQ \text{ end (pa, A)} \mid \square) \overset{\circ}{\rightarrow} (\text{if In } T \text{ then } AP' \text{ else } AQ \text{ end (pa, A)} \mid \square')} \\
\text{[I3a]} \quad \frac{(AQ \mid \square) \overset{\circ}{\rightarrow} (AQ' \mid \square')}{(\text{if In } F \text{ then } AP \text{ else } AQ \text{ end (pa, A)} \mid \square) \overset{\circ}{\rightarrow} (\text{if In } F \text{ then } AP \text{ else } AQ' \text{ end (pa, A)} \mid \square')} \\
\text{[I4a]} \quad \frac{m = next()}{(\text{if In } T \text{ then skip } I \text{ else } AQ \text{ end (pa, A)} \mid \delta, \square) \xrightarrow{m} (\text{skip } m:A \mid \delta[(m, T) \rightarrow B], \square)} \\
\text{[I5a]} \quad \frac{m = next()}{(\text{if In } F \text{ then } AP \text{ else skip } I \text{ end (pa, A)} \mid \delta, \square) \xrightarrow{m} (\text{skip } m:A \mid \delta[(m, F) \rightarrow B], \square)}
\end{array}$$

**While Loop** The first two rules are m-rules, saving the next available identifier  $m$  (via  $next()$ ) and an element of the boolean sequence onto stack  $W$  on  $\delta$ . [W1a] handles the first iteration of a loop, creating a mapping on  $\beta$  as in Section 3 (but with annotated programs), and saving the first element  $F$  of the boolean sequence (see [10]). [W2a] handles any other iteration, updating the current mapping as before and saving the next element  $T$  of the boolean sequence (see [10]). Both rules rename the loop body. The

body executes via [W3a], now reflecting all annotation changes of  $AR'$  into the stored copy, written using  $\beta'[refW(Wn, AR')]$ . Finally, a loop either continues via [W4a], or finishes via the m-rule [W5a]. This final rule stores the next available identifier  $m$  and all annotation information ( $getAI()$ ) onto stack  $WI$ , before removing the mapping, written  $\beta[Wn]$ .

$$[W1a] \frac{m = next() \quad \beta(Wn) = und \quad (b \ pa \ | \ \beta, \square) \hookrightarrow_b^* (V \ | \ \beta, \square)}{(S \ | \ \delta, \beta, \square) \xrightarrow{m} (\text{while } Wn \ V \ \text{do } reL(AP) \ \text{end} \ (pa, m:A) \ | \ \delta[(m, F) \rightarrow W], \beta[Wn \Rightarrow AR], \square)}$$

where  $S = \text{while } Wn \ b \ \text{do } AP \ \text{end} \ (pa, A)$  and  $AR = \text{while } Wn \ b \ \text{do } reL(AP) \ \text{end} \ (pa, m:A)$

$$[W2a] \frac{m = next() \quad \beta(Wn) = \text{while } Wn \ b \ \text{do } AQ \ \text{end} \ (pa, A) \quad (b \ pa \ | \ \beta, \square) \hookrightarrow_b^* (V \ | \ \beta, \square)}{(S \ | \ \delta, \beta, \square) \xrightarrow{m} (\text{while } Wn \ V \ \text{do } reL(AQ) \ \text{end} \ (pa, m:A) \ | \ \delta[(m, T) \rightarrow W], \beta[Wn \Rightarrow AR'], \square)}$$

where  $S = \text{while } Wn \ b \ \text{do } AP \ \text{end} \ (pa, A)$  and  $AR' = \text{while } Wn \ b \ \text{do } reL(AQ) \ \text{end} \ (pa, m:A)$

$$[W3a] \frac{(AR \ | \ \beta, \square) \overset{\circ}{\rightarrow} (AR' \ | \ \beta', \square')}{(\text{while } Wn \ T \ \text{do } AR \ \text{end} \ (pa, A) \ | \ \beta, \square) \overset{\circ}{\rightarrow} (\text{while } Wn \ T \ \text{do } AR' \ \text{end} \ (pa, A) \ | \ \beta'', \square')}$$

where  $\beta'' = \beta'[refW(Wn, AR')]$

$$[W4a] \frac{\beta(Wn) = AP}{(\text{while } Wn \ T \ \text{do } skip \ I \ \text{end} \ (pa, A) \ | \ \beta, \square) \rightarrow (AP \ | \ \beta, \square)}$$

$$[W5a] \frac{m = next() \quad \beta(Wn) = AR}{(\text{while } Wn \ F \ \text{do } AP \ \text{end} \ (pa, A) \ | \ \delta, \beta, \square) \xrightarrow{m} (\text{skip } m:A \ | \ \delta[(m, getAI(\beta(Wn))) \rightarrow WI], \beta[Wn], \square)}$$

**Block** These are identical to before, but using the annotated syntax.

$$[B1a] \frac{}{(\text{begin } Bn \ AP \ \text{end} \ | \ \square) \rightarrow (\text{runB } AP \ \text{end} \ | \ \square)} \quad \text{where } AP = ADV; ADP; AQ; ARP; ARV$$

$$[B2a] \frac{(AP \ | \ \square) \overset{\circ}{\rightarrow} (AP' \ | \ \square')}{(\text{runB } AP \ \text{end} \ | \ \square) \overset{\circ}{\rightarrow} (\text{runB } AP' \ \text{end} \ | \ \square')} \quad [B3a] \frac{}{(\text{runB } skip \ I \ \text{end} \ | \ \square) \rightarrow (\text{skip} \ | \ \square)}$$

**Variable and Procedure Declaration** Variable declarations [L1a] are as before, but are now m-rules without state-saving. Procedure declarations [L2a] are also as before, but is also an m-rule without state-saving, and all programs are now annotated.

