Modular Verification of SPARCv8 Code

Abstract   Inline assembly code is common in system software to interact with the underlying hardware platforms. Safety and correctness of the assembly code is crucial to guarantee the safety of the whole system. In this paper we propose a practical Hoare-style program logic for verifying SPARC assembly code. The logic supports modular reasoning about the main features of SPARCv8 ISA, including delayed control transfers, delayed writes to special registers, and register windows. We extend to it support contextual refinement verification, and apply it to verify that there is a contextual refinement between a context switch routine in SPARCv8 and `switch` primitive. The program logic and its soundness proof have been mechanized in Coq.

Keywords   SPARCv8, assembly code verification, context switch, Coq, refinement verification

1   Introduction

Operating system kernels are at the most foundational layer of computer software systems. To interact directly with hardware, many important components in OS kernels are implemented in assembly, such as the context switch code or the code that manages interrupts. And some other codes that are not required to be written in assembly are also implemented in assembly (*e.g.* memcpy in linux v2.6.17.10 [1]), in order to achieve high performance. Their correctness is crucial to ensure the safety and security of the whole system. However, assembly code verification remains a challenging task in existing work on OS kernel verification (*e.g.* [2, 3, 4]), where the assembly code is either unverified or verified based on operational semantics without a general program logic. SPARC (Scalable Processor ARChitecture) is a CPU instruction set architecture (ISA) with high-performance and great flexibility [5]. It has been widely used in various processors for workstations and embedded systems. The SPARCv8 ISA has some interesting features, which make it a non-trivial task to design a Hoare-style program logic for assembly code.

- *Delayed control transfers*. SPARCv8 has two program counters `pc` and `npc`. The `npc` register points to the next instruction to run. Control-transfer instructions in SPARCv8 change `npc` instead of `pc` to the target program point, while `pc` takes the original value of `npc`. This makes the control transfer to happen one cycle later than the execution of the control transfer instructions.

- *Delayed writes*. The `wr` instruction that writes a special class of registers does not take effect immediately. Instead the write operation is buffered and then executed X cycles later, where X is a predefined system parameter which usually ranges from 0 to 3.

- *Register windows*. SPARCv8 uses register windows and the window rotation mechanism to avoid saving contexts in the stack directly and achieves high performance in context management.

We use a simple example in Fig. 1 to show these three features. The function CALLER calls ChangeY,

which updates the special register Y and returns its original value.

```
CALLER :                        ChangeY :
    ...                       5  rd  Y, %l_0
 1  mov  1, %o_0              6  wr  %i_0, 0, Y
 2  call  ChangeY            7  nop
 3  save  %sp, −64, %sp      8  nop
 4  mov  %o_0, %l_0          9  nop
    ...                      10 ret
                             11 restore  %l_0, 0, %o_0
```

Fig.1. An Example for SPARC Code

ChangeY requires an input parameter as the new value for the special register Y. CALLER calls ChangeY at line 2, and pc and npc point to line 2 and 3 respectively at this moment. The call instruction changes the value of pc to npc and let npc points to the entry of ChangeY at line 5, which means the control-flow will not transfer to ChangeY in the next cycle, but in the cycle after the execution of the save instruction following the call. Similarly, when ChangeY returns (at line 10), the control is transferred back to the caller after executing the restore instruction at line 11. We call this feature "delayed control transfers".

SPARCv8 uses the save instruction (at line 3 in the example) to save the current context and restore (at line 10) to restore it. Its 32 general registers are split into four logic groups as global ($r_0 \sim r_7$), out ($r_8 \sim r_{15}$), local ($r_{16} \sim r_{23}$) and in ($r_{24} \sim r_{31}$) registers. Correspondingly, we give aliases "$\%g_0 \sim \%g_7$", "$\%o_0 \sim \%o_7$", "$\%l_0 \sim \%l_7$" and "$\%i_0 \sim \%i_7$" for these groups respectively. The out, local and in registers form the *current register window*. The local registers are for private use in the current context. The in and out registers are shared with adjacent register windows for parameters passing. The save instruction rotates the register window from the current one to the next. Then the local and in registers in the original window are no longer accessible, and the original out

registers becomes the in registers in the current window. The restore instruction does the inverse. The arguments taken by the save and restore instructions are irrelevant here and can be ignored.

At line 6, the wr instruction tries to update the special register Y with the value of $\%i_0 \oplus 0$ (bitwise exclusive OR). However, the write is delayed for X cycles, where X is some predefined system parameter that ranges from 0 to 3. For portability, programmers usually do not rely on the exact value of X and assume it takes the maximum value 3. Therefore three nop instructions are inserted. Reading of Y earlier than line 9 may give us the old value. This feature is called "delayed writes".

These features make the semantics of the SPARCv8 code context-dependent. For instance, a read of a special register (*e.g.* the register Y in the above example) needs to make sure there are enough instructions executed since the most recent *delayed* write. As another example, the instruction following the call can be any instruction in general, but it is not supposed to update the register r_{15}, which contains the return address saved by the call instruction. In addition, the delayed control transfer and the register windows also allow highly flexible calling conventions. Together, they make it a challenging task to have a Hoare-style program logic for local and modular reasoning of SPARCv8 assembly code.

Working towards a fully certified OS kernel for aerospace crafts whose inline assembly is written in SPARCv8, we try to address these challenges and propose a practical program logic for realistically modelled SPARCv8 code. However, the OS kernel is implemented as C language mainly and SPARCv8 as inline assembly. Just having a traditional Hoare-style program logic for SPARCv8, which can only make sure the partial correctness of SPARCv8 program, is insufficient.

Xu *et al.* [2] propose a program logic for verifying the correctness of OS kernel implemented in C language with inline assembly, but they use abstract assembly primitives, like switch primitive for task switching, to substitute the inline assembly in their verification work. As a supplement to their work, we need to extend our program logic so that it can make sure the contextual refinement relation between the implementations and their corresponding abstract assembly primitives, which can be presented as the following form. Here, we use $\mathbb{C}$, A, and C_{as} to represent the C language program, the set of abstract assembly primitive and the implementations of abstract assembly primtives respectively. It means C_{as} refines A under *any context* $\mathbb{C}$.

$$\forall \mathbb{C}. \mathbb{C}[C_{as}] \sqsubseteq \mathbb{C}[A]$$

However, if we use C program as a client code to call inline assembly code, we need to define the semantics of C-assembly linking. It may be possible to use *interaction semantics* [6] to implement multi-language linking. However, there is a contraint in interaction semantics that it requires that callee must return to caller when finished. So, it can't handle calling context switch routine, whose correctness we hope to prove in our work.

$$[\![\mathbb{C}[A]]\!]^{\mathsf{C}}$$

$$① \sqcup | \impliedby \boxed{C = Comp(\mathbb{C})}$$

$$[\![C[\Omega]]\!]^{\mathsf{P\text{-}SPARCv8}}$$

$$② \sqcup | \impliedby \boxed{\vdash C_{as} : \Omega}$$

$$[\![C[C_{as}]]\!]^{\mathsf{SPARCv8}}$$

Fig.2. Idea to establish contextual refinement

We consider a method to establish the contextual refinement between implementations and their corresponding abstract assembly primitives and avoid multi-language linking. We use Fig. 2 to illustrate our idea.

The source program of OS kernel shown as $\mathbb{C}[A]$, which executes under C language semantics (shown as $[\![\_]\!]^{\mathsf{C}}$), is implementated as C language with a set of abstract assembly primtives A. The compiler (*Comp*) translates the C program $\mathbb{C}$ to SPARCv8 code C. As ① shown, we hope the compilation makes sure the refinement relation between $\mathbb{C}[A]$ and $C[\Omega]$ that executes under Pseudo-SPARCv8 semantics shown as $[\![\_]\!]^{\mathsf{P\text{-}SPARCv8}}$. Here, the Ω represents the set of abstract assembly primitives in the middle layer. We use distinguished notations to represent the set of abstract assembly primitives in source and intermediate level, because they execute on different program states and have different semantics. The Pseudo-SPARCv8 language $C[\Omega]$, which uses SPARCv8 as client code and is able to call abstract assembly primtive in Ω, will be defined in the following. In step ②, we extend our program logic to support refinement verification and get the totally SPARCv8 program $C[C_{as}]$ executing under the realistically modelled SPARCv8 semantics, represented as $[\![\_]\!]^{\mathsf{SPARCv8}}$, refines the program $C[\Omega]$ executing under Pseudo-SPARCv8 semantics. Finally, we can get $[\![\mathbb{C}[\Omega]]\!]^{\mathsf{C}} \sqsubseteq [\![C[C_{as}]]\!]^{\mathsf{SPARCv8}}$. In this work, we focus on step ②, and remain step ① as a future work.

Our work is based on earlier work on assembly code verification but makes the following contributions [1]:

- We propose a traditional Hoare-style program logic for SPARCv8 and support all the above features of SPARCv8We redefine basic blocks to include the instruction following the jump or return as the tail of a block, which models the delayed control transfer. To reason about delayed writes, we introduce a modal assertion $\rhd_t \mathtt{sr} \mapsto w$, saying that the special register $\mathtt{sr}$ will hold the value w in up to t cycles. We also give logic rules for save

---

[1] We list six contributions here. The first four contributions have already presented in the preliminary version [7] of this paper. And the additional two contributions about refinement verification are our new contributions.

and `restore` instructions that do register window rotation.

- Following SCAP [8], our logic supports modular reasoning of function calls in a direct-style. We use the standard pre- and post-conditions as function specifications, instead of the binary assertion g used in SCAP. This allows us to reuse existing techniques (*e.g.* Coq tactics) to simplify the program verification process. The logic rules for function call and return is general and independent of any specific calling convention.

- In order to support refinement verification, we define a Pseudo-SPARCv8 program as our high-level program, which is multi-threaded and can call abstract assembly primitives. It also simplifies the sophisticated register window mechanism in SPARCv8, and makes the program state of Pseudo-SPARCv8 program simpler than the physical SPARCv8 program (defined in Sec. 2). So, it can provide some convenience to write the abstract assembly primitive and reason in high-level program.

- We extend our program logic defined previously to support refinement verification. The extended logic can ensure the *contextual refinement* between SPARCv8 functions and their corresponding abstract assembly primitives.

- We give direct-style semantic interpretation for the logic judgments, based on which we establish the soundness. This is different from previous work, which either does syntactic-based soundness proof (*e.g.* SCAP [8]) or treats return code pointers as first-class code pointers and gives

CPS-style (continuation-passing style) semantics. Those approaches for soundness make it difficult to verify the interaction between the inline assembly and the C code in the kernel, the latter being verified following a direct-style program logic.

- Context switch of concurrent tasks is an important component in OS kernels. It is usually implemented as inline assembly because of the need to access registers and the stack. We verify that there is a contextual refinement between a context switch routine implementated in SPARCv8 and `switch` primitive for task switching[2].

The program logic, including the extended program logic for refinement verification, and its soundness proof have been mechanized in Coq [9].

This article extends the conference paper in APLAS 2018 [7]. We extend the program logic, which can only prove the partial correctness of SPARCv8 code in conference paper, to support refinement verification and use it to verify that there is a contextual refinement between a context switch routine implementated in SPARCv8 and `switch` primitive for task switching. And in order to support refinement verification, we define Pseudo-SPARCv8 language as the high-level.

In the rest of paper, we present the program model and operational semantics of SPARCv8 in Sec. 2. Then we propose our traditional Hoare-style program logic for SPARCv8 in Sec. 3 and show how our logic supprts the three main features of SPARCv8. The Pseudo-SPARCv8 program and extended program logic for refinement verification is presented in Sec. 4. We show the verification of a context switch routine in SPARCv8 in Sec. 5. Finally we discuss more on related work and conclude in Sec. 6.

---

[2]We also apply our program logic to verify the main body of the context switch routine in a realistic embedded OS kernel for aerospace crafts, which consists of around 250 lines of SPARCv8 code, by 6690 lines of Coq proof scripts. But this part of code cannot be published due to copyright issues.

$$
\begin{array}{llll}
\text{(Word)} \ w, \mathtt{f} \ \in \ \text{Int32} & \text{(Block)} \ b \in \mathbb{Z} & \text{(Addr)} \ l \ \in \ \text{Block} \times \text{Word} & \text{(Val)} \ v \ ::= w \mid l
\end{array}
$$

$$
\begin{array}{llll}
\text{(Prog)} & \text{P} & ::= & (C, S, \mathtt{pc}, \mathtt{npc}) & \text{(CodeHeap)} & C & \in \ \text{Word} \rightharpoonup \text{Comm} \\
\text{(State)} & S & ::= & (M, Q, D) & \text{(RState)} & Q & ::= (R, F) \\
\text{(Memory)} & M & \in & \text{Addr} \rightharpoonup \text{Val} & \text{(ProgCount)} \ \mathtt{pc}, \mathtt{npc} & \in \ \text{Word} \\
\text{(OpExp)} & \mathtt{o} & ::= & \mathtt{r} \mid w & \text{(AddrExp)} & \mathtt{a} & ::= \mathtt{o} \mid \mathtt{r} + \mathtt{o}
\end{array}
$$

$$
\begin{array}{llll}
\text{(Comm)} & c & ::= & \mathtt{i} \mid \mathtt{call\ f} \mid \mathtt{jmp\ a} \mid \mathtt{retl} \mid \mathtt{be\ f} \\
\text{(SimpIns)} & \mathtt{i} & ::= & \mathtt{ld\ a\ r}_d \mid \mathtt{st\ r}_s\ \mathtt{a} \mid \mathtt{nop} \mid \mathtt{add\ r}_s\ \mathtt{o\ r}_d \mid \mathtt{save\ r}_s\ \mathtt{o\ r}_d \mid \mathtt{restore\ r}_s\ \mathtt{o\ r}_d \\
& & & \mid \mathtt{rd\ sr\ r}_d \mid \mathtt{wr\ r}_s\ \mathtt{o\ sr} \mid \ldots \\
\text{(InstrSeq)} & \mathbb{I} & ::= & \mathtt{i}; \mathbb{I} \mid \mathtt{jmp\ a}; \mathtt{i} \mid \mathtt{call\ f}; \mathtt{i}; \mathbb{I} \mid \mathtt{retl\ i} \mid \mathtt{be\ f}; \mathtt{i}; \mathbb{I}
\end{array}
$$

Fig.3. Machine States and Language for SPARCv8 Code

2   The SPARCv8 Assembly Language

We introduce the key SPARCv8 instructions, the model of machine states, and the operational semantics in this section.

2.1   Language syntax and states

The machine model and syntax of SPARCv8 assembly language are defined in Fig. 3. Here, we follow the block-based memory [10] introduced in CompCert to define the memory in our work. The memory address l is defined as a pair of its block id and the offset. The type of block is the integer in mathemantics represent as $\mathbb{Z}$, and the type of offset is a 32-bit integer, which we called *words* in our work. So, the value here is either a word w or address l. The whole program configuration P consists of the code heap C, the machine state S, and the program counters $\mathtt{pc}$ and $\mathtt{npc}$. The code heap C is a partial function from labels $\mathtt{f}$ to commands c. Labels are also 32-bit integers (called *words*), which can be viewed as addresses or locations where the commands are saved in code heap. The operand expression $\mathtt{o}$, which is either a general register $\mathtt{r}$ or a word w, and address expression $\mathtt{a}$, which is either a operand expression or a sum of the value of register $\mathtt{r}$ and an operation expression, are auxiliary definitions used as paramters of commands. Commands in SPARCv8 can be classi-fied into two categories, the simple instructions $\mathtt{i}$ and the control-transfer instructions like $\mathtt{call}$ and $\mathtt{jmp}$.

The machine state S consists of three parts: the memory M, the register state Q which is a pair of register file R and frame list F, and the delay buffer D. As defined in Fig. 4, R is a partial mapping from register names to values. Registers include the general registers $\mathtt{r}$, the processor state register $\mathtt{psr}$ and the special registers $\mathtt{sr}$. The processor state register $\mathtt{psr}$ contains the integer condition code fields $\mathtt{n}$, $\mathtt{z}$, $\mathtt{v}$ and $\mathtt{c}$, which can be modified by the arithmetic and logical instructions and used for conditional control-transfer, and $\mathtt{cwp}$ recording the id of the current register window. We explain the frame list F and the delay buffer D below.

Register windows and frame List.   SPARCv8 provides 32 general registers, which are split into four groups as global ($\mathtt{r}_0 \sim \mathtt{r}_7$), out ($\mathtt{r}_8 \sim \mathtt{r}_{15}$), local($\mathtt{r}_{16} \sim \mathtt{r}_{23}$) and in ($\mathtt{r}_{24} \sim \mathtt{r}_{31}$) registers. The latter three groups (out, local and in) form the current *register window*.

At the entry and exit of functions and traps, one may need to save and restore some of the general registers as execution contexts. Instead of saving them into stacks in memory, SPARCv8 uses multiple register windows to form a circular stack, and does window rotation for efficient context save and restore. As

(RegFile)	R	$\in$	$RegName \rightharpoonup Val$	(RegName)	$\mathtt{rn}$	$::=$	$\mathtt{r_0} \mid \ldots \mid \mathtt{r_{31}} \mid \mathtt{psr} \mid \mathtt{sr}$
(PsrReg)	$\mathtt{psr}$	$::=$	$\mathtt{n} \mid \mathtt{z} \mid \mathtt{v} \mid \mathtt{c} \mid \mathtt{cwp}$	(SpeReg)	$\mathtt{sr}$	$::=$	$\mathtt{wim} \mid \mathtt{Y} \mid \mathtt{asr_0} \mid \ldots \mid \mathtt{asr_{31}}$
(FrameList)	F	$::=$	$\mathrm{nil} \mid \mathrm{fm} :: F$	(Frame)	fm	$:=$	$[v_0, \ldots, v_7]$
(DelayBuff)	D	$::=$	$\mathrm{nil} \mid (t, \mathtt{sr}, w) :: D$	(DelayCycle)	t	$\in$	$\{0, 1, \ldots, X\}$

Fig.4. Register File, Frame List and DelayBuffer

Fig.5. Register Windows (figure taken from [5])

shown in Fig. 5, there are N register windows ($N = 8$ here) consisting of $2 \times N$ groups of registers (each group containing 8 registers). The $\mathtt{cwp}$ register (part of $\mathtt{psr}$) records the id number of the current window ($\mathtt{cwp} = 0$ in this example).

The in and out registers of each window are shared with its adjacent windows for parameter passing. For example, the in registers of the w_0 is the out registers of the w_1, and the out registers of the w_0 is the in registers of the w_7. This explains why we need only $2 \times N$ groups of registers for N windows, while each window consisting of three groups (out, local and in).

To save the context, the $\mathtt{save}$ instruction rotates the window by decrements the $\mathtt{cwp}$ pointer (modulo N). So w_7 becomes the current window. The out registers of w_0 becomes the in registers of w_7. The in and local reg-

isters of w_0 become inaccessible. This is like pushing them onto the circular stack. The $\mathtt{restore}$ instruction does the inverse, which is like a stack pop.

The $\mathtt{wim}$ register is used as a bit vector to record the end of the stack. Each bit in $\mathtt{wim}$ corresponds to a register window. The bit corresponding to the last available window is set to 1, which means *invalid*. All other bits are 0 (*i.e. valid*). When executing $\mathtt{save}$ (and $\mathtt{restore}$), we need to ensure the next window is valid, in order to avoid the overflow of register window because of the limitation of the number of windows. We use the assertion **win_valid**(w_{id}, R) defined in Fig. 6 to say the window pointed to by w_{id} is valid, given the value of $\mathtt{wim}$ in R.

We use the frame list F to model the circular stack consisting of register windows. As defined in Fig. 4, a

$$\text{out} ::= [\mathbf{r}_8, \ldots, \mathbf{r}_{15}] \qquad \text{local} ::= [\mathbf{r}_{16}, \ldots, \mathbf{r}_{23}] \qquad \text{in} ::= [\mathbf{r}_{24}, \ldots, \mathbf{r}_{31}]$$

$$R([\mathbf{r}_i, \ldots, \mathbf{r}_{i+k}]) ::= [R(\mathbf{r}_i), \ldots, R(\mathbf{r}_{i+k})]$$

$$R\{[\mathbf{r}_i, \ldots, \mathbf{r}_{i+7}] \rightsquigarrow \text{fm}\} ::= R\{\mathbf{r}_i \rightsquigarrow v_0\} \ldots \{\mathbf{r}_{i+7} \rightsquigarrow v_7\}$$
$$\text{where} \quad \text{fm} = [v_0, \ldots, v_7]$$

$$\mathbf{win\_valid}(w_{id}, R) ::= 2^{w_{id}} \,\&\, R(\texttt{wim}) = 0$$
$$\text{where } \& \text{ is the bitwise AND operation.}$$

$$\mathbf{next\_cwp}(w_{id}) ::= (w_{id} + N - 1)\%N \qquad \mathbf{prev\_cwp}(w_{id}) ::= (w_{id} + 1)\%N$$

$$\mathbf{save}(R, F) ::= \begin{cases} (R', F') & \text{if } w'_{id} = \mathbf{next\_cwp}(R(\texttt{cwp})), \mathbf{win\_valid}(w'_{id}, R), \\ & \quad F = F'' \cdot \text{fm}_1 \cdot \text{fm}_2, \ F' = R(\text{local}) :: R(\text{in}) :: F'', \\ & \quad R'' = R\{\text{in} \rightsquigarrow R(\text{out}), \text{local} \rightsquigarrow \text{fm}_2, \text{out} \rightsquigarrow \text{fm}_1\}, \\ & \quad R' = R''\{\texttt{cwp} \rightsquigarrow w'_{id}\}, \\ \bot & \text{if } \neg\mathbf{win\_valid}(\mathbf{next\_cwp}(R(\texttt{cwp})), R) \end{cases}$$

$$\mathbf{restore}(R, F) ::= \begin{cases} (R', F') & \text{if } w'_{id} = \mathbf{prev\_cwp}(R(\texttt{cwp})), \mathbf{win\_valid}(w'_{id}, R), \\ & \quad F = \text{fm}_1 :: \text{fm}_2 :: F'', \ F' = F'' \cdot R(\text{out}) \cdot R(\text{local}), \\ & \quad R'' = R\{\text{in} \rightsquigarrow \text{fm}_2, \text{local} \rightsquigarrow \text{fm}_1, \text{out} \rightsquigarrow R(\text{in})\}, \\ & \quad R' = R''\{\texttt{cwp} \rightsquigarrow w'_{id}\}, \\ \bot & \text{if } \neg\mathbf{win\_valid}(\mathbf{prev\_cwp}(R(\texttt{cwp})), R) \end{cases}$$

Fig.6. Auxiliary Definitions for Instruction `save` and `restore`

frame is an array of 8 words, modeling a group of 8 registers. F consists of a sequence of frames corresponding to all the register windows except the out, local and in registers in the current window. Then `save` saves the local and in registers onto the head of F and loads the two groups of register at the *tail* of F to the local and out registers (and the original out registers becomes the in group). The `restore` instruction does the inverse. The operations are defined formally in Fig. 6.

The delay buffer. The delay buffer D is a sequence of delayed writes. Because the `wr` instruction does not update the target register immediately, we put the write operation onto the delay buffer. A delayed write is recorded as a triple consisting of the remaining cycles t to be delayed, the target special register `sr` and the value w to be written. Note that we restrict that the value of a special register can only be a word, because the special registers are used to record the state of processor, and there is impossible to store memory

addresses in them.

