

A Logic for Information Flow Analysis with an Application to Forward Slicing of Simple Imperative Programs

Torben Amtoft¹, Anindya Banerjee²

*Department of Computing and Information Sciences
Kansas State University, Manhattan KS 66506, USA*

Abstract

We specify an information flow analysis for a simple imperative language, using a Hoare-like logic. The logic facilitates static checking of a larger class of programs than can be checked by extant type-based approaches in which a program is deemed insecure when it contains an insecure subprogram. The logic is based on an abstract interpretation of a “prelude” semantics which makes independence between program variables explicit. Unlike other, more precise, approaches based on Hoare logics, our approach does not require a theorem prover to generate invariants. We demonstrate the modularity of our approach by showing that a frame rule holds in our logic. Finally, we show how our logic can be applied to a program transformation, namely, forward slicing: given a derivation of a program in the logic, with the information that variable l is independent of variable h , the slicing transformation systematically creates the forward l -slice of the program: the slice contains all the commands independent of h . We show that the slicing transformation is semantics preserving.

Key words: Abstract interpretation, denotational semantics, frame rule, Hoare logic, information flow analysis, program slicing, strongest postcondition.

1 Introduction

This article specifies an information flow analysis using a Hoare-like logic and considers an application of the logic to forward slicing in simple imperative

Email addresses: tamtsoft@cis.ksu.edu (Torben Amtoft), ab@cis.ksu.edu (Anindya Banerjee).

¹ Supported by NSF grants CCR-0296182 and CCR-0209205

² Supported by NSF grants CCR-0296182 and CCR-0209205 and ITR-0326577

programs.

Given a system with high, or secret (H), and low, or public (L) inputs and outputs, where $L \leq H$ is a security lattice, a classic security problem is how to enforce the following end-to-end *confidentiality* policy: protect secret data, i.e., prevent leaks of secrets at public output channels. An information flow analysis checks if a program satisfies the policy. Denning and Denning (1977) were the first to formulate an information flow analysis for confidentiality. Subsequent advances have been comprehensively summarized in the recent survey by Sabelfeld and Myers (2003). An oft-used approach for specifying static analyses for information flow is *security type systems* (Ørbæk and Palsberg, 1997; Volpano and Smith, 1997; Volpano et al., 1996). Security types can be assigned to program variables and expressions annotated with security levels. Security typing rules prevent leaks of secret information to public channels. For example, the security typing rule for assignment prevents H data from being assigned to a L variable. A well-typed program “protects secrets”, i.e., no information flows from H to L during program execution.

In the security literature, “protects secrets” is formalized as *noninterference* (Goguen and Meseguer, 1982) and is described in terms of an “indistinguishability” relation on states. Two program states are indistinguishable for L if they agree on values of L variables. The noninterference property says that any two runs of a program starting from two initial states that are indistinguishable for L, yield two final states that are indistinguishable for L. The two initial states may differ on values of H variables but not on values of L variables; the two final states must agree on the current values of L variables. One reading of the noninterference property is as a form of independence (Cohen, 1978): L output is independent of H inputs. It is this notion of independence that is made explicit in the information flow analysis specified in this article.

A shortcoming of usual type-based approaches for information flow (Banerjee and Naumann, 2005; Heintze and Riecke, 1998; Pottier and Simonet, 2003; Volpano and Smith, 1997) is that a type system can be too imprecise. Consider the sequential program $l := h ; l := 0$, where l has type L and h has type H. Although a programmer would never write such a program, it may arise naturally as a result of program transformation. The program is rejected by a security type system on account of the first assignment. But the program obviously satisfies noninterference – final states of any two runs of the program will always have the same value, 0, for l and are thus indistinguishable for L. Similarly, the program $h := l ; l := h$ is secure, yet is rejected by a security type system.

How can we admit such programs? Our inspiration comes from abstract interpretation (Cousot and Cousot, 1977a), which can be viewed as a method for statically computing approximations of program invariants (Cousot and

Cousot, 1977b). A benefit of this view is that the static abstraction of a program invariant can be used to annotate a program with pre- and postconditions and the annotated program can be checked against a Hoare-like logic. In information flow analysis, the invariant of interest is *independence*, for which we use the notation $[x \times w]$ to denote that x is independent of variable w . The intuition is this: a command denotes a *prelude transformer* (formalized in Section 2) that transforms a pre-prelude (i.e., a prelude before execution of the command) to a post-prelude (i.e., a prelude after execution of the command). A prelude, in turn, is a store transformer, that transforms an initial store to either a current store or \perp where \perp denotes nontermination. If x is a variable, then $[x \times w]$ holds for a prelude T provided for any two initial stores that differ only on the value of w , if T transforms them into current stores that are non- \perp then the current stores agree on the value of x . Alternatively, if x is \perp , then $[\perp \times w]$ holds for T provided for any two initial stores that differ only on the value of w , if T transforms them into two current stores then the one store is \perp if and only if the other store is. The intuition above is just a convenient restatement of noninterference but we tie it to the static notion of independence.

Our approach statically computes finite abstractions of the concrete preludes before and after the execution of a command. The notation $T^\#$ will be used to describe an abstraction of a concrete prelude T . This is formalized in Section 3. We formulate (in Section 4) a Hoare-like logic for checking independences and show (Section 5) that a checked program satisfies noninterference. The assertion language of the logic is decidable since it is just the language of finite sets of independences with subset inclusion. Specifications in the logic have the form $\{T_0^\#\} C \{T^\#\}$. Given precondition $T^\#$, we show in Section 7 how to compute strongest postconditions; for programs with loops, this necessitates a fixpoint computation³. The logic deems the program $l := h; l := 0$ secure (Example 4.1); the strongest postcondition of the program contains the independence $[l \times h]$.

Our approach falls in between type-based analysis and full verification. In the latter, verification conditions for loops depend on loop invariants generated by a theorem prover, typically using a fixpoint computation. Also in our setting, a fixpoint computation is employed to approximate loop invariants; these are lightweight in the sense that they do not put any constraints on the actual *values* of the program variables. Our approach is modular and we show that our logic satisfies a frame rule (Section 8). The frame rule permits local reasoning about a program: the relevant independences for a program are only those $[x \times w]$ where x occurs in the program. Moreover, in a larger context, the frame rule allows the following inference (in analogy with (O’Hearn et al.,

³ The set of independences is a finite lattice, hence the fixpoint computation will terminate.

2001)): start with a specification $\{\mathbb{T}_0^\#\} \mathbf{C} \{\mathbb{T}^\#\}$ describing independences before and after store modifications; then, $\{\mathbb{T}_0^\# \cup \mathbb{T}_1^\#\} \mathbf{C} \{\mathbb{T}^\# \cup \mathbb{T}_1^\#\}$ holds provided \mathbf{C} does not modify any variable \mathbf{y} where $[\mathbf{y} \times \mathbf{w}]$ appears in $\mathbb{T}_1^\#$. The initial specification $\{\mathbb{T}_0^\#\} \mathbf{C} \{\mathbb{T}^\#\}$, can reason with only the slice of store that \mathbf{C} touches.

Contributions. To summarize, this article makes three contributions. First and foremost, we formulate information flow analysis in a logical form via a Hoare-like logic. The approach deems more programs secure than extant type-based approaches. In Section 9, we show how our logic conservatively extends the security type system of Smith and Volpano (1997), by showing that any well-typed program in their system satisfies the invariant $[\mathbf{l} \times \mathbf{h}]$. Secondly, we describe the relationship between information flow and program dependence, explored in (Abadi et al., 1999; Hunt and Sands, 1991), in a more direct manner by computing independences between program variables. The independences themselves are static descriptions of the noninterference property. The development in this article considers *nontermination sensitive* noninterference: we assume that an attacker can observe nontermination. Finally, in Section 6, we show an application of our logic to forward slicing. Given a derivation of a program in the logic, with the information that variable \mathbf{l} is independent of variable \mathbf{h} , the slicing transformation systematically creates the forward \mathbf{l} -slice of the program: the slice contains all the commands independent of \mathbf{h} . We show that the slicing transformation is semantics preserving.

2 Language: Syntax, Preludes, Semantics

This section gives the syntax of a simple imperative language, formalizes the notion of preludes, and gives the semantics of the language in terms of preludes.

Syntax. We consider a simple imperative language with assignment, sequencing, conditionals and loops as formalized by the following BNF. Commands $\mathbf{C} \in \mathbf{Cmd}$ are given by the syntax

$$\mathbf{C} ::= \mathbf{x} := \mathbf{E} \mid \mathbf{C}_1 ; \mathbf{C}_2 \mid \mathbf{if} \ \mathbf{E} \ \mathbf{then} \ \mathbf{C}_1 \ \mathbf{else} \ \mathbf{C}_2 \mid \mathbf{while} \ \mathbf{E} \ \mathbf{do} \ \mathbf{C}$$

where \mathbf{Var} is an infinite set of variables, $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w} \in \mathbf{Var}$ range over variables, and where $\mathbf{E} \in \mathbf{Exp}$ ranges over expressions. Expressions are left unspecified but we shall assume the existence of a function $\text{fv}(\mathbf{E})$ that computes the free variables of expression \mathbf{E} . For commands, $\text{fv}(\mathbf{C})$ is defined in the obvious way.

For partial correctness – i.e., when we do not care about nontermination – we define, for $x \neq \perp$:

$$s_1 \stackrel{x}{=} s_2 \iff ((s_1 \neq \perp \wedge s_2 \neq \perp) \Rightarrow s_1 x = s_2 x)$$

For total correctness – i.e., when we are interested in nontermination – we additionally define, for $x = \perp$:

$$s_1 \stackrel{x}{=} s_2 \iff (s_1 = \perp \iff s_2 = \perp)$$

Fig. 1. When are two stores equal on x ?

We also define a function $modified : \mathbf{Cmd} \rightarrow \mathcal{P}(\mathbf{Var})$ that given a command, returns the set of variables potentially assigned to by the command.

$$\begin{aligned} modified(x := E) &= \{x\} \\ modified(C_1 ; C_2) &= modified(C_1) \cup modified(C_2) \\ modified(\mathbf{if } E \mathbf{ then } C_1 \mathbf{ else } C_2) &= modified(C_1) \cup modified(C_2) \\ modified(\mathbf{while } E \mathbf{ do } C) &= modified(C) \end{aligned}$$

In our examples, we shall often use a **skip** command; the formal treatment of that command (wrt. the semantics and the logic) should be obvious and we shall not bother to write down the rules explicitly.