$$[L1a] \frac{m = next() \quad nextLoc() = 1 \quad pa = Bn * pa'}{(\text{var } X = v \ (pa, A) \ | \ \sigma, \gamma, \square) \xrightarrow{m} (\text{skip } m:A \ | \ \sigma[1 \mapsto v], \gamma[(X, Bn) \Rightarrow 1], \square)}$$

$$[L2a] \frac{m = next()}{(\text{proc } Pn \ n \ \text{is } AP \ (pa, A) \ | \ \mu, \square) \xrightarrow{m} (\text{skip } m:A \ | \ \mu[Pn \Rightarrow (n, AP)], \square)}$$

**Procedure Call** [G1a] inserts a renamed copy of the basis entry onto  $\mu$ , exactly as in Section 3, but with an annotated program. [G2a] uses the  $runC$  construct to execute this renamed copy, but now reflects any annotation changes to the stored copy, written  $\mu'[refC(Cn, AP')]$ . [G3a] is now an m-rule, saving all annotation changes from the copy to the stack  $Pr$ , alongside the next available identifier  $m$  (via  $next()$ ). The renamed copy is removed from  $\mu$ , written as  $\mu[Cn]$ .

$$[G1a] \frac{evalP(n, pa) = Pn \quad \mu(Pn) = (n, AP) \quad reP(AP, Cn) = AP'}{(\text{call } Cn \ n \ (pa, A) \ | \ \mu, \square) \rightarrow (\text{runC } Cn \ AP' \ \text{end} \ A \ | \ \mu[Cn \Rightarrow (n, AP')], \square)}$$

$$[G2a] \frac{(AP \ | \ \mu, \square) \overset{\circ}{\rightarrow} (AP' \ | \ \mu', \square')}{(\text{runC } Cn \ AP \ \text{end} \ A \ | \ \mu, \square) \overset{\circ}{\rightarrow} (\text{runC } Cn \ AP' \ \text{end} \ A \ | \ \mu'[refC(Cn, AP')], \square')}$$

$$[\text{G3a}] \frac{m = \text{next}() \quad \mu(\text{Cn}) = \text{AP}}{(\text{runC Cn skip I end A} \mid \delta, \mu, \square) \xrightarrow{m} (\text{skip } m:A \mid \delta[(m, \text{getAI}(\mu(\text{Cn})) \rightarrow \text{Pr}], \mu[\text{Cn}], \square)}$$

**Variable and Procedure Removal** Variable removal [H1a] is now an m-rule, saving the final value of that variable and the next available identifier  $m$  (via  $\text{next}()$ ) onto that variables stack on  $\delta$ . The mapping is removed as in Section 3. Procedure removal [H2a] is as before, but an m-rule without state-saving.

$$[\text{H1a}] \frac{m = \text{next}() \quad \text{pa} = \text{Bn} * \text{pa}' \quad \gamma(X, \text{Bn}) = 1}{(\text{remove } X = v \text{ (pa, A)} \mid \delta, \sigma, \gamma, \square) \xrightarrow{m} (\text{skip } m:A \mid \delta[(m, \sigma(1)) \rightarrow X], \sigma[1 \mapsto 0], \gamma[(X, \text{Bn})], \square)}$$

$$[\text{H2a}] \frac{m = \text{next}() \quad \mu(\text{Pn}) = \text{AQ}}{(\text{remove Pn n is AP (pa, A)} \mid \mu, \square) \xrightarrow{m} (\text{skip } m:A \mid \mu[\text{Pn}], \square)}$$

## 4.1 Results

We first define equivalence between traditional and annotated environments.

**Definition 4.1.** Let  $\sigma$  be a data store,  $\sigma_1$  be an annotated data store,  $\gamma$  be a variable environment and  $\gamma_1$  be an annotated variable environment. We have  $(\sigma, \gamma)$  is *equivalent* to  $(\sigma_1, \gamma_1)$ , written  $(\sigma, \gamma) \approx_S (\sigma_1, \gamma_1)$ , if and only if  $\text{dom}(\gamma) = \text{dom}(\gamma_1)$  and  $\sigma(\gamma(X, \text{Bn})) = \sigma_1(\gamma_1(X, \text{Bn}))$  for all  $X \in \text{dom}(\gamma)$  and block names  $\text{Bn}$ .

**Definition 4.2.** Let  $\mu$  be a procedure environment, and  $\mu_1$  be an annotated procedure environment. We have that  $\mu$  is *equivalent* to  $\mu_1$ , written  $\mu \approx_P \mu_1$ , if and only if  $\text{dom}(\mu) = \text{dom}(\mu_1)$ ,  $\mu(\text{Pn}) = (n, \text{P})$ ,  $\mu_1(\text{Pn}) = (n, \text{AP})$  and  $\text{removeAnn}(\text{AP}) = \text{P}$  for all  $\text{Pn} \in \text{dom}(\mu)$ . (Note  $\text{Pn}$  could be  $\text{Cn}$  here).

**Definition 4.3.** Let  $\beta$  be a while environment, and  $\beta_1$  be an annotated while environment. We have that  $\beta$  is *equivalent* to  $\beta_1$ , written  $\beta \approx_W \beta_1$ , if and only if  $\text{dom}(\beta) = \text{dom}(\beta_1)$ ,  $\beta(\text{Wn}) = \text{P}$ ,  $\beta_1(\text{Wn}) = \text{AP}$  and  $\text{removeAnn}(\text{AP}) = \text{P}$  for all  $\text{Wn} \in \text{dom}(\beta)$ .