Instruction sequences. We use an instruction sequence $\mathbb{I}$ to model a basic block, *i.e.* a sequence of commands ending with a control transfer. As defined in Fig. 3, we require that a delayed control-transfer instruction must be followed by a simple instruction `i`, because the actual control-transfer occurs after the execution of `i`. The end of each instruction sequence can only be `jmp` or `retl` followed by a simple instruction `i`. Note that we do not view the `call` instruction as the end of a basic block, since the callee is expected to return, following our direct-style semantics for function calls. We define $C[\mathbf{f}]$ to extract an instruction sequence starting from `f` in C below.

$$C[\mathbf{f}] = \begin{cases} \mathbf{i}; \mathbb{I} & C(\mathbf{f}) = \mathbf{i} \text{ and } C[\mathbf{f} + 4] = \mathbb{I} \\ c; \mathbf{i} & c = C(\mathbf{f}) \text{ and } c = \texttt{jmp a} \text{ or } \texttt{retl} \\ & \text{and } C(\mathbf{f} + 4) = \mathbf{i} \\ c; \mathbf{i}; \mathbb{I} & c = C(\mathbf{f}) \text{ and } c = \texttt{call f} \text{ or } \texttt{be f} \\ & \text{and } C(\mathbf{f} + 4) = \mathbf{i} \text{ and } C[\mathbf{f} + 8] = \mathbb{I} \\ \mathbf{undefined} & \text{otherwise} \end{cases}$$

$$(R, D) \rightrightarrows (R', D')$$
$$C \vdash ((M, (R', F), D'), \mathtt{pc}, \mathtt{npc}) \circ\!\!\longrightarrow ((M', (R'', F'), D''), \mathtt{pc}', \mathtt{npc}')$$
$$\overline{C \vdash ((M, (R, F), D), \mathtt{pc}, \mathtt{npc}) \longmapsto ((M', (R'', F'), D''), \mathtt{pc}', \mathtt{npc}')}$$

(a) Program Transistion

$$C(\mathtt{pc}) = \mathtt{i} \qquad (M, (R, F), D) \bullet\!\!\xrightarrow{\mathtt{i}} (M', (R', F'), D')$$
$$\overline{C \vdash ((M, (R, F), D), \mathtt{pc}, \mathtt{npc}) \circ\!\!\longrightarrow ((M', (R', F'), D'), \mathtt{npc}, \mathtt{npc} + 4)}$$

$$C(\mathtt{pc}) = \mathtt{jmp\ a} \qquad [\![\mathtt{a}]\!]_R = \mathtt{f}$$
$$\overline{C \vdash ((M, (R, F), D), \mathtt{pc}, \mathtt{npc}) \circ\!\!\longrightarrow ((M, (R, F), D), \mathtt{npc}, \mathtt{f})}$$

$$C(\mathtt{pc}) = \mathtt{call\ f} \qquad \mathtt{r}_{15} \in \mathrm{dom}(R)$$
$$\overline{C \vdash ((M, (R, F), D), \mathtt{pc}, \mathtt{npc}) \circ\!\!\longrightarrow ((M, (R\{\mathtt{r}_{15} \rightsquigarrow \mathtt{pc}\}, F), D), \mathtt{npc}, \mathtt{f})}$$

$$C(\mathtt{pc}) = \mathtt{retl} \qquad R(\mathtt{r}_{15}) = \mathtt{f}$$
$$\overline{C \vdash ((M, (R, F), D), \mathtt{pc}, \mathtt{npc}) \circ\!\!\longrightarrow ((M, (R, F), D), \mathrm{npc}, \mathtt{f} + 8)}$$

(b) Control Transfer Instruction Transition

$$\frac{(M, R) \xrightarrow{\mathtt{i}} (M', R')}{(M, (R, F), D) \bullet\!\!\xrightarrow{\mathtt{i}} (M', (R', F), D)}$$

$$R(\mathtt{r}_s) = w_1 \qquad [\![\mathtt{o}]\!]_R = w_2 \qquad w = w_1 \oplus w_2$$
$$\mathtt{sr} \in \mathrm{dom}(R) \qquad D' = \mathbf{set\_delay}(\mathtt{sr}, w, D)$$
$$\overline{(M, (R, F), D) \bullet\!\!\xrightarrow{\mathtt{wr\ r_s\ o\ sr}} (M, (R, F), D')}$$

$$\mathbf{save}(R, F) = (R', F') \qquad [\![\mathtt{o}]\!]_R = v \qquad v' = R(\mathtt{r}_s) + v$$
$$\overline{(M, (R, F), D) \bullet\!\!\xrightarrow{\mathtt{save\ r_s\ o\ r_d}} (M, (R\{\mathtt{r}_d \rightsquigarrow v'\}, F'), D)}$$

$$\mathbf{restore}(R, F) = (R', F') \qquad [\![\mathtt{o}]\!]_R = v \qquad v' = R(\mathtt{r}_s) + v$$
$$\overline{(M, (R, F), D) \bullet\!\!\xrightarrow{\mathtt{restore\ r_s\ o\ r_d}} (M, (R\{\mathtt{r}_s \rightsquigarrow v'\}, F'), D)}$$

(c) Save, Restore and Wr instruction Transition

$$\frac{R(\mathtt{sr}) = w \qquad \mathtt{r}_d \in \mathrm{dom}(R)}{(M, R) \xrightarrow{\mathtt{rd\ sr\ r_d}} (M, R\{\mathtt{r}_d \rightsquigarrow w\})}$$

$$\frac{R(\mathtt{r}_s) = v_1 \quad [\![\mathtt{o}]\!]_R = v_2 \quad v = v_1 + v_2 \quad \mathtt{r}_d \in \mathrm{dom}(R)}{(M, R) \xrightarrow{\mathtt{add\ r_s\ o\ r_d}} (M, R\{\mathtt{r}_d \rightsquigarrow v\})}$$

$$\frac{[\![\mathtt{a}]\!]_R = l \quad M(l) = v' \quad \mathtt{r}_d \in \mathrm{dom}(R)}{(M, R) \xrightarrow{\mathtt{ld\ a\ r_d}} (M, R\{\mathtt{r}_d \rightsquigarrow v'\})}$$

(d) Simple Instruction Transition

$$[\![\mathtt{o}]\!]_R ::= \begin{cases} R(r) & \textbf{if } \mathtt{o} = r \\ w & \textbf{if } \mathtt{o} = w, \\ & \quad -4096 \leq w \leq 4095 \\ \bot & \textbf{otherwise} \end{cases} \qquad [\![\mathtt{a}]\!]_R ::= \begin{cases} [\![\mathtt{o}]\!]_R & \textbf{if } \mathtt{a} = \mathtt{o} \\ v_1 + v_2 & \textbf{if } \mathtt{a} = \mathtt{r} + \mathtt{o},\ R(\mathtt{r}) = v_1 \\ & \quad \textbf{and } [\![\mathtt{o}]\!]_R = v_2 \\ \bot & \textbf{otherwise} \end{cases}$$

(e) Expression Semantics

Fig.7. Selected operational semantics rules

2.2 Operational Semantics

The operational semantics is taken from Wang *et al.* [11], but we use block-based memory model and omit features like interrupts and traps. We show the selected rules in Fig. 7. The program transition relation $C \vdash (S, \mathtt{pc}, \mathtt{npc}) \longmapsto (S', \mathtt{pc}', \mathtt{npc}')$ is defined in Fig. 7 (a). Before the execution of the instruction pointed by $\mathtt{pc}$, the delayed writes in D with 0 delay cycles are executed first. The execution of the delayed writes are defined in the form of $(R, D) \rightrightarrows (R', D')$ below:

$$\overline{(R, \mathrm{nil}) \rightrightarrows (R, \mathrm{nil})}$$

$$\frac{(R, D) \rightrightarrows (R', D')}{(R, (t{+}1, \mathtt{sr}, w) {::} D) \rightrightarrows (R', (t, \mathtt{sr}, w) {::} D')}$$

$$\frac{(R, D) \rightrightarrows (R', D') \qquad \mathtt{sr} \in \mathrm{dom}(R)}{(R, (0, \mathtt{sr}, w) {::} D) \rightrightarrows (R'\{\mathtt{sr} \rightsquigarrow w\}, D')}$$

$$\frac{(R, D) \rightrightarrows (R', D') \qquad \mathtt{sr} \notin \mathrm{dom}(R)}{(R, (0, \mathtt{sr}, w) {::} D) \rightrightarrows (R', D')}$$

Note that the write of $\mathtt{sr}$ has no effect if $\mathtt{sr}$ is not in the domain of R. Since R is defined as a partial map, we can prove the following lemma.

Lemma 1. $(R, D) \rightrightarrows (R', D')$ and $R = R_1 \uplus R_2$, if and only if there exists R'_1 and R'_2, such that $(R_1, D) \rightrightarrows (R'_1, D')$, $(R_2, D) \rightrightarrows (R'_2, D')$, and $R' = R'_1 \uplus R'_2$.

Here the disjoint union $R_1 \uplus R_2$ represents the union of R_1 and R_2 if they have disjoint domains, and undefined otherwise. This lemma is important to give sound semantics to delay buffer related assertions, as discussed in Sec. 3.

The transition steps for individual instructions are classified into three categories: the control transfer steps ($\_ \vdash \_ \ \circ\!\longrightarrow \_$), the steps for $\mathtt{save}$, $\mathtt{restore}$ and $\mathtt{wr}$ instructions ($\_ \ \bullet\!\!=\!\!\Rightarrow \_$), and the steps for other simple instructions ($\_ \ =\!\!\Rightarrow \_$). The corresponding step transition relations are defined inductively in Fig. 7 (b), (c) and (d) respectively.

Note that, after the control-transfer instructions, $\mathtt{pc}$ is set to $\mathtt{npc}$ and $\mathtt{npc}$ contains the target code pointer. This explains the one cycle delay for the control transfer. The $\mathtt{call}$ instruction saves $\mathtt{pc}$ into the register $\mathtt{r}_{15}$, while $\mathtt{retl}$ uses $\mathtt{r}_{15} {+} 8$ as the return address (which is the address for the second instruction following the $\mathtt{call}$). Evaluation of expressions $\mathtt{a}$ and $\mathtt{o}$ is defined as and in Fig. 7 (e). Here, we define the sum of two values v_1 and v_2 below. The result of $v_1 {+} v_2$ is legel, if both of the v_1 and v_2 are words (Int32), or v_1 is an address and v_2 is an offset. The offset is a word, which acts as an immediate value in the calculation of address.

$$v_1 {+} v_2 \ ::= \begin{cases} w_1 + w_2 & \textbf{if } v_1 = w_1, \text{ and } v_2 = w_2 \\ (b, w_1 + w_2) & \textbf{if } v_1 = (b, w_1), \text{ and } v_2 = w_2 \\ \bot & \textbf{otherwise} \end{cases}$$

The $\mathtt{wr}$ wants to save the bitwise exclusive OR of the operands into the special register $\mathtt{sr}$, but it puts the write into the delay buffer D instead of updating R immediately. The operation $\mathbf{set\_delay}(\mathtt{sr}, w, D)$ is defined below:

$$\mathbf{set\_delay}(\mathtt{sr}, w, D) ::= (X, \mathtt{sr}, w) {::} D$$

where X $(0 \le X \le 3)$ is a predefined system parameter for the delay cycle.

The $\mathtt{save}$ and $\mathtt{restore}$ instruction rotate the register windows and update the register file. Their operations over F and R are defined in Fig. 6.

3 Program Logic

We first design a program logic for SPARCv8 program, which doesn't consider refinement verification. In this section, we use such program logic to present how our logic handles the features of SPARCv8.

3.1 Assertions

$$\begin{aligned} (Asrt) \ p, q \ ::= \ &\mathtt{emp} \mid l \mapsto v \mid \mathtt{rn} \mapsto v \mid \rhd_t \mathtt{sr} \mapsto w \mid p{\downarrow} \\ &\mid \mathtt{cwp} \mapsto (\!| w_{id}, F |\!) \mid p \wedge q \mid p \vee q \mid p * q \\ &\mid \mathtt{a} =_a v \mid \mathtt{o} = v \mid \forall x.\, p \mid \exists x.\, q \mid \ldots \end{aligned}$$

Fig.8. Syntax of Assertions

$$S \models \mathtt{emp} \quad\quad ::= S.M = \emptyset \wedge S.Q.R = \emptyset$$

$$S \models l \mapsto v \quad\quad ::= S.M = \{l \rightsquigarrow v\} \wedge S.Q.R = \emptyset$$

$$S \models \mathtt{rn} \mapsto v \quad\quad ::= S.Q.R = \{\mathtt{rn} \rightsquigarrow v\} \wedge \mathtt{rn} \notin \mathrm{dom}(S.D) \wedge S.M = \emptyset$$

$$S \models \rhd_t \mathtt{sr} \mapsto w \quad\quad ::= \exists k, R', D'. \, 0 \le k \le t{+}1 \wedge (R, D) \rightrightarrows^k (R', D') \wedge$$
$$((M, (R', F), D') \models \mathtt{sr} \mapsto w) \wedge \mathbf{noDup}(D, \mathtt{sr})$$
$$\text{where } S = (M, (R, F), D)$$

$$S \models p \downarrow \quad\quad ::= \exists R', D'. \, ((M, (R', F), D') \models p) \wedge (R', D') \rightrightarrows (R, D)$$
$$\text{where } S = (M, (R, F), D)$$

$$S \models \mathtt{cwp} \mapsto (\!| w_{id}, F |\!) ::= (S \models \mathtt{cwp} \mapsto w_{id}) \wedge \exists F'. \, F \cdot F' = S.Q.F$$

$$S \models \mathtt{a} =_a v \quad\quad ::= [\![\mathtt{a}]\!]_{S.Q.R} = v \wedge \mathbf{word\_align}(v)$$

$$S \models \mathtt{o} = v \quad\quad ::= [\![\mathtt{o}]\!]_{S.Q.R} = v$$

$$S \models p_1 * p_2 \quad\quad ::= \exists S_1, S_2. \, S_1 \models p_1 \wedge S_2 \models p_2 \wedge S = S_1 \uplus S_2$$

$$S_1 \uplus S_2 \;\; ::= \begin{cases} (M_1 \cup M_2, (R_1 \cup R_2, F), D) & \text{if } M_1 \perp M_2 \wedge R_1 \perp R_2 \wedge \\ & \quad S_1 = (M_1, (R_1, F), D) \wedge S_2 = (M_2, (R_2, F), D) \\ \mathbf{undefined} & \text{otherwise} \end{cases}$$

$$\mathrm{dom}(D) \;\; ::= \begin{cases} \{\mathtt{sr}\} \cup \mathrm{dom}(D') & \text{if } D = (t, \mathtt{sr}, w) :: D' \\ \emptyset & \text{if } D = \mathrm{nil} \end{cases}$$

$$\mathbf{noDup}(D, \mathtt{sr}) \;\; ::= \begin{cases} \mathtt{sr} \notin \mathrm{dom}(D') & \text{if } D = (t, \mathtt{sr}, w) :: D' \\ \mathtt{sr} \neq \mathtt{sr}' \wedge \mathbf{noDup}(D', \mathtt{sr}) & \text{if } D = (t, \mathtt{sr}', w) :: D' \\ \mathbf{True} & \text{if } D = \mathrm{nil} \end{cases}$$

Fig.9. Semantics of Assertions

We define syntax of assertions in Fig. 8, and their semantics in Fig. 9. We extend separation logic assertions with specifications of delay buffers and register windows. Registers are like variables in separation logic, but are treated as resources. The assertion $\mathtt{emp}$ says that the memory and the register file are both empty. $l \mapsto v$ specifies a singleton memory cell with value v stored in the address l. $\mathtt{rn} \mapsto v$ says that $\mathtt{rn}$ is the only register in the register file and it contains the value v. Also $\mathtt{rn}$ is *not* in the delay buffer. Separating conjunction $p * q$ has the standard semantics as in separation logic [12].

The assertion $\rhd_t \mathtt{sr} \mapsto w$ describes a delayed write in the delay buffer D. It describes the uncertainty of $\mathtt{sr}$'s value in R, which is unknown for now but will become w in up to $t{+}1$ cycles. We use $\_ \rightrightarrows^k \_$ to represent k-step execution of the delayed writes in D. It also requires that there be at most one delayed write for a specific special register $\mathtt{sr}$ in D (*i.e.* $\mathbf{noDup}(\mathtt{sr}, D)$). This pre-

vents more than one delayed writes to the same register within 4 instruction cycles, which practically have no restrictions on programming. By the semantics we have

$$\mathtt{sr} \mapsto w \implies \rhd_t \mathtt{sr} \mapsto w \quad\quad \rhd_t \, \mathtt{sr} \mapsto w \implies \rhd_{t+k} \mathtt{sr} \mapsto w$$

The assertion $p \downarrow$ allows us to reduce the uncertainty by executing one step of the delayed writes. It specifies states reachable after executing one step of delayed writes from those states satisfying p. Therefore we know:

$$(\rhd_0 \mathtt{sr} \mapsto w) \downarrow \implies \mathtt{sr} \mapsto w \quad\quad (\rhd_{t+1} \mathtt{sr} \mapsto w) \downarrow \implies \rhd_t \mathtt{sr} \mapsto w$$

Also it's easy to see that if p syntactically does not contain sub-terms in the form of $\rhd_t \mathtt{sr} \mapsto w$, then $(p \downarrow) \iff p$.

The following lemma shows $(\_) \downarrow$ is distributive over separating conjunction.

Lemma 2. $(p * q) \downarrow \iff (p \downarrow) * (q \downarrow)$.

The lemma can be proved following Lemma 1.

We use $\text{cwp} \mapsto (\!|w_{id}, F|\!)$ to describe the pointer cwp of the current register window and the frame list as a circular stack. Note that F is just a prefix of the frame list, since usually we do not need to know contents of the full list. Here we use $F \cdot F'$ to represent the concatenation of lists F and F'. Therefore we have $\text{cwp} \mapsto (\!|w_{id}, F \cdot F'|\!) \implies \text{cwp} \mapsto (\!|w_{id}, F|\!)$.

The assertions $\text{a} =_a v$ and $\text{o} = v$ describe the value of a and o respectively. They are intuitionistic assertions. Since a is used as an address, we also require it to be properly aligned on a 4-byte boundary. We define **word_align** to represent this restriction below. The result of the address expression a may be a word, if it's a pointer in code heap, or a memory address, if it's a location of memory.

$$\textbf{word\_align}(v) ::= \exists\, w, b.\, (v = w \vee v = (b, w)) \wedge w \% 4 = 0$$

3.2 Inference Rules

The code specification θ and code heap specification Ψ are defined below:

(valList) $\iota \in$ list value   (pAsrt) $\text{fp}, \text{fq} \in$ valList $\rightarrow Asrt$
(CdSpec) $\theta ::= (\text{fp}, \text{fq})$ (CdHpSpec) $\Psi ::= \{\text{f} \rightsquigarrow \theta\}^*$

The code heap specification Ψ maps the code labels for basic blocks to their specifications θ, which is a pair of pre- and post-conditions. Instead of using normal assertions, the pre- and post-conditions are assertions parameterized over a list of values $lgvl$. They play the role of auxiliary variables — Feeding the pre- and the post-conditions with the same $lgvl$ allows us to establish relationship of states specified in the pre- and post-conditions.

Although we assign a θ to each basic block, the post-condition does not specify the states reached at the end of the block. Instead, it specifies the condition that needs to be specified in the future when the

current function returns. This follows the idea developed in SCAP [8], but we use the standard unary state assertion instead of the binary state assertions used in SCAP, so that existing proof techniques (such as Coq tactics) for standard Hoare-triples can be applied to simplify the verification process.

```
–  {(fp, fq)}
     add   %i₀, %i₁, %l₇
     add   %l₇, %i₂, %l₇
     retl
     nop
```

$$\text{fp} ::= \lambda\, lv.\, ((\%\text{i}_0 \mapsto lv[0]) * (\%\text{i}_1 \mapsto lv[1]) * (\%\text{i}_2 \mapsto lv[2])$$
$$* \%\text{l}_7 \mapsto \_ * (\text{r}_{15} \mapsto lv[3]))$$
$$\wedge (lv[1], lv[2], lv[3] \in \text{Word})$$
$$\text{fq} ::= \lambda\, lv.\, (\%\text{i}_0 \mapsto lv[0]) * (\%\text{i}_1 \mapsto lv[1]) * (\%\text{i}_2 \mapsto lv[2])$$
$$* (\%\text{l}_7 \mapsto lv[0] + lv[1] + lv[2]) * (\text{r}_{15} \mapsto lv[3])$$

Fig.10. Example for Function Specification

We give a simple example in Fig. 10 to show a specification for a function, which simply sums the values of the registers $\%\text{i}_0$, $\%\text{i}_1$ and $\%\text{i}_2$ and writes the result into the register $\%\text{l}_7$. The specification (fp, fq) says that, when provided with the same lv as argument, the function preserves the value of $\%\text{i}_0$, $\%\text{i}_1$ and $\%\text{i}_2$, $\%\text{l}_7$ at the end contains the sum of $\%\text{i}_0$, $\%\text{i}_1$ and $\%\text{i}_2$, and the function also preserves the value of r_{15}, which it uses as the return address. To verify the function, we need to prove that it satisfies $(\text{fp}\ lv, \text{fq}\ lv)$ for all lv. Here, $lv[1]$ and $lv[2]$ cannot be a memory address, because a value plus a memory address is illegal. $lv[3]$ also should be a word, because it records a word as a return code pointer.

Fig. 11 shows selected inference rules in our logic. The top rule **CDHP** verifies the code heap C. It requires that every basic block specified in Ψ can be verified with respect to the specification, with any argument ι used to instantiate the pre- and post-conditions.

The **SEQ** rule is applied when meeting an instruction sequence starting with a simple instruction i.