Stores. A store, $s \in \mathbf{Sto}$, associates each variable with its current value; here values $v \in \mathbf{Val}$ are yet unspecified but we assume that there exists a predicate $true?$ on \mathbf{Val} . (For instance, we could have \mathbf{Val} as the set of integers and let $true?(v)$ be defined as $v \neq 0$). The following notation is used to denote a store update:

- $[s \mid y \mapsto v]$ returns a store s' with the property: for all $x \in \mathbf{Var}$, if $x \neq y$ then $s' x = s x$; but $s'(y) = v$.

Stores $s_1, s_2 \in \mathbf{Sto}_\perp$ agree on $x \in \mathbf{Var} \cup \{\perp\}$, written, $s_1 \stackrel{x}{=} s_2$, when either of the following conditions hold: (a) s_1, s_2, x are all non- \perp and then $s_1 x = s_2 x$; (b) $x \neq \perp$ and either $s_1 = \perp$ or $s_2 = \perp$ (or both); (c) $x = \perp$ and then $s_1 = \perp$ if and only if $s_2 = \perp$. This is made precise in Fig. 1 and motivated shortly in Section 3. Note that for $s_1, s_2 \in \mathbf{Sto}$ and $x \in \mathbf{Var}$, $s_1 \stackrel{x}{=} s_2$ amounts to $s_1 x = s_2 x$.

We write $s_1 \stackrel{w}{=} s_2$, where $s_1, s_2 \in \mathbf{Sto}$ and $w \in \mathbf{Var}$, to denote that for all variables $y \neq w$, $s_1 y = s_2 y$ holds. That is, the values of all variables, *except of* w , are equal in stores s_1 and s_2 .

Preludes. A prelude⁴, $T \in \mathbf{Prelude}$, maps an initial store to either a current store or \perp , where \perp denotes nontermination. Thus $\mathbf{Prelude} = \mathbf{Sto} \rightarrow \mathbf{Sto}_\perp$. Note that \mathbf{Sto}_\perp is a complete partial order (CPO) with the ordering \sqsubseteq defined as: $s_1 \sqsubseteq s_2$ iff either $s_1 = \perp$ or s_1 equals s_2 . Thus $\mathbf{Prelude}$ is a CPO, under the pointwise ordering: $T_1 \sqsubseteq T_2$ iff for all $s \in \mathbf{Sto}$, $T_1 s \sqsubseteq T_2 s$.

Semantics. For expressions, we assume there exists a semantic function $\llbracket E \rrbracket : \mathbf{Sto} \rightarrow \mathbf{Val}$ which satisfies the following property:

Property 2.1 *If for all $x \in \text{fv}(E)$ we have $s_1 x = s_2 x$, with $s_1, s_2 \in \mathbf{Sto}$, then $\llbracket E \rrbracket s_1 = \llbracket E \rrbracket s_2$.*

The definition of $\llbracket E \rrbracket$ would contain the clause $\llbracket x \rrbracket s = s x$. The semantics of a command has functionality $\llbracket C \rrbracket : \mathbf{Prelude} \rightarrow \mathbf{Prelude}$, and is defined in Fig. 2. To streamline the treatment of \perp , the metalanguage expression “let $\alpha = \beta$ in ...” denotes \perp if β is \perp .

Observe that since $\mathbf{Prelude}$ is a CPO also $\mathbf{Prelude} \rightarrow \mathbf{Prelude}$ is a CPO, with the following pointwise ordering: $f_1 \sqsubseteq f_2$ iff $f_1(T) \sqsubseteq f_2(T)$ for all $T \in \mathbf{Prelude}$. Let C be of the form **while** E **do** C_0 , and let \mathcal{F} be the corresponding function (called the *functor* of the **while** command) as defined in Fig. 2. The following calculation shows that \mathcal{F} is continuous on $\mathbf{Prelude} \rightarrow \mathbf{Prelude}$.

$$\begin{aligned} \mathcal{F}(\sqcup_i f_i) &= \lambda T. \lambda s. \text{let } s' = T s \text{ in (if } \text{true?}(\llbracket E \rrbracket s') \text{ then } (\sqcup_i f_i)(\llbracket C_0 \rrbracket T) s \text{ else } s') \\ &= \lambda T. \lambda s. \text{let } s' = T s \text{ in (if } \text{true?}(\llbracket E \rrbracket s') \text{ then } \sqcup_i (f_i(\llbracket C_0 \rrbracket T) s) \text{ else } s') \\ &= \sqcup_i (\lambda T. \lambda s. \text{let } s' = T s \text{ in (if } \text{true?}(\llbracket E \rrbracket s') \text{ then } f_i(\llbracket C_0 \rrbracket T) s \text{ else } s')) \\ &= \sqcup_i (F(f_i)) \end{aligned}$$

Hence the least fixpoint of \mathcal{F} is indeed well-defined, and

$$\llbracket C \rrbracket = \text{lfp}(\mathcal{F}) = \bigsqcup_{i \in \mathcal{N}} f_i$$

where f_i (called an *iterand* of the **while** command) is defined by:

⁴ We have learned from Dave Schmidt that the nomenclature is due to Bob Tennent.

$$\begin{aligned}
\llbracket x := E \rrbracket &= \lambda T. \lambda s. \text{let } s' = T s \text{ in } [s' \mid x \mapsto \llbracket E \rrbracket s'] \\
\llbracket C_1 ; C_2 \rrbracket &= \lambda T. \llbracket C_2 \rrbracket (\llbracket C_1 \rrbracket T) \\
\llbracket \text{if } E \text{ then } C_1 \text{ else } C_2 \rrbracket &= \lambda T. \lambda s. \text{let } s' = T s \text{ in} \\
&\quad \text{if } \text{true?}(\llbracket E \rrbracket s') \text{ then } \llbracket C_1 \rrbracket T s \text{ else } \llbracket C_2 \rrbracket T s \\
\llbracket \text{while } E \text{ do } C_0 \rrbracket &= \text{lf}_p(\mathcal{F}) \quad \text{where} \\
&\quad \mathcal{F} : \rightarrow (\mathbf{Prelude} \rightarrow \mathbf{Prelude}) \text{ is given by} \\
&\quad \mathcal{F}(f) = \lambda T. \lambda s. \text{let } s' = T s \text{ in} \\
&\quad \quad \text{if } \text{true?}(\llbracket E \rrbracket s') \text{ then } f(\llbracket C_0 \rrbracket T) s \text{ else } s'
\end{aligned}$$

Fig. 2. The Prelude Semantics.

$$\begin{aligned}
f_0 &= \lambda T. \lambda s. \perp \\
f_{i+1} &= \mathcal{F}(f_i)
\end{aligned}$$

Fact 2.2 *Let C be a **while** command, with $\{f_i\}$ the iterands of C . Then for all $T \in \mathbf{Prelude}$ and $s \in \mathbf{Sto}$ there exists j_0 such that for all $j \geq j_0$ we have $f_j T s = \llbracket C \rrbracket T s$.*

The intuition behind having a prelude semantics is as follows: Independences, as to be defined in Section 3, are with respect to the initial state of the whole computation. So to provide the meaning of a program fragment, we want to consider how a prelude (i.e., a function that describes how the state before execution of a command is reached from an initial state) is mapped to another (i.e., a function that describes how the state after execution of the command is reached from the initial state). For a concrete example, consider the computation of the program $x := x + y ; y := x + 1$ in an initial state that maps x to 0 and y to 2.

- Under a standard semantics, after execution of $x := x + y$, the new state maps x to 2 and leaves y unchanged. After execution of $y := x + 1$, the new state maps y to 3 and leaves x unchanged.
- In the prelude semantics, we have the following situation. Suppose the effect of executing $x := x + y$ is the prelude that maps the initial state $[x = 0, y = 2]$ to the state $[x = 2, y = 2]$. Then the effect of executing $y := x + 1$ under this prelude is a new prelude that maps the initial state $[x = 0, y = 2]$ to the state $[x = 2, y = 3]$.

Next some technical results about the semantic functions.

Definition 2.3 We say that a function $f \in \mathbf{Prelude} \rightarrow \mathbf{Prelude}$ is fully strict if for all $T \in \mathbf{Prelude}$ and $s \in \mathbf{Sto}$, $T s = \perp$ implies $f T s = \perp$.

We say that a fully strict function f preserves $z \in \mathbf{Var}$ if whenever $f T s \neq \perp$ (and thus $T s \neq \perp$) then $f T s z = T s z$.

Fact 2.4 For all commands C , the function $\llbracket C \rrbracket$ is monotone and fully strict. Also, all iterands f_i of **while** commands are monotone and fully strict.

PROOF. We go by structural induction on C . For **while**, we show that \mathcal{F} maps monotone functions into monotone functions and fully strict functions into fully strict functions; the result then follows (as f_0 is clearly monotone and fully strict) since the limit of monotone functions is itself monotone, and the limit of fully strict functions is itself fully strict.

Lemma 2.5 For all commands C , and all $z \in \mathbf{Var}$ with $z \notin \text{modified}(C)$,

- $\llbracket C \rrbracket$ preserves z ;
- if C is a **while** command then all its iterands f_i preserve z .

PROOF. Structural induction in C , with a case analysis. In all cases we are given T and s , and assume that $\llbracket C \rrbracket T s \neq \perp$ (and by Fact 2.4 thus $T s \neq \perp$); we must show that $\llbracket C \rrbracket T s z = T s z$.

$C = x := E$. From $z \notin \text{modified}(C)$ we infer that $z \neq x$, and the claim is trivial.

$C = C_1 ; C_2$. Since $z \notin \text{modified}(C_2)$, we infer inductively that $\llbracket C_2 \rrbracket (\llbracket C_1 \rrbracket T) s z = \llbracket C_1 \rrbracket T s z$. Since $z \notin \text{modified}(C_1)$, we infer inductively that $\llbracket C_1 \rrbracket T s z = T s z$. We thus get $\llbracket C \rrbracket T s z = \llbracket C_2 \rrbracket (\llbracket C_1 \rrbracket T) s z = T s z$, as desired.