**Definition 4.4.** Let  $\delta$  be an auxiliary store, and  $\delta_1$  be an annotated auxiliary store. Firstly, we define the equivalence of stacks  $\text{St}$  and  $\text{St}'$ , written as  $\text{St} \approx_{ST} \text{St}'$ , as true if both stacks have matching elements. We have that  $\delta$  is *equivalent* to  $\delta_1$ , written  $\delta \approx_A \delta_1$ , if and only if for each stack  $\text{St} \in \text{dom}(\delta)$ , we have  $\delta(\text{St}) \approx_{ST} \delta_1(\text{ST})$ .

**Definition 4.5.** Let  $\square$  represent the set of environments  $\{\sigma, \gamma, \mu, \beta\}$  and  $\square_1$  represent the set of annotated environments  $\{\sigma_1, \gamma_1, \mu_1, \beta_1\}$ . We have that  $\square$  is *equivalent* to  $\square_1$ , written  $\square \approx \square_1$ , if and only if  $(\sigma, \gamma) \approx_S (\sigma_1, \gamma_1)$ ,  $\mu \approx_P \mu_1$  and  $\beta \approx_W \beta_1$ .

We now present our results. Theorem 1 states that identifiers are used in ascending order. Theorem 2 below states if an original program terminates, the annotated version will also, and annotation does not change the behaviour of the program w.r.t the stores  $\square$ , but does produce a populated auxiliary store  $\delta'$ .

**Theorem 1.** *Let AP and AQ be annotated programs,  $\square$  be the set of environments and  $\delta$  be an auxiliary store. If  $(\text{AP} \mid \square, \delta) \xrightarrow{\circ}^* (\text{AP}' \mid \square', \delta') \xrightarrow{n} (\text{AQ} \mid \square'', \delta'') \rightarrow^* (\text{AQ}' \mid \square''', \delta''') \xrightarrow{m} (\text{AQ}'' \mid \square''', \delta''''),$  and the computation  $(\text{AQ} \mid \square'', \delta'') \rightarrow^* (\text{AQ}' \mid \square''', \delta''')$  does not have any identifier transitions, then  $m = n + 1$ .*

*Proof.* The order of identifiers used during execution is maintained using  $\text{next}()$ . The program AP will begin with any number of steps. At some point, a transition occurs that will use the next available identifier  $n$ , while simultaneously incrementing  $\text{next}()$  by one to  $m$ . Any number of non m-rules can then apply, before the next m-action uses  $\text{next}()$  to get  $m$ . Hence,  $m = n + 1$ .  $\square$

**Lemma 1.** *Let P be a program,  $\square$  be the set  $\{\sigma, \gamma, \mu, \beta\}$  of all environments,  $\square_1$  be the set  $\{\sigma_1, \gamma_1, \mu_1, \beta_1\}$  of annotated environments such that  $\square \approx \square_1$  and  $\delta$  be the auxiliary store. If  $(\text{P} \mid \square, \delta) \hookrightarrow (\text{P}' \mid \square', \delta)$ , for some  $\square'$ , then there exists an execution  $(\text{ann}(\text{P}) \mid \square_1, \delta) \xrightarrow{\circ} (\text{P}'' \mid \square'_1, \delta')$ , for some  $\square'_1, \delta'$  and  $\text{P}'' = \text{ann}(\text{P}')$  such that  $\square' \approx \square'_1$ .*

**Theorem 2.** *Let  $P$  be an original program,  $\square$  be the set  $\{\sigma, \gamma, \mu, \beta\}$  of all environments,  $\square_1$  be the set  $\{\sigma_1, \gamma_1, \mu_1, \beta_1\}$  of annotated environments such that  $\square \approx \square_1$  and  $\delta$  be the auxiliary store. If  $(P \mid \square, \delta) \hookrightarrow^* (\text{skip} \mid \square', \delta)$ , for some  $\square'$ , then there exists an execution  $(\text{ann}(P) \mid \square_1, \delta) \xrightarrow{\circ}^* (\text{skip} \mid \square_1, \delta')$ , for some  $\square_1$ ,  $\square'_1$  and  $\delta'$ , such that  $\square' \approx \square'_1$ .*

*Proof.* The proof is by induction on the length of the sequence  $(P \mid \square, \delta) \hookrightarrow^* (\text{skip} \mid \square', \delta)$ . We consider  $P$  being either a sequential or parallel composition of programs. Both cases hold using Lemma 1.  $\square$

We note the implication holds in the opposite direction, but defer proof to future work.

## 5 Reverse Semantics of Inverted Programs

Inversion is the process of generating the inverted version of a given program, produced from the executed annotated version, as the populated identifier stacks are necessary. As the inverted version is a program that executes forwards, inversion inverts the overall statement order. The functions that performs this, namely  $\text{inv}()$  and the supplementary  $i()$ , are defined below.

$$\begin{aligned}
\text{inv}(\varepsilon) &= \varepsilon & \text{inv}(\text{AS}; \text{AP}) &= \text{inv}(\text{AP}); i(\text{AS}) & \text{inv}(\text{AP} \text{ par } \text{AQ}) &= \text{inv}(\text{AP}) \text{ par } \text{inv}(\text{AQ}) \\
i(\text{skip } \text{I}) &= \text{skip } \text{I} \\
i(\text{X} = \text{e} \text{ (pa, A)}) &= \text{X} = \text{e} \text{ (pa, A)} \\
i(\text{if In b then AP else AQ end (pa, A)}) &= \text{if In b then } \text{inv}(\text{AP}) \text{ else } \text{inv}(\text{AQ}) \text{ end (pa, A)} \\
i(\text{while Wn b do AP end (pa, A)}) &= \text{while Wn b do } \text{inv}(\text{AP}) \text{ end (pa, A)} \\
i(\text{begin Bn ADV ADP AP ARP ARV end}) &= \text{begin Bn } \text{inv}(\text{ARV}) \text{ } \text{inv}(\text{ARP}) \text{ } \text{inv}(\text{AP}) \text{ } \text{inv}(\text{ADP}) \\
&\quad \text{inv}(\text{ADV}) \text{ end} \\
i(\text{var X} = \text{v} \text{ (pa, A)}) &= \text{remove X} = \text{v} \text{ (pa, A)} \\
i(\text{proc Pn n is AP (pa, A)}) &= \text{remove Pn n is } \text{inv}(\text{AP}) \text{ (pa, A)} \\
i(\text{call Cn n (pa, A)}) &= \text{call Cn n (pa, A)} \\
i(\text{remove X} = \text{v} \text{ (pa, A)}) &= \text{var X} = \text{v} \text{ (pa, A)} \\
i(\text{remove Pn n is AP (pa, A)}) &= \text{proc Pn n is } \text{inv}(\text{AP}) \text{ (pa, A)}
\end{aligned}$$