$\boxed{\vdash C : \Psi}$ **(Well-Formed Code Heap)**

$$\frac{\text{for all } \mathtt{f} \in \text{dom}(\Psi),\ \iota : \ \ \Psi(\mathtt{f}) = (\text{fp}, \text{fq}) \quad \Psi \vdash \{(\text{fp }\iota, \text{fq }\iota)\}\ \mathtt{f} : C[\mathtt{f}]}{\vdash C : \Psi} \ \textbf{(CDHP)}$$

$\boxed{\Psi \vdash \{(p,q)\}\ \mathtt{f} : \mathbb{I}}$ **(Well-Formed Instruction Sequences)**

$$\frac{\vdash \{p\downarrow\}\ \mathtt{i}\ \{p'\} \quad \Psi \vdash \{(p',q)\}\ \mathtt{f}\!+\!4 : \mathbb{I}}{\Psi \vdash \{(p,q)\}\ \mathtt{f} : \mathtt{i}; \mathbb{I}} \ \textbf{(SEQ)}$$

$$\frac{\begin{array}{c} p\downarrow \Rightarrow (\mathtt{a} =_a \mathtt{f}') \quad \mathtt{f}' \in \text{dom}(\Psi) \quad \Psi(\mathtt{f}') = (\text{fp}, \text{fq}) \\ \vdash \{p\downarrow\downarrow\}\ \mathtt{i}\ \{p'\} \quad \exists \iota, p_r.\ (p' \Rightarrow \text{fp }\iota * p_r) \wedge (\text{fq }\iota * p_r \Rightarrow q) \end{array}}{\Psi \vdash \{(p,q)\}\ \mathtt{f} : \mathtt{jmp}\ \mathtt{a}; \mathtt{i}} \ \textbf{(JMP)}$$

$$\frac{\begin{array}{c} \mathtt{f}' \in \text{dom}(\Psi) \quad \Psi(\mathtt{f}') = (\text{fp}, \text{fq}) \quad \Psi \vdash \{(p',q)\}\ \mathtt{f}\!+\!8 : \mathbb{I} \\ p\downarrow \Rightarrow (\mathtt{r}_{15} \mapsto \_) * p_1 \quad \vdash \{(\mathtt{r}_{15} \mapsto \mathtt{f} * p_1)\downarrow\}\ \mathtt{i}\ \{p_2\} \\ \exists \iota, p_r.\ (p_2 \Rightarrow \text{fp }\iota * p_r) \wedge (\text{fq }\iota * p_r \Rightarrow p') \wedge (\text{fq }\iota \Rightarrow \mathtt{r}_{15} = \mathtt{f}) \end{array}}{\Psi \vdash \{(p,q)\}\ \mathtt{f} : \mathtt{call}\ \mathtt{f}'; \mathtt{i}; \mathbb{I}} \ \textbf{(CALL)}$$

$$\frac{p\downarrow\downarrow \Rightarrow (\mathtt{r}_{15} \mapsto \mathtt{f}') * p_1 \quad \vdash \{p_1\}\ \mathtt{i}\ \{p_2\} \quad (\mathtt{r}_{15} \mapsto \mathtt{f}') * p_2 \Rightarrow q}{\Psi \vdash \{(p,q)\}\ \mathtt{f} : \mathtt{retl}; \mathtt{i}} \ \textbf{(RETL)}$$

$\boxed{\vdash \{p\}\ \mathtt{i}\ \{q\}}$ **(Well-Formed Instructions)**

$$\frac{\mathtt{sr} \mapsto \_ * p \Rightarrow (\mathtt{r}_s = w_1 \wedge \mathtt{o} = w_2)}{\vdash \{\mathtt{sr} \mapsto \_ * p\}\ \mathtt{wr}\ \mathtt{r}_s\ \mathtt{o}\ \mathtt{sr}\ \{(\rhd_3 \mathtt{sr} \mapsto (w_1 \oplus w_2)) * p\}} \ \textbf{(WR)}$$

$$\frac{}{\vdash \{\mathtt{sr} \mapsto w * \mathtt{r}_d \mapsto \_\}\ \mathtt{rd}\ \mathtt{sr}\ \mathtt{r}_d\ \{\mathtt{sr} \mapsto w * \mathtt{r}_d \mapsto w\}} \ \textbf{(RD)}$$

$$\frac{\begin{array}{c} p \Rightarrow (\mathtt{r}_s = v_1 \wedge \mathtt{o} = v_2) \quad w'_{id} = \textbf{next\_cwp}(w_{id}) \quad w\ \&\ 2^{w'_{id}} = 0 \quad v = v_1 + v_2 \\ p \Rightarrow (\mathtt{cwp} \mapsto (\!| w_{id}, F \cdot \_ \cdot \_ |\!)) * (\mathtt{out} \mapsto \text{fm}_o) * (\mathtt{local} \mapsto \text{fm}_l) * (\mathtt{in} \mapsto \text{fm}_i) * p_1 \\ (\mathtt{cwp} \mapsto (\!| w'_{id}, \text{fm}_l :: \text{fm}_i :: F |\!)) * (\mathtt{out} \mapsto \_) * (\mathtt{local} \mapsto \_) * (\mathtt{in} \mapsto \text{fm}_o) * p_1 \Rightarrow \mathtt{r}_d \mapsto \_ * p_2 \end{array}}{\vdash \{(\mathtt{wim} \mapsto w) * p\}\ \mathtt{save}\ \mathtt{r}_s\ \mathtt{o}\ \mathtt{r}_d\ \{(\mathtt{wim} \mapsto w) * (\mathtt{r}_d \mapsto v) * p_2\}} \ \textbf{(SAVE)}$$
where $[\mathtt{r}_i, \dots, \mathtt{r}_{i+7}] \mapsto [w_0, \dots, w_7] \ ::= \mathtt{r}_i \mapsto w_0 * \cdots * \mathtt{r}_{i+7} \mapsto w_7$
and out, local and in are defined in Fig. 6.

$$\frac{\begin{array}{c} p \Rightarrow (\mathtt{r}_s = v_1 \wedge \mathtt{o} = v_2) \quad w'_{id} = \textbf{prev\_cwp}(w_{id}) \quad w\ \&\ 2^{w'_{id}} = 0 \quad v = v_1 + v_2 \\ p \Rightarrow (\mathtt{cwp} \mapsto (\!| w_{id}, \text{fm}_1 :: \text{fm}_2 :: F |\!)) * (\mathtt{out} \mapsto \_) * (\mathtt{local} \mapsto \_) * (\mathtt{in} \mapsto \text{fm}_i) * p_1 \\ (\mathtt{cwp} \mapsto (\!| w'_{id}, F \cdot \_ \cdot \_ |\!)) * (\mathtt{out} \mapsto \text{fm}_i) * (\mathtt{local} \mapsto \text{fm}_1) * (\mathtt{in} \mapsto \text{fm}_2) * p_1 \Rightarrow \mathtt{r}_d \mapsto \_ * p_2 \end{array}}{\vdash \{(\mathtt{wim} \mapsto w) * p\}\ \mathtt{restore}\ \mathtt{r}_s\ \mathtt{o}\ \mathtt{r}_d\ \{(\mathtt{wim} \mapsto w) * (\mathtt{r}_d \mapsto v) * p_2\}} \ \textbf{(RESTORE)}$$

Fig.11. Seleted Inference Rules

The instruction i is verified by the corresponding well-formed instruction rules, with the precondition $p{\downarrow}$ and some post-condition p'. We use $p{\downarrow}$ because there is an implicit step executing delayed writes before executing every instruction. The post-condition p' for i is then used as the precondition to verify the remaining part of the instruction sequence.

Delayed control transfers. We distinguish the jmp and call instructions — The former makes an *intra-function* control transfer, while the latter makes function calls. The **JMP** rule requires that the target address is a valid one specified in Ψ. Starting from the precondition p, after executing the instruction i following **JMP** and the corresponding delayed writes, the post-condition p' of i should satisfy the precondition of the target instruction sequence, with some instantiation ι of the logical variables and a frame assertion p_r. Since the target instruction sequence of jmp is in the same function as the jmp instruction itself, the post-condition fq specified at the target address (with the same instantiation ι of the logical variables and the frame assertion p_r) should meet the post-condition q of the current function. As we explained before, the post-condition q does not specify the states reached at the end of the instruction sequence (which are specified by p' instead).

The **CALL** rule is similar to the **JMP** rule in that it also requires the post-condition p_2 of the instruction i following the call satisfy the precondition of the target instruction sequence, with some instantiation ι of the logical variables and a frame assertion p_r. Here we need to record that the code label f is saved in r_{15} by the call instruction. When the callee returns, its post-condition fq (with the same instantiation of auxiliary variables ι) needs to ensure r_{15} still contains f, so that the callee returns to the correct address. Also the fq

with the frame p_r needs to satisfy the precondition p' for the remaining instruction sequences of the caller.

The **RETL** rule simply requires that the post-condition q holds at the end of the instruction i following retl. Also i cannot touch the register r_{15}, therefore r_{15} specified in p must be the same as in q. Since at the calling point we already required that the post-condition of the callee guarantees r_{15} contains the correct return address, we know r_{15} contains the correct value before retl.

Delayed writes and register windows. The bottom layer of our logic is for well-formed instructions. The **WR** rule requires the ownership of the target register sr in the precondition $(\text{sr} \mapsto \_)$. Also it implies there is no delayed writes to sr in the delay buffer (see the semantics defined in Fig. 9). At the end of the delayed write, we use $\triangleright_3 \text{sr} \mapsto w_1 \oplus w_2$ to indicate the new value will be ready in up to 3 cycles. Since the maximum delay cycle X cannot be bigger than 3 and the value of X may vary in different systems, programmers usually take a conservative approach to assume $X = 3$ for portability of code. Our rule reflects this conservative view. The **RD** rule says the special register can be read only if it is not in the delay buffer. The **SAVE** and **RESTORE** rules reflect the save and recovery of the execution contexts, which is consistent with the operational semantics of the save and restore instructions given in Figs. 6 and 7.

4 Refinement Verification of SPARCv8

In this section, we present *relational* program logic for refinement verification of SPARCv8 code. We first define the high- and low-level program in Sec. 4.1 and Sec. 4.2 respectively. The refinement relation is represented in Sec. 4.3, and program logic is shown in Sec. 4.4 Finally, logic soundness is shown in Sec. 4.5.

4.1 High-level Pseudo-SPARCv8 Language

The Pseudo-SPARCv8 language in this contains two part: the SPARCv8 code as client code, and the set of abstract assembly primitives. Here, we require that the execution of client SPARCv8 code preserves a restriction between register window and stack in memory, shown as the left side in Fig. 12 (`cwp` points to the current window and `wim` marks the invalid window, the details of overlapping of adjacent windows are omitted in the figure).
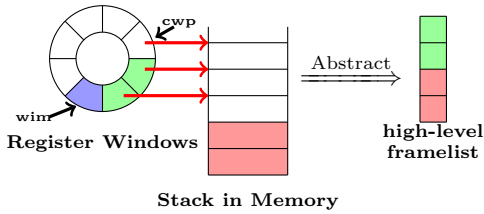

Fig.12. Abstraction of context management

During the SPARCv8 program's execution, part of previous procedures' contexts (the green part in the left side of the Fig. 12) are saving in register window, the others (the pink part in the left side of the Fig. 12) are stored in stack in memory, because the number of windows is limited. The restriction is that the stack pointer (%sp) of each procedure, including the current one and the previous one, whose contexts are saved in register window currently, should point to the top of its stack frame (shown as the red arrow in Fig. 12), so that the contexts in these windows can be stored correctly in memory when needed. For instance, the context switch routine will check whether the previous window is valid (in clockwise direction in Fig. 12), and use instruction `restore` to set it as the current one and save its contents into stack (in memory) until the previous one is invalid (marked in blue in Fig. 12). We require the execution of client code preserving such restriction. Otherwise, some SPARCv8 functions like context switch routine whose implementations will

store the contexts saved in register window into stack in memory cannot be verified if it's unclear where to save the contents of some windows. We do the following when defining Pseudo-SPARCv8 program to make the execution of client code preserves such restriction:

- In order to ensure that the stack pointer (%sp) always point to the top of its stack frame, we require that each instruction, like `add` and `ld`, whose execution will not rotate register window, is not allowed to update the value of %sp; and as for the `save` and `restore`, whose executions will rotate register window, we restrict them to be used in specific forms. We introduce "`Psave` w" as a macro of "`save` %sp, $-w$, %sp", whose execution makes sure that a new %sp will be generated for the next window and point to the stack frame size w allocated newly. We also introduce "`Prestore`" as a macro of "`restore` %g$_0$, %g$_0$, %g$_0$" [3], whose execution just restores the previous window and doesn't modify the value of any register in the previous window restored. The original `save` and `restore` instructions have *no* semantics in high-level client code.

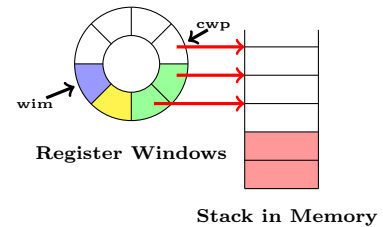- The special registers in SPARCv8 usually act specific roles and modifying them should be carefully,

Fig.13. Problem of modifying `wim` arbitrary

for example, `wim` marks which window is invalid. If we change its value shown in Fig. 12, to mark another window invalid, as shown in Fig. 13, and

---

[3]In SPARCv8, %g$_0$ is always equal to 0, and usually used as parameters when instructions do not require specific parameters.

call context switch routine, which will save the contents of previous windows into memory until the invalid one, at this moment. There will be a problem that we don't know where to save the contents of window that is marked invalid originally (window in yellow in Fig. 13). So, we don't allow client code to modify special registers and do *not* give semantics to `wr` instruction in high-level client code. Modifying them is hidden in the implementation of the abstract assembly primitives in low-level. The delay buffer can be omitted in high-level program state.

- As shown in Fig. 12, we find that we can abstract the register window and the memory in stack for storing contexts as a list (defined as HFrmList formally in Fig. 15). After this abstraction, We don't need to care about whether the contexts are saved in register window or memory and describe the contents of windows unused (the windows in white color in Fig. 12, but excluding the current one pointed by `cwp`) in the Pseudo-SPARCv8 level. The `cwp` register is no longer needed in Pseudo-SPARCv8 program because the register window is abstracted away. The low-level program in our work doesn't use this abstraction, because the low-level program should be realistically modelled, and the implementations of some primitives need to know the existance of register window, for instance, the context switch routine needs to save the contents of the register window into stack (in memory).

We define the syntax of the high-level Pseudo-SPARCv8 language in Fig. 14. The code Π has two parts : the code heap C and a set of abstract primitives Ω, which is a partial mapping from labels to abstract assembly primitive. The code heap C in Π acts as the client code to call abstract assembly primitive. The abstract assembly primitive Υ is defined as a relation that takes a list of values as arguments and maps a high-level program state (defined in Fig. 15) to another. Comparing the simple instruction with the one shown in Fig. 3, we add three pseudo instructions. The `Psave` w and `Prestore` restrict the `save` and `restore` instructions can only be used in specific form as mentioned before. We also introduce `print r`, whose execution will output the value v in `r` and occur an message $\mathsf{out}(v)$, to generate observable event. The high-level message α can be either an empty message τ, or an output $\mathsf{out}(v)$, or a $\mathsf{call}(\mathtt{f}, \overline{v})$ meaning to call a primitive labelled $\mathtt{f}$ with arguments $\overline{v}$.

The machine states of high-level Pseudo-SPARCv8 program is defined in Fig. 15. The high-level program $\mathbb{P}$ is a a pair of high-level code Π and high-level state $\mathbb{S}$. High-level program state is a tuple including: a thread pool T, current thread id $\mathtt{t}$, the thread local state $\mathcal{K}$ of the current thread, and the memory M.

Thread Local State. The thread local state $\mathcal{K}$ is a triple of high-level register state $\mathbb{Q}$, and program conunters `pc` and `npc`. The high-level register state $\mathbb{Q}$ consists: the high-level register file $\mathbb{R}$, and the high-level frame list $\mathbb{F}$. $\hat{\mathrm{rn}}$ is the high-level register names, where the `cwp` is omitted as introduced before, and we also omit special register for simplicity, because we don't allow the high-level client code to modify them. The high-level frame list $\mathbb{F}$ is a list of pair $(\mathrm{fm}_1, \mathrm{fm}_2)$, which is used to save the contexts (`local` and `in` registers) fm_1 and fm_2 of the previous procedure. After introducing the state of high-level program, We define the primitive `switch` as an instantiation of Υ following:

$$
\begin{aligned}
\mathsf{switch} ::= \\
\lambda \overline{v}, \mathbb{S}, \mathbb{S}'.\ \exists \mathtt{t}'.\ M(\mathsf{TaskNew}) = (\mathtt{t}', 0) \wedge T(\mathtt{t}') = (\mathbb{Q}', \mathtt{pc}', \mathtt{npc}') \\
\wedge T' = T\{\mathtt{t} \rightsquigarrow (\mathbb{Q}, \mathtt{pc}, \mathtt{npc})\} \wedge \mathtt{t} \neq \mathtt{t}' \wedge \overline{v} = \mathrm{nil} \\
\text{where } \mathbb{S} = (T, \mathtt{t}, (\mathbb{Q}, \mathtt{pc}, \mathtt{npc}), M), \\
\mathbb{S}' = (T', \mathtt{t}', (\mathbb{Q}', \mathtt{f}{+}8, \mathtt{f}{+}12), M), \mathtt{f} = \mathbb{Q}'.\mathbb{R}(\mathbf{r}_{15}).
\end{aligned}
$$

(HCode) Π ::= (C, Ω) (CodeHeap) C $\in$ Word $\rightharpoonup$ Comm

(PrimSet) Ω ::= $\{f_1 \rightsquigarrow \Upsilon_1, \ldots, f_n \rightsquigarrow \Upsilon_n\}$ (Prim) Υ $\in$ List Val $\rightarrow$ HState $\rightarrow$ HState $\rightarrow$ Prop

(Comm) c ::= $\mathtt{i} \mid \mathtt{call\ f} \mid \mathtt{jmp\ a} \mid \mathtt{retl} \mid \mathtt{be\ f}$

(SimpIns) $\mathtt{i}$ ::= $\mathtt{Psave}\ w \mid \mathtt{Prestore} \mid \mathtt{print\ r} \mid \mathtt{ld\ a\ r}_d \mid \mathtt{st\ r}_s\ \mathtt{a} \mid \mathtt{add\ r}_s\ \mathtt{o\ r}_d \mid \mathtt{rd\ sr\ r}_d \mid \mathtt{wr\ r}_s\ \mathtt{o\ sr}$
               $\mid \mathtt{save\ r}_s\ \mathtt{o\ r}_d \mid \mathtt{restore\ r}_s\ \mathtt{o\ r}_d \mid \ldots$

(HMsg) α ::= $\tau \mid \mathsf{out}(v) \mid \mathsf{call}(\mathtt{f}, \overline{v})$

Fig.14. Syntax of Pseudo-SPARCv8 Code

(HProg) $\mathbb{P}$ ::= $(\Pi, \mathbb{S})$ (HState) $\mathbb{S}$ ::= $(T, \mathsf{t}, \mathcal{K}, M)$

(ThrdPool) T ::= $\{\mathsf{t} \rightsquigarrow \mathcal{K}\}^*$ (ThrdLcSt) $\mathcal{K}$ ::= $(\mathbb{Q}, \mathsf{pc}, \mathsf{npc})$ (Tid) $\mathsf{t} \in \mathbb{Z}$

(HRstate) $\mathbb{Q}$ ::= $(\mathbb{R}, \mathbb{F})$

(HRegFile) $\mathbb{R}$ $\in$ HRegName $\rightarrow$ Val   (HRegName) $\hat{\mathtt{rn}}$ ::= $\mathtt{r}_0 \mid \ldots \mid \mathtt{r}_{31} \mid \mathtt{n} \mid \mathtt{z} \mid \mathtt{c} \mid \mathtt{v}$

(HFrmList) $\mathbb{F}$ ::= $\mathrm{nil} \mid (\mathrm{fm}_1, \mathrm{fm}_2) :: \mathbb{F}$ (HFrame) fm ::= $[v_0, \ldots, v_7]$

Fig.15. Machine States for Pseudo-SPARCv8 Code

The execution of switch primitive takes no arguments ($\overline{v} = \mathrm{nil}$), and change the identifier of current thread according to the pointer saved in location TaskNew. We use parameters $\mathbb{S}$ and $\mathbb{S}'$ to represent the machine state before and after execution of switch respectively.

Operational Semantics in High-level. The operational semantics for high-level Pseudo-SPARCv8 program is defined in Fig. 16. The high-level program transition relation $(\Pi, \mathbb{S}) :\stackrel{\alpha}{\Longrightarrow} (\Pi, \mathbb{S}')$ is defined in Fig. 16 (a). In each step, the program may either execute the instruction pointed by pc, and occur empty message τ or an output $\mathsf{out}(v)$, or call an abtract assembly primitive in primitive set. When calling an abstract assembly primitive, the execution of current thread (defined as $(\_ \Vdash \_ \circ \longrightarrow \_)$ in Fig. 16 (b)) will occur a message $\mathsf{call}(\mathtt{f}, \overline{v})$, which means that it hopes to call the abstract assembly primitive Υ labelled $\mathtt{f}$, which is *not* in the domain of code heap C, with arguments $\overline{v}$ (we use $\mathbf{args}(\mathbb{Q}, M, \overline{v})$ to get arguments $\overline{v}$ from high-level state, and its definition is omitted here).

The control transfer step is defined in Fig. 16 (b). Here, the step for simple instruction i is represented as "$\mathbf{execi}(\mathtt{i}, \_) =_{\mathsf{H}} \_$". We show the state transition relation for pseudo instructions $\mathtt{Psave}\ w$ and $\mathtt{Prestore}$

in Fig. 16 (c). Supposing the current register state $\mathbb{Q}$ is $(\mathbb{R}, \mathbb{F})$, executing instruction $\mathtt{Psave}\ w$ will save the local and in registers into high-level frame list $\mathbb{F}$. It also allocates a new block b, size from 64 byte to w byte, as a new stack frame in memory (represented as $\boldsymbol{alloc}(M, b, 64, w) = M'$). The reason why it starts from 64 byte is that the 0 to 64 bytes (16 words) in a stack frame are usually reserved to save the contexts in window (local and in registers) in convention [5]. However this part of memory is abstracted away in Pseudo-SPARCv8 program as we have explained and shown in Fig. 12. The instruction $\mathtt{Prestore}$ does the reverse, freeing the block of current stack frame (represented as $\boldsymbol{free}(b, M) = M'$), and restoring the contexts of the previous procedure saved in $\mathbb{F}$.

4.2   Low-level SPARCv8 Program

The low-level SPARCv8 program are very closed to the SPARCv8 program defined in Fig. 3. The only difference here is that we use simple instructions and commands defined in Fig. 14. So, the global program transition of the low-level SPARCv8 program is defined

$$\frac{\Pi = (C, \Omega) \quad C \Vdash (\mathcal{K}, M) \circ\xrightarrow{\tau} (\mathcal{K}', M')}{(\Pi, (T, \mathsf{t}, \mathcal{K}, M)) :\Longrightarrow (\Pi, (T, \mathsf{t}, \mathcal{K}', M'))} \qquad \frac{\Pi = (C, \Omega) \quad C \Vdash (\mathcal{K}, M) \circ\xrightarrow{\mathsf{out}(v)} (\mathcal{K}', M)}{(\Pi, (T, \mathsf{t}, \mathcal{K}, M)) :\xrightarrow{\mathsf{out}(v)} (\Pi, (T, \mathsf{t}, \mathcal{K}', M))}$$

$$\frac{\begin{array}{c} \Pi = (C, \Omega) \quad C \Vdash (\mathcal{K}, M) \circ\xrightarrow{\mathsf{call}(\mathtt{f}, \overline{v})} (\mathcal{K}', M) \quad \Omega(\mathtt{f}) = \Upsilon \\ \Upsilon(\overline{v})(T, \mathsf{t}, \mathcal{K}', M)(T', \mathsf{t}', \mathcal{K}'', M') \end{array}}{(\Pi, (T, \mathsf{t}, \mathcal{K}, M)) :\Longrightarrow (\Pi, (T', \mathsf{t}', \mathcal{K}'', M'))}$$

(a) High-level Program Transition

$$\frac{C(\mathtt{pc}) = \mathtt{i} \quad \mathbf{execi}(\mathtt{i}, (\mathbb{Q}, M)) =_{\mathsf{H}} (\mathbb{Q}', M')}{C \Vdash ((\mathbb{Q}, \mathtt{pc}, \mathtt{npc}), M) \circ\xrightarrow{\tau} ((\mathbb{Q}', \mathtt{npc}, \mathtt{npc} + 4), M')}$$

$$\frac{C(\mathtt{pc}) = \mathtt{call}\ \mathtt{f} \quad \mathtt{r}_{15} \in \mathrm{dom}(\mathbb{R})}{C \Vdash (((\mathbb{R}, \mathbb{F}), \mathtt{pc}, \mathtt{npc}), M) \circ\xrightarrow{\tau} (((\mathbb{R}\{\mathtt{r}_{15} \rightsquigarrow \mathtt{pc}\}, \mathbb{F}), \mathtt{npc}, \mathtt{f}), M)}$$

$$\frac{C(\mathtt{pc}) = \mathtt{retl} \quad \mathbb{R}(\mathtt{r}_{15}) = \mathtt{f}}{C \Vdash (((\mathbb{R}, \mathbb{F}), \mathtt{pc}, \mathtt{npc}), M) \circ\xrightarrow{\tau} (((\mathbb{R}, \mathbb{F}), \mathtt{npc}, \mathtt{f} + 8), M)}$$

$$\frac{C(\mathtt{pc}) = \mathtt{print}\ \mathtt{r} \quad \mathbb{R}(\mathtt{r}) = v}{C \Vdash (((\mathbb{R}, \mathbb{F}), \mathtt{pc}, \mathtt{npc}), M) \circ\xrightarrow{\mathsf{out}(v)} (((\mathbb{R}, \mathbb{F}), \mathtt{npc}, \mathtt{npc} + 4), M)}$$

$$\frac{\mathtt{pc} \notin \mathrm{dom}(C) \quad \mathtt{npc} = \mathtt{pc} + 4 \quad \mathbf{args}(\mathbb{Q}, M, \overline{v})}{C \Vdash ((\mathbb{Q}, \mathtt{pc}, \mathtt{npc}), M) \circ\xrightarrow{\mathsf{call}(\mathtt{pc}, \overline{v})} ((\mathbb{Q}, \mathtt{pc}, \mathtt{npc}), M)}$$

(b) High-level Control Transfer Instruction Transition

$$\frac{\begin{array}{c} \mathbb{Q} = (\mathbb{R}, \mathbb{F}) \quad \mathbb{R}' = \mathbb{R}\{\mathsf{out} \rightsquigarrow \_, \mathsf{local} \rightsquigarrow \_, \mathsf{in} \rightsquigarrow \mathbb{R}([\mathsf{out}])\}\{\%\mathtt{sp} \rightsquigarrow (b, 0)\} \\ \boldsymbol{alloc}(M, b, 64, w) = M' \quad \mathbb{Q}' = (\mathbb{R}', (\mathbb{R}([\mathsf{local}]), \mathbb{R}([\mathsf{in}]) :: \mathbb{F}) \end{array}}{\mathbf{execi}(\mathtt{Psave}\ w, (\mathbb{Q}, M)) =_{\mathsf{H}} (\mathbb{Q}', M')}$$

$$\frac{\begin{array}{c} \mathbb{Q} = (\mathbb{R}, (\mathrm{fm}_1, \mathrm{fm}_2 :: \mathbb{F}) \quad \mathbb{R}(\%\mathtt{sp}) = (b, 0) \quad \boldsymbol{free}(b, M) = M' \\ \mathbb{R}' = \mathbb{R}\{\mathsf{out} \rightsquigarrow \mathbb{R}([\mathsf{in}]), \mathsf{local} \rightsquigarrow \mathrm{fm}_1, \mathsf{in} \rightsquigarrow \mathrm{fm}_2\} \quad \mathbb{Q}' = (\mathbb{R}', \mathbb{F}) \end{array}}{\mathbf{execi}(\mathtt{Prestore}, (\mathbb{Q}, M)) =_{\mathsf{H}} (\mathbb{Q}', M')}$$

(c) High-level Instruction Transition

Fig.16. Seletcted operational semantics rules for high-level program

as the following form :

$$(R, D) \rightrightarrows (R', D')$$
$$\frac{C \vdash ((M, (R', F), D'), \mathtt{pc}, \mathtt{npc}) \circ\xrightarrow{\tau/\mathsf{out}(v)} (M', (R'', F'), D'')}{(C, (M, (R, F), D), \mathtt{pc}, \mathtt{npc}) ::\xrightarrow{\tau/\mathsf{out}(v)} (C, (M, (R'', F'), D''))}$$

Each step of the program produces either an empty message τ, or an output $\mathsf{out}(v)$ that is produced by the instruction $\mathtt{print}$ and acts as an observable behavior.