$C = \text{if } E \text{ then } C_1 \text{ else } C_2$. Wlog. we can assume that $\text{true}?(\llbracket E \rrbracket (T s))$, in which case $\llbracket C \rrbracket T s z = \llbracket C_1 \rrbracket T s z$. Since $z \notin \text{modified}(C_1)$, we infer inductively that $\llbracket C_1 \rrbracket T s z = T s z$. This yields the claim.

$C = \text{while } E \text{ do } C_0$. Let f_i be the iterands of C ; our first task is to prove by induction in i that each f_i preserves z . For $i = 0$, the claim follows vacuously since $f_i T s = \perp$. For the inductive case, we split into two cases, in both cases assuming $f_{i+1} T s \neq \perp$ (and by Fact 2.4 thus $T s \neq \perp$):

- if $\text{true}?(\llbracket E \rrbracket (T s))$ then

$$f_{i+1} T s z = f_i (\llbracket C_0 \rrbracket T) s z = \llbracket C_0 \rrbracket T s z = T s z$$

where the second equality follows from the induction hypothesis on f_i , and the third equality follows from the overall induction hypothesis on C_0 (applicable since $z \notin \text{modified}(C_0)$).

- if $\text{true?}(\llbracket E \rrbracket(T s))$ does not hold, then $f_{i+1} T s z = T s z$ follows directly.

We are left with showing that also $\llbracket C \rrbracket$, given as the least upper bound of the iterands, preserves z . But since $\llbracket C \rrbracket T s \neq \perp$, we infer from Fact 2.2 that there exists a natural number j such that $\llbracket C \rrbracket T s = f_j T s$, yielding the desired $\llbracket C \rrbracket T s z = f_j T s z = T s z$.

3 Independences

The prelude semantics says that the meaning of a command is a prelude transformer: it transforms initial preludes into current preludes. In this section we will be interested in a finite abstraction of the preludes relevant to the execution of a command. The abstract preludes are termed *independences*: an independence $T^\# \in \mathbf{Independ} = \mathcal{P}((\mathbf{Var} \cup \{\perp\}) \times \mathbf{Var})$ is a set of pairs of the form $[x \times w]$. If x is a variable, then $[x \times w]$ denotes that the *current* value of x is independent of the *initial* value of w . If x is \perp , then *nontermination*, i.e., whether control reaches the current program point, is independent of w . This is formalized by the following definition which states the condition under which an independence correctly describes a concrete prelude.

Definition 3.1 *For all $T \in \mathbf{Prelude}$, for all $x \in \mathbf{Var} \cup \{\perp\}$, for all $w \in \mathbf{Var}$, $T \models [x \times w]$ holds iff for all $s_1, s_2 \in \mathbf{Sto}$: $s_1 \stackrel{w}{=} s_2$ implies $T s_1 \stackrel{x}{=} T s_2$.*

$T \models T^\#$ holds iff for all $[x \times w] \in T^\#$ it holds that $T \models [x \times w]$.

In the definition of $T \models [x \times w]$, note that s_1, s_2 are the “initial” stores and the antecedent of the implication asserts that these stores are equal except for w . The “current” stores are obtained via $T s_1$ and $T s_2$ and the consequent of the implication demands that the current stores be equal on x . Because x can be \perp , and because of the way $[\perp \times w]$ is defined, an analysis based on Definition 3.1 is *nontermination sensitive*. In particular, observe the following result:

Fact 3.2 *Assume that $T \models \{[x \times w], [\perp \times w]\}$. Then for all $s_1, s_2 \in \mathbf{Sto}$, $s_1 \stackrel{w}{=} s_2$ implies that either $T s_1 = T s_2 = \perp$, or $T s_1 \neq \perp$ and $T s_2 \neq \perp$ with $T s_1 x = T s_2 x$.*

Definition 3.3 *The ordering $T_1^\# \preceq T_2^\#$ holds iff $T_2^\# \subseteq T_1^\#$.*

The intuition behind the definition is that in $\mathbb{T}_1^\# \preceq \mathbb{T}_2^\#$, $\mathbb{T}_1^\#$ is more precise than $\mathbb{T}_2^\#$ because $\mathbb{T}_1^\#$ deems more variables independent than $\mathbb{T}_2^\#$. A motivation for the definition is the desire for a subtyping rule, stating that if $\mathbb{T}_1^\# \preceq \mathbb{T}_2^\#$ then $\mathbb{T}_1^\#$ can be replaced by $\mathbb{T}_2^\#$ (c.f. Fact 3.4). The subtyping rule is sound provided $\mathbb{T}_2^\#$ is a subset of $\mathbb{T}_1^\#$ and therefore obtainable from $\mathbb{T}_1^\#$ by removing information.

Clearly, **Independ** forms a complete lattice wrt. the ordering \preceq ; let $\sqcap_i \mathbb{T}_i^\#$ denote the greatest lower bound (which is the set union). The greatest element is the empty set of independences and the least element is **Independ**; the least element is obtained when all variables have constant values. We have some expected properties, where Fact 3.4 is used in the proof of the correctness theorem (Section 5), and where Fact 3.5 follows since if $[x \times w]$ belongs to $\sqcap_i \mathbb{T}_i^\#$ then it also belongs to some $\mathbb{T}_i^\#$.

Fact 3.4 *If $\mathbb{T} \models \mathbb{T}_1^\#$ and $\mathbb{T}_1^\# \preceq \mathbb{T}_2^\#$ then $\mathbb{T} \models \mathbb{T}_2^\#$.*

Fact 3.5 *If for all $i \in I$ it holds that $\mathbb{T} \models \mathbb{T}_i^\#$, then $\mathbb{T} \models \sqcap_{i \in I} \mathbb{T}_i^\#$.*

We have the following fact about the identity prelude.

Fact 3.6 *For all $x \in \mathbf{Var} \cup \{\perp\}$, for all $y \in \mathbf{Var}$, if $x \neq y$ then $\text{ls.s} \models [x \times y]$.*

An Abstract Interpretation. Facts 3.4 and 3.5 lead us to explore the connection of our framework with abstract interpretation (Cousot and Cousot, 1977a). We can write a function $\gamma : \mathbf{Independ} \rightarrow \mathcal{P}(\mathbf{Prelude})$:

$$\gamma(\mathbb{T}^\#) = \{\mathbb{T} \in \mathbf{Prelude} \mid \mathbb{T} \models \mathbb{T}^\#\}$$

and demonstrate that γ is completely multiplicative. We calculate:

$$\begin{aligned} \mathbb{T} \in \gamma(\sqcap_i \mathbb{T}_i^\#) &\Leftrightarrow \forall [x \times y] \in \sqcap_i \mathbb{T}_i^\# \bullet \mathbb{T} \models [x \times y] \\ &\Leftrightarrow \forall i \bullet \forall [x \times y] \in \mathbb{T}_i^\# \bullet \mathbb{T} \models [x \times y] \\ &\Leftrightarrow \forall i \bullet \mathbb{T} \in \gamma(\mathbb{T}_i^\#) \\ &\Leftrightarrow \mathbb{T} \in \bigcap_i \gamma(\mathbb{T}_i^\#) \end{aligned}$$

Therefore, γ can be considered a concretization function so that, by properties of Galois Connections, e.g., (Nielsen et al., 1999, Lemma 4.23), there exists an abstraction function $\alpha : \mathcal{P}(\mathbf{Prelude}) \rightarrow \mathbf{Independ}$:

$$\alpha(\mathbb{U}) = \bigcup \{\mathbb{T}^\# \mid \mathbb{U} \subseteq \gamma(\mathbb{T}^\#)\}$$

such that $(\mathcal{P}(\mathbf{Prelude}), \alpha, \gamma, \mathbf{Independ})$ is a Galois connection.

Equivalently, Facts 3.4 and 3.5 state that relation \models is \mathbf{U} -closed (upwards closed) and \mathbf{G} -closed (greatest lower bound closed) respectively. Therefore, (see, e.g., (Schmidt, 2002)) \models defines a Galois connection. It follows that by defining $\gamma(\mathbf{T}^\#)$ as above, we have the properties:

- γ is a monotone function.
- For all $\mathbf{T}^\# \in \mathbf{Independ}$, for all $\mathbf{U} \in \mathcal{P}(\mathbf{Prelude})$, $\mathbf{U} \subseteq \gamma(\mathbf{T}^\#)$ iff $\alpha(\mathbf{U}) \preceq \mathbf{T}^\#$.

The second item above is the definition of a Galois connection and in this sense, \models “is” a Galois connection.

4 Static Checking of Independences

To statically check independences we define, in Fig. 3, a Hoare-like logic where judgements are of the form

$$\mathbf{G} \vdash \{\mathbf{T}_0^\#\} \mathbf{C} \{\mathbf{T}^\#\}$$

The judgement is interpreted as saying that if the independences in $\mathbf{T}_0^\#$ hold *before* execution of \mathbf{C} then the independences in $\mathbf{T}^\#$ will hold *after* execution of \mathbf{C} . The context $\mathbf{G} \in \mathbf{Context} = \mathcal{P}(\mathbf{Var})$ is a *control dependence*, denoting (a superset of) the variables that at least one test surrounding \mathbf{C} depends on. For example, in **if** x **then** $y := 0$ **else** $z := 1$, the static checking of $y := 0$ takes place in the context that contains all variables that x is dependent on. This is crucial, especially since x may depend on a high variable.

We now explain a few of the rules in Fig. 3. Checking an assignment, $x := E$, in context \mathbf{G} , involves checking any $[y \times w]$ in the postcondition $\mathbf{T}^\#$. There are two cases. If $x \neq y$, then $[y \times w]$ must also appear in the precondition $\mathbf{T}_0^\#$. Likewise, if $[\perp \times w]$ appears in the postcondition $\mathbf{T}^\#$, then it must also appear in the precondition $\mathbf{T}_0^\#$. Otherwise, if $x = y$ then $[x \times w]$ appears in the postcondition provided all variables referenced in E are independent of w ; moreover, w must not appear in \mathbf{G} , as otherwise, x would be control dependent on w .