All inverted programs are of the annotated syntax in Section 4, but with IP and IS used for inverted programs and statements respectively. The inverse of runB and runC constructs simply invert the body.

Starting with the final state of all environments from the forwards execution, the inverted program will *no longer perform any expression evaluation*, offering potential time saving when compared to traditional cyclic debugging. The result of any expression evaluation that happened during forwards execution is retrieved from the appropriate stack on  $\delta$ . For example, a while loop will iterate until the top element of stack W on  $\delta$  is no longer true. The non-determinism that possibly occurred during the forwards execution will also not feature in the inverted execution. The identifiers assigned to statements ensure that any statement can only execute provided it has the highest unseen identifier. Using the function  $\text{previous}()$ , and starting it with the final value of  $\text{next}()$ , all m-rules will be reversed in *backtracking order*. There is freedom in the order of non m-rules, with examples being parallel skip operations, or parallel block closings. Reversing these in any order produces no adverse effects.

We recognise that all environments will now store inverted programs wherever necessary. Since an inverted program is of the same syntax as an annotated program, our current environments are sufficient. We now return to our example discussed in Section 4.

```

1 begin b1
2  proc p1 fib is
3    begin b2
4      var T = 0 (b2*b1,A);
5      if i1 (N - 2 > 0) then
6        T = F + S (b2*b1,A);
7        F = S (b2*b1,A);
8        S = T (b2*b1,A);
9        N = N - 1 (b2*b1,A);
10     call c2.0 fib (b2*b1,A);
11     end (b2*b1,A)
12     remove T = 0 (b2,A);
13   end
14 end (b1, [1])
15 call c1 fib (b1, [21]);
16 remove p1 fib is AP (b1, [22]);
17 end

```

Figure 5: Executed annotated version

```

1 begin b1
2  proc p1 fib is
3    begin b2
4      var T = 0 (b2*b1,A);
5      if i1 (N - 2 > 0) then
6        call c2 fib (b2*b1,A);
7        N = N - 1 (b2*b1,A);
8        S = T (b2*b1,A);
9        F = S (b2*b1,A);
10     T = F + S (b2*b1,A);
11     end (b2*b1,A)
12     remove T = 0 (b2,A);
13   end
14 end (b1, [22])
15 call c1 fib (b1, [21]);
16 remove p1 fib is IP (b1, [1]);
17 end

```

Figure 6: Inverted version

**Example 2.** We now execute the annotated version of our original program (Figure 4), producing the final annotated version shown in Figure 5. The initial value of  $\text{next}()$  is 1. We enter block  $b_1$ , and perform the procedure declaration, assigning the identifier 1 to its stack. The call statement then happens, performing a renamed copy of the procedure body. This will execute lines 3-9 using identifiers 2-6, before hitting the recursive call. This call will then execute another renamed copy of the procedure body. This renamed copy for the second call is shown in Figure 7, where the unique call name  $c_1:c_2$  has been used to rename all constructs. Lines 1-7 are executed using identifiers 7-11, before again hitting a recursive call. This renamed version is then executed, with the conditional on line 5 evaluating to false, meaning the recursion is now finished. This version uses the identifiers 12-14, before the recursive calls begin to close. The second call (Figure 7) then concludes, executing lines 8-11 using identifiers 15-17. The first call then concludes using the identifiers 18-20. Finally, the original program concludes, using identifier 21 to finish the call statement, and then the last identifier 22 (meaning  $\text{next}() = 23$ ) to remove the procedure declaration. This concludes our execution, producing the final state  $F=7$ ,  $S=11$  and  $N=2$ .

We now consider the inverted execution. Application of the function  $\text{inv}()$  to the executed annotated version (Figure 5) produces the inverted version shown in Figure 6. The initial value of  $\text{previous}()$  will be 22 ( $\text{next}() - 1$ ). Inverse execution begins by opening the outer block and performing the procedure declaration (the inverse of the procedure removal) using identifier 22. The call statement is then entered using identifier 21. A renamed copy of the procedure body is then executed, performing lines 3-6 using identifiers 20-18. The recursive call is then hit, beginning the execution of another renamed copy, shown in Figure 8. Lines 1-4 are executed using identifiers 17-15, before the recursive call is hit again. The third renamed version now executes, and since the condition will be false (retrieved from the stack  $B$  on  $\delta$ ), recursion now stops. This call concludes using identifiers 14-12, before the second call (Figure 8) now concludes lines 5-11 using identifiers 11-7. The first call then concludes using identifiers 6-2. Finally, the original program concludes with the removal of the procedure (inverse of declaration) using identifier 1. This execution order restores the initial program state.