4.3 Primitive Correctness

We first establish a relation as an *invariant* between low- and high-level program state. We define this rela-

tion below, shown as "$S \sim \mathbb{S}$".

$$\frac{M = M_c \uplus M_T \uplus \{\mathsf{TaskCur} \rightsquigarrow (\mathsf{t}, 0)\} \uplus M'}{(M_c, Q) \Downarrow_{\mathsf{c}} (\mathsf{t}, \mathcal{K}) \quad M_T \Downarrow_{\mathsf{r}} T \backslash \{\mathsf{t}\} \quad D = \mathrm{nil}}{(M, Q, D) \sim (T, \mathsf{t}, \mathcal{K}, M')}$$

The low-level memory M is splitted into four parts: M_c used to save the context of the current thread t; M_T saving the contexts of the ready threads, except the current thread t; a singleton memory cell located $\mathsf{TaskCur}$ saving the current thread id; and shared memory M'. The delayed buffer D is nil, because client code is not permitted to modify special register by $\mathtt{wr}$

instruction. The memory M_T used to save the contexts of the ready threads is *abstracted* as a thread pool in high-level program. Their relation is represented as "$M_T \Downarrow_r T\backslash\{\mathsf{t}\}$". We use "$(M_c, Q) \Downarrow_\mathsf{c} (\mathsf{t}, \mathcal{K})$" to represent the state relation of current thread t in low- and high-level program. Full definitions of the state relation can be found in Appendix C.

The correctness of abstract assembly primitive can be defined in terms of *contextual refinement*. Below we will give the formal definition in Def. 1. And we use *event trace refinement* proposed by Liang *et al.* [13].

Definition 1 (Primitive Correctness). $C_{as} \sqsubseteq \Omega$ iff for any C, S, $\mathbb{S}$, $\mathtt{pc}$ and $\mathtt{npc}$, if $S \sim \mathbb{S}$, and $\mathsf{ProgSafe}(\mathbb{P})$, then $\mathrm{P} \sqsubseteq \mathbb{P}$ holds. (where $\mathrm{P} = (C \uplus C_{as}, S, \mathtt{pc}, \mathtt{npc})$, $\mathbb{P} = ((C, \Omega), \mathbb{S})$, and $\mathbb{S}.\mathcal{K} = (\_, \mathtt{pc}, \mathtt{npc})$).

We use a code heap C_{as} to represent the implementations and Ω to represent the set of corresponding abstract assembly primitives. The contextual refinement between C_{as} and Ω, denoted as $C_{as} \sqsubseteq \Omega$, says that if and only if for any client code (or context) C, low-level program state S, high-level program state $\mathbb{S}$, program counters $\mathtt{pc}$ and $\mathtt{npc}$, if the low- and high-level program states satisfy the state relation $S \sim \mathbb{S}$ and the high-level program will never get stuck (shown as $\mathsf{ProgSafe}(\mathbb{P})$), then there is an *event trace refinement* relation, which means P produce no more observable behaviors than $\mathbb{P}$ and is denoted as $\mathrm{P} \sqsubseteq \mathbb{P}$, between low- and high-level program. $\mathsf{ProgSafe}(\mathbb{P})$ is defined below formally :

$$\mathsf{ProgSafe}(\mathbb{P}) ::= \forall \mathbb{P}'. (\mathbb{P} :\Longrightarrow^* \mathbb{P}') \Longrightarrow (\exists \mathbb{P}''. \mathbb{P}' :\Longrightarrow \mathbb{P}'')$$

4.4 Relational Program Logic for Refinement Verification

$$
\begin{aligned}
(RelAsrt) \quad \mathsf{p}, \mathsf{q} \quad ::=\ & \mathtt{Emp} \mid p \mid \hat{\mathsf{rn}} \rightarrowtail v \mid l \rightarrowtail v \\
& \mid\ \mathsf{t} \rightsquigarrow_\mathsf{c} \mathcal{K} \mid \mathsf{t} \rightsquigarrow_\mathsf{r} \mathcal{K} \mid (\!|A|\!) \mid \blacklozenge(w) \\
& \mid\ \mathsf{p}{\downarrow} \mid \mathsf{p} \wedge \mathsf{q} \mid \mathsf{p} \vee \mathsf{q} \mid \mathsf{p} * \mathsf{q} \mid \dots
\end{aligned}
$$

Fig.17. Syntax of Relational Assertion

Relational Assertion Fig. 17 gives the *relational* assertion language, and its semantics is given in Fig. 18. The relational assertions are interpreted over relational states $(S, \mathbb{S}, A, w)$, which contains the low-level state S, the high-level state $\mathbb{S}$, the abstract assembly primitive command A defined in Fig. 18, and a word w recording the number of the tokens. The high-level primitive command A is either an abstract assembly primitive Υ parametered with its arguments $\overline{v}$, or a $\perp$ meaning the primitive has already been executed. The relational assertion p reserves original assertion p describing the low-level state S.

We define $\hat{\mathsf{rn}} \rightarrowtail v$ and $l \rightarrowtail v$ to describe the state of register file and memory in high-level. The assertion $\mathtt{Emp}$ says that the high-level memory and thread pool are both empty, and the low-level state satisfies emp defined in Fig. 9. The assertion $\mathsf{t} \rightsquigarrow_\mathsf{c} \mathcal{K}$ and $\mathsf{t} \rightsquigarrow_\mathsf{r} \mathcal{K}$ represent the thread local state of current thread and ready thread respectively. Note the threads in thread pool are viewed as resources and can be separated by separation conjunction.

The assertion $(\!|A|\!)$ means the current high-level primitive command is A. And the assertion $\blacklozenge(w)$ takes a word w to record lower bound of the number of the tokens, which can also be separated by separation conjunction, in current state. In the introduction of the inference rules following, we use tokens to avoid infinite loops and recursive calls to make sure the termination preserving refinement.

Inference Rules in Relational Program Logic
The code specification $\hat{\theta}$ and code heap specification Ψ for refinement verification are defined below :

$$
\begin{aligned}
&(\text{valList}) \quad \iota \ \in\ \text{list value} \quad (\text{pAsrt}) \ \ \mathbb{fp}, \mathbb{fq} \in \text{valList} \rightarrow RelAsrt \\
&(\text{CdSpec}) \ \ \hat{\theta} ::= (\mathbb{fp}, \mathbb{fq}) \quad\quad (\text{CdHpSpec}) \ \Psi \ ::= \{\mathsf{f} \rightsquigarrow \hat{\theta}\}^*
\end{aligned}
$$

Here, $\mathbb{fp}$ and $\mathbb{fq}$ are relational assertions paramterized over a list of values ι. Fig. 19 shows selected inference rules for refinement verification in our logic.

$$(S, \mathbb{S}, A, w) \models \mathtt{Emp} \quad ::= \quad \mathbb{S}.M = \emptyset \land \mathbb{S}.T = \emptyset \land S \models \mathsf{emp}$$

$$(S, \mathbb{S}, A, w) \models p \quad ::= \quad S \models p \land \mathbb{S}.M = \emptyset \land \mathbb{S}.T = \emptyset$$

$$(S, \mathbb{S}, A, w) \models \hat{\mathfrak{rn}} \rightarrowtail v \quad ::= \quad \mathbb{S}.\mathcal{K}.\mathbb{Q}.\mathbb{R}(\hat{\mathfrak{rn}}) = v \land (S, \mathbb{S}, A, w) \models \mathtt{Emp}$$

$$(S, \mathbb{S}, A, w) \models l \rightarrowtail v \quad ::= \quad \mathbb{S}.M = \{l \rightsquigarrow v\} \land \mathbb{S}.T = \emptyset \land S \models \mathsf{emp}$$

$$(S, \mathbb{S}, A, w) \models \mathtt{t} \rightsquigarrow_c \mathcal{K} \quad ::= \quad \mathbb{S}.T \backslash \{\mathtt{t}\} = \emptyset \land \mathbb{S}.\mathtt{t} = \mathtt{t} \land \mathbb{S}.\mathcal{K} = \mathcal{K} \land \mathbb{S}.M = \emptyset \land S \models \mathsf{emp}$$

$$(S, \mathbb{S}, A, w) \models \mathtt{t} \rightsquigarrow_r \mathcal{K} \quad ::= \quad \mathbb{S}.T = \{\mathtt{t} \rightsquigarrow \mathcal{K}\} \land \mathbb{S}.M = \emptyset \land \mathtt{t} \neq \mathbb{S}.\mathtt{t} \land S \models \mathsf{emp}$$

$$(S, \mathbb{S}, A, w) \models (\!|A'|\!) \quad ::= \quad A = A' \land (S, \mathbb{S}, A, w) \models \mathtt{Emp}$$

$$(S, \mathbb{S}, A, w) \models \blacklozenge(w') \quad ::= \quad w' \leq w \land (S, \mathbb{S}, A, w) \models \mathtt{Emp}$$

$$(S, \mathbb{S}, A, w) \models \mathfrak{p}\!\downarrow \quad ::= \quad \exists S'. ((S', \mathbb{S}, A, w) \models \mathfrak{p}) \land (R', D') \rightrightarrows (R, D)$$
$$\text{where} \quad S = (M, (R, F), D), \; S' = (M, (R', F), D')$$

$$(S, \mathbb{S}, A, w) \models \mathfrak{p} * \mathfrak{q} \quad ::= \quad \exists S_1, S_2, \mathbb{S}_1, \mathbb{S}_2, w_1, w_2. \, S = S_1 \uplus S_2 \land \mathbb{S} = \mathbb{S}_1 \uplus \mathbb{S}_2 \land$$
$$w = w_1 + w_2 \land (S_1, \mathbb{S}_1, A, w_1) \models \mathfrak{p} \land (S_2, \mathbb{S}_2, A, w_2) \models \mathfrak{q}$$

$$\mathbb{S}_1 \uplus \mathbb{S}_2 \; ::= \; \begin{cases} (T_1 \cup T_2, \mathtt{t}, \mathcal{K}, M_1 \cup M_2) & \text{if } T_1 \perp T_2 \land M_1 \perp M_2 \land \\ & \quad S_1 = (T_1, \mathtt{t}, \mathcal{K}, M_1) \land S_2 = (T_2, \mathtt{t}, \mathcal{K}, M_2) \\ \mathbf{undefined} & \mathbf{otherwise} \end{cases}$$

$$(\text{HPrimCom}) \quad A \; ::= \; \Upsilon(\overline{v}) \mid \perp \qquad \frac{\Upsilon(\overline{v})(\mathbb{S})(\mathbb{S}')}{(\Upsilon(\overline{v}), \mathbb{S}) \dashrightarrow (\perp, \mathbb{S}')}$$

Fig.18. Semantics of Relation Assertion

The top rule **WfPrim** verifies the contextual refinement between the code heap C_{as} and the corresponding abstract assembly primitive set Ω. It requires that each code block specified in Ψ can be verified with respect to the specification, shown as ($\vdash C_{\mathrm{as}} : \Psi$), and the specification of the implementation of the abstract assembly primitive needs to meet some restrictions, shown as $\mathsf{wdSpec}(\mathfrak{fp}, \mathfrak{fq}, \Upsilon)$, which we will discuss in more details following. Most of the inference rules for verifying the instruction sequence ($\Psi \vdash \{(\mathfrak{p}, \mathfrak{q})\} \, \mathtt{f} : \mathbb{I}$) are similar with the rules shown in Fig. 11. Here, we require that verifying the instruction $\mathtt{jmp}$ and $\mathtt{call}$ will consume a token, shown as $\blacklozenge(1)$, in order to avoid infinite loops and recursive function calls. The $\mathsf{wf}(\mathfrak{p}_r)$, whose definition is omitted here, means there is no sub-term in form of $(\mathtt{t} \rightsquigarrow_c \mathcal{K})$, $(\hat{\mathfrak{rn}} \rightarrowtail v)$ and $(\!|A|\!)$ in frame $\mathfrak{p}_r$, because they are not separated by separation conjunction $*$. The **ABSCSQ** rule allows us to execute the high-level primitive command described in precondi-

tion. The implication $\mathfrak{p} \Rightarrow \mathfrak{p}'$ is defined below formally:

$$(\mathfrak{p} \Rightarrow \mathfrak{p}') \vee$$
$$(\forall S, \mathbb{S}, A, w. \, ((S, \mathbb{S}, A, w) \models \mathfrak{p}) \Longrightarrow$$
$$((\exists \mathbb{S}', A', w'. \, ((A, \mathbb{S}) \dashrightarrow (A', \mathbb{S}')) \land ((S, \mathbb{S}', A', w') \models \mathfrak{p}')))$$

The inference rules for verifying instructions are not presented here, because they have no difference with the rules shown in Fig. 11.

Well-defined Specification. The $\mathsf{wdSpec}(\mathfrak{fp}, \mathfrak{fq}, \Upsilon)$ defined formally in Def. 2. It contains three properies that the specifications need to satisfy, and we explain them in turn in the following.

Definition 2 (Well-defined Specification). $\mathsf{wdSpec}(\mathfrak{fp}, \mathfrak{fq}, \Upsilon)$ holds, iff

1 for any $\overline{v}, \mathbb{S}, \mathbb{S}', \mathbb{S}_r$. if $\Upsilon(\overline{v})(\mathbb{S})(\mathbb{S}')$, and $\mathbb{S} \perp \mathbb{S}_r$, then the following holds :

- $\mathbb{S}'.\mathcal{K}.\mathtt{pc} = \mathtt{f} + 8$, $\mathbb{S}'.\mathcal{K}.\mathtt{npc} = \mathtt{f} + 12$ (where $\mathbb{S}'.\mathcal{K}.\mathbb{Q}.\mathbb{R}(\mathbf{r}_{15}) = \mathtt{f}$);

- there exists $\mathbb{S}'', \mathbb{S}'_r$, $\Upsilon(\overline{v})(\mathbb{S} \uplus \mathbb{S}_r)(\mathbb{S}'')$, $\mathbb{S}'' = \mathbb{S}' \uplus \mathbb{S}'_r$, and $\mathbb{S}_r.T = \mathbb{S}'_r.T$, $\mathbb{S}_r.M = \mathbb{S}'_r.M$;

2 for any ι, there exists $\overline{v}$, $\mathfrak{fp} \, \iota \Longrightarrow (\!|\Upsilon(\overline{v})|\!) * \mathsf{true}$, and $\mathfrak{fq} \, \iota \Longrightarrow (\!|\perp|\!) * \mathsf{true}$;

$\boxed{\Psi \vdash C_{\mathrm{as}} : \Psi}$ **(Well-formed Primitive)**

$$\vdash C_{\mathrm{as}} : \Psi$$
$$\text{for all } \mathtt{f} \in \mathrm{dom}(\Omega):$$
$$\frac{\Psi(\mathtt{f}) = (\mathbb{fp}, \mathbb{fq}) \qquad \Omega(\mathtt{f}) = \Upsilon \qquad \mathsf{wdSpec}(\mathbb{fp}, \mathbb{fq}, \Upsilon)}{\Psi \vdash C_{\mathrm{as}} : \Omega} \text{ (WfPrim)}$$

$\boxed{\vdash C_{\mathrm{as}} : \Psi}$ **(Well-formed Code Heap)**

$$\frac{\text{for all } \mathtt{f} \in \mathrm{dom}(\Psi), \iota: \ \Psi(\mathtt{f}) = (\mathbb{fp}, \mathbb{fq}) \qquad \Psi \vdash \{(\mathbb{fp} \ \iota, \mathbb{fq} \ \iota)\} \mathtt{f} : C_{\mathrm{as}}[\mathtt{f}]}{\vdash C_{\mathrm{as}} : \Psi} \text{ (WfInt)}$$

$\boxed{\Psi \vdash \{(\mathbb{p}, \mathbb{q})\} \mathtt{f} : \mathbb{I}}$ **(Well-formed Instruction Sequences)**

$$\frac{\Vdash \{\mathbb{p}\downarrow\} \mathtt{i} \{\mathbb{p}'\} \qquad \Psi \vdash \{(\mathbb{p}', \mathbb{q})\} \mathtt{f} + 4 : \mathbb{I}}{\Psi \vdash \{(\mathbb{p}, \mathbb{q})\} \mathtt{f} : \mathtt{i}; \mathbb{I}} \text{ (SEQ)}$$

$$\frac{\mathbb{p}\downarrow \Rightarrow (\mathtt{a} =_a \mathtt{f}') \quad \mathtt{f}' \in \mathrm{dom}(\Psi) \quad \Psi(\mathtt{f}') = (\mathbb{fp}, \mathbb{fq})}{\Psi \vdash \{(\mathbb{p}, \mathbb{q})\} \mathtt{f} : \mathtt{jmp} \ \mathtt{a}; \mathtt{i}} \text{ (JMP)}$$
$$\frac{\Vdash \{\mathbb{p}\downarrow\downarrow\} \mathtt{i} \{\mathbb{p}' * \blacklozenge(1)\} \quad \exists \iota, \mathbb{p}_r. \ (\mathbb{p}' \Rightarrow \mathbb{fp} \ \iota * \mathbb{p}_r) \wedge (\mathbb{fq} \ \iota * \mathbb{p}_r \Rightarrow \mathbb{q}) \wedge \mathsf{wf}(\mathbb{p}_r)}{}$$

$$\frac{\begin{array}{c} \mathtt{f}' \in \mathrm{dom}(\Psi) \quad \Psi(\mathtt{f}') = (\mathbb{fp}, \mathbb{fq}) \quad \Psi \vdash \{(\mathbb{p}', \mathbb{q})\} \mathtt{f} + 8 : \mathbb{I} \\ \mathbb{p}\downarrow \Rightarrow (\mathtt{r}_{15} \mapsto \_) * \mathbb{p}_1 \quad \Vdash \{(\mathtt{r}_{15} \mapsto \mathtt{f} * \mathbb{p}_1)\downarrow\} \mathtt{i} \{\mathbb{p}_2 * \blacklozenge(1)\} \\ \exists \iota, \mathbb{p}_r. \ (\mathbb{p}_2 \Rightarrow \mathbb{fp} \ \iota * \mathbb{p}_r) \wedge (\mathbb{fq} \ \iota * \mathbb{p}_r \Rightarrow \mathbb{p}') \wedge (\mathbb{fq} \ \iota \Rightarrow \mathtt{r}_{15} = \mathtt{f}) \wedge \mathsf{wf}(\mathbb{p}_r) \end{array}}{\Psi \vdash \{(\mathbb{p}, \mathbb{q})\} \mathtt{f} : \mathtt{call} \ \mathtt{f}'; \mathtt{i}; \mathbb{I}} \text{ (CALL)}$$

$$\frac{\mathbb{p}\downarrow\downarrow \Rightarrow (\mathtt{r}_{15} \mapsto \mathtt{f}') * \mathbb{p}_1 \quad \Vdash \{\mathbb{p}_1\} \mathtt{i} \{\mathbb{p}_2\} \quad (\mathtt{r}_{15} \mapsto \mathtt{f}') * \mathbb{p}_2 \Rightarrow \mathbb{q}}{\Psi \vdash \{(\mathbb{p}, \mathbb{q})\} \mathtt{f} : \mathtt{retl}; \mathtt{i}} \text{ (RETL)}$$

$$\frac{\mathbb{p} \Rrightarrow \mathbb{p}' \qquad \Psi \vdash \{(\mathbb{p}', \mathbb{q}')\} \mathtt{f} : \mathbb{I} \qquad \mathbb{q}' \Rrightarrow \mathbb{q}}{\Psi \vdash \{(\mathbb{p}, \mathbb{q})\} \mathtt{f} : \mathbb{I}} \text{ (ABSCSQ)}$$

Fig.19. Selected Inference Rules for Refinement Verification

3 for any $\overline{v}, S, \mathbb{S}$, if $(S, \mathbb{S}, \_, \_) \in \mathsf{INV}(\Upsilon(\overline{v}), \overline{v})$, there exists $\iota, \mathbb{p}_r$ and w, such that $(S, \mathbb{S}, \Upsilon(\overline{v}), w) \models (\mathbb{fp} \ \iota * \mathbb{p}_r), (\mathbb{fq} \ \iota * \mathbb{p}_r) \Longrightarrow \mathsf{INV}(\bot, \_)$, and $\mathsf{Sta}(\Upsilon(\overline{v}), \mathbb{p}_r)$ hold.