Checking a conditional, **if** E **then** C_1 **else** C_2 , involves checking C_1 and C_2 in a context \mathbf{G}_0 that includes not only the “old” context \mathbf{G} but also the variables that E depends on (as variables modified in C_1 or C_2 will be control dependent on the variables that E depends on.) Equivalently, if w is not in \mathbf{G}_0 , then all free variables x in E must be independent of w , that is, $[x \times w]$ must appear in the precondition $\mathbf{T}_0^\#$.

$$\begin{array}{c}
\text{if } \forall [\mathbf{y} \times \mathbf{w}] \in \mathsf{T}^\# \bullet \\
[\text{Assign}] \quad \frac{G \vdash \{\mathsf{T}_0^\#\} \mathbf{x} := \mathsf{E} \{\mathsf{T}^\#\} \quad (x \neq \mathbf{y} \vee \mathbf{y} = \perp) \Rightarrow [\mathbf{y} \times \mathbf{w}] \in \mathsf{T}_0^\#}{x = \mathbf{y} \Rightarrow (\mathbf{w} \notin G \wedge \forall z \in \text{fv}(\mathsf{E}) \bullet [z \times \mathbf{w}] \in \mathsf{T}_0^\#)} \\
\\
[\text{Seq}] \quad \frac{G \vdash \{\mathsf{T}_0^\#\} \mathsf{C}_1 \{\mathsf{T}_1^\#\} \quad G \vdash \{\mathsf{T}_1^\#\} \mathsf{C}_2 \{\mathsf{T}_2^\#\}}{G \vdash \{\mathsf{T}_0^\#\} \mathsf{C}_1 ; \mathsf{C}_2 \{\mathsf{T}_2^\#\}} \\
\\
[\text{If}] \quad \frac{G_0 \vdash \{\mathsf{T}_0^\#\} \mathsf{C}_1 \{\mathsf{T}^\#\} \quad G_0 \vdash \{\mathsf{T}_0^\#\} \mathsf{C}_2 \{\mathsf{T}^\#\}}{G \vdash \{\mathsf{T}_0^\#\} \text{if } \mathsf{E} \text{ then } \mathsf{C}_1 \text{ else } \mathsf{C}_2 \{\mathsf{T}^\#\}} \quad \text{if } \quad G \subseteq G_0 \wedge \\
\quad (\mathbf{w} \notin G_0 \Rightarrow \forall x \in \text{fv}(\mathsf{E}) \bullet [x \times \mathbf{w}] \in \mathsf{T}_0^\#) \\
\\
[\text{While}] \quad \frac{G_0 \vdash \{\mathsf{T}^\#\} \mathsf{C} \{\mathsf{T}^\#\}}{G \vdash \{\mathsf{T}^\#\} \text{while } \mathsf{E} \text{ do } \mathsf{C} \{\mathsf{T}^\#\}} \quad \text{if } \quad G \subseteq G_0 \wedge \\
\quad (\mathbf{w} \notin G_0 \Rightarrow \forall x \in \text{fv}(\mathsf{E}) \bullet [x \times \mathbf{w}] \in \mathsf{T}^\#) \wedge \\
\quad ([\perp \times \mathbf{w}] \in \mathsf{T}^\# \Rightarrow \mathbf{w} \notin G_0) \\
\\
[\text{Sub}] \quad \frac{G_1 \vdash \{\mathsf{T}_1^\#\} \mathsf{C} \{\mathsf{T}_2^\#\}}{G_0 \vdash \{\mathsf{T}_0^\#\} \mathsf{C} \{\mathsf{T}_3^\#\}} \quad \text{if } (\mathsf{T}_0^\# \preceq \mathsf{T}_1^\#) \wedge (\mathsf{T}_2^\# \preceq \mathsf{T}_3^\#) \wedge (G_0 \subseteq G_1)
\end{array}$$

Fig. 3. The Logic.

Checking a while loop is similar to checking a conditional. The only difference is that it requires guessing an “invariant” $\mathsf{T}^\#$ that is both the precondition and the postcondition of the loop and its body. With respect to non-termination, note that if $\mathbf{w} \in G_0$ then \mathbf{w} may influence the termination of the program: either directly, as in **while** $w > 7$ **do** $w := w + 1$ (where the side condition for [While] forces \mathbf{w} to be added to G_0); or indirectly, as in **if** $w > 7$ **then skip else while true do skip** (where the side condition for [If] has already added \mathbf{w} to G). Therefore we demand that if $[\perp \times \mathbf{w}] \in \mathsf{T}^\#$ then \mathbf{w} must not be in G_0 .

In Section 7, when we define *strongest postcondition*, we will select $G_0 = G \cup \{\mathbf{w} \mid \exists x \in \text{fv}(\mathsf{E}) \bullet [x \times \mathbf{w}] \notin \mathsf{T}_0^\#\}$ for the conditional and the while loop. Instead of guessing the invariant, we will show how to compute it using fix-points.

Example 4.1 (*Illustrating “recovery” of independences.*) We have the derivations

$$\begin{array}{l}
\emptyset \vdash \{[\mathbf{l} \times \mathbf{h}], [\mathbf{h} \times \mathbf{l}]\} \mathbf{l} := \mathbf{h} \{[\mathbf{h} \times \mathbf{l}], [\mathbf{l} \times \mathbf{l}]\} \text{ and} \\
\emptyset \vdash \{[\mathbf{h} \times \mathbf{l}], [\mathbf{l} \times \mathbf{l}]\} \mathbf{l} := \mathbf{0} \{[\mathbf{h} \times \mathbf{l}], [\mathbf{l} \times \mathbf{l}], [\mathbf{l} \times \mathbf{h}]\}
\end{array}$$

and therefore also

$$\emptyset \vdash \{[l \times h], [h \times l]\} l := h ; l := 0 \{[h \times l], [l \times l], [l \times h]\}.$$

With the intuition that l stands for “low” or “public” and h stands for “high” or “sensitive”, the derivation asserts that if l is independent of h before execution, then l is independent of h after execution. Thus $[l \times h]$ is an *invariant* of the computation. By Definition 3.1, any prelude of the program applied to initial stores that differ only in the value for h , creates new stores that agree on the current value for l . Thus the program is secure, although it contains an insecure sub-program: the independence $[l \times h]$ is lost after $l := h$, but recovered after $l := 0$.

Example 4.2 (*Illustrating control dependence.*) The reader may check that the following informally annotated program gives rise to a derivation in our logic. Initially, G is empty, and all variables are pairwise independent; we write $[x \times y, z]$ to abbreviate $[x \times y], [x \times z]$.

$$\begin{array}{ll} & \{[l \times h, x], [h \times l, x], [x \times l, h]\} \\ x := h & \{[l \times h, x], [h \times l, x], [x \times l, x]\} \\ \text{if } x > 0 & (G \text{ is now } \{h\}) \\ \quad \text{then } l := 7 & \{[l \times x, l], [h \times l, x], [x \times l, x]\} \\ \quad \text{else } x := 0 & \{[l \times h, x], [h \times l, x], [x \times l, x]\} \\ \text{end of if} & \{[l \times x], [h \times l, x], [x \times l, x]\} \end{array}$$

A few remarks:

- in the preamble, only x is assigned, so the independences for l and h are carried through, but $[x \times l, x]$ holds afterwards, as $[h \times l, x]$ holds beforehand;
- the free variable in the guard is independent of l and x but not of h , implying that h has to be in G .

Example 4.3 (*Illustrating how the logic works with nontermination.*)

With $P = \mathbf{while} \ l \neq 0 \ \mathbf{do} \ h := 7$ and $T^\# = \{[\perp \times h], [l \times h]\}$ we have the judgement $\emptyset \vdash \{T^\#\} P \{T^\#\}$ (since $\{l\} \vdash \{T^\#\} h := 7 \{T^\#\}$) showing that P is deemed secure by our logic; an observer able to observe even nontermination cannot detect the initial value of h . (The reason why $[h \times l]$ and $[\perp \times l]$ are not in $T^\#$ is that h clearly depends on l and nontermination depends on l .)

However, with $P = \mathbf{while} \ h \neq 0 \ \mathbf{do} \ h := 7$ and $T^\# = \{[\perp \times h], [l \times h], [h \times l]\}$ we do not have a derivation, even though $T^\#$ is an invariant for $h := 7$. But be-

cause the derivation requires $\mathbf{h} \in \mathbf{G}_0$, $[\perp \times \mathbf{h}]$ can no longer be in $\mathbf{T}^\#$. This suggests that \mathbf{P} is insecure when nontermination is observable: indeed, an observer able to observe nontermination can detect whether \mathbf{h} was initially 0 or not. Interestingly, if nontermination is *not* observable, then $\mathbf{T}^\# = \{[\perp \times \mathbf{h}], [\mathbf{h} \times \perp]\}$ remains invariant, as expected. In particular, the current value of \perp is independent of the initial value of \mathbf{h} . Similarly, we do not have a derivation for $\mathbf{P} = \text{if } \mathbf{h} \neq 0 \text{ then skip else while true do skip}$ with $\mathbf{T}^\#$ containing $[\perp \times \mathbf{h}]$.

A couple of simple results can be proven about the logic in Fig. 3:

Fact 4.4 *Assume $\mathbf{G} \vdash \{\mathbf{T}_0^\#\} \mathbf{C} \{\mathbf{T}^\#\}$. If $[\perp \times \mathbf{w}] \in \mathbf{T}^\#$ then $[\perp \times \mathbf{w}] \in \mathbf{T}_0^\#$.*

The proof is an easy induction on a derivation of $\mathbf{G} \vdash \{\mathbf{T}_0^\#\} \mathbf{C} \{\mathbf{T}^\#\}$.

Lemma 4.5 *Assume $\mathbf{G} \vdash \{\mathbf{T}_0^\#\} \mathbf{C} \{\mathbf{T}^\#\}$ with $[\perp \times \mathbf{w}] \in \mathbf{T}^\#$ and $\mathbf{w} \in \mathbf{G}$. Then $\mathbf{T} \mathbf{s} \neq \perp$ implies $\llbracket \mathbf{C} \rrbracket \mathbf{T} \mathbf{s} \neq \perp$.*

PROOF. An easy induction in the derivation, where for [Seq] we use Fact 4.4; for [While] the result follows vacuously since having both $[\perp \times \mathbf{w}] \in \mathbf{T}^\#$ and $\mathbf{w} \in \mathbf{G}$ violates the side conditions of that rule.

To restate Fact 4.4, if $[\perp \times \mathbf{w}] \notin \mathbf{T}_0^\#$, then $[\perp \times \mathbf{w}]$ cannot be recovered in $\mathbf{T}^\#$. This is in contrast to Example 4.1, which shows that although $[\perp \times \mathbf{h}]$ does not appear in the precondition of $\perp := 0$, it could nonetheless be recovered in the postcondition of $\perp := 0$. In other words, once it is established that nontermination may depend on \mathbf{w} , it continues to depend on \mathbf{w} . This is what we should expect, as otherwise, with \mathbf{w} interpreted “high”, nontermination can be used to leak the high value.