Prior to defining the inverse operational semantics, we must introduce the function  $\text{setAI}()$ . This takes

```

1 begin c1:c2:b2
2 var T = 0 (c1:c2:b2*b1, [7]);
3 if c1:c2:i1 (N - 2 > 0) then
4 T = F + S (c1:c2:b2*b1, [8]);
5 F = S (c1:c2:b2*b1, [9]);
6 S = T (c1:c2:b2*b1, [10]);
7 N = N - 1 (c1:c2:b2*b1, [11]);
8 call c1:c2:c2 fib (c1:c2:b2*b1, [15]);
9 end (c1:c2:b2*b1, [16])
10 remove T = 0 (c1:c2:b2*b1, [17]);
11 end

```

Figure 7: Executed annotated version of 2nd call

```

1 begin c1:c2:b2
2 var T = 0 (c1:c2:b2*b1, [17]);
3 if c1:c2:i1 (N - 2 > 0) then
4 call c1:c2:c2 fib (c1:c2:b2*b1, [15]);
5 N = N - 1 (c1:c2:b2*b1, [11]);
6 S = T (c1:c2:b2*b1, [10]);
7 F = S (c1:c2:b2*b1, [9]);
8 T = F + S (c1:c2:b2*b1, [8]);
9 end (c1:c2:b2*b1, [16])
10 remove T = 0 (c1:c2:b2*b1, [7]);
11 end

```

Figure 8: Inverted version of 2nd call

the output of the function *getAI()* from Section 4 and a program, and returns a copy of this program with the given annotation information inserted. Recall the functions *IreP* and *IreL* from Section 2.2.

We now give the operational semantics of inverted programs. We introduce *reverse m-rules*, the name given to all statements of the inverse execution that use an identifier. Transitions  $\overset{m}{\rightsquigarrow}$  are also called reverse identifier transitions. All other rules remain non m-rules. We note a correspondence between each m-rule and the matching reverse m-rule.

**Sequential and Parallel Composition** The inverted program is still executed forwards, meaning these are like those in Section 4, but with  $\rightsquigarrow$  replacing  $\rightarrow$ , and IP and IS replacing AP and AS respectively. Each rule is named correspondingly to Section 4, but with the appended ‘a’ replaced with ‘r’. For example, rule [S1a] is now [S1r].

**Assignment** The inverse of an assignment will be a reverse m-rule, allowed to execute provided the top element of stack A is m (written  $A = m:A'$ ), and m is the last used identifier (via *previous()*). This retrieves the old value, with matching identifier m, from the stack on  $\delta$  for this variable, and assigns it to the corresponding location (evaluated as in Section 4).

$$[D1r] \frac{A = m:A' \quad m = \text{previous()} \quad \text{evalVar}(\gamma, pa, X) = 1 \quad \delta(X) = (m, v):X'}{(X = e \ (pa, A) \mid \delta, \sigma, \square) \overset{m}{\rightsquigarrow} (\text{skip } A' \mid \delta[X/X'], \sigma[1 \mapsto v], \square)}$$

**Conditional** This will begin with the reverse m-rule [I1r] that, provided this statement has the next identifier to invert (via *previous()* and  $A = m:A'$ ) retrieves the boolean value from the stack B with matching identifier m on  $\delta$ . Rules [I2r] and [I3r] follow Section 4, but with inverted programs. Finally, [I4r] and [I5r] simply concludes the statement.

$$[I1r] \frac{A = m:A' \quad m = \text{previous()} \quad \delta(B) = (m, V):B'}{(S \mid \delta, \square) \overset{m}{\rightsquigarrow} (\text{if In } V \text{ then IP else IQ end } (pa, A') \mid \delta[B/B'], \square)}$$

where  $S = \text{if In } b \text{ then IP else IQ end } (pa, A)$

$$[I2r] \frac{(IP \mid \square) \overset{\circ}{\rightsquigarrow} (IP' \mid \square')}{(\text{if In } T \text{ then IP else IQ end } (pa, A) \mid \square) \overset{\circ}{\rightsquigarrow} (\text{if In } T \text{ then IP' else IQ end } (pa, A) \mid \square')}$$

$$[I3r] \frac{(IQ \mid \square) \overset{\circ}{\rightsquigarrow} (IQ' \mid \square')}{(\text{if In } F \text{ then IP else IQ end } (pa, A), \square) \overset{\circ}{\rightsquigarrow} (\text{if In } F \text{ then IP else IQ' end } (pa, A) \mid \square')}$$

$$[I4r] \frac{}{(\text{if In } T \text{ then skip I else IQ end } (pa, A) \mid \square) \rightsquigarrow (\text{skip} \mid \square')}$$

$$[I5r] \frac{}{(\text{if In F then IP else skip I end (pa,A)} \mid \square) \rightsquigarrow (\text{skip} \mid \square)}$$

**While Loop** The reverse m-rule [W1r] handles the first iteration of a loop, retrieving either the T or F from the stack B on  $\delta$ , and creating a mapping on  $\beta$ . This mapping is similar to that of [W1a] but with an inverted copy of the loop IP' that is both renamed and updated with annotated information C, retrieved from the stack WI (via *setAI()*). The reverse m-rule [W2r] handles all iterations except the first, meaning the renaming is applied to the current mapping, written  $\beta[Wn \Rightarrow IR']$ . Both rules only execute provided they have the last used identifier (via *previous()* and  $A = m:A'$ ). The body is then executed repeatedly with changes reflected to the stored copy, written  $\beta'[Wn \Rightarrow IR']$ , via rule [W3r]. The loop continues through the rule [W4r], until the condition is false and thus the mapping removed by rule [W5r].