First, we should give some restrictions for the execution of abstract assembly primitive. The return code pointers should be equal to $\mathtt{f} + 8$ and $\mathtt{f} + 12$, where $\mathtt{f}$ is contained in $\mathtt{r}_{15}$ register after the execution of abstract assembly primitive. This restriction ensures that the low-level function and high-level abstract assembly primtive will return to the same code pointers after executions, because the **RETL** rule in our logic also restricts that the execution of low-level program will return to code pointers according to $\mathtt{r}_{15}$ register in low-level state when function returns. We also restrict that if an abstract assembly primitive can execute safely on a subset of program state, it can also execute safely on the whole program state, and additional program state keeps unchanged. **Second**, the abtract assembly primitive should be specifid in the precondition, and it's execution should be done in the final state. **Third**, there is an *invariant* between low- and high-level program, holding at the entry of the function, and our logic needs to ensure that such invariant can be reestablished at the exit of function. We define INV to represent this invariant below formally:

$$\mathsf{INV}(A, \overline{v}) ::= \{(S, \mathbb{S}, A, w) \mid S \sim \mathbb{S} \wedge \exists \mathbb{S}'. (A, \mathbb{S}) \dashrightarrow^* (\bot, \mathbb{S}') \\ \wedge \ \mathbf{args}(\mathbb{S}.\mathcal{K}.\mathbb{Q}, \mathbb{S}.M, \overline{v})\}$$

The invariant consists of the state relation between low- and high-level program state, shown as $S \sim \mathbb{S}$, and the safe execution of the primitive command ($\dashrightarrow^*$ means zero or one step). Including the safe execution of the primitive command is essential because we can get some knowledges of high-level program state from the safe execution of primitive command A. For example, if $\mathsf{INV}(\mathsf{switch}(\mathrm{nil}), \mathrm{nil})$ holds, we can know that the location $\mathsf{TaskNew}$ must save a pointer pointing to a ready thread in thread pool from the safe execution of primitive switch. And we can know that the memory location $\mathsf{TaskNew}$ in low-level state also saves such pointer according to the state relation between low- and high-level program state.

$$\mathsf{INV}(\mathsf{switch}(\mathrm{nil}), \mathrm{nil}) \Longrightarrow$$
$$\exists\, \mathsf{t}, \mathcal{K}.\,(\mathsf{TaskNew} \rightarrowtail (\mathsf{t}, 0)) * (\mathsf{t} \rightsquigarrow_r \mathcal{K}) * (\mathsf{TaskNew} \mapsto (\mathsf{t}, 0)) * \mathrm{true}$$

We introduce frame $\mathbb{p}_r$ for local reasoning, and it should be stable under the execution of the abstract assembly primitive (shown as $\mathsf{Sta}(\Upsilon(\overline{v}), \mathbb{p}_r)$) defined below.

$$\mathsf{Sta}(\Upsilon(\overline{v}), \mathbb{p}_r) ::= \forall\, S, \mathbb{S}, \mathbb{S}', w.$$
$$(((S, \mathbb{S}, \Upsilon(\overline{v}), w) \models \mathbb{p}_r * \mathrm{true}) \wedge \Upsilon(\overline{v})(\mathbb{S})(\mathbb{S}')) \Longrightarrow$$
$$(S, \mathbb{S}', \bot, w) \models \mathbb{p}_r * \mathrm{true}$$

4.5 Semantics and Soundness

We first define the simulation relation for instruction sequence. It says C_{as} can execute safely from S, $\mathtt{pc}$ and $\mathtt{npc}$ until reaching the end of the current instruction sequence ($C_{\mathrm{as}}[\mathtt{pc}]$), and q holds if $C_{\mathrm{as}}[\mathtt{pc}]$ ends with the return instruction $\mathtt{retl}$, and for each step of low-level execution, the high-level program will execute zero or one step. It is formally defined in Def. 3. Here we use "$\_ \longmapsto^n \_$" to represent n-step execution. The defintion can ensure the *progress* and *preservations* of the execution of the instruction sequence. *Progress* property means the program can execute a step, when meeting simple instructions, or two steps, when meeting the delayed control transfer instructions, *e.g.*$\mathtt{call}$ and $\mathtt{jmp}$. *Preservations* property means that if the program can

execute one or two steps, the remaining part of instruction sequence and high-level primitive can still execute safely. The w in simulation records the number of tokens. The tokens will be consumed when meeting $\mathtt{jmp}$ and $\mathtt{call}$ instructions, so as to avoid infinite loop and recursive function. The execution of high-level abstract assembly primitive can reset the number of tokens.

Definition 3 (Simulation for Instruction Sequence).
$\mathbb{q}; \Psi \models (C_{\mathrm{as}}, S, \mathtt{pc}, \mathtt{npc}) \preccurlyeq_w (A, \mathbb{S})$ holds if and only if the following are true (we omit the case for $\mathtt{be}$ here, which is similar to $\mathtt{jmp}$):

1. if $C_{\mathrm{as}}(\mathtt{pc}) = \mathtt{i}$ then:
   - there exist $S', \mathtt{pc}, \mathtt{npc}'$, such that $C \vdash (S, \mathtt{pc}, \mathtt{npc}) \longmapsto (S', \mathtt{pc}', \mathtt{npc}')$,
   - for any $S', \mathtt{pc}', \mathtt{npc}'$, if $C \vdash (S, \mathtt{pc}, \mathtt{npc}) \longmapsto (S', \mathtt{pc}', \mathtt{npc}')$, then there exists A' and $\mathbb{S}'$, such that:
     (1) either $A' = A$, $\mathbb{S}' = \mathbb{S}$ and $w' = w$; or $(A, \mathbb{S}) \dashrightarrow (A', \mathbb{S}')$,
     (2) $\mathbb{q}; \Psi \models (C_{\mathrm{as}}, S', \mathtt{pc}', \mathtt{npc}') \preccurlyeq_{w'} (A', \mathbb{S}')$.

2. if $C_{\mathrm{as}}(\mathtt{pc}) = \mathtt{jmp\ a}$ then:
   - there exist $S', \mathtt{pc}', \mathtt{npc}'$, such that $C \vdash (S, \mathtt{pc}, \mathtt{npc}) \longmapsto^2 (S', \mathtt{pc}', \mathtt{npc}')$,
   - for any $S', \mathtt{pc}', \mathtt{npc}'$, if $C \vdash (S, \mathtt{pc}, \mathtt{npc}) \longmapsto^2 (S', \mathtt{pc}', \mathtt{npc}')$, then there exists $\mathbb{fp}, \mathbb{fq}, \iota, A', \mathbb{S}', w', w'' < w'$ and $\mathbb{p}_r$ such that the following hold:
     (1) $\mathtt{npc}' = \mathtt{pc}'+4$, $\Psi(\mathtt{pc}') = (\mathbb{fp}, \mathbb{fq})$,
     (2) either $A' = A$, $\mathbb{S}' = \mathbb{S}$ and $w' = w$; or $(A, \mathbb{S}) \dashrightarrow (A', \mathbb{S}')$,
     (3) $(S', \mathbb{S}', A', w'') \models (\mathrm{fp}\ \iota) * \mathbb{p}_r$, $(\mathbb{fq}\ \iota) * \mathbb{p}_r \Rightarrow \mathbb{q}$, $\mathsf{wf}(\mathbb{p}_r)$.

3. if $C_{\mathrm{as}}(\mathtt{pc}) = \mathtt{be\ f}$ then $\ldots$

4. if $C_{\mathrm{as}}(\mathtt{pc}) = \mathtt{call\ f}$ then:
   - there exist $S', \mathtt{pc}', \mathtt{npc}'$, such that $C \vdash (S, \mathtt{pc}, \mathtt{npc}) \longmapsto^2 (S', \mathtt{pc}', \mathtt{npc}')$,
   - for any $S', \mathtt{pc}'$ and $\mathtt{npc}'$, if $C \vdash (S, \mathtt{pc}, \mathtt{npc}) \longmapsto^2 (S', \mathtt{pc}', \mathtt{npc}')$, then there exist $\mathbb{fp}, \mathbb{fq}, \iota, A', \mathbb{S}', w', w'' < w'$ and $\mathbb{p}_r$, such that the following hold:
     (1) $\mathtt{npc}' = \mathtt{pc}'+4$, $\Psi(\mathtt{pc}') = (\mathbb{fp}, \mathbb{fq})$,
     (2) either $A' = A$, $\mathbb{S}' = \mathbb{S}$ and $w' = w$; or $(A, \mathbb{S}) \dashrightarrow (A', \mathbb{S}')$,
     (3) $(S', \mathbb{S}', A', w'') \models (\mathbb{fp}\ \iota) * \mathbb{p}_r$, $\mathsf{wf}(\mathbb{p}_r)$,
     (4) for any $S', \mathbb{S}', A', w'$, if $(S', \mathbb{S}', A', w') \models (\mathbb{fq}\ \iota) * \mathbb{p}_r$, then $\mathbb{q}; \Psi \models (C_{\mathrm{as}}, S', \mathtt{pc}+8, \mathtt{npc}+12) \preccurlyeq_{w'} (A', \mathbb{S}')$,

    (5) $(\mathbb{f}\mathbb{q}\,\iota)*\mathbb{p}_r \Rightarrow (\mathbf{r}_{15} = \mathbf{pc})$.

5. if $C_{\mathrm{as}}(\mathbf{pc}) = \mathtt{retl}$ then :

- there exist $S', \mathbf{pc}', \mathbf{npc}'$, such that
$C \vdash (S, \mathbf{pc}, \mathbf{npc}) \longmapsto^2 (S', \mathbf{pc}', \mathbf{npc}')$,

- for any $S', \mathbf{pc}'$ and $\mathbf{npc}'$,
if $C \vdash (S, \mathbf{pc}, \mathbf{npc}) \longmapsto^2 (S', \mathbf{pc}', \mathbf{npc}')$, then there exists $A', \mathbb{S}', w'$, such that:

    (1) either $A' = A$, $\mathbb{S}' = \mathbb{S}$ and $w' = w$; or $(A, \mathbb{S}) \dashrightarrow (A', \mathbb{S}')$,

    (2) $(S', \mathbb{S}', A', w') \models \mathbb{q}$, $\mathbf{pc}' = S'.Q.R(\mathbf{r}_{15}){+}8$, and $\mathbf{npc}' = S'.Q.R(\mathbf{r}_{15}){+}12$.

Then we can define the semantics for well-formed instruction sequences and well-formed internal function. The semantics of well-formed instruction sequences tells us that for any C_{as}, if the instruction sequence starting from label $\mathbf{f}$ is $\mathbb{I}$, then the instruction sequence $\mathbb{I}$ can execute safely if the relational state satisfies the precondition $\mathbb{p}$. And the semantics of well-formed code heap says that all the instruction sequences in code heap C_{as} given specifications can execute safely.

Definition 4 (Judgment Semantics).

- $\Psi \models \{(\mathbb{p}, \mathbb{q})\}\, \mathbf{f} : \mathbb{I}$ if and only if, for all $C_{\mathrm{as}}, S, \mathbb{S}, A$ and w such that $C_{\mathrm{as}}[\mathbf{f}] = \mathbb{I}$ and $(S, \mathbb{S}, A, w) \models \mathbb{p}$, we have $\mathbb{q}; \Psi \models (C_{\mathrm{as}}, S, \mathbf{f}, \mathbf{f}{+}4) \preccurlyeq_w (A, \mathbb{S})$.

- $\models C_{\mathrm{as}} : \Psi$ if and only if, for all $\mathbf{f}$, $\mathbb{f}\mathbb{p}$ and $\mathbb{f}\mathbb{q}$ such that $\Psi(\mathbf{f}) = (\mathbb{f}\mathbb{p}, \mathbb{f}\mathbb{q})$, we have $\Psi \models \{(\mathbb{f}\mathbb{p}\,\iota, \mathbb{f}\mathbb{q}\,\iota)\}\, \mathbf{f} : C_{\mathrm{as}}[\mathbf{f}]$ for all ι.

Then we define the simulation for function in Def. 5. It means that if there exists a relational state $(S, \mathbb{S}, A, w)$ satisfies the precondition $\mathbb{p}$, then we have the simulation $\mathbb{q} \models (C_{\mathrm{as}}, S, \mathbf{f}, \mathbf{f} + 4) \preccurlyeq_i^0 (A, \mathbb{S})$ defined in Def. 6.

Definition 5 (Simulation for Function).

$(C_{\mathrm{as}}, \mathbf{f}) \preccurlyeq^{(\mathbb{p}, \mathbb{q})} A ::= \forall S, \mathbb{S}, w.\, (S, \mathbb{S}, A, w) \models \mathbb{p} \Longrightarrow$
$\qquad\qquad \exists i \in \mathit{Index}.\, \mathbb{q} \models (C_{\mathrm{as}}, S, \mathbf{f}, \mathbf{f} + 4) \preccurlyeq_i^0 (A, \mathbb{S})$
   where $\mathbb{q} \models (C_{\mathrm{as}}, S, \mathbf{pc}, \mathbf{npc}) \preccurlyeq_i^k (A, \mathbb{S})$ is defined in Def. 6.

Definition 6. Whenever $\mathbb{q} \models (C_{\mathrm{as}}, S, \mathbf{pc}, \mathbf{npc}) \preccurlyeq_i^k (A, \mathbb{S})$ holds, we have the following holds :

1. if $C_{\mathrm{as}}(\mathbf{pc}) = \mathtt{i}$, then:

- there exists $S', \mathbf{pc}', \mathbf{npc}'$, such that
$(C_{\mathrm{as}}, S, \mathbf{pc}, \mathbf{npc}) ::\!\stackrel{\tau}{\Longrightarrow} (C_{\mathrm{as}}, S', \mathbf{pc}', \mathbf{npc}')$;

- for any $S', \mathbf{pc}', \mathbf{npc}'$,
if $(C_{\mathrm{as}}, S, \mathbf{pc}, \mathbf{npc}) ::\!\stackrel{\tau}{\Longrightarrow} (C_{\mathrm{as}}, S', \mathbf{pc}', \mathbf{npc}')$, then one of the following holds:

    (a) $\exists j < i.\, \mathbb{q} \models (C_{\mathrm{as}}, S', \mathbf{pc}', \mathbf{npc}') \preccurlyeq_j^k (A, \mathbb{S})$;

    (b) there exists $\mathbb{S}', j \in \mathit{Index}$, such that $(A, \mathbb{S}) \dashrightarrow (\bot, \mathbb{S}')$ and $\mathbb{q} \models (C_{\mathrm{as}}, S', \mathbf{pc}', \mathbf{npc}') \preccurlyeq_j^k (\bot, \mathbb{S}')$ holds;

2. if $C_{\mathrm{as}}(\mathbf{pc}) = \mathtt{call\ f}$, then:

- there exists $S', \mathbf{pc}', \mathbf{npc}'$, such that
$(C_{\mathrm{as}}, S, \mathbf{pc}, \mathbf{npc}) ::\!\stackrel{\tau}{\Longrightarrow}^2 (C_{\mathrm{as}}, S', \mathbf{pc}', \mathbf{npc}')$;

- for any $S', \mathbf{pc}', \mathbf{npc}'$,
if $(C_{\mathrm{as}}, S, \mathbf{pc}, \mathbf{npc}) ::\!\stackrel{\tau}{\Longrightarrow}^2 (C_{\mathrm{as}}, S', \mathbf{pc}', \mathbf{npc}')$, then one of the following holds:

    (a) $\exists j < i.\, \mathbb{q} \models (C_{\mathrm{as}}, S', \mathbf{pc}', \mathbf{npc}') \preccurlyeq_j^{k+1} (A, \mathbb{S})$;

    (b) there exists $\mathbb{S}', j \in \mathit{Index}$, such that $(A, \mathbb{S}) \dashrightarrow (\bot, \mathbb{S}')$ and $\mathbb{q} \models (C_{\mathrm{as}}, S', \mathbf{pc}', \mathbf{npc}') \preccurlyeq_j^{k+1} (\bot, \mathbb{S}')$ holds;

3. if $C_{\mathrm{as}}(\mathbf{pc}) = \mathtt{retl}$, then:

- there exists $S', \mathbf{pc}', \mathbf{npc}'$, such that
$(C_{\mathrm{as}}, S, \mathbf{pc}, \mathbf{npc}) ::\!\stackrel{\tau}{\Longrightarrow}^2 (C_{\mathrm{as}}, S', \mathbf{pc}', \mathbf{npc}')$;

- for any $S', \mathbf{pc}', \mathbf{npc}'$,
if $(C_{\mathrm{as}}, S, \mathbf{pc}, \mathbf{npc}) ::\!\stackrel{\tau}{\Longrightarrow}^2 (C_{\mathrm{as}}, S', \mathbf{pc}', \mathbf{npc}')$ then there exists $j \in \mathit{Index}$, $\mathbb{S}'$ and A', such that the following holds:

    (a) either $j < i$, $\mathbb{S}' = \mathbb{S}$ and $A' = A$; or $(A, \mathbb{S}) \dashrightarrow (A', \mathbb{S}')$;

    (b) if $k = 0$, then there exists w': (where $S'.Q.R(\mathbf{r}_{15}) = \mathbf{f}$)
        $(S', \mathbb{S}', A', w') \models \mathbb{q}$, $A' = \bot$,
        $\mathbf{pc}' = \mathbf{f}{+}8$, and $\mathbf{npc}' = \mathbf{f}{+}12$;
    else
        $\mathbb{q} \models (C_{\mathrm{as}}, S', \mathbf{pc}', \mathbf{npc}') \preccurlyeq_j^{k-1} (A', \mathbb{S}')$.

The definition of simulation $\mathbb{q} \models (C_{\mathrm{as}}, S, \mathbf{pc}, \mathbf{npc}) \preccurlyeq_i^k (A, \mathbb{S})$ carries an index i, which is used to ensure the termination preserving, and the depth k of function call, which increases by the $\mathtt{call}$ instruction and decreases by $\mathtt{retl}$ (unless $k = 0$). The simulation relation make sure the safe execution of low-level SPARCv8 function and the corresponding high-level abstract assembly primitive. The we give the semantics of well-formed primitive in Def. 7.

Definition 7 (Well-defined Primitive Set Semantics).

$$\Psi \models C_{as} : \Omega ::= \forall f \in \text{dom}(\Omega), \iota. \; \exists \Upsilon, \overline{v}, fp, fq.$$
$$\text{wdSpec}(fp, fq, \Upsilon) \wedge (fp \; \iota \Longrightarrow (\!| \Upsilon(\overline{v}) |\!) * \text{true})$$
$$\wedge (C_{as}, f) \preccurlyeq^{(fp \; \iota, \; fq \; \iota)} \Upsilon(\overline{v})$$
$$\text{where } \Omega(f) = \Upsilon, \Psi(f) = (fp, fq).$$

It says that any high-level abstract assembly primitive in primitive set Ω, we can establish a simulation relation defined in Def. 5 between its low-level implementation in code heap C_{as} and the high-level abstract assembly primitive. Theorem 1, whose correctness can be derived from Lemmas 3 and 4, shows the soundness of our logic, which means that the extended program logic can imply the contextual refinement between implementation C_{as} and abstract assembly primitives Ω.

Lemma 3 (Logic Ensures Simulation).

- $\vdash C_{as} : \Psi \Longrightarrow \models C_{as} : \Psi$

- $\Psi \vdash C_{as} : \Omega \Longrightarrow \Psi \models C_{as} : \Omega$

Lemma 4 (Simulation Implies Primitive Correctness).

$$\Psi \models C_{as} : \Omega \Longrightarrow C_{as} \sqsubseteq \Omega$$

Theorem 1 (Logic Soundness).

$$\Psi \vdash C_{as} : \Omega \Longrightarrow C_{as} \sqsubseteq \Omega$$

5 Verifying Context Switch Routine

We apply our relational program logic to verify that a context switch routine implemented in SPARCv8, which is used to save the current task's context and restore the new task's context, have contextual refinement with the switch primitive defined in Sec. 4.
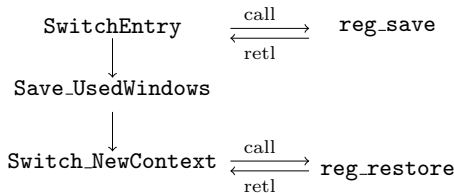
Fig. 20 shows the structure of the code.

Fig.20. The Structure of Context Switch Routine

- `SwitchEntry` is the entry of the context switch routine. It saves local and in registers of current window into stack (in memory), and call `reg_save` to save other registers into TCB.

- `Save_UsedWindows` saves the register windows (except the current one) into the current task's stack in memory. It checks whether the previous window is valid. If it's valid, use the instruction `restore` to set the previous window as the current one, and save its contents into stack (in memory), then check the previous one continuously.

- `Switch_NewContext` restores the general registers from the new task's TCB (by calling `reg_restore`) and its stack in memory, respectively. Then it sets the new task as the current one.

The main complexity of the verification lies in the code manages the register windows. To save all the register windows, `Save_UsedWindows` repetitively restores the next window into general registers (as the current window) and then saves them into memory, until all the windows are saved.

Specification. Below we give the pre- and post-conditions (a_{pre} and a_{post}) of the verified module. Each of them takes 6 arguments, the id of the current task t_c, the id of the new task t_n, the values env of general registers and all other register windows, the new task's context nst that needs to be restored, and the thread local state of current task $\mathcal{K}_c$ and the new task $\mathcal{K}_n$.

$$a_{pre}(t_c, t_n, env, nst, \mathcal{K}_c, \mathcal{K}_n) ::=$$
$$\text{Env}(env) * (\text{TaskNew} \mapsto (t_n, 0)) * \blacklozenge(10) *$$
$$\text{CurT}(t_c, \_, env, \mathcal{K}_c) * \text{RdyT}(t_n, nst, \mathcal{K}_n) * (\!| \text{switch}(\text{nil}) |\!)$$

$$a_{post}(t_c, t_n, env, nst, \mathcal{K}_c, \mathcal{K}_n) ::=$$
$$\exists env', \mathcal{K}'. \text{Env}(env') * (\text{TaskNew} \mapsto (t_n, 0)) *$$
$$(\text{CurT}(t_n, nst, env', \mathcal{K}') \wedge \text{p\_env}(env') = nst) *$$
$$\text{RdyT}(t_c, \text{p\_env}(env), \mathcal{K}_c) * (\!| \perp |\!)$$

In the specification, we use $\text{Env}(env)$ to specify the values of general registers and the register windows. We

describe the state of the current task (its TCB and stack in memory) using $\mathsf{CurT}(\mathsf{t}_c, \_, env, \mathsf{t}_c)$. It describes the memory of current task's TCB and stack for saving contexts in low-level, the thread local state $\mathcal{K}_n$ in high-level, and state relation of t_c in low- and high-level program. Similarly, the $\mathsf{RdyT}(\mathsf{t}_n, nst, \mathcal{K}_n)$, describes states of new task t_n in low- and high-level program state and their relation. The memory location $\mathsf{TaskNew}$ records the thread id of the new task. Here, we use $\mathsf{TaskNew} \Mapsto (\mathsf{t}_n, 0)$ to present $\mathsf{TaskNew}$ in both low- and high-level memory saving pointer $(\mathsf{t}_n, 0)$.

$$l \Mapsto v \ ::= (l \mapsto v) * (l \rightarrowtail v)$$

The precondition takes 10 tokens ($\blacklozenge(10)$). As we have explained, verifying instruction `call` and `jmp` will consume a token. So, verifying calling `reg_save` and `reg_restore` will both consume a token. And `Save_UsedWindows`, which checks each previous window and saves its contexts into memory until the invalid one, will execute at most eight times, because upper bound of the number of windows is eight. So, 10 tokens is sufficient.

If we compare a_{pre} and a_{post}, we can see that t_n becomes the current task ($\mathsf{CurT}(\mathsf{t}_n, nst, env', \mathcal{K}')$), and its general registers and stack, specified by $\mathsf{Env}(env')$, are loaded from the saved context nst (*i.e.* $\mathsf{p\_env}(env') = nst$). Here $\mathsf{p\_env}(env')$ refers to the part of the environment that we want to save or restore as context. Correspondingly, t_c becomes non-current-thread, and part of its environment env at the entry of the context switch is saved, as specified $\mathsf{RdyT}(\mathsf{t}_c, \mathsf{p\_env}(env), \mathcal{K}_c)$. The execution of `switch` should be done in the final state. We use $\mathcal{K}'$ to represent the thread local state of t_n instead of $\mathcal{K}_n$ in final state, because the execution of `switch` will modify the code pointer in $\mathcal{K}_n$. More details about verifying context switch routine can be found in appendix D.

6   Related Work and Conclusion

There has been much work on assembly or machine code verification. Most of them do not support function calls or simply treat function calls in the continuation-passing style where return addresses are viewed as first class code pointers [14, 15, 16, 17, 18, 19, 20]. SCAP [8] supports assembly code verification with various stack-based control abstractions, including function call and return. We follow the same idea here. However, SCAP gives a syntactic-based soundness proof by establishing the preservation of the syntactic judgment, which makes it difficult to interact with other modules verified in different logic. Since our goal is to verify inline assembly and link the verified code with the verified C programs, we give a direct-style semantic model of the logic judgments. And it allows us to extend our program logic to support verifying contextual refinement without meet much challenges. Also SCAP is based on a simplified subset of assembly instructions, while our work is focused on a realistically modeled subset of SPARCv8 instructions.

In terms of the support of realistic instruction sets, previous work on proof-carrying code (PCC) and typed assembly language (TAL) mostly supports subsets of x86. Myreen's work [21] presents a framework for ARM verification based on a realistic model (but it doesn't support function call and return).