In a judgement $\mathbf{G} \vdash \{\mathbf{T}_0^\#\} \mathbf{C} \{\mathbf{T}^\#\}$, suppose $\mathbf{w} \in \mathbf{G}$. This means that any assignment in \mathbf{C} is control dependent on \mathbf{w} . Suppose now that \mathbf{y} is a variable and is independent of \mathbf{w} in the postcondition $\mathbf{T}^\#$. This implies that \mathbf{y} cannot be assigned to in \mathbf{C} — otherwise, it would be dependent on \mathbf{w} . If \mathbf{y} is not assigned to in \mathbf{C} , then \mathbf{y} must be independent of \mathbf{w} in the precondition too. These intuitions are collected together in Lemma 4.6 below. Note that with \mathbf{y} interpreted as “low” and \mathbf{w} as “high”, the lemma essentially says that low variables may not be written to under a high guard. Thus the lemma is the counterpart of the “no write down” rule that underlies information flow control; the term “*-property” (Bell and LaPadula, 1973) is also used. The value of low variables remains the same after execution of \mathbf{C} .

Lemma 4.6 (Write Confinement) *Assume that $\mathbf{G} \vdash \{\mathbf{T}_0^\#\} \mathbf{C} \{\mathbf{T}^\#\}$. Then the following holds:*

If $[\mathbf{y} \times \mathbf{w}] \in \mathsf{T}^\#$ and $\mathbf{y} \in \mathbf{Var}$ and $\mathbf{w} \in \mathsf{G}$
then $\mathbf{y} \notin \text{modified}(\mathsf{C})$ and $[\mathbf{y} \times \mathbf{w}] \in \mathsf{T}_0^\#$.

PROOF. We perform induction in the derivation of $\mathsf{G} \vdash \{\mathsf{T}_0^\#\} \mathsf{C} \{\mathsf{T}^\#\}$, and do a case analysis on the last rule applied:

[Assign], with $\mathsf{C} = \mathbf{x} := \mathbf{E}$. If $\mathbf{x} = \mathbf{y}$, then $\mathbf{w} \notin \mathsf{G}$, contradicting our assumptions. If $\mathbf{x} \neq \mathbf{y}$, then $\mathbf{y} \notin \text{modified}(\mathsf{C})$ and $[\mathbf{y} \times \mathbf{w}] \in \mathsf{T}_0^\#$.

[Seq], with $\mathsf{C} = \mathsf{C}_1 ; \mathsf{C}_2$. Assume that $\mathsf{G} \vdash \{\mathsf{T}_0^\#\} \mathsf{C}_1 \{\mathsf{T}_1^\#\}$ and $\mathsf{G} \vdash \{\mathsf{T}_1^\#\} \mathsf{C}_2 \{\mathsf{T}^\#\}$ and also assume that $\mathbf{w} \in \mathsf{G}$. By applying the induction hypothesis to the latter judgement, we see that $\mathbf{y} \notin \text{modified}(\mathsf{C}_2)$ and that $[\mathbf{y} \times \mathbf{w}] \in \mathsf{T}_1^\#$. By then applying the induction hypothesis to the former judgement, we see that $\mathbf{y} \notin \text{modified}(\mathsf{C}_1)$ and that $[\mathbf{y} \times \mathbf{w}] \in \mathsf{T}_0^\#$. Therefore $\mathbf{y} \notin \text{modified}(\mathsf{C})$, as desired.

[If], with $\mathsf{C} = \text{if } \mathbf{E} \text{ then } \mathsf{C}_1 \text{ else } \mathsf{C}_2$. Assume that

$$\mathsf{G}_0 \vdash \{\mathsf{T}_0^\#\} \mathsf{C}_1 \{\mathsf{T}^\#\} \text{ and } \mathsf{G}_0 \vdash \{\mathsf{T}_0^\#\} \mathsf{C}_2 \{\mathsf{T}^\#\}$$

where $\mathsf{G} \subseteq \mathsf{G}_0$. Let $[\mathbf{y} \times \mathbf{w}] \in \mathsf{T}^\#$ with $\mathbf{w} \in \mathsf{G}$, then $\mathbf{w} \in \mathsf{G}_0$ so by applying the induction hypothesis we get $\mathbf{y} \notin \text{modified}(\mathsf{C}_1)$ and $\mathbf{y} \notin \text{modified}(\mathsf{C}_2)$ and $[\mathbf{y} \times \mathbf{w}] \in \mathsf{T}_0^\#$. This implies $\mathbf{y} \notin \text{modified}(\mathsf{C})$, and thereby the desired result.

[While], with $\mathsf{C} = \text{while } \mathbf{E} \text{ do } \mathsf{C}_0$. Our assumptions are that $\mathsf{G} \vdash \{\mathsf{T}^\#\} \mathsf{C} \{\mathsf{T}^\#\}$ because with $\mathsf{G} \subseteq \mathsf{G}_0$ we have $\mathsf{G}_0 \vdash \{\mathsf{T}^\#\} \mathsf{C}_0 \{\mathsf{T}^\#\}$. Let $[\mathbf{y} \times \mathbf{w}] \in \mathsf{T}^\#$ with $\mathbf{w} \in \mathsf{G}$, then $\mathbf{w} \in \mathsf{G}_0$ so by applying the induction hypothesis we get $\mathbf{y} \notin \text{modified}(\mathsf{C}_0)$ (and trivially $[\mathbf{y} \times \mathbf{w}] \in \mathsf{T}^\#$) which is as desired.

[Sub]. Assume that $\mathsf{G} \vdash \{\mathsf{T}_0^\#\} \mathsf{C} \{\mathsf{T}^\#\}$ because with $\mathsf{G} \subseteq \mathsf{G}_0$ and $\mathsf{T}_0^\# \preceq \mathsf{T}_1^\#$ and $\mathsf{T}_2^\# \preceq \mathsf{T}^\#$ we have $\mathsf{G}_0 \vdash \{\mathsf{T}_1^\#\} \mathsf{C} \{\mathsf{T}_2^\#\}$. Also assume that $[\mathbf{y} \times \mathbf{w}] \in \mathsf{T}^\#$ and that $\mathbf{w} \in \mathsf{G}$. Then $[\mathbf{y} \times \mathbf{w}] \in \mathsf{T}_2^\#$ and $\mathbf{w} \in \mathsf{G}_0$, so inductively we obtain $\mathbf{y} \notin \text{modified}(\mathsf{C})$ and $[\mathbf{y} \times \mathbf{w}] \in \mathsf{T}_1^\#$. Then also $[\mathbf{y} \times \mathbf{w}] \in \mathsf{T}_0^\#$, as desired.

5 Correctness

We are now in a position to prove the correctness of the logic with respect to the prelude semantics.

Theorem 5.1 *Assume that*

$G \vdash \{\mathbb{T}_0^\#\} C \{\mathbb{T}^\#\}$ where for all $[x \times y] \in \mathbb{T}_0^\#$ it is the case that $x \neq y$.
Then $\llbracket C \rrbracket(\lambda s.s) \models \mathbb{T}^\#$.

That is, $\mathbb{T}^\#$ correctly describes the concrete prelude obtained by executing command C on $\lambda s.s$, the initial prelude.

The correctness theorem can be seen as the noninterference theorem for information flow. Indeed, consider the prelude, \mathbb{T} , obtained from executing the command C with the initial prelude, $\lambda s.s$. With l and h interpreted as “low” and “high” respectively, suppose $[l \times h]$ appears in $\mathbb{T}^\#$. Then for any two stores s_1, s_2 that differ only on h , the stores $\mathbb{T} s_1$ and $\mathbb{T} s_2$ must agree on the value of l , meaning that if both $\mathbb{T} s_1$ and $\mathbb{T} s_2$ terminate then they must not give different values for l . On the other hand, it is possible for either $\mathbb{T} s_1$ or $\mathbb{T} s_2$ or both to be \perp , as in the example

if h then $h := 7$ else (while true do skip).

which has $[l \times h]$ as an invariant.

Moreover, the correctness theorem says that if $[\perp \times h]$ appears in $\mathbb{T}^\#$, then $\mathbb{T} s_1 = \perp$ iff $\mathbb{T} s_2 = \perp$: thus noninterference is preserved under nontermination.

To prove Theorem 5.1, we need the following main lemma. Then the theorem follows by substituting \mathbb{T} by $\lambda s.s$, and then using Fact 3.6.

Lemma 5.2 *If $G \vdash \{\mathbb{T}_0^\#\} C \{\mathbb{T}^\#\}$ and $\mathbb{T} \models \mathbb{T}_0^\#$, then $\llbracket C \rrbracket \mathbb{T} \models \mathbb{T}^\#$.*

PROOF. We perform induction on a derivation of $G \vdash \{\mathbb{T}_0^\#\} C \{\mathbb{T}^\#\}$, and do a case analysis on the last rule applied.

[Assign], with $C = x := E$. We are given $[z \times w] \in \mathbb{T}^\#$, and we consider $s_1, s_2 \in \mathbf{Sto}$ such that $s_1 \stackrel{w}{=} s_2$. With $\mathbb{T}' = \llbracket C \rrbracket \mathbb{T}$, we must show that $\mathbb{T}' s_1 \stackrel{z}{=} \mathbb{T}' s_2$.

If $z = \perp$, then $[z \times w] \in \mathbb{T}_0^\#$, so $\mathbb{T} s_1 \stackrel{\perp}{=} \mathbb{T} s_2$. Since for all s , $\mathbb{T}' s = \perp$ iff $\mathbb{T} s = \perp$, this shows that also $\mathbb{T}' s_1 \stackrel{\perp}{=} \mathbb{T}' s_2$, as desired.

So now consider $z \in \mathbf{Var}$. If either $\mathbb{T} s_1 = \perp$ or $\mathbb{T} s_2 = \perp$, also $\mathbb{T}' s_1 = \perp$ or $\mathbb{T}' s_2 = \perp$, and the claim is trivial. We thus assume that $\mathbb{T} s_1 \neq \perp$ and $\mathbb{T} s_2 \neq \perp$. We have two subcases:

- (1) $z \neq x$: Then $[z \times w] \in \mathbb{T}_0^\#$, so $\mathbb{T} s_1 \stackrel{z}{=} \mathbb{T} s_2$, that is, $\mathbb{T} s_1 z = \mathbb{T} s_2 z$. Since for all s we have $\mathbb{T}' s z = \mathbb{T} s z$, this shows that also $\mathbb{T}' s_1 z = \mathbb{T}' s_2 z$, as desired.