$$[W1r] \frac{m = \text{previous()} \quad A = m:A' \quad \beta(Wn) = \text{und} \quad \delta(WI) = (m,C):WI' \quad IP' = \text{IreL}(\text{setAI}(IP,C))}{(S \mid \delta, \beta, \square) \rightsquigarrow (\text{while Wn b do IP}' \text{ end (pa,A')} \mid \delta[WI/WI'], \beta[Wn \Rightarrow IR'], \square)}$$

where  $S = \text{while Wn b do IP end (pa,A)}$  and  $IR = \text{while Wn b do IP}' \text{ end (pa,A')}$

$$[W2r] \frac{m = \text{previous()} \quad A = m:A' \quad \beta(Wn) = \text{while Wn b do IQ end (pa,A)} \quad \delta(W) = (m,V):W'}{(S \mid \delta, \beta, \square) \rightsquigarrow (\text{while Wn V do IreL}(IQ) \text{ end (pa,A')} \mid \delta[W/W'], \beta[Wn \Rightarrow IR'], \square)}$$

where  $S = \text{while Wn b do IP end (pa,A)}$  and  $IR' = \text{while Wn b do IreL}(IQ) \text{ end (pa,A')}$

$$[W3r] \frac{(\text{IR} \mid \delta, \beta, \square) \overset{\circ}{\rightsquigarrow} (\text{IR}' \mid \delta', \beta', \square')}{(\text{while Wn T do IR end (pa,A)} \mid \delta, \beta, \square) \overset{\circ}{\rightsquigarrow} (\text{while Wn T do IR}' \text{ end (pa,A)} \mid \delta', \beta'', \square')}$$

where  $\beta'' = \beta'[\text{refW}(Wn, IR')]$

$$[W4r] \frac{\beta(Wn) = IP}{(\text{while Wn T do skip I end (pa,A)} \mid \beta, \square) \rightsquigarrow (IP \mid \beta, \square)}$$

$$[W5r] \frac{\beta(Wn) = IR}{(\text{while Wn F do IP end (pa,A)} \mid \beta, \square) \rightsquigarrow (\text{skip} \mid \beta[Wn], \square)}$$

**Block** The inversion of a block is very similar to that of Section 4, but with the inverted syntax. Recall that a variable declaration in the inverted syntax is the inverse of a variable removal, and similarly for procedures. This means an inverted block has a body of the form shown below as IP in [B1r].

$$[B1r] \frac{}{(\text{begin Bn IP end} \mid \square) \rightsquigarrow (\text{runB IP end} \mid \square)} \quad \text{where IP} = \text{IDV}; \text{IDP}; \text{IQ}; \text{IRP}; \text{IRV}$$

$$[B2r] \frac{(\text{IP} \mid \square) \overset{\circ}{\rightsquigarrow} (\text{IP}' \mid \square')}{(\text{runB IP end} \mid \square) \overset{\circ}{\rightsquigarrow} (\text{runB IP}' \text{ end} \mid \square')} \quad [B3r] \frac{}{(\text{runB skip I end} \mid \square) \rightsquigarrow (\text{skip} \mid \square)}$$

**Variable and Procedure Declaration** Reverse variable declaration [L1r] is a reverse m-rule, allowed to execute provided it has the last used identifier (via *previous()* and  $A = m:A'$ ). The given value is ignored, and the variable is instead set to its final value retrieved from the corresponding stack on  $\delta$  (written  $\delta(X) = (m, v'):X'$ ). The location 1 is used as in Section 4. A procedure declaration [L2r] creates the basis mapping as in Section 4, but with inverted programs, provided its identifiers allow this.

$$[L1r] \frac{A = m:A' \quad m = \text{previous()} \quad \delta(X) = (m, v'):X' \quad \text{nextLoc}() = 1 \quad \text{pa} = \text{Bn} * \text{pa}'}{(\text{var X} = v \text{ (pa,A)} \mid \delta, \sigma, \gamma, \square) \rightsquigarrow (\text{skip } A' \mid \delta[X/X'], \sigma[1 \mapsto v'], \gamma[(X, \text{Bn}) \Rightarrow 1], \square)}$$

$$[L2r] \frac{A = m:A' \quad m = \text{previous}()}{(\text{proc } Pn \ n \text{ is } IP \ (pa,A) \mid \mu, \square) \overset{m}{\rightsquigarrow} (\text{skip } A' \mid \mu[Pn \Rightarrow (n, IP)], \square)}$$

**Procedure Call** The reverse m-rule [G1r] creates a renamed copy  $IP'$  of the basis procedure body  $IP$ , that has annotated changes  $C$  from stack  $WI$  inserted. This is inserted into  $\mu$  via  $\mu[Cn \Rightarrow (n, IP')]$ . Rule [G2r] follows Section 4 but uses inverted programs, while [G3r] removes the mapping.

$$[G1r] \frac{m = \text{previous}() \quad A = m:A' \quad \mu(\text{eval}P(n, pa)) = (n, IP) \quad \delta(Pr) = (m, C) : Pr'}{(\text{call } Cn \ n \ (pa,A) \mid \mu, \square) \overset{m}{\rightsquigarrow} (\text{runC } Cn \ IP' \ \text{end } A' \mid \mu[Cn \Rightarrow (n, IP')], \square)}$$

where  $IP' = \text{Ire}P(\text{setAI}(IP, C), Cn)$

$$[G2r] \frac{(\text{IP} \mid \mu, \square) \overset{\circ}{\rightsquigarrow} (\text{IP}' \mid \mu', \square')}{(\text{runC } Cn \ IP \ \text{end } A \mid \mu, \square) \overset{\circ}{\rightsquigarrow} (\text{runC } Cn \ IP' \ \text{end } A \mid \mu'[\text{ref}C(Cn, IP')], \square')}$$

$$[G3r] \frac{\mu(Cn) = IP}{(\text{runC } Cn \ \text{skip } I \ \text{end } A \mid \mu, \square) \rightsquigarrow (\text{skip } A \mid \mu[Cn], \square)}$$

**Variable and Procedure Removal** Variable removal [H1r] is similar to Section 3, as no state-saving is required. But it is a reverse m-rule and can execute provided  $A = m:A'$  and  $m = \text{previous}()$ . Procedure removal [H2r] is a reverse m-rule with no state-saving, removing the mapping.