As part of the Foundational Proof-Carrying Code (FPCC) project [15], Tan and Appel present a program logic $\mathcal{L}_c$ for reasoning about control flow in assembly code [20]. Although $\mathcal{L}_c$ is implemented on top of SPARC machine language, the underlying logic is a type system instead of a full-blown program logic for functional correctness. It reasons about functions in the continuation-passing style. Also handling SPARC features such as delayed writes or delayed control transfers

is not the focus of $\mathcal{L}_c$. There has been work on mechanized semantics of the SPARCv8 ISA. Hou *et al.* [22] model the SPARCv8 ISA in Isabelle/HOL, and test their formal model against LENON3 simulation board [23], which is a synthesisable VHDL model of a 32-bit processor compliant with the SPARCv8 architecture, through more than 100,000 instruction instances. Wang *et al.* [11] formalize its semantics in Coq. Our operational semantics of SPARCv8 follows Wang *et al.* [11]. But Wang *et al.* do not validate their formalization against actual hardware, we remain it as a future work.

Ni *et al.* [24] verify a context switch module of 19 lines in x86 code to show case the support of embedded code pointers (ECP) in XCAP [19]. We use our extended program logic to verify the contextual refinement between a context switch routine in SPARCv8 and switch primitive. The context switch routine implemented in SPARCv8 is more complicated then implemented in x86, because of the requirement to save the contexts stored in register window in memory.

Yang and Hawblitzel [25] verify Verve, an x86 implementation of an experimental operating system. Verve has two levels, the high-level TAL code and the low-level "Nucleus" that provides primitive access to hardware and memory. The Nucleus code is verified automatically using the Z3 SMT solver, while the goal of our work is to generate machine checkable proofs. Another key difference is the use of different ISAs. Here we give details to verify specific features of SPARCv8 programs.

There have been many techniques and tools proposed for automated program verification (*e.g.* [26, 27]). It is possible to adapt them to verify SPARCv8 code. We propose a new program logic and do the verification in Coq mainly because the work is part of a big project for a fully certified OS kernel for aerospace crafts whose inline assembly is written in SPARCv8.

We already have a program logic implemented in Coq for C programs, which allows us to verify C code with Coq proofs. Therefore we want to have a program logic for SPARCv8 so that it can be linked with the logic for C and can generate machine-checkable Coq proofs too. That said, many of the automated verification techniques can be applied to reduce the manual efforts to write Coq proofs, which we would like to study in the future work.

Conclusion. We present a program logic for SPARCv8. Our logic is based on a realistic semantics model and supports main features of SPARCv8, including delayed control transfer, delayed writes, and register windows. And we also extend the program logic to support refinement verification and apply the extended program logic to verify that there is a contextual refinement between a context switch routine implementated in SPARCv8 and switch primitive for task switching. Our current work can only handle sequential SPARCv8 program verification and do not consider interrupt in machine model. We will extend it for concurrency verification and finish the step ① shown in Fig. 2 that the compilant can ensure the behaviors of the Pseudo-SPARCv8 code calling abstraction assembly primitives in intermediate level refines the behaviors of the client C code calling abstract assembly primitives in source level in the future.

References

[1] Linux v2.6.17.10. `https://elixir.bootlin.com/linux/v2.6.17.10/source/arch`.

[2] Fengwei Xu, Ming Fu, Xinyu Feng, Xiaoran Zhang, Hui Zhang, and Zhaohui Li. A practical verification framework for preemptive os kernels. In *CAV*, pages 59–79, July 2016.

[3] Gerwin Klein, Kevin Elphinstone, Gernot Heiser, June Andronick, David Cock, Philip Derrin, Dhammika Elkaduwe, Kai Engelhardt, Rafal Kolanski, Michael Norrish, Thomas Sewell, Harvey Tuch, and Simon Winwood. seL4: Formal

Verification of an OS Kernel. In *SOSP*, pages 207–220, Oct 2009.

[4] Ronghui Gu, Jérémie Koenig, Tahina Ramananandro, Zhong Shao, Xiongnan (Newman) Wu, Shu-Chun Weng, Haozhong Zhang, and Yu Guo. Deep specifications and certified abstraction layers. In *POPL*, pages 595–608, Jan 2015.

[5] SPARC. `https://gaisler.com/doc/sparcv8.pdf`.

[6] Gordon Stewart, Lennart Beringer, Santiago Cuellar, and Andrew W. Appel. Compositional compcert. In *POPL*, pages 275–287, 2015.

[7] Junpeng Zha, Xinyu Feng, and Lei Qiao. Modular Verification of SPARCv8 Code. In *APLAS*, pages 245–263, December 2018.

[8] Xinyu Feng, Zhong Shao, Alexander Vaynberg, Sen Xiang, and Zhaozhong Ni. Modular Verification of Assembly Code with Stack-Based Control Abstractions. In *PLDI*, June 2006.

[9] Program logic for SPARCv8 implementation in Coq (project code). `https://github.com/jpzha/VeriSparc`.

[10] Xavier Leroy and Sandrine Blazy. Formal verification of a C-like memory model and its uses for verifying program transformations. *Journal of Automated Reasoning*, 41(1):1–31, July 2008.

[11] Jiawei Wang, Ming Fu, Lei Qiao, and Xinyu Feng. Formalizing SPARCv8 Instruction Set Architecture in Coq. In *SETTA*, Oct 2017.

[12] J.C. Reynolds. Separation logic: a logic for shared mutable data structures. July 2002.

[13] Hongjin Liang, Xinyu Feng, and Zhong Shao. Compositional verification of termination-preserving refinement of concurrent programs. In *LICS*, July 2014.

[14] George C. Necula and Peter Lee. Safe Kernel Extensions Without Run-Time Checking. In *Proc.2nd USENIX Symp. on Operating System Design and Impl*, pages 229–243, 1996.

[15] Andrew W. Appel. Foundational proof-carrying code. In *Proc. 16th Annual IEEE Symposium on Logic in Computer Science*, pages 85–97, Jan 1998.

[16] Greg Morrisett, Karl Crary, Neal Glew, Dan Grossman, Richard Samuels, Frederick Smith, David Walker, Stephanie Weirich, and Steve Zdancewic. Talx86:a realistic typed assembly language. In *1999 ACM SIGPLAN Workshop on Compiler Support for System Software*, pages 25–35, May 1996.

[17] Greg Morrisett, David Walker, Karl Crary, and Neal Glew. From System F to typed assembly language. In *POPL*, pages 85–97, Jan 1998.

[18] Dachuan Yu, A. Hamid Nadeem, and Zhong Shao. Building certified libraries for PCC : Dynamic storage allocation. *Science of Computer Programming*, 50(1-3):101–127, Mar 2004.

[19] Zhaozhong Ni and Zhong Shao. Certified Assembly Programming with Embedded Code Pointers. In *POPL*, pages 320–333, 2006.

[20] Gang Tan and Andrew W. Appel. A compositional logic for control flow. In *VMCAI*, Jan 2006.

[21] Magnus O. Myreen and Michael J.C. Gordon. Hoare logic for realistically modelled machine code. In *Proc. 13th International Conference on Tools and Algorithms for Construction and Analysis of Systems*, 2007.

[22] Hou Zhe, David Sanan, Alwen Tiu, Yang Liu, and Koh Chuen Hoa. An Executable Formalisation of the SPARCv8 Instruction Set Architecture: A Case Study for the LEON3 Processor. In *FM*, 2016.

[23] Leon3. `https://www.gaisler.com/index.php/products/processors/leon3`.

[24] Zhaozhong Ni, Dachuan Yu, and Zhong Shao. Using XCAP to Certify Realistic Systems code: Machine context management. In *TPHOLs*, Sept 2007.

[25] Jean Yang and Chris Hawblitzel. Safe to the last instruction: automated verification of a type-safe operating system. In *PLDI*, pages 99–110, 2010.

[26] Josh Berdine, Cristiano Calcagno, and PeterW. O'Hearn. Symbolic execution with separation logic. In *APLAS*, 2005.

[27] Josh Berdine, Cristiano Calcagno, and PeterW. O'Hearn. Smallfoot: Modular automatic assertion checking with separation logic. In *FMCO*, 2005.

A   More about High-level Instructions Execution

We give some supplements about the execution of high-level instructions. As we have explained in Sec. 4.1, the register windows and delayed buffer in pyhsical SPARCv8 program state are omitted in high-level Pseudo-SPARCv8 program state. So, we do not define state transtion rules for instructions `save`, `restore`, `rd`, and `wr`. The instruction transition rules for the rest of instructions, like `ld` and `add`, have no much difference with the rules

in pyhical SPARCv8 program.

$$\frac{[\![a]\!]_{\mathbb{R}} = l \qquad M(l) = v \qquad \mathbb{R}' = \mathbb{R}\{\!|\, \mathbf{r}_d \rightsquigarrow v\, |\!\}}{\mathbf{execi}(\mathtt{ld}\ \mathtt{a}\ \mathbf{r}_d, ((\mathbb{R}, \mathbb{F}), M)) =_{\mathsf{H}} ((\mathbb{R}', \mathbb{F}), M)}$$

$$\frac{\mathbb{R}(\mathbf{r}_s) = v_1 \qquad [\![o]\!]_{\mathbb{R}} = v_2 \qquad \mathbf{r}_d = \mathrm{dom}(\mathbb{R}) \qquad \mathbb{R}' = \mathbb{R}\{\!|\, \mathbf{r}_d \rightsquigarrow v\, |\!\}}{\mathbf{execi}(\mathtt{add}\ \mathbf{r}_s, \mathtt{o}, \mathbf{r}_d, ((\mathbb{R}, \mathbb{F}), M)) =_{\mathsf{H}} ((\mathbb{R}', \mathbb{F}), M)}$$

Fig.A1. Transition rules for instructions `ld` and `add` in high-level

We show the state transition rules for instructions `ld` and `add` in high-level in Fig. A1. The register file updating operation is defined formally below :

$$\mathbb{R}\{\!|\, \hat{\mathbf{rn}} \rightsquigarrow v\, |\!\} \ ::= \ \mathbb{R}\{\hat{\mathbf{rn}} \rightsquigarrow v\} \qquad \text{where } \hat{\mathbf{rn}} \notin \{\%\mathtt{sp}, \%\mathtt{fp}\}$$

According to the definition, we can find that updating the register $\%\mathtt{sp}$ (alias of $\mathbf{r}_{14}$), which is used to point to the top of the current stack frame, and $\%\mathtt{fp}$ (alias of $\mathbf{r}_{14}$), which is used to point to the top of the previous stack frame is not allowed. Only the execution of instructions `Psave`, which is used to allocate a new stack frame, and `Prestore`, which is used to free the current stack frame, can modify them. The evaluation of the opand and address expression in high-level is defined formally below :

$$[\![o]\!]_{\mathbb{R}} ::= \begin{cases} R(r) & \text{if } \mathtt{o} = r \\ w & \text{if } \mathtt{o} = w, \\ & \quad -4096 \le w \le 4095 \\ \bot & \text{otherwise} \end{cases} \qquad [\![a]\!]_{\mathbb{R}} ::= \begin{cases} [\![o]\!]_{\mathbb{R}} & \text{if } \mathtt{a} = \mathtt{o} \\ v_1 + v_2 & \text{if } \mathtt{a} = \mathtt{r} + \mathtt{o},\ \mathbb{R}(\mathtt{r}) = v_1 \\ & \quad \text{and } [\![o]\!]_{\mathbb{R}} = v_2 \\ \bot & \text{otherwise} \end{cases}$$

The " `Psave` w " can be viewed as a macro of " `save` $\%\mathtt{sp}, -w, \%\mathtt{sp}$ ", and " `Prestore` " can be viewed as a marco of " `restore` $\%\mathtt{g}_0, \%\mathtt{g}_0, \%\mathtt{g}_0$ ".

save	$\%\mathtt{sp}, -128, \%\mathtt{sp}$		Psave	128	
add	$\%\mathtt{i}_0, \%\mathtt{i}_1, \%\mathtt{i}_0$		add	$\%\mathtt{i}_0, \%\mathtt{i}_1, \%\mathtt{i}_0$	
ret			ret		
restore	$\%\mathtt{g}_0, \%\mathtt{g}_0, \%\mathtt{g}_0$		Prestore		
	(a)			(b)	

Fig.A2. Realistic SPARCv8 Code and Pseudo-SPARCv8 Code

Fig. A2 gives a simple comparision with the realistic SPARCv8 code and our Pseudo SPARCv8 code in high-level. Fig. A2 (a) is the realistic SPARCv8 code. It uses instruction " `save` $\%\mathtt{sp}, -128, \%\mathtt{sp}$ " to store the caller's context and allocate a new stack frame size 128 bytes for the current procedure, and use instruction " `restore` $\%\mathtt{g}_0, \%\mathtt{g}_0, \%\mathtt{g}_0$ " to restore the caller's context at the exitance of the current procedure. Fig. A2 (b) is the same function in Pseudo-SPARCv8 code, and we can find that the instructions that is responsible for saving and restoring the context of caller is replaced by " `Psave` 128 " and " `Prestore` ".

B   More about Low-level Language

The machine states and syntax low-level SPARCv8 language (defined in Fig. A3) are taken from the model of SPARCv8 defined in Fig. 3. So, we omit some definitions, like RegName and DelayCycle, which are same as ones defined in Fig. 3 here.

$$
\begin{array}{llll}
\text{(LProg)} & \text{P} & ::= (C, S, \texttt{pc}, \texttt{npc}) & \qquad \text{(LState)} \quad S \ ::= (M, Q, D) \\
\text{(LRstate)} & Q & ::= (R, F) & \qquad \text{(LRegFile)} \quad R \ ::= \text{RegName} \rightharpoonup \text{Val} \\
\text{(LFrmList)} & F & ::= \text{nil} \mid \text{fm} :: F & \qquad \text{(LFrame)} \quad \text{fm} \ ::= [v_0, \ldots, v_7] \\
\text{(DBuf)} & D & ::= \text{nil} \mid (t, X, v) & \qquad \text{(LMsg)} \quad \beta \ ::= \tau \mid \textsf{out}(v)
\end{array}
$$

Fig.A3. Machine States and Syntax for Low-level SPARCv8 Language

The low-level program P is a tuple including the code heap C, low-level program state S, program counter $\texttt{pc}$ and $\texttt{npc}$. The code heap C is defined in Fig. 14. The low-level program state S uses the block-based memory model M, which is the same as the high-level program. The low-level message does not need $\textsf{call}(\texttt{f}, \overline{v})$, because the low-level program does not call abstract assembly primitive, but call its corresponding implementation, which is a function.

Operational Semantics for Low-level Code.   The operational semantics for low-level program is defined in Fig. A4. Most of the state transition rules are taken from Fig. 7. Here, we use "$\textbf{execi}(\texttt{i}, \_) =_{\textsf{L}} \_$" to represent the step for simple instruction $\texttt{i}$.

The execution of instruction Psave is discussed in divided into two cases : (1) if we can successfully set the next register window as the current one (represented as $\textbf{save}(R, F) = (R', F')$), a new stack frame in memory will be allocated (shown as $\textbf{\textit{alloc}}(M, b, 0, w) = M'$); (2) if we can't set the next register window as the current one (represented as $\textbf{save}(R, F) = \textbf{undefined}$), a windows overflow trap will be triggered, and we redo the instruction Psave. The execution of instruction Prestore does the reverse. Here, we use "$\_ \Uparrow \_$" to represent the state transition caused by window overflow trap and use "$\_ \Downarrow \_$" to represent the state transition caused by window underflow trap. Their formal definitions are shown in Fig. A6.

$$(R, F) \rightrightarrows (R', F')$$
$$\frac{C \vdash ((M, (R', F), D'), \mathtt{pc}, \mathtt{npc}) \circ\!\xrightarrow{\beta} (M', (R'', F'), D'')}{(C, (M, (R, F), D), \mathtt{pc}, \mathtt{npc}) ::\!\!\Longrightarrow (C, (M, (R'', F'), D''))}$$

(a) Low-level Program Transition

$$\frac{C(\mathtt{pc}) = \mathtt{i} \qquad \mathbf{execi}(\mathtt{i}, (M, Q, D)) =_{\mathsf{L}} (M', Q', D')}{C \vdash ((M, Q, D), \mathtt{pc}, \mathtt{npc}) \circ\!\xrightarrow{\tau} ((M', Q', D'), \mathtt{npc}, \mathtt{npc} + 4)}$$

$$\frac{C(\mathtt{pc}) = \mathtt{jmp\ a} \qquad [\![\mathtt{a}]\!]_R = \mathtt{f}}{C \vdash ((M, (R, F), D), \mathtt{pc}, \mathtt{npc}) \circ\!\xrightarrow{\tau} ((M, (R, F), D), \mathtt{npc}, \mathtt{f})}$$

$$\frac{C(\mathtt{pc}) = \mathtt{call\ f} \qquad \mathbf{r}_{15} \in \mathrm{dom}(R)}{C \vdash ((M, (R, F), D), \mathtt{pc}, \mathtt{npc}) \circ\!\xrightarrow{\tau} ((M, (R\{\mathbf{r}_{15} \rightsquigarrow \mathtt{pc}\}, F), D), \mathtt{npc}, \mathtt{f})}$$

$$\frac{C(\mathtt{pc}) = \mathtt{retl} \qquad R(\mathbf{r}_{15}) = \mathtt{f}}{C \vdash ((M, (R, F), D), \mathtt{pc}, \mathtt{npc}) \circ\!\xrightarrow{\tau} ((M, (R, F), D), \mathtt{npc}, \mathtt{f} + 8)}$$

$$\frac{C(\mathtt{pc}) = \mathtt{print} \qquad R(\%\mathtt{o}_0) = v}{C \vdash ((M, (R, F), D), \mathtt{pc}, \mathtt{npc}) \circ\!\xrightarrow{\mathsf{out}(v)} ((M, (R, F), D), \mathtt{npc}, \mathtt{npc} + 4)}$$

$$\frac{C(\mathtt{pc}) = \mathtt{Psave\ } w \qquad \mathbf{save}(R, F) = \mathbf{undefined} \qquad (M, R, F) \Uparrow (M', R', F')}{C \vdash ((M, (R, F), D), \mathtt{pc}, \mathtt{npc}) \circ\!\xrightarrow{\tau} ((M', (R', F'), D), \mathtt{pc}, \mathtt{npc})}$$

$$\frac{C(\mathtt{pc}) = \mathtt{Prestore} \qquad \mathbf{restore}(R, F) = \mathbf{undefined} \qquad (M, R, F) \Downarrow (M', R', F')}{C \vdash ((M, (R, F), D), \mathtt{pc}, \mathtt{npc}) \circ\!\xrightarrow{\tau} ((M', (R', F'), D), \mathtt{pc}, \mathtt{npc})}$$

(b) Low-level Control Transfer Instruction Transition

$$\frac{\mathbf{\mathit{alloc}}(M, b, 0, w) = M' \qquad \mathbf{save}(R, F) = (R', F') \qquad R'' = R'\{\%\mathtt{sp} \rightsquigarrow (b, 0)\}}{\mathbf{execi}(\mathtt{Psave\ } w, (M, (R, F), D)) =_{\mathsf{L}} (M', (R', F'), D)}$$

$$\frac{\mathbf{\mathit{free}}(b, M) = M' \qquad \mathbf{restore}(R, F) = (R', F')}{\mathbf{execi}(\mathtt{Prestore}, (M, (R, F), D)) =_{\mathsf{L}} (M', (R'', F'), D)}$$

$$\frac{\mathbf{save}(R, F) = (R'', F') \qquad [\![\mathtt{o}]\!]_R = v \qquad R'' = R'\{\mathbf{r}_d \rightsquigarrow R(\mathbf{r}_s) + v\}}{\mathbf{execi}(\mathtt{save\ } \mathbf{r}_s\ \mathtt{o\ } \mathbf{r}_d, (M, (R, F), D)) =_{\mathsf{L}} (M', (R', F'), D)}$$

$$\frac{\mathbf{restore}(R, F) = (R', F') \qquad [\![\mathtt{o}]\!]_R = v \qquad R'' = R'\{\mathbf{r}_d \rightsquigarrow R(\mathbf{r}_s) + v\}}{\mathbf{execi}(\mathtt{restore\ } \mathbf{r}_s\ \mathtt{o\ } \mathbf{r}_d, (M, (R, F), D)) =_{\mathsf{L}} (M, (R'', F'), D')}$$

(c) Low-level Instruction Transition

$$[\![\mathtt{o}]\!]_R ::= \begin{cases} R(r) & \textbf{if } \mathtt{o} = r \\ w & \textbf{if } \mathtt{o} = (b, w') \textbf{ or } \mathtt{o} = w, \\ & \quad -4096 \leq w \leq 4095 \\ \bot & \textbf{otherwise} \end{cases} \qquad [\![\mathtt{a}]\!]_R ::= \begin{cases} [\![\mathtt{o}]\!]_R & \textbf{if } \mathtt{a} = \mathtt{o} \\ v_1 + v_2 & \textbf{if } \mathtt{a} = \mathtt{r} + \mathtt{o},\ R(\mathbf{r}) = v_1 \\ & \quad \textbf{and } [\![\mathtt{o}]\!]_R = v_2 \\ \bot & \textbf{otherwise} \end{cases}$$

(d) Low-level Expression Semantics

Fig.A4. Selected operational semantics rules for low-level program

$$\mathbf{fresh}(b, M) ::= \forall\, w.\, (b, w) \notin \mathrm{dom}(M)$$

$$\boldsymbol{alloc}(M, b, w_l, w_h) = M' ::= (M' = M \wedge w_l = w_h)\, \vee$$
$$(M' = M\{(b, w_l) \rightsquigarrow \_, \ldots, (b, w_h - 1) \rightsquigarrow \_\} \wedge \mathbf{fresh}(b, M) \wedge w_l < w_h)$$

$$\boldsymbol{free}(b, M) = M' ::= \forall\, b' \neq b, w'.\, M'(b', w') = M(b', w') \wedge \nexists w.\, (b, w) \in \mathrm{dom}(M)$$

Fig.A5. Auxiliary Definitions for Memory Operation

$$\frac{\begin{array}{c} F = F_1 \cdot \mathrm{fm}_1 \cdot \mathrm{fm}_2 \cdot \mathrm{fm}_3 \cdot \mathrm{fm}_4 \qquad \mathrm{fm}_1[6] = (b, 0) \\ R(\mathtt{wim}) = 2^n \qquad \{(b, 0), \ldots, (b, 15)\} \subseteq \mathrm{dom}(M) \qquad R' = R''\{\mathtt{wim} \rightsquigarrow 2^{\mathbf{next\_cwp}(n)}\} \\ M' = M\{[(b, 0), \ldots, (b, 7)] \rightsquigarrow \mathrm{fm}_2\}\{[(b, 8), \ldots, (b, 15)] \rightsquigarrow \mathrm{fm}_3\} \end{array}}{(M, (R, F)) \upuparrows (M', (R', F))}$$

$$\frac{\begin{array}{c} F = \mathrm{fm}_1 :: \mathrm{fm}_2 :: F'' \qquad R(\mathtt{r}_{30}) = (b, 0) \qquad R(\mathtt{wim}) = 2^n \\ \{[(b, 0), \ldots, (b, 7)] \rightsquigarrow \mathrm{fm}_1', [(b, 8), \ldots, (b, 15)] \rightsquigarrow \mathrm{fm}_2'\} \subseteq M' \\ R' = R''\{\mathtt{wim} \rightsquigarrow 2^{\mathbf{prev\_cwp}(n)}\} \qquad F' = \mathrm{fm}_1' :: \mathrm{fm}_2' :: F'' \end{array}}{(M, (R, F)) \downdownarrows (M, (R', F'))}$$

Fig.A6. Windows Over- and UnderFlow

C More about State Relation Between Low- and High-level Program

After introducing the definitions of low- and high-level program, we establish the state relation between low- and high-level program in this section. Establishing their state relation is not a trivial task, because there are two major differences low- and high-level program states. **First**, all the procedures' contexts of a specific thread are saved in high-level frame list $\mathbb{F}$. However, for low-level program, part of the contexts are saved in register windows (modeled as low-level frame list F), the other part of the contexts are saved in corresponding stack frame in memory, because the number of register windows is limited; **Second**, the high-level concurrent Pseudo-SPARCv8 program is multithreaded, but the low-level SPARCv8 program does not have the concept of thread pool.