- (2) $z = x$: Then for all $\mathbf{y} \in \text{fv}(E)$, we have $[\mathbf{y} \times \mathbf{w}] \in T_0^\#$ and therefore $T s_1 \mathbf{y} = T s_2 \mathbf{y}$. By Property 2.1 then $\llbracket E \rrbracket(T s_1) = \llbracket E \rrbracket(T s_2)$, showing that $T' s_1 x = T' s_2 x$, as desired.

[Seq], with $C = C_1 ; C_2$. Assume that

$$G \vdash \{T_0^\#\} C_1 \{T_1^\#\} \text{ and that } G \vdash \{T_1^\#\} C_2 \{T_2^\#\}.$$

By applying the induction hypothesis to the first judgement, we get

$$\llbracket C_1 \rrbracket(T) \models T_1^\#.$$

We then apply the induction hypothesis to the second judgement and get the desired result:

$$\llbracket C_2 \rrbracket(\llbracket C_1 \rrbracket(T)) \models T^\#.$$

[If], with $C = \text{if } E \text{ then } C_1 \text{ else } C_2$. Assume that

$$G_0 \vdash \{T_0^\#\} C_1 \{T^\#\} \text{ and } G_0 \vdash \{T_0^\#\} C_2 \{T^\#\}$$

where $w \notin G_0$ implies that $\forall x \in \text{fv}(E) \bullet [x \times w] \in T_0^\#$. Inductively, we can assume that $\llbracket C_1 \rrbracket T \models T^\#$ and $\llbracket C_2 \rrbracket T \models T^\#$. We are given $[z \times w] \in T^\#$ (where z might be \perp), and we consider $s_1, s_2 \in \mathbf{Sto}$ such that $s_1 \xrightarrow{w} s_2$. With $T' = \llbracket C \rrbracket T$, we must show that $T' s_1 \stackrel{z}{\cong} T' s_2$.

First assume that at least one of $T s_1$ or $T s_2$ is \perp . By Fact 2.4, at least as many of $T' s_1$ and $T' s_2$ will be \perp . So if $z \in \mathbf{Var}$, then trivially $T' s_1 \stackrel{z}{\cong} T' s_2$. If $z = \perp$, then by Fact 4.4 we know that $[\perp \times w] \in T_0^\#$, so from $T \models T_0^\#$ we get $T s_1 \stackrel{\perp}{=} T s_2$ and therefore $T s_1 = T s_2 = \perp$. But then also $T' s_1 = T' s_2 = \perp$, showing $T' s_1 \stackrel{\perp}{=} T' s_2$.

We can thus assume that there exists stores $s'_1, s'_2 \neq \perp$ such that $s'_1 = T s_1$ and $s'_2 = T s_2$. Apart from symmetry, there are two cases:

true?($\llbracket E \rrbracket s'_1$) and ***true?***($\llbracket E \rrbracket s'_2$). Then $T' s_1 = \llbracket C_1 \rrbracket T s_1$ and $T' s_2 = \llbracket C_1 \rrbracket T s_2$. From $\llbracket C_1 \rrbracket T \models T^\#$ and $[z \times w] \in T^\#$ we infer that $\llbracket C_1 \rrbracket T s_1 \stackrel{z}{\cong} \llbracket C_1 \rrbracket T s_2$, which amounts to the desired $T' s_1 \stackrel{z}{\cong} T' s_2$.

true?($\llbracket E \rrbracket s'_1$) but not ***true?***($\llbracket E \rrbracket s'_2$). We claim

$$w \in G_0$$

and prove the claim by contradiction: suppose $w \notin G_0$. Then by the logic, for all $x \in \text{fv}(E)$, $[x \times w] \in T_0^\#$ implying $T s_1 \stackrel{x}{\cong} T s_2$, that is $s'_1(x) = s'_2(x)$. By

Property 2.1 this implies $\llbracket E \rrbracket s'_1 = \llbracket E \rrbracket s'_2$ – a contradiction.

Having established $w \in G_0$, consider two cases. First assume $z = \perp$. Then (by Lemma 4.5 applied to the sub-derivations) we infer that $\llbracket C_1 \rrbracket T s_1 \neq \perp$ and $\llbracket C_2 \rrbracket T s_2 \neq \perp$, and thus $T' s_1 \neq \perp$ and $T' s_2 \neq \perp$. But this yields the desired relation $T' s_1 \stackrel{\perp}{=} T' s_2$.

Now assume that $z \in \mathbf{Var}$. By Lemma 4.6 applied to the sub-derivations, we infer that

$$[z \times w] \in T_0^\# \text{ and } z \notin \text{modified}(C_1) \text{ and } z \notin \text{modified}(C_2).$$

Since $T \models T_0^\#$, this shows $T s_1 \stackrel{z}{=} T s_2$, that is, $T s_1 z = T s_2 z$. We know that $T' s_1 = \llbracket C_1 \rrbracket T s_1$ and $T' s_2 = \llbracket C_2 \rrbracket T s_2$. If either is \perp , the claim vacuously holds; so assume neither is \perp . By Lemma 2.5, $T s_1 z = \llbracket C_1 \rrbracket T s_1 z$ and $T s_2 z = \llbracket C_2 \rrbracket T s_2 z$. This shows the desired $T' s_1 z = T s_1 z = T s_2 z = T' s_2 z$.

[While], with $C = \mathbf{while} E \mathbf{do} C_0$. Our assumptions are that

$$G \vdash \{T^\#\} C \{T^\#\}$$

because with G_0 such that $w \notin G_0$ implies that $\forall x \in \text{fv}(E) \bullet [x \times w] \in T^\#$, and such that $[\perp \times w] \in T^\#$ implies $w \notin G_0$, we have

$$G_0 \vdash \{T^\#\} C_0 \{T^\#\}.$$

We define an auxiliary predicate \mathcal{P} :

$$\mathcal{P}(f) \Leftrightarrow \forall T \bullet (T \models T^\# \Rightarrow f T \models T^\#)$$

We shall establish

$$\forall i \geq 0 \bullet \mathcal{P}(f_i) \tag{1}$$

and do so by induction in i . For the base case, note that $f_0(T) = \lambda s. \perp$, and that $\lambda s. \perp \models T^\#$ always holds because no matter whether $z = \perp$ or not, $\perp \stackrel{z}{=} \perp$.

For the inductive case, we assume that $T \models T^\#$ and must prove $\mathcal{F} f T \models T^\#$, with f an iterand. So let $[z \times w] \in T^\#$ and $s_1 \stackrel{w}{=} s_2$; we know that $T s_1 \stackrel{z}{=} T s_2$ and must prove $\mathcal{F} f T s_1 \stackrel{z}{=} \mathcal{F} f T s_2$.

First assume that at least one of $T s_1$ or $T s_2$ is \perp . By Fact 2.4, at least as many of $\mathcal{F} f T s_1$ and $\mathcal{F} f T s_2$ will be \perp . So if $z \in \mathbf{Var}$, then trivially $\mathcal{F} f T s_1 \stackrel{z}{=} \mathcal{F} f T s_2$. If $z = \perp$, then from $[\perp \times w] \in T^\#$ and $T \models T^\#$ we get $T s_1 \stackrel{\perp}{=} T s_2$ and therefore $T s_1 = T s_2 = \perp$. But then also $\mathcal{F} f T s_1 =$

$\mathcal{F} \text{ f t } s_2 = \perp$, showing $\mathcal{F} \text{ f T } s_1 \stackrel{\perp}{=} \mathcal{F} \text{ f T } s_2$.

We can thus assume that there exists stores $s'_1, s'_2 \neq \perp$ such that $s'_1 = \text{T } s_1$ and $s'_2 = \text{T } s_2$. Apart from symmetry, there are three cases:

***true?*($\llbracket \text{E} \rrbracket s'_1$) and *true?*($\llbracket \text{E} \rrbracket s'_2$).** Then $\mathcal{F} \text{ f T } s_1 = f(\llbracket \text{C}_0 \rrbracket \text{T})s_1$ and $\mathcal{F} \text{ f T } s_2 = f(\llbracket \text{C}_0 \rrbracket \text{T})s_2$. By applying the overall induction hypothesis on $G_0 \vdash \{\text{T}^\#\} \text{C}_0 \{\text{T}^\#\}$, we get $\llbracket \text{C}_0 \rrbracket \text{T} \models \text{T}^\#$; by applying the innermost induction hypothesis, we then get $f(\llbracket \text{C}_0 \rrbracket \text{T}) \models \text{T}^\#$, so that $f(\llbracket \text{C}_0 \rrbracket \text{T})s_1 \stackrel{\cong}{=} f(\llbracket \text{C}_0 \rrbracket \text{T})s_2$ which amounts to the desired $\mathcal{F} \text{ f T } s_1 \stackrel{\cong}{=} \mathcal{F} \text{ f T } s_2$.

not *true?*($\llbracket \text{E} \rrbracket s'_1$) and not *true?*($\llbracket \text{E} \rrbracket s'_2$). Then $\mathcal{F} \text{ f T } s_1 = \text{T } s_1$ and $\mathcal{F} \text{ f T } s_2 = \text{T } s_2$, and the claim is trivial.

***true?*($\llbracket \text{E} \rrbracket s'_1$) but not *true?*($\llbracket \text{E} \rrbracket s'_2$).** If $w \notin G_0$, then by the logic, for all $x \in \text{fv}(\text{E})$, $[x \times w] \in \text{T}^\#$ implying $\text{T } s_1 \stackrel{\cong}{=} \text{T } s_2$, that is $s'_1(x) = s'_2(x)$. By Property 2.1 this implies $\llbracket \text{E} \rrbracket s'_1 = \llbracket \text{E} \rrbracket s'_2$ – a contradiction. Thus $w \in G_0$. Since $[z \times w] \in \text{T}^\#$, we infer from the logic that $z \neq \perp$. Thus, z is a variable.