$$[H1r] \frac{A = m:A' \quad m = \text{previous}() \quad pa = Bn * pa' \quad \gamma(X, Bn) = 1}{(\text{remove } X = v \ (pa,A) \mid \delta, \sigma, \gamma, \square) \overset{m}{\rightsquigarrow} (\text{skip } A' \mid \delta, \sigma[1 \mapsto 0], \gamma[(X, Bn)], \square)}$$

$$[H2r] \frac{A = m:A' \quad m = \text{previous}() \quad \mu(Pn) = IQ}{(\text{remove } Pn \ n \text{ is } IP \ (pa,A) \mid \mu, \square) \overset{m}{\rightsquigarrow} (\text{skip } A' \mid \mu[Pn], \square)}$$

## 5.1 Results

Prior to describing our inversion results, we first note that the definitions in Section 4.1 are now used to relate annotated environments with inverted environments, instead of traditional. As an example, Definition 4.2 would no longer have that  $P = \text{removeAnn}(AP)$ , but instead that  $IP = \text{inv}(P)$ .

Theorem 3 states that identifiers are used in descending order throughout a reverse execution (opposite of Theorem 1). Theorem 4 shows that if an original *sequential* program and its annotated execution terminate, then the reverse execution will also, and that the reverse execution beginning in the final state can restore the initial state.

**Theorem 3.** *Let  $P$  and  $Q$  be original programs,  $AP$  and  $AQ$  be the annotated versions producing the executed versions  $AP'$  and  $AQ'$  respectively, and  $IP$  and  $IQ$  be the inverted versions  $\text{inv}(AP')$  and  $\text{inv}(AQ')$  respectively. Further let  $\square$  be the set of all environments and  $\delta$  be the auxiliary store. If  $(IP \mid \square, \delta) \overset{\circ}{\rightsquigarrow}^* (IP' \mid \square', \delta') \overset{n}{\rightsquigarrow} (IQ \mid \square'', \delta'') \rightsquigarrow^* (IQ' \mid \square''', \delta''') \overset{m}{\rightsquigarrow} (IQ'' \mid \square''', \delta''')$ , and provided that the computation  $(IQ \mid \square'', \delta'') \rightsquigarrow^* (IQ' \mid \square''', \delta''')$  does not include any identifier transitions, then  $m = n - 1$ .*

*Proof.* The order of identifiers is maintained using the function  $\text{previous}()$ . This proof follows closely to the argument within the proof of Theorem 1, but uses  $\overset{\circ}{\rightsquigarrow}$  in place of  $\overset{\circ}{\rightarrow}$ .  $\square$

**Theorem 4.** *Let  $P$  be a sequential program (does not contain  $\text{par}$ ) and  $AP$  be  $\text{ann}(P)$ . Further let  $\square$  be the set  $\{\sigma, \gamma, \mu, \beta\}$  of all environments,  $\square_1$  be the set  $\{\sigma_1, \gamma_1, \mu_1, \beta_1\}$  of annotated environments such that  $\square \approx \square_1$ ,  $\square'_1$  be the set  $\{\sigma'_1, \gamma'_1, \mu'_1, \beta'_1\}$  of final annotated environments,  $\square_2$  be the set  $\{\sigma_2, \gamma_2, \mu_2, \beta_2\}$  of inverted environments such that  $\square_2 \approx \square'_1$ ,  $\delta$  be the auxiliary store,  $\delta'$  be the final auxiliary store and  $\delta_2$  be the inverted auxiliary store such that  $\delta_2 \approx_A \delta'$ .*



- 1) If  $(P \mid \square, \delta) \hookrightarrow^* (\text{skip} \mid \square', \delta)$ , for some  $\square'$ , and there exists an annotated execution  $(AP \mid \square_1, \delta) \overset{\circ}{\rightarrow}^* (\text{skip } I \mid \square'_1, \delta')$ , for some  $I$ ,  $\square'_1$  and  $\delta'$  such that the executed annotated version of  $AP$  produced by its execution is  $AP'$ , then there also exists  $(IP \mid \square_2, \delta_2) \overset{\circ}{\rightsquigarrow}^* (\text{skip } I' \mid \square'_2, \delta'_2)$ , for  $IP = \text{inv}(AP')$  and some  $I'$ ,  $\square'_2$  and  $\delta'_2$ . (Termination)
- 2) If  $(P \mid \square, \delta) \hookrightarrow^* (\text{skip} \mid \square', \delta)$ , for some  $\square'$ , and there exists an annotated execution  $(AP \mid \square_1, \delta) \overset{\circ}{\rightarrow}^* (\text{skip } I \mid \square'_1, \delta')$ , for some  $I$ ,  $\square'_1$  and  $\delta'$ , such that  $\square' \approx \square'_1$  and that the executed annotated version of  $AP$  produced by its execution is  $AP'$ , then there also exists  $(IP \mid \square_2, \delta_2) \overset{\circ}{\rightsquigarrow}^* (\text{skip } I' \mid \square'_2, \delta'_2)$ , for  $IP = \text{inv}(AP')$  and some  $I'$ ,  $\square'_2$  and  $\delta'_2$ , such that  $\square'_2 \approx \square$  and  $\delta'_2 \approx_A \delta$ .

*Proof.* This proof is by induction on the length of the sequence  $(P \mid \square, \delta) \hookrightarrow^* (\text{skip} \mid \square', \delta)$ .  $\square$

The full version of Theorem 4, where programs can contain parallel composition, is currently being considered. We note the implication in the other direction would be valid, but defer proof to future work.