$$\frac{}{(b, \mathrm{nil}, \emptyset) \Downarrow \mathrm{nil}} \qquad \frac{M_K = \{(b, \mathrm{fm}_1, \mathrm{fm}_2)\} \uplus M_K' \qquad \mathrm{fm}_2[6] = (b', 0) \qquad (b', \mathrm{nil}, M_K') \Downarrow \mathbb{F}}{(b, \mathrm{nil}, M_K) \Downarrow (\mathrm{fm}_1, \mathrm{fm}_2) :: \mathbb{F}}$$

$$\frac{M_K = \{(b, \_, \_)\} \uplus M_K' \qquad \mathrm{fm}_2[6] = (b', 0) \qquad (b', F, M_K') \Downarrow \mathbb{F}}{(b, \mathrm{fm}_1 :: \mathrm{fm}_2 :: F, M_K) \Downarrow (\mathrm{fm}_1, \mathrm{fm}_2) :: \mathbb{F}}$$

Fig.A7. Relation for low- and high-level FrameList

Relation for low- and high-level FrameList. The relation between low- and high-level frame list is defined in Fig. A7. We represent this relation as form "$(b, F, M_K) \Downarrow \mathbb{F}$", The tuple of b, F and M_K is the state of stack in low-level program, because, in the low-level program, part of the produces' contexts are saved in frame list F, which can also be understand as a prefix the whole frame list describe in assertion $\mathtt{cwp} \mapsto (\!|\_, F|\!)$, the other part of the contexts are saved in corresponding frame list represent as M_K. The high-level frame list $\mathbb{F}$ represents the state of stack in high-level program. Fig. 12 gives a more intuition understanding of this relation. Here, some part

$$\mathsf{ctxfm}(R, F) ::= \begin{cases} F_1 & \textbf{if } R(\mathtt{cwp}) = w_{id},\ R(\mathtt{wim}) = 2^n,\ \mathtt{cwp} \neq n, \\ & \qquad F = F_1 \cdot F_2,\ 0 \leq w_{id}, n \leq N,\ |F_1| = 2 \times (N + n - w_{id} - 1)\%N \\ \bot & \textbf{otherwise} \end{cases}$$

$$R \hookrightarrow \mathbb{R} \quad ::= (\forall\, i \in \{0, \ldots, 31\}.\, R(\mathtt{r}_i) = \mathbb{R}(\mathtt{r}_i)) \,\wedge\, (\forall\, \mathtt{sr} \neq \mathtt{wim}.\, R(\mathtt{sr}) = \mathbb{R}(\mathtt{sr}))$$
$$\wedge\, R(\mathtt{n}) = \mathbb{R}(\mathtt{n}) \,\wedge\, R(\mathtt{z}) = \mathbb{R}(\mathtt{z}) \,\wedge\, R(\mathtt{c}) = \mathbb{R}(\mathtt{c}) \,\wedge\, R(\mathtt{v}) = \mathbb{R}(\mathtt{v})$$

$$\frac{\begin{array}{c} M_c = M_{\mathrm{ctx}} \uplus M_K \quad \mathrm{dom}(M_{\mathrm{ctx}}) = \mathsf{DomCtxM}(\mathtt{t}) \quad R(\%\mathtt{sp}) = (b, 0) \\ \mathsf{ctxfm}(R, F) = F' \quad R(\%\mathtt{fp}) = (b', 0) \quad (b', F', M_K) \Downarrow \mathbb{F} \quad R \hookrightarrow \mathbb{R} \end{array}}{(M_c, (R, F)) \Downarrow_{\mathsf{c}} (\mathtt{t}, ((\mathbb{R}, \mathbb{F}), \mathtt{pc}, \mathtt{npc}))}$$

$$\frac{M_1 \Downarrow_{\mathsf{r}} T_1 \quad M_2 \Downarrow_{\mathsf{r}} T_2}{M_1 \uplus M_2 \Downarrow_{\mathsf{r}} T_1 \uplus T_2} \qquad \frac{M \blacktriangleright_{\mathsf{t}} Q \quad (M, Q) \Downarrow_{\mathsf{c}} (\mathtt{t}, \mathcal{K})}{M \Downarrow_{\mathsf{r}} \{\mathtt{t} \rightsquigarrow \mathcal{K}\}}$$

Fig.A8. Relation for Thread Pool and low-level Memory

of the contexts F (the pink part in the left side of the Fig. 12) are saved in register windows, and the other part of contexts M_K (the green part in the left side of the Fig. 12) are saved in stack frame in memory. However, in high-level state, they are abstracted as list named high-level frame list $\mathbb{F}$.

As shown in Fig. A7, if the low-level frame list F is nil and the memory is $\emptyset$, and the high-level frame list $\mathbb{F}$ is nil, it means there is no context stored. If the frame list is nil but the high-level frame list is $(b, \mathrm{fm}_1, \mathrm{fm}_2) :: \mathbb{F}$, it means that the contexts fm_1 and fm_2 are saved in stack frame in memory, whose block identifier is b. Here, we use "$\{(b, \mathrm{fm}_1, \mathrm{fm}_2)\}$" defined below to represent the part of memory saving fm_1 and fm_2. This memory contains only one block b.

$$\{(b, \mathrm{fm}_1, \mathrm{fm}_2)\} ::= \{(b, 0) \rightsquigarrow v_0, (b, 4) \rightsquigarrow v_1, \ldots, (b, 28) \rightsquigarrow v_7\}$$
$$\uplus \{(b, 32) \rightsquigarrow v_0', (b, 36) \rightsquigarrow v_1', \ldots, (b, 60) \rightsquigarrow v_7'\}$$
$$\text{where } \mathrm{fm}_1 = [v_0, \ldots, v_7],\ \mathrm{fm}_2 = [v_0', \ldots, v_7'].$$

If the frame list is $\mathrm{fm}_1 :: \mathrm{fm}_2 :: F$ and the high-level frame list is $(b', \mathrm{fm}_1', \mathrm{fm}_2') :: \mathbb{F}$, it means that the contexts fm_1 and fm_2 have not been saved in block b'. So, we require the contexts fm_1 and fm_2 saved in low-level frame list and the fm_1' and fm_2' saved in high-level frame list are equal. The block b' used to save fm_1 and fm_2 has not been used yet, so we don't care about its contents.

Relation for ThreadPool and low-level Memory. In high-level program, the thread local state of each thread is saved in a thread pool T. However, in low-level program, the local state of each thread is saved in memory (TCB and stack). For example, in Sec. 5, we introduce that the execution of the context switch module will save the register state of current thread into its TCB and stack in memory. So, the thread pool in high-level program can be viewed as an abstraction of low-level memory used to store the contexts of threads.

We define the relation between high-level thread pool and the memory used to save context in Fig. A8 formally. We use "$(M_c, (R, F)) \Downarrow_{\mathsf{c}} (\mathtt{t}, ((\mathbb{R}, b, \mathbb{F}), \mathtt{pc}, \mathtt{npc}))$" to represent the relation between the thread local states of *current thread* of low- and high-level program. The memory M_c owned the current thread $\mathtt{t}$ can be splitted into two parts M_{ctx} and M_K. The M_{ctx} are use the register file, whose domain is represented as $\mathsf{DomCtxM}(\mathtt{t}, b)$. It takes two arguments : the identifier $\mathtt{t}$ of the current thread and the block b of the stack frame at the top of the stack. Because

the context switch module may save the register file in TCB and the stack frame of the current procedure. The other part of the memory M_K is used to save the contexts of the previous procedures, which is abstracted as $\mathbb{F}$ in high-level program. We define $R \hookrightarrow \mathbb{R}$ to represent the relation between the register file R in low-level and $\mathbb{R}$ in high-level program. The operation $\mathsf{ctxfm}(R, F)$ is used to exact the prefix F_1 of the frame list F, which saves the contexts of the previous procedures. Supping the value of the $\mathtt{cwp}$ is w_{id}, meaning that the id of the current window is w_{id}, and the value of the $\mathtt{wim}$ is 2^n, meaning the id n register window is invalid. According to the introduction in Fig. 2.1, we usually set a window invalid to avoid over- and underflow of the register windows. So, we known that register windows id from $(w_{id}+1)\%N$ to $(n-1+N)\%N$ save the contexts of the previous procedures. So, we extract the contents F_1 of them from the whole frame list F.

We define "$M \Downarrow_r \{\mathsf{t} \rightsquigarrow \mathcal{K}\}$" to represent the relation between the thread local states of *ready thread* of low- and high-level program. The operation "$M \blacktriangleright_t Q$" means that we can restore the register state Q from memory M. When the context of the ready thread has been restored, we can establish a relation "$(M, Q) \Downarrow_c (\mathsf{t}, \mathcal{K})$" between low- and high-level thread local states of thread t. Here, we don't represent the definitions of $\mathsf{DomCtxM}(\mathsf{t}, b)$ and $M \blacktriangleright_t Q$ here, because their definitions are based on the implementation of the context switch routine in OS kernel. And the soundness of our extended program logic does not rely on their concrete definition.

Relation for Whole Program State. Finally, we introduce the state relation for whole program states between low- and high-level program below :

$$\frac{M = M_c \uplus M_T \uplus \{\mathsf{TaskCur} \rightsquigarrow (\mathsf{t}, 0)\} \uplus M' \quad (M_c, Q) \Downarrow_c (\mathsf{t}, \mathcal{K}) \quad M_T \Downarrow_r T\backslash\{\mathsf{t}\} \quad D = \mathrm{nil}}{(M, Q, D) \sim (T, \mathsf{t}, \mathcal{K}, M')}$$

D Application of Extended Program Logic : Verifying a Simplified Version of Context Switch Routine

In this section, we give a simplified version of context switch routine in Fig. A9. It reserves the main functionalities of the context switch routine introduced in Sec. 5, *e.g.*saving the contexts of currrent thread and restoring the new one. We omit some details like judging whether the current thread is a valid thread. We give a simple introduction to the function shown in Fig. A9, and show how to verify its correctness by applying our extended program logic for SPARCv8.

D.1 Simplified Context Switch Routine

At the entry of the context switch rountine shown in Fig. A9, we first save the local and in registers into the stack in memory, and we omit this part of code in presentation. Then, as shown in line 3 to 7, we call the $\mathtt{reg\_save}$ to store the out and global registers into the TCB of the current thread. As for the line 8 to 11, we get the identity of the current register window and the value of the $\mathtt{wim}$. The block $\mathtt{save\_usedwindow}$(line 12 to 22) saves the register windows (except the current one) into the stack of the current task in memory. It checks whether the previous window is valid. If it's valid, it uses the instruction $\mathtt{restore}$ to set the previous window as the current one, and calls function $\mathtt{window\_save}$ to save its contents into stack (in memory), then check the previous one continuously.

```
      SwitchEntry :
1     / * save the in and local registers of current window into stack frame * /
2     nop
3     set     TaskCur, %l₁
4     ld      [%l₁], %l₁
5     call    reg_save
6     nop
7     get     cwp, %g₄
8     rd      wim, %g₇
9     set     1, %g₆
10    sll     %g₆, %g₄, %g₄

     Save_Usedwindow :
11    sll     %g₄, 1, %g₅
12    srl     %g₄, (OS_WINDOWS − 1), %g₄
13    or      %g₄, %g₅, %g₄
14    andcc   %g₄, %g₇, %g₀
15    bne     switch_new_task
16    nop
17    restore %g₀, %g₀, %g₀
18    / * save the in and local registers of current window into stack frame * /
19    nop
20    jmp     Save_Usedwindow
21    nop

     Switch_NewContext :
22    set     TaskCur, %l₀
23    set     TaskNew, %l₁
24    ld      [%l₁], %l₁
25    st      %l₁, [%l₀]
26    call    reg_restore
27    nop
28    / * restore the in and local registers of current window from stack frame * /
29    nop
30    retl
31    nop
```

Fig.A9. Main function of context switch routine

The block `switch_new_task` is responsible for restoring the context of the new task. From line 23 to 26, it sets the new task as the current one. Then, it calls function `reg_restore` to restore the out and global registers from the new task's TCB, and calls function `window_restore` restore the local and in registers from the new task's stack in memory. The implementations of the internal functions are omitted here.

D.2 Specification of the Simplified Context Switch Routine

First, we define the abstract assembly primitive switch, which is already introduced in the Sec. 4.1.

$$\mathsf{switch} ::= \lambda\,\overline{v}, \mathbb{S}, \mathbb{S}'.\ \exists\,\mathsf{t}'.\ M(\mathsf{TaskNew}) = (\mathsf{t}', 0)\ \wedge\ T(\mathsf{t}') = (\mathbb{Q}', \mathtt{pc}', \mathtt{npc}')$$
$$\wedge\, T' = T\{\mathsf{t} \rightsquigarrow (\mathbb{Q}, \mathtt{pc}, \mathtt{npc})\}\ \wedge\ \mathsf{t} \neq \mathsf{t}'\ \wedge\ \overline{v} = \mathrm{nil}$$
$$\text{where } \mathbb{S} = (T, \mathsf{t}, (\mathbb{Q}, \mathtt{pc}, \mathtt{npc}), M),\ \mathbb{S}' = (T', \mathsf{t}', (\mathbb{Q}', \mathtt{f} + 8, \mathtt{f} + 12), M), \mathtt{f} = \mathbb{Q}'.\mathbb{R}(\mathbf{r}_{15}).$$

$$\mathsf{StkFrm}(b, \mathrm{fm}_1, \mathrm{fm}_2) ::= ((b,0) \Mapsto \mathrm{fm}_1[0]) * \ldots * ((b,28) \Mapsto \mathrm{fm}_1[7])$$
$$* ((b,32) \Mapsto \mathrm{fm}_2[0]) * \ldots * ((b,60) \Mapsto \mathrm{fm}_2[7])$$

$$\mathsf{RelStk}(b, F, \mathbb{F}) ::= \begin{cases} \mathsf{StkFrm}(b, \_, \_) * \mathsf{RelStk}(b', F', \mathbb{F}') & \text{if } \mathrm{fm}_2[6] = (b', 0),\ F = \mathrm{fm}_1 :: \mathrm{fm}_2 :: F' \\ & \mathbb{F} = (b, \mathrm{fm}_1, \mathrm{fm}_2) :: \mathbb{F}' \\[6pt] \mathsf{StkFrm}(b, \mathrm{fm}_1, \mathrm{fm}_2) * \mathsf{RelStk}(b', F', \mathbb{F}') & \text{if } \mathrm{fm}_2[6] = (b', 0),\ F = \mathrm{nil} \\ & \mathbb{F} = (b, \mathrm{fm}_1, \mathrm{fm}_2) :: \mathbb{F}' \\[6pt] \mathtt{Emp} & \text{if } F = \mathrm{nil}, \mathbb{F} = \mathrm{nil} \\[4pt] \mathrm{false} & \textbf{otherwise} \end{cases}$$

$$\mathsf{LRegs}(R) ::= \mathsf{global} \mapsto R(\mathsf{global}) * \mathsf{out} \mapsto R(\mathsf{out}) * \mathsf{local} \mapsto R(\mathsf{local}) * \mathsf{in} \mapsto R(\mathsf{in}) *$$
$$\mathtt{n} \mapsto R(\mathtt{n}) * \mathtt{z} \mapsto R(\mathtt{z}) * \mathtt{c} \mapsto R(\mathtt{z}) * \mathtt{v} \mapsto R(\mathtt{v}) * \ldots$$

$$\mathsf{wfwin}(R, F) ::= (\mathtt{cwp} \mapsto (\!| R(\mathtt{cwp}), F |\!) * \mathtt{wim} \mapsto R(\mathtt{wim})) \wedge \mathsf{ctxfm}(R, F)$$

$$\mathsf{context}(\mathsf{t}, b, \mathit{nst}) ::= ((\mathsf{t}, \mathtt{G0\_OFFSET}) \Mapsto \mathit{nst}(\%\mathsf{g}_0)) * \ldots * ((\mathsf{t}, \mathtt{G7\_OFFSET}) \Mapsto \mathit{nst}(\%\mathsf{g}_7))$$
$$* ((\mathsf{t}, \mathtt{O0\_OFFSET}) \Mapsto \mathit{nst}(\%\mathsf{o}_0)) * \ldots * ((\mathsf{t}, \mathtt{O7\_OFFSET}) \Mapsto \mathit{nst}(\%\mathsf{o}_7))$$
$$* ((\mathsf{t}, \mathtt{N\_OFFSET}) \Mapsto \mathit{nst}(\mathtt{n})) * \ldots * ((\mathsf{t}, \mathtt{V\_OFFSET}) \Mapsto \mathit{nst}(\mathtt{v}))$$
$$* \mathsf{StkFrm}(b, \mathit{nst}[\mathsf{local}], \mathit{nst}[\mathsf{in}])$$

$$\mathsf{Env}(\mathit{env}) ::= \mathsf{LRegs}(R) * \mathsf{wfwin}(R, F) \qquad \text{where } \mathit{env} = (R, F)$$

$$\mathsf{CurT'}(\mathsf{t}_c, \mathit{nst}, \mathit{env}, \mathcal{K}) ::= (\mathsf{context}(\mathsf{t}_c, b, \mathit{nst}) * \mathsf{RelStk}(b', F, \mathbb{F}) * (\mathsf{t}_c \leadsto_\mathsf{c} \mathcal{K}))$$
$$\wedge R(\%\mathtt{sp}) = (b, 0) \wedge R \hookrightarrow \mathbb{R}$$
$$\text{where } \mathit{env} = (R, F),\ \mathcal{K} = ((\mathbb{R}, b, \mathbb{F}), \mathtt{pc}, \mathtt{npc}),\ R(\%\mathtt{fp}) = (b', 0),\ \mathit{nst} \in \mathrm{RegFile}$$

$$\mathsf{CurT}(\mathsf{t}_c, \mathit{nst}, \mathit{env}, \mathcal{K}) ::= (\mathsf{TaskCur} \Mapsto (\mathsf{t}_c, 0)) * \mathsf{CurT'}(\mathsf{t}_c, \mathit{nst}, \mathit{env}, \mathcal{K})$$

$$\mathsf{RdyT}(\mathsf{t}_n, \mathit{nst}, \mathcal{K}) ::= (\mathsf{context}(\mathsf{t}_n, b, \mathit{nst}) * \mathsf{RelStk}(b', \mathrm{nil}, \mathbb{F}) * (\mathsf{t}_n \leadsto_\mathsf{r} \mathcal{K}))$$
$$\wedge \mathit{nst} \hookrightarrow \mathbb{R} \wedge R(\%\mathtt{sp}) = (b, 0)$$
$$\text{where } \mathcal{K} = ((\mathbb{R}, b, \mathbb{F}), \mathtt{pc}, \mathtt{npc}),\ \mathit{nst}(\%\mathtt{fp}) = (b', 0),\ \mathit{nst} \in \mathrm{RegFile}$$

$$\mathsf{p\_env}(\mathit{env}) ::= R \qquad \text{where } \mathit{env} = (R, F) \qquad l \Mapsto v ::= l \mapsto v * l \rightarrowtail v$$

Fig.A10. Auxiliary Definitions for Specification

Then we show the specification of the simplified context switch routine below, and some auxiliary definitions used in specification can be found in Fig. A10:

$$a_{pre}(\mathsf{t}_c, \mathsf{t}_n, \mathit{env}, \mathit{nst}, \mathcal{K}_c, \mathcal{K}_n) ::= \mathsf{Env}(\mathit{env}) * (\mathsf{TaskNew} \Mapsto (\mathsf{t}_n, 0) \wedge \mathsf{t}_c \neq \mathsf{t}_n) * \blacklozenge(10) *$$
$$\mathsf{CurT}(\mathsf{t}_c, \_, \mathit{env}, \mathcal{K}_c) * \mathsf{RdyT}(\mathsf{t}_n, \mathit{nst}, \mathcal{K}_n) * (\!| \mathsf{switch}(\mathrm{nil}) |\!)$$

$$a_{post}(\mathsf{t}_c, \mathsf{t}_n, \mathit{env}, \mathit{nst}, \mathcal{K}_c, \mathcal{K}_n) ::= \exists\, \mathit{env}', \mathcal{K}'.\, \mathsf{Env}(\mathit{env}') * (\mathsf{TaskNew} \Mapsto (\mathsf{t}_n, 0) \wedge \mathsf{t}_c \neq \mathsf{t}_n) *$$
$$\mathsf{CurT}(\mathsf{t}_n, \mathit{nst}, \mathit{env}', \mathcal{K}') * \mathsf{RdyT}(\mathsf{t}_c, \mathsf{p\_env}(\mathit{env}), \mathcal{K}_c) * (\!| \bot |\!)$$

Note that the execution of context switch routine will call function `reg_save`, `reg_restore`, and `window_restore` once, and call function `window_save` and jump to block `save_usedwindow` no more than 8 times separately, because the number of the register windows is 8. So, assigning 10 tokens to the precondition of the context switch routine is enough. According to the logic rules of extended program logic shown in Fig. 19, we need to check whether the specification of context switch routine is well-defined.

Lemma 5. $\mathsf{wdSpec}(a_{pre}, a_{post}, \mathsf{switch})$.

Proof. We unfold $\mathsf{wdSpec}(a_{pre}, a_{post}, \mathsf{switch})$ by Def. 2, and we need to prove three properities about the specification and abstract assembly primitive switch.

1. for any $\overline{v}, \mathbb{S}, \mathbb{S}', \mathbb{S}_r$. if $\mathsf{switch}(\overline{v})(\mathbb{S})(\mathbb{S}')$, and $\mathbb{S} \perp \mathbb{S}_r$, then the following holds :

$$
\mathsf{Mem}(M) \quad ::= \quad
\begin{cases}
\texttt{Emp} & \textbf{if } M = \emptyset \\
(l \mapsto v) * (l \rightarrowtail v) & \textbf{if } M = \{l \rightsquigarrow v\} \\
\exists M_1, M_2.\, \mathsf{Mem}(M_1) * \mathsf{Mem}(M_2) & \textbf{otherwise}
\end{cases}
$$

$$
\mathsf{RdyTs}(T) \quad ::= \quad
\begin{cases}
\texttt{Emp} & \textbf{if } T = \emptyset \\
\mathsf{RdyT}(\texttt{t}, \_, \mathcal{K}) & \textbf{if } T = \{\texttt{t} \rightsquigarrow \mathcal{K}\} \\
\exists T_1, T_2.\, \mathsf{RdyTs}(T_1) * \mathsf{RdyTs}(T_2) & \textbf{otherwise}
\end{cases}
$$

$$
\begin{aligned}
M \blacktriangleright_{\texttt{t}} (R, F) ::=\ & \exists b.\, R(\%\texttt{g}_0) = M(\texttt{t}, \texttt{G0\_OFFSET}) \wedge \ldots \wedge R(\%\texttt{g}_7) = M(\texttt{t}, \texttt{G7\_OFFSET}) \\
& \wedge R(\%\texttt{o}_0) = M(\texttt{t}, \texttt{O0\_OFFSET}) \wedge \ldots \wedge R(\%\texttt{o}_7) = M(\texttt{t}, \texttt{O7\_OFFSET}) \\
& \wedge R(\texttt{n}) = M(\texttt{t}, \texttt{N\_OFFSET}) \wedge \ldots \wedge R(\texttt{v}) = M(\texttt{t}, \texttt{V\_OFFSET}) \\
& \wedge R(\%\texttt{l}_0) = M(b, 0) \wedge \ldots \wedge R(\%\texttt{l}_7) = M(b, 28) \\
& \wedge R(\%\texttt{i}_0) = M(b, 32) \wedge \ldots \wedge R(\%\texttt{i}_7) = M(b, 60) \wedge R(\%\texttt{sp}) = (b, 0) \\
& \wedge (\exists w_{id}, n.\, R(\texttt{cwp}) = w_{id} \wedge R(\texttt{wim}) = 2^n \wedge \mathbf{prev\_cwp}(w_{id}) = n)
\end{aligned}
$$

$$
\mathsf{rRegs} \quad ::= \quad \texttt{asr}_0 \mapsto \_ * \ldots \texttt{asr}_{31} \mapsto \_ * \texttt{Y} \mapsto \_
$$

Fig.A11. Auxiliary Definitions About Frame Assertion

- $\mathbb{S}'.\mathcal{K}.\texttt{pc} = \texttt{f} + 8$, $\mathbb{S}'.\mathcal{K}.\texttt{npc} = \texttt{f} + 12$ (where $\mathbb{S}'.\mathcal{K}.\mathbb{Q}.\mathbb{R}(\mathbf{r}_{15}) = \texttt{f}$);

- there exists $\mathbb{S}'', \mathbb{S}'_r$, $\mathsf{switch}(\overline{v})(\mathbb{S} \uplus \mathbb{S}_r)(\mathbb{S}'')$, $\mathbb{S}'' = \mathbb{S}' \uplus \mathbb{S}'_r$, and $\mathbb{S}_r.T = \mathbb{S}'_r.T$, $\mathbb{S}_r.M = \mathbb{S}'_r.M$;

The correctness proof of this property can be achieved directly from the definition of the switch.