By Lemma 4.6 (applied to the sub-derivation), we infer that $z \notin \text{modified}(\text{C}_0)$. We know that $\mathcal{F} \text{ f T } s_1 = f(\llbracket \text{C}_0 \rrbracket \text{T})s_1$ and $\mathcal{F} \text{ f T } s_2 = \text{T } s_2$. We can assume that $\mathcal{F} \text{ f T } s_1 \neq \perp$, as otherwise the claim vacuously holds. By Lemma 2.5, $\llbracket \text{C}_0 \rrbracket \text{T } s_1 z = \text{T } s_1 z$, and, by the same lemma, $f(\llbracket \text{C}_0 \rrbracket \text{T}) s_1 z = \llbracket \text{C}_0 \rrbracket \text{T } s_1 z$. So we have the desired relation:

$$\mathcal{F} \text{ f T } s_1 z = f(\llbracket \text{C}_0 \rrbracket \text{T})s_1 z = \text{T } s_1 z = \text{T } s_2 z = \mathcal{F} \text{ f T } s_2 z.$$

We have proved (1). We must now prove $\mathcal{P}(\llbracket \text{C} \rrbracket)$, that is $\mathcal{P}(\sqcup_i f_i)$. So assume that $\text{T} \models \text{T}^\#$ and that $[z \times w] \in \text{T}^\#$ and that $s_1 \stackrel{w}{=} s_2$; we must prove that $\llbracket \text{C} \rrbracket \text{T } s_1 \stackrel{\cong}{=} \llbracket \text{C} \rrbracket \text{T } s_2$. By Fact 2.2, there exists j such that $\llbracket \text{C} \rrbracket \text{T } s_1 = f_j \text{T } s_1$ and $\llbracket \text{C} \rrbracket \text{T } s_2 = f_j \text{T } s_2$. The claim now follows from (1).

[Sub]. Assume that

$$G \vdash \{\text{T}_0^\#\} \text{C} \{\text{T}^\#\}$$

because with $G \subseteq G_0$ and $\text{T}_0^\# \preceq \text{T}_1^\#$ and $\text{T}_2^\# \preceq \text{T}^\#$ we have

$$G_0 \vdash \{\text{T}_1^\#\} \text{C} \{\text{T}_2^\#\}.$$

From our assumption $\text{T} \models \text{T}_0^\#$ we infer by Fact 3.4 that $\text{T} \models \text{T}_1^\#$, so inductively we can assume that $\llbracket \text{C} \rrbracket \text{T} \models \text{T}_2^\#$. One more application of Fact 3.4 then yields the desired result.

6 An Application of the Logic: Forward Slicing

In this section we shall see how to compute the “low” forward slice of a program P . That is, we focus on one particular “high” variable h , and eliminate all commands in P that may depend on h , yielding a “slice” P' . With $l_1 \dots l_n$ the variables not depending on h in P (and thus considered “low”), our aim is that P' should be equivalent to P on $l_1 \dots l_n$. Then a user, wanting to compute the low variables but (for security reasons) not given clearance to view h , could be given P' to run, rather than P .

The slicing function \mathcal{S} is defined in Fig. 4, inductively on derivations in the logic (Fig. 3.) The idea is to replace by **skip** (i) assignments to variables that may depend on h ; (ii) conditionals and loops whose test may depend on h .

With $s_1, s_2 \in \mathbf{Sto}_\perp$, we write $s_1 \stackrel{T^\#}{=} s_2$ to denote that $s_1 \stackrel{z}{=} s_2$ holds for all $z \in \mathbf{Var} \cup \{\perp\}$ such that $[z \times h] \in T^\#$. We can now formulate correctness of slicing:

Theorem 6.1 *Let D be a derivation for the judgement $G \vdash \{T_0^\#\} C \{T^\#\}$, and let $\mathcal{S}(D) = C'$. Then $\llbracket C \rrbracket T s \stackrel{T^\#}{=} \llbracket C' \rrbracket T s$.*

In particular, if $T^\# = \{[l \times h], [\perp \times h]\}$ we can infer that either $\llbracket C \rrbracket T s$ and $\llbracket C' \rrbracket T s$ are both \perp , or they are both $\neq \perp$ and agree on l .

On the other hand, if $T^\#$ contains only $[l \times h]$ we can infer only partial correctness; it may happen that $\llbracket C \rrbracket T s = \perp$ but $\llbracket C' \rrbracket T s \neq \perp$. For an example of that, consider the program **while** $h > 0$ **do skip**; it can be given a derivation

$$\frac{\{h\} \vdash \{[l \times h], [h \times l]\} \text{ skip } \{[l \times h], [h \times l]\}}{\emptyset \vdash \{[l \times h], [h \times l]\} \text{ while } h > 0 \text{ do skip } \{[l \times h], [h \times l]\}}$$

which by \mathcal{S} is transformed into **skip**.

Theorem 6.1 follows immediately from the following main lemma, facilitating a proof by induction.

Lemma 6.2 *Let D be a derivation for the judgement $G \vdash \{T_0^\#\} C \{T^\#\}$, and let $\mathcal{S}(D) = C'$.*

If $T_1 s \stackrel{T_0^\#}{=} T_2 s$ then $\llbracket C \rrbracket T_1 s \stackrel{T^\#}{=} \llbracket C' \rrbracket T_2 s$.

PROOF. Induction in the derivation, where in all cases, we can assume that

$$T_1 s \neq \perp \text{ and } T_2 s \neq \perp.$$

$$\begin{array}{c}
\text{[Assign]} \quad G \vdash \{T_0^\#\} x := E \{T^\#\} \quad \Rightarrow \\
\left\{ \begin{array}{l} x := E, \text{ if } [x \times h] \in T^\# \\ \text{skip}, \text{ if } [x \times h] \notin T^\# \end{array} \right. \\
\\
\text{[Seq]} \quad \frac{\overbrace{G \vdash \{T_0^\#\} C_1 \{T_1^\#\}}^{D_1} \quad \overbrace{G \vdash \{T_1^\#\} C_2 \{T_2^\#\}}^{D_2}}{G \vdash \{T_0^\#\} C_1 ; C_2 \{T_2^\#\}} \quad \Rightarrow \mathcal{S}(D_1) ; \mathcal{S}(D_2) \\
\\
\text{[If]} \quad \frac{\overbrace{G_0 \vdash \{T_0^\#\} C_1 \{T^\#\}}^{D_1} \quad \overbrace{G_0 \vdash \{T_0^\#\} C_2 \{T^\#\}}^{D_2}}{G \vdash \{T_0^\#\} \text{if } E \text{ then } C_1 \text{ else } C_2 \{T^\#\}} \quad \Rightarrow \\
\left\{ \begin{array}{l} \text{if } E \text{ then } \mathcal{S}(D_1) \text{ else } \mathcal{S}(D_2), \text{ if } h \notin G_0 \\ \text{skip}, \quad \quad \quad \text{if } h \in G_0 \end{array} \right. \\
\\
\text{[While]} \quad \frac{\overbrace{G_0 \vdash \{T^\#\} C \{T^\#\}}^{D_0}}{G \vdash \{T^\#\} \text{while } E \text{ do } C \{T^\#\}} \quad \Rightarrow \\
\left\{ \begin{array}{l} \text{while } E \text{ do } \mathcal{S}(D_0), \text{ if } h \notin G_0 \\ \text{skip}, \quad \quad \quad \text{if } h \in G_0 \end{array} \right. \\
\\
\text{[Sub]} \quad \frac{\overbrace{G_1 \vdash \{T_1^\#\} C \{T_2^\#\}}^{D_0}}{G_0 \vdash \{T_0^\#\} C \{T_3^\#\}} \quad \Rightarrow \mathcal{S}(D_0)
\end{array}$$

Fig. 4. Forward Slicing: the slicing function \mathcal{S} .

For if that is not the case, then (by Fact 2.4) we have $\llbracket C \rrbracket T_1 s = \perp$ or $\llbracket C' \rrbracket T_2 s = \perp$. We would therefore have the desired relation $\llbracket C \rrbracket T_1 s \stackrel{z}{\approx} \llbracket C' \rrbracket T_2 s$ for all $z \in \mathbf{Var}$, but must still consider the case $[\perp \times h] \in T^\#$. Then by Fact 4.4 we infer that $[\perp \times h] \in T_0^\#$ and therefore $T_1 s \stackrel{\perp}{\approx} T_2 s$ implying $T_1 s = T_2 s = \perp$; by Fact 2.4 this implies $\llbracket C \rrbracket T_1 s = \llbracket C' \rrbracket T_2 s = \perp$ and therefore the desired relation $\llbracket C \rrbracket T_1 s \stackrel{\perp}{\approx} \llbracket C' \rrbracket T_2 s$.

We now embark on a case analysis.

[Assign]. Given $[z \times h] \in T^\#$, we must show

$$\llbracket C \rrbracket T_1 s \stackrel{z}{=} \llbracket C' \rrbracket T_2 s.$$

If $z = \perp$, the claim is trivial, since $\llbracket C \rrbracket T_1 s \neq \perp$ and $\llbracket C' \rrbracket T_2 s \neq \perp$. So assume that $z \neq \perp$.

If $z \neq x$, then $[z \times h] \in T_0^\#$, implying $T_1 s z = T_2 s z$. So no matter whether C' is $x := E$ or **skip**, we have the desired equality

$$\llbracket C \rrbracket T_1 s z = T_1 s z = T_2 s z = \llbracket C' \rrbracket T_2 s z.$$

Now assume that $z = x$, in which case Fig. 4 tells us that $C' = C$, and where the side conditions from Fig. 3 tell us that for all $y \in \text{fv}(E)$, $[y \times h] \in T_0^\#$ and thus $T_1 s y = T_2 s y$. Therefore, using Property 2.1, $\llbracket E \rrbracket (T_1 s) = \llbracket E \rrbracket (T_2 s)$. Again we therefore get the desired equality

$$\llbracket C \rrbracket T_1 s z = \llbracket E \rrbracket (T_1 s) = \llbracket E \rrbracket (T_2 s) = \llbracket C' \rrbracket T_2 s z.$$

[Seq]. Assume that with C of the form $C_1 ; C_2$, the derivation D of $G \vdash \{T_0^\#\} C \{T^\#\}$ has two children: a derivation D_1 of a judgement $G \vdash \{T_0^\#\} C_1 \{T_1^\#\}$, and a derivation D_2 of a judgement $G \vdash \{T_1^\#\} C_2 \{T^\#\}$. With $C'_1 = \mathcal{S}(D_1)$ and $C'_2 = \mathcal{S}(D_2)$, we have $C' = \mathcal{S}(D) = C'_1 ; C'_2$.