## 6 Conclusion

We have presented an approach to reversing a language containing blocks, local variables, procedures and the interleaving parallel composition. We defined annotation, the process of creating a state-saving annotated version capable of assigning identifiers to capture the interleaving order. This was proved to not alter the behaviour of the original program and to populate the auxiliary store. Inversion creates an inverted version that uses this saved information to restore the initial program state, with this being proved to hold. The auxiliary store is also restored, meaning it is garbage free.

We are also currently developing a simulator capable of implementing this approach, with one application being to aid the proof of our results. The current prototype is capable of simulating the annotated forwards execution, with the simulation of the inverted execution currently being worked on. This will be used to evaluate the performance overhead and costs associated with our approach to reversibility.

Our future work will continue the development of this simulator, as well as modify the approach allowing for causal-consistent reversibility.

## Acknowledgements

We are grateful to the referees for their detailed and helpful comments and suggestions. The authors acknowledge partial support of COST Action IC1405 on Reversible Computation - extending horizons of computing. The third author is supported by JSPS KAKENHI grant numbers 17H01722 and 17K19969.

## References

- [1] V. Danos & J. Krivine (2004): *Reversible Communicating Systems*. In: *CONCUR 2004, Proceedings*, pp. 292–307, doi:10.1007/978-3-540-28644-8\_19.
- [2] R. Fujimoto (1990): *Parallel Discrete Event Simulation*. *Communications of the ACM* 33(10), pp. 30–53, doi:10.1145/84537.84545.
- [3] E. Giachino, I. Lanese & C.A. Mezzina (2014): *Causal-Consistent Reversible Debugging*. In: *Proceedings of FASE 2014*, pp. 370–384, doi:10.1007/978-3-642-54804-8\_26.
- [4] R. Glück & M. Kawabe (2004): *Derivation of Deterministic Inverse Programs Based on LR Parsing*. In: *FLOPS 2004, LNCS 2998*, Springer, pp. 291–306, doi:10.1007/978-3-540-24754-8\_21.

- [5] R. Glück & M. Kawabe (2005): *Revisiting an Automatic Program Inverter for LISP*. *SIGPLAN Notices* 40(5), pp. 8–17, doi:10.1145/1071221.1071222.
- [6] R. Glück & T. Yokoyama (2016): *A Linear-Time Self-Interpreter of a Reversible Imperative Language*. *Computer Software* 33(3), pp. 3\_108–3\_128, doi:10.11309/jssst.33.3\_108.
- [7] R. Glück & T. Yokoyama (2017): *A Minimalist's Reversible While Language* E100.D, pp. 1026–1034. doi:10.1587/transinf.2016EDP7274.
- [8] D. Gries (1981): *The Science of Programming*. Springer, doi:10.1007/978-1-4612-5983-1.
- [9] T. Haulund (2017): *Design and Implementation of a Reversible Object-Oriented Programming Language*. CoRR abs/1707.07845. Available at <http://arxiv.org/abs/1707.07845>.
- [10] J. Hoey, I. Ulidowski & S. Yuen (2017): *Reversing Imperative Parallel Programs*. In: *Proceedings of Express/SOS, 2017, EPTCS 255*, pp. 51–66, doi:10.4204/EPTCS.255.4.
- [11] H. Hüttel (2010): *Transitions and Trees - An Introduction to Structural Operational Semantics*. Cambridge University Press, doi:10.1017/CBO9780511840449.
- [12] D.R. Jefferson (1985): *Virtual Time*. *ACM Transactions on Programming Languages and Systems* 7(3), pp. 404–425, doi:10.1145/3916.3988.
- [13] I. Lanese, C.A. Mezzina & F. Tiezzi (2014): *Causal-Consistent Reversibility*. *Bulletin of the EATCS* 114.
- [14] K. Perumalla (2014): *Introduction to Reversible Computing*. CRC Press, doi:10.1201/b15719.
- [15] I.C.C. Phillips & I. Ulidowski (2007): *Reversing Algebraic Process Calculi*. *J. Log. Algebr. Program.* 73(1-2), pp. 70–96, doi:10.1016/j.jlap.2006.11.002.
- [16] I.C.C. Phillips, I. Ulidowski & S. Yuen (2012): *A Reversible Process Calculus and the Modelling of the ERK Signalling Pathway*. In: *RC2012, LNCS 7581*, Springer, pp. 218–232, doi:10.1007/978-3-642-36315-3\_18.
- [17] M. Schordan, D.R. Jefferson, P.D. Barnes Jr., T. Oppelstrup & D.J. Quinlan (2015): *Reverse Code Generation for Parallel Discrete Event Simulation*. In: *RC 2015, LNCS 9138*, Springer, pp. 95–110, doi:10.1007/978-3-319-20860-2\_6.
- [18] U.P. Schultz & H.B. Axelsen (2016): *Elements of a Reversible Object-Oriented Language - Work-in-Progress Report*. In: *Reversible Computation - 8th International Conference, RC 2016, Proceedings*, pp. 153–159, doi:10.1007/978-3-319-40578-0\_10.
- [19] G. Vulov, C. Hou, R.W. Vuduc, R. Fujimoto, D.J. Quinlan & D.R. Jefferson (2011): *The Backstroke Framework for Source Level Reverse Computation Applied to Parallel Discrete Event Simulation*. In: *WSC 2011, WSC*, doi:10.1109/WSC.2011.6147998.
- [20] T. Yokoyama, H.B. Axelsen & R. Glück (2008): *Principles of a Reversible Programming Language*. In: *Proceedings of the 5th Conference on Computing Frontiers*, ACM, pp. 43–54, doi:10.1145/1366230.1366239.
- [21] T. Yokoyama & R. Glück (2007): *A Reversible Programming Language and its Invertible Self-Interpreter*. In: *Proceedings of the 2007 ACM SIGPLAN Workshop on Partial Evaluation and Semantics-based Program Manipulation*, ACM, pp. 144–153, doi:10.1145/1244381.1244404.