2. for any $\texttt{t}_c, \texttt{t}_n, \mathit{env}, \mathit{nst}, \mathcal{K}_c, \mathcal{K}_n$,

   - $a_{pre}(\texttt{t}_c, \texttt{t}_n, \mathit{env}, \mathit{nst}, \mathcal{K}_c, \mathcal{K}_n) \implies (\!|\mathsf{switch}|\!) * \mathrm{true}$;

   - $a_{post}(\texttt{t}_c, \texttt{t}_n, \mathit{env}, \mathit{nst}, \mathcal{K}_c, \mathcal{K}_n) \implies (\!|\bot|\!) * \mathrm{true}$;

According the definition of a_{pre} and a_{post}, this property's proof is trivial.

3. for any $\overline{v}, S, \mathbb{S}$, if $(S, \mathbb{S}, \_, \_) \in \mathsf{INV}(\mathsf{switch}(\overline{v}), \overline{v})$, then there exists $\texttt{t}_c, \texttt{t}_n, \mathit{env}, \mathit{nst}, \mathcal{K}_c, \mathcal{K}_n, \mathbb{p}_r$ and w, such that:

   - $(S, \mathbb{S}, \mathsf{switch}(\overline{v}), w) \models a_{pre}(\texttt{t}_c, \texttt{t}_n, \mathit{env}, \mathit{nst}, \mathcal{K}_c, \mathcal{K}_n) * \mathbb{p}_r$;

   - $a_{post}(\texttt{t}_c, \texttt{t}_n, \mathit{env}, \mathit{nst}, \mathcal{K}_c, \mathcal{K}_n) * \mathbb{p}_r \implies \mathsf{INV}(\bot, \_)$;

   - $\mathsf{Sta}(\mathsf{switch}(\overline{v}), \mathbb{p}_r)$.

The key to prove this case is to find $\texttt{t}_c$, $\texttt{t}_n$, env, nst, $\mathcal{K}_c$, $\mathcal{K}_n$, $\mathbb{p}_r$ and w. Because we have $(S, \mathbb{S}, \_, \_) \in \mathsf{INV}(\mathsf{switch}(\overline{v}), \overline{v})$, we can know that there exists a ready thread $\texttt{t}'$, a prefix of the frame list F' and a register state Q', where $\mathbb{S}.T(\texttt{t}') = \mathcal{K}'$, $\texttt{t} \neq \texttt{t}'$, $\mathbb{S}.M(\mathsf{TaskNew}) = (\texttt{t}', 0)$, $\mathsf{ctxfm}(S.Q) = F'$ and $S.M \blacktriangleright_{\texttt{t}'} Q'$ hold. And we require $\texttt{t}_c = \mathbb{S}.\texttt{t}$, $\texttt{t}_n = \texttt{t}'$, $\mathit{env} = (S.Q.R, F')$, $\mathit{nst} = Q'.R$, $w = 10$, and $\mathbb{p}_r = \exists M, T.\, \mathsf{Mem}(M) * \mathsf{RdyTs}(T) * \mathsf{rRegs}$. Then, we can finish the proof.

$\square$

Loop invariant I:

$\mathsf{wptr}(R_0) ::= (R_0(\%\mathsf{g}_7) = R_0(\mathtt{wim})) \wedge$
$\qquad\qquad ((R_0(\%\mathsf{g}_4) = (1 \ll R_0(\mathtt{cwp}))) \vee (R_0(\%\mathsf{g}_4) = ((1 \ll R_0(\mathtt{cwp})) \ll 8)))$

$\mathsf{linkF}((b_1, \mathbb{F}_1), (b_2, \mathbb{F}_2), \mathbb{F}) ::= \mathbb{F}_1 \cdot \mathbb{F}_2 = \mathbb{F} \wedge (\mathbb{F}_1 = \mathrm{nil} \to b_1 = b_2)$
$\qquad\qquad\qquad\qquad\qquad (\forall b, \mathrm{fm}_1, \mathrm{fm}_2, \mathbb{F}'. \, \mathbb{F}_1 = (b, \mathrm{fm}_1, \mathrm{fm}_2) :: \mathbb{F}' \to \mathrm{fm}_2[6] = (b_2, 0))$

$I(\mathsf{t}_c, R, \mathcal{K}_c) ::= \exists R_0, F_0. \, (\mathsf{Env}(R_0, F_0) \wedge \mathsf{wptr}(R_0)) * \blacklozenge(|F_0| + 2)$
$\qquad\qquad\quad * ((\mathsf{TaskCur} \mapsto (\mathsf{t}_c, 0) * \mathsf{context}(\mathsf{t}_c, b, R) * \mathsf{t}_c \rightsquigarrow_{\mathsf{c}} \mathcal{K}_c) \wedge R \hookrightarrow \mathbb{R})$
$\qquad\qquad\quad * (\exists b'', \mathbb{F}_1, \mathbb{F}_2. \, (\mathsf{RelStk}(b', \mathrm{nil}, \mathbb{F}_1) * \mathsf{RelStk}(b'', F_0, \mathbb{F}_2)) \wedge R_0(\%\mathsf{sp}) = (b'', 0)$
$\qquad\qquad\qquad \wedge \mathsf{linkF}((b', \mathbb{F}_1), (b'', \mathbb{F}_2), \mathbb{F}))$
$\qquad\quad$ where $\mathcal{K}_c = (\mathbb{R}, b, \mathbb{F})$, $R(\%\mathsf{sp}) = (b, 0)$, and $R(\%\mathtt{fp}) = (b', 0)$

$\mathtt{reg\_save}:\; (\iota = (\mathsf{t}, \mathcal{K}, R, F, A, nst))$
$\quad \mathsf{fp}_{rs} \, \iota ::= \mathsf{Env}(R, F) * \mathsf{context}(\mathsf{t}, b, nst) * (\mathsf{t} \rightsquigarrow_{\mathsf{c}} \mathcal{K}) * (\!|A|\!)$
$\quad \mathsf{fq}_{rs} \, \iota ::= (\exists nst'. \, \mathsf{Env}(R, F) * \mathsf{context}(\mathsf{t}, b, nst')) * (\mathsf{t} \rightsquigarrow_{\mathsf{c}} \mathcal{K} * (\!|A|\!))$
$\qquad\qquad \wedge nst' = nst\{\mathsf{global} \rightsquigarrow R(\mathsf{global}), \mathsf{out} \rightsquigarrow R(\mathsf{out}), \mathsf{n} \rightsquigarrow R(\mathsf{n}), \ldots, \mathsf{v} \rightsquigarrow R(\mathsf{v})\}$

$\mathtt{window\_save}:\; (\iota = (\mathsf{t}, \mathcal{K}, R, F, A, b))$
$\quad \mathsf{fp}_{ws} \, \iota ::= (\mathsf{Env}(R, F) * \mathsf{StkFrm}(b, \_, \_) * (\mathsf{t} \rightsquigarrow_{\mathsf{c}} \mathcal{K}) * (\!|A|\!)) \wedge R(\%\mathsf{sp}) = (b, 0)$
$\quad \mathsf{fq}_{ws} \, \iota ::= (\mathsf{Env}(R, F) * \mathsf{StkFrm}(b, R(\mathsf{local}), R(\mathsf{in})) * (\mathsf{t} \rightsquigarrow_{\mathsf{c}} \mathcal{K}) * (\!|A|\!)) \wedge R(\%\mathsf{sp}) = (b, 0)$

$\mathtt{reg\_restore}:\; (\iota = (\mathsf{t}, \mathcal{K}, R, F, b, nst))$
$\quad \mathsf{fp}_{rr} \, \iota ::= \mathsf{Env}(R, F) * \mathsf{context}(\mathsf{t}, b, nst) * (\mathsf{t} \rightsquigarrow_{\mathsf{c}} \mathcal{K}) * (\!|A|\!)$
$\quad \mathsf{fq}_{rr} \, \iota ::= \exists R'. \, (\mathsf{Env}(R', F) * \mathsf{context}(\mathsf{t}, b, nst) * (\mathsf{t} \rightsquigarrow_{\mathsf{c}} \mathcal{K}) * (\!|A|\!))$
$\qquad\qquad \wedge R' = R\{\mathsf{global} \rightsquigarrow nst(\mathsf{global}), \mathsf{out} \rightsquigarrow nst(\mathsf{out}), \mathsf{n} \rightsquigarrow nst(\mathsf{n}), \ldots, \mathsf{v} \rightsquigarrow nst(\mathsf{v})\}$

$\mathtt{window\_restore}:\; (\iota = (\mathsf{t}, \mathcal{K}, R, F, A, b, \mathrm{fm}_1, \mathrm{fm}_2))$
$\quad \mathsf{fp}_{wr} \, \iota ::= (\mathsf{Env}(R, F) * \mathsf{StkFrm}(b, \mathrm{fm}_1, \mathrm{fm}_2) * (\mathsf{t} \rightsquigarrow_{\mathsf{c}} \mathcal{K}) * (\!|A|\!)) \wedge R(\%\mathsf{sp}) = (b, 0)$
$\quad \mathsf{fq}_{wr} \, \iota ::= (\mathsf{Env}(R\{\mathsf{local} \rightsquigarrow \mathrm{fm}_1, \mathsf{in} \rightsquigarrow \mathrm{fm}_2\}, F) * \mathsf{StkFrm}(b, \mathrm{fm}_1, \mathrm{fm}_2) * (\mathsf{t} \rightsquigarrow_{\mathsf{c}} \mathcal{K}) * (\!|A|\!))$
$\qquad\qquad \wedge R(\%\mathsf{sp}) = (b, 0)$

$\mathtt{save\_usedwindow}:\; (\iota = (\mathsf{t}_c, \mathsf{t}_n, \mathcal{K}_c, \mathcal{K}_n, \mathcal{K}_c, \mathcal{K}_n, nst))$
$\quad \mathsf{fp}_{su} \, \iota ::= I(\mathsf{t}_c, R, \mathcal{K}_c) * (\mathsf{TaskNew} \mapsto (\mathsf{t}_n, 0) \wedge \mathsf{t}_c \neq \mathsf{t}_n) * \mathsf{RdyT}(\mathsf{t}_n, nst, \mathcal{K}_n) * (\!|\mathsf{switch}(\mathrm{nil})|\!)$
$\quad \mathsf{fq}_{su} \, \iota ::= \exists \mathcal{K}'. \, \mathsf{Env}(nst, \mathrm{nil}) * (\mathsf{TaskNew} \mapsto (\mathsf{t}_n, 0) \wedge \mathsf{t}_c \neq \mathsf{t}_n) *$
$\qquad\qquad \mathsf{CurT}(\mathsf{t}_n, nst, (nst, \mathrm{nil}), \mathcal{K}') * \mathsf{RdyT}(\mathsf{t}_c, R, \mathcal{K}_c) * (\!|\bot|\!)$

$\mathtt{switch\_newtask}:\; (\iota = (\mathsf{t}_c, \mathsf{t}_n, \mathcal{K}_c, \mathcal{K}_n, R, nst))$
$\quad \mathsf{fp}_{sn} \, \iota ::= \exists R_0. \, \mathsf{Env}((R_0, \mathrm{nil})) * (\mathsf{TaskNew} \mapsto (\mathsf{t}_n, 0) \wedge \mathsf{t}_c \neq \mathsf{t}_n) * \blacklozenge(1) *$
$\qquad\qquad \mathsf{CurT}(\mathsf{t}_c, R, (R, \mathrm{nil}), \mathcal{K}_c) * \mathsf{RdyT}(\mathsf{t}_n, nst, \mathcal{K}_n) * (\!|\mathsf{switch}(\mathrm{nil})|\!)$
$\quad \mathsf{fq}_{sn} \, \iota ::= \exists \mathcal{K}'. \, \mathsf{Env}(nst, \mathrm{nil}) * (\mathsf{TaskNew} \mapsto (\mathsf{t}_n, 0) \wedge \mathsf{t}_c \neq \mathsf{t}_n) *$
$\qquad\qquad \mathsf{CurT}(\mathsf{t}_n, nst, (nst, \mathrm{nil}), \mathcal{K}') * \mathsf{RdyT}(\mathsf{t}_c, R, \mathcal{K}_c) * (\!|\bot|\!)$

<p align="center">Fig.A12. Specifications of Internal Functions</p>

We use $C_{\mathtt{switch}}$ to represent the code heap that stores the code of context switch routine shown in Fig. A9. The specifications of the internal function can be found in Fig. A12. The function $\mathtt{reg\_save}$ is responsible for saving the local, in and integer condition code fields n, z, c and v registers into TCB in memory. The $\mathtt{window\_save}$ stores the local and in registers into current task's stack in memory. Function $\mathtt{reg\_restore}$ does the reverse of $\mathtt{reg\_save}$, and function $\mathtt{window\_restore}$ does the reverse of $\mathtt{window\_save}$. The specification of code block $\mathtt{save\_usedwindow}$ is a little complicated. We can find its implementation is a loop, which checks whether the previous window is valid and saving the contents of the valid previous window until the previous one is invalid. We need to define the loop invariant I here.

```
    SwitchEntry :
```
$\{a_{pre}(\mathsf{t}_c, \mathsf{t}_n, (R, F), \mathit{nst}, \mathcal{K}_c, \mathcal{K}_n)\}$

$\left\{ \begin{array}{l} \mathsf{Env}(R, F) * (\mathsf{TaskNew} \mapsto (\mathsf{t}_n, 0) \wedge \mathsf{t}_c \neq \mathsf{t}_n) * \blacklozenge(10) * \\ \quad \mathsf{CurT}(\mathsf{t}_c, \_, (R, F), \mathcal{K}_c) * \mathsf{RdyT}(\mathsf{t}_n, \mathit{nst}, \mathcal{K}_n) * (\!|\, \mathsf{switch}(\mathrm{nil}) \,|\!) \end{array} \right\}$ ①

```
1      / * saving the in and local registers of current window into stack frame * /
2      nop
       ...
5      call reg_save
6      nop
```

$\left\{ \begin{array}{l} \mathsf{Env}(R\{\%\mathtt{l}_1 \rightsquigarrow \_\}, F) * (\mathsf{TaskNew} \mapsto (\mathsf{t}_n, 0) \wedge \mathsf{t}_c \neq \mathsf{t}_n) * \blacklozenge(9) * \\ \quad \boxed{\mathsf{CurT}(\mathsf{t}_c, R, (R, F), \mathcal{K}_c)} * \mathsf{RdyT}(\mathsf{t}_n, \mathit{nst}, \mathcal{K}_n) * (\!|\, \mathsf{switch}(\mathrm{nil}) \,|\!) \end{array} \right\}$ ②

```
7      get   cwp, %g_4
8      rd    wim, %g_7
9      set   1, %g_6
10     sll   %g_6, %g_4, %g_4
```

$\left\{ \begin{array}{l} \exists R_0.\, (\mathsf{Env}(R_0, F) \wedge \mathsf{wptr}(R_0)) * (\mathsf{TaskNew} \mapsto (\mathsf{t}_n, 0) \wedge \mathsf{t}_c \neq \mathsf{t}_n) * \blacklozenge(9) * \\ \quad \mathsf{CurT}(\mathsf{t}_c, R, (R, F), \mathcal{K}_c) * \mathsf{RdyT}(\mathsf{t}_n, \mathit{nst}, \mathcal{K}_n) * (\!|\, \mathsf{switch}(\mathrm{nil}) \,|\!) \end{array} \right\}$

$\{I(\mathsf{t}_c, R, \mathcal{K}_c) * (\mathsf{TaskNew} \mapsto (\mathsf{t}_n, 0) \wedge \mathsf{t}_c \neq \mathsf{t}_n) * \mathsf{RdyT}(\mathsf{t}_n, \mathit{nst}, \mathcal{K}_n) * (\!|\, \mathsf{switch}(\mathrm{nil}) \,|\!)\}$

```
   Save_Usedwindow :
11     sll   %g_4, 1, %g_5
       ...
20     jmp   Save_UsedWindow
21     nop
```

$\left\{ \begin{array}{l} \exists R_0.\, \boxed{\mathsf{Env}(R_0, \mathrm{nil})} * (\mathsf{TaskNew} \mapsto (\mathsf{t}_n, 0) \wedge \mathsf{t}_c \neq \mathsf{t}_n) * \blacklozenge(1) * \\ \quad \mathsf{CurT}(\mathsf{t}_c, R, (R, \mathrm{nil}), \mathcal{K}_c) * \mathsf{RdyT}(\mathsf{t}_n, \mathit{nst}, \mathcal{K}_n) * (\!|\, \mathsf{switch}(\mathrm{nil}) \,|\!) \end{array} \right\}$ ③

$\left\{ \begin{array}{l} \exists R_0.\, \mathsf{Env}(R_0, \mathrm{nil}) * (\mathsf{TaskNew} \mapsto (\mathsf{t}_n, 0) \wedge \mathsf{t}_c \neq \mathsf{t}_n) * \blacklozenge(1) * (\!|\, \mathsf{switch}(\mathrm{nil}) \,|\!) \\ \quad * \boxed{\mathsf{TaskCur} \mapsto (\mathsf{t}_c, 0) * \mathsf{CurT'}(\mathsf{t}_c, R, (R, \mathrm{nil}), \mathcal{K}_c)} * \mathsf{RdyT}(\mathsf{t}_n, \mathit{nst}, \mathcal{K}_n) \end{array} \right\}$

```
   Switch_NewContext :
22     set   TaskCur, %l_0
23     set   TaskNew, %l_1
       ...
```

$\left\{ \begin{array}{l} \boxed{\mathsf{Env}(\mathit{nst}, \mathrm{nil})} * (\mathsf{TaskNew} \mapsto (\mathsf{t}_n, 0) \wedge \mathsf{t}_c \neq \mathsf{t}_n) * (\!|\, \mathsf{switch}(\mathrm{nil}) \,|\!) \\ \quad * \mathsf{TaskCur} \mapsto (\mathsf{t}_n, 0) * \boxed{\mathsf{CurT'}(\mathsf{t}_c, R, (R, \mathrm{nil}), \mathcal{K}_c)} * \boxed{\mathsf{RdyT}(\mathsf{t}_n, \mathit{nst}, \mathcal{K}_n)} \end{array} \right\}$ ④

$\Downarrow$

$\left\{ \begin{array}{l} \mathsf{Env}(\mathit{nst}, \mathrm{nil}) * (\mathsf{TaskNew} \mapsto (\mathsf{t}_n, 0) \wedge \mathsf{t}_c \neq \mathsf{t}_n) * (\!|\, \bot \,|\!) \\ \quad * \mathsf{TaskCur} \mapsto (\mathsf{t}_n, 0) * \boxed{\mathsf{RdyT}(\mathsf{t}_c, R, \mathcal{K}_c)} * \boxed{(\exists \mathcal{K}'.\, \mathsf{RdyT}(\mathsf{t}_n, \mathit{nst}, \mathcal{K}'))} \end{array} \right\}$ ⑤

```
30     retl
31     nop
```

$\{a_{post}(\mathsf{t}_c, \mathsf{t}_n, (R, F), \mathcal{K}_c, \mathcal{K}_n)\}$

Fig.A13. Proof Sketch of the Context Switch Routine

Lemma 6. There exists Ψ_i, for any ι, where $\iota = (\mathsf{t}_c, \mathsf{t}_n, \mathit{env}, \mathit{nst}, \mathcal{K}_c, \mathcal{K}_n)$, $\vdash C_{\mathtt{switch}} : \Psi_i$ and $\Psi_i \vdash \{(a_{pre}\, \iota, a_{post}\, \iota)\}\, \mathtt{f}_{\mathtt{switch}} : C_{\mathtt{switch}}[\mathtt{f}_{\mathtt{switch}}]$ hold.

Proof. The code heap specification Ψ_i is the collection of the specifications of each code block shown in Fig. A12, and the verification of internal functions has no differences with the original proof of context switch routine introduced in Sec. 5. So, here we just give a proof sketch of the verifying of the main function, which can be found in Fig. A13.

```
st %l0, [%sp + L0_OFFSET]              ld [%sp + L0_OFFSET], %l0
st %l1, [%sp + L1_OFFSET]              ld [%sp + L1_OFFSET], %l1
st %l2, [%sp + L2_OFFSET]              ld [%sp + L2_OFFSET], %l2
st %l3, [%sp + L3_OFFSET]              ld [%sp + L3_OFFSET], %l3
st %l4, [%sp + L4_OFFSET]              ld [%sp + L4_OFFSET], %l4
st %l5, [%sp + L5_OFFSET]              ld [%sp + L5_OFFSET], %l5
st %l6, [%sp + L6_OFFSET]              ld [%sp + L6_OFFSET], %l6
st %l7, [%sp + L7_OFFSET]              ld [%sp + L7_OFFSET], %l7
st %i0, [%sp + I0_OFFSET]              ld [%sp + I0_OFFSET], %i0
st %i1, [%sp + I1_OFFSET]              ld [%sp + I1_OFFSET], %i1
st %i2, [%sp + I2_OFFSET]              ld [%sp + I2_OFFSET], %i2
st %i3, [%sp + I3_OFFSET]              ld [%sp + I3_OFFSET], %i3
st %i4, [%sp + I4_OFFSET]              ld [%sp + I4_OFFSET], %i4
st %i5, [%sp + I5_OFFSET]              ld [%sp + I5_OFFSET], %i5
st %i6, [%sp + I6_OFFSET]              ld [%sp + I6_OFFSET], %i6
st %i7, [%sp + I7_OFFSET]              ld [%sp + I7_OFFSET], %i7
```
(a) Save local and out into memory   (b) Restore local and in from memory

Fig.A14. Code for saving and restoring local and in registers

Supposing in the inital state (described as assertion marked ①), the register file is R, and the part of the frame list, which is waitting for saving into the stack in memory, is F. The code segment from line 1 to line 6 is responsible for saving the register file R into current task's TCB, and we achieve assertion marked ②.

The codes from line 7 to 21 saves the prefix F of the frame list into current task's stack in memory. After execution of this segment. The part of the frame list, waitting for storing into memory, becomes empty (nil). And the assertion marked ③ holds.

Then, we prove the code block `switch_new_task`, which restores the context nst of the new task t_n. After executing the codes from line 22 to line 29, the context nst of the new task t_n is restored (shown as $\mathsf{Env}((nst, \mathrm{nil})))$, and the assertion marked ④ holds.

Finally, we apply **ABSCSQ** rule, shown in Fig. 19, to execute the abstract assembly primitive switch, and the assertion marked ⑤ holds. By applying **RETL** rule shown in Fig. 19, we finish the proof.   □

Theorem 2. $\{\mathtt{f_{switch}} \rightsquigarrow (a_{pre}, a_{post})\} \vdash C_{\mathtt{switch}} : \{\mathtt{f_{switch}} \rightsquigarrow \mathsf{switch}\}$.

Proof. The proof can be finished by applying **Lemma** 5 and **Lemma** 6.   □