Inductively on D_1 , we from $T_1 s \stackrel{T_0^\#}{=} T_2 s$ infer $\llbracket C_1 \rrbracket T_1 s \stackrel{T_1^\#}{=} \llbracket C'_1 \rrbracket T_2 s$; inductively on D_2 , we further infer $\llbracket C_2 \rrbracket (\llbracket C_1 \rrbracket T_1) s \stackrel{T^\#}{=} \llbracket C'_2 \rrbracket (\llbracket C'_1 \rrbracket T_2) s$ which amounts to the desired equality $\llbracket C \rrbracket T_1 s \stackrel{T^\#}{=} \llbracket C' \rrbracket T_2 s$.

[Sub]. Assume that the derivation D of $G \vdash \{T_0^\#\} C \{T^\#\}$ has as child a derivation D_0 of a judgment $G_0 \vdash \{T_1^\#\} C \{T_2^\#\}$, where $G \subseteq G_0$ and $T_0^\# \preceq T_1^\#$ and $T_2^\# \preceq T^\#$. Here $C' = \mathcal{S}(D) = \mathcal{S}(D_0)$. Our assumption $T_1 s \stackrel{T_0^\#}{=} T_2 s$ clearly implies $T_1 s \stackrel{T_1^\#}{=} T_2 s$ (as $T_1^\# \subseteq T_0^\#$); inductively on D_0 we can thus infer that $\llbracket C \rrbracket T_1 s \stackrel{T_2^\#}{=} \llbracket C' \rrbracket T_2 s$ which (as $T^\# \subseteq T_2^\#$) implies the desired equality $\llbracket C \rrbracket T_1 s \stackrel{T^\#}{=} \llbracket C' \rrbracket T_2 s$.

[If]. Assume that with C of the form **if** E **then** C_1 **else** C_2 , the derivation D of $G \vdash \{T_0^\#\} C \{T^\#\}$ has two children: a derivation D_1 of a judgement $G_0 \vdash \{T_0^\#\} C_1 \{T^\#\}$, and a derivation D_2 of a judgement $G_0 \vdash \{T_0^\#\} C_2 \{T^\#\}$. Let $C'_1 = \mathcal{S}(D_1)$, and $C'_2 = \mathcal{S}(D_2)$. Looking at Fig. 4, we see that there are two subcases:

$C' = \text{if } E \text{ then } C'_1 \text{ else } C'_2$, with $h \notin G_0$. The side condition in Fig. 3 tells

us that for all $x \in \text{fv}(E)$ it holds that $[x \times h] \in T_0^\#$ and thus $T_1 s x = T_2 s x$. Therefore, using Property 2.1, $\llbracket E \rrbracket(T_1 s) = \llbracket E \rrbracket(T_2 s)$; we can assume wlog. that this value satisfies *true?*. Thus $\llbracket C \rrbracket T_1 s = \llbracket C_1 \rrbracket T_1 s$ and $\llbracket C' \rrbracket T_2 s = \llbracket C'_1 \rrbracket T_2 s$; this amounts to the desired equality $\llbracket C \rrbracket T_1 s \stackrel{T^\#}{=} \llbracket C' \rrbracket T_2 s$ since inductively on D_1 we infer that $\llbracket C_1 \rrbracket T_1 s \stackrel{T^\#}{=} \llbracket C'_1 \rrbracket T_2 s$.

$C' = \mathbf{skip}$, with $h \in G_0$. Let $[z \times h] \in T^\#$ be given; we must show that

$$\llbracket C \rrbracket T_1 s \stackrel{z}{=} \llbracket C' \rrbracket T_2 s.$$

If $z = \perp$, Lemma 4.5 (applied to D_1 and D_2) tells us that $\llbracket C_1 \rrbracket T_1 s \neq \perp$ and $\llbracket C_2 \rrbracket T_1 s \neq \perp$; thus $\llbracket C \rrbracket T_1 s \neq \perp$, yielding the claim (since trivially, $\llbracket C' \rrbracket T_2 s \neq \perp$).

We thus now assume that $z \in \mathbf{Var}$; from Lemma 4.6 (applied to D_1 and D_2) we then infer that

- $z \notin \text{modified}(C_1)$ and $z \notin \text{modified}(C_2)$, and therefore $z \notin \text{modified}(C)$ which by Lemma 2.5 implies that $\llbracket C \rrbracket$ preserves z ;
- $[z \times h] \in T_0^\#$, implying that $T_1 s z = T_2 s z$.

If $\llbracket C \rrbracket T_1 s = \perp$, the claim is trivial; otherwise we from the above infer the desired equality:

$$\llbracket C \rrbracket T_1 s z = T_1 s z = T_2 s z = \llbracket C' \rrbracket T_2 s z.$$

[While]. Assume that with C of the form **while** E **do** C_0 , and with $T_0^\# = T^\#$, the derivation D of $G \vdash \{T_0^\#\} C \{T^\#\}$ has one child: a derivation D_0 of a judgement $G_0 \vdash \{T^\#\} C_0 \{T^\#\}$. Let $C'_0 = \mathcal{S}(D_0)$. Looking at Fig. 4, we see that there are two subcases:

$C' = \mathbf{while}$ E **do** C'_0 , with $h \notin G_0$. Let \mathcal{F} be the functor of C , let \mathcal{F}' be the functor of C' , let f_i ($i \geq 0$) be the iterands of C , and let f'_i ($i \geq 0$) be the iterands of C' . We shall prove by induction in i that

$$\text{for all } T_1, T_2, s: T_1 s \stackrel{T^\#}{=} T_2 s \text{ implies } f_i T_1 s \stackrel{T^\#}{=} f'_i T_2 s. \quad (1)$$

In all cases, we can assume that $T_1 s \neq \perp$ and $T_2 s \neq \perp$. For if that is not the case, then (by Fact 2.4) we have $f_i T_1 s = \perp$ or $f'_i T_2 s = \perp$, which would imply the desired judgement, except if $[\perp \times h] \in T^\#$ in which case we must show that $f_i T_1 s = f'_i T_2 s = \perp$ which (by Fact 2.4) can be done by showing $T_1 s = T_2 s = \perp$. But this follows from $T_1 s \stackrel{\perp}{=} T_2 s$.

For $i = 0$, the claim is obvious, since $\perp \stackrel{T^\#}{=} \perp$ always holds. Now consider

the inductive step, where we must prove that $T_1 s \stackrel{T^\#}{=} T_2 s$ implies $\mathcal{F} f_i T_1 s \stackrel{T^\#}{=} \mathcal{F}' f'_i T_2 s$. Since $h \notin G_0$, the side condition in Fig. 3 tells us that for all $x \in \text{fv}(E)$ it holds that $[x \times h] \in T^\#$ and thus $T_1 s x = T_2 s x$. Therefore, using Property 2.1, $\llbracket E \rrbracket(T_1 s) = \llbracket E \rrbracket(T_2 s)$.

If this value does not satisfy *true?*, then $\mathcal{F} f_i T_1 s = T_1 s$ and $\mathcal{F}' f'_i T_2 s = T_2 s$, trivially implying the claim. Otherwise, our task is to prove that

$$f_i (\llbracket C_0 \rrbracket T_1) s \stackrel{T^\#}{=} f'_i (\llbracket C'_0 \rrbracket T_2) s.$$

By applying the overall induction hypothesis on D_0 , we infer that $\llbracket C_0 \rrbracket T_1 s \stackrel{T^\#}{=} \llbracket C'_0 \rrbracket T_2 s$; the claim then follows by applying the innermost induction hypothesis.

We can now return to our main task, proving that $\llbracket C \rrbracket T_1 s \stackrel{T^\#}{=} \llbracket C' \rrbracket T_2 s$ (where we are given that $T_1 s \stackrel{T_0^\#}{=} T_2 s$). But by Fact 2.2 there exists j such that $\llbracket C \rrbracket T_1 s = f_j T_1 s$ and $\llbracket C' \rrbracket T_2 s = f'_j T_2 s$; therefore (1) yields the claim.

$C' = \mathbf{skip}$, with $h \in G_0$. Let $[z \times h] \in T^\#$ be given; we must show that

$$\llbracket C \rrbracket T_1 s \stackrel{z}{=} \llbracket C' \rrbracket T_2 s.$$

The side condition in Fig. 3 tells us that $z \in \mathbf{Var}$; from Lemma 4.6 (applied to D_0) we then infer that $z \notin \text{modified}(C_0) = \text{modified}(C)$ which by Lemma 2.5 implies that $\llbracket C \rrbracket$ preserves z . If $\llbracket C \rrbracket T_1 s = \perp$, the claim is trivial; otherwise, the desired equality follows from the calculation

$$\llbracket C \rrbracket T_1 s z = T_1 s z = T_2 s z = \llbracket C' \rrbracket T_2 s z$$

where the second equality follows from $[z \times h] \in T^\#$.

Example 6.3 (*Illustrating forward slicing.*) For the program $l := h ; l := 0$, the forward slice is $\mathbf{skip} ; l := 0$, because we lose the independence $[l \times h]$ after $l := h$. For the program $h := l ; l := h$, the forward slice is $h := l ; l := h$ itself, since we retain the independence $[l \times h]$ after each of the commands in the sequence, and since the first command establishes the independence $[h \times h]$.

7 Computing Independences

In Fig. 5 we define a function

$$sp : \mathbf{Context} \times \mathbf{Cmd} \times \mathbf{Independ} \rightarrow \mathbf{Independ}$$

with the intuition (formalized below) that given a control dependence G , a command C and a precondition $T^\#$, $sp(G, C, T^\#)$ computes a postcondition $T_1^\#$ such that $G \vdash \{T^\#\} C \{T_1^\#\}$ holds, and $T_1^\#$ is the “largest” set (wrt. the subset ordering) that makes the judgement hold. Thus we compute the “strongest provable postcondition”, which might differ⁵ from the strongest *semantic* postcondition, that is, the largest set $T_1^\#$ such that for all T , if $T \models T^\#$ then $\llbracket C \rrbracket(T) \models T_1^\#$.

In a companion technical report (Amtoft and Banerjee, 2004a, Appendix A), we show how to also compute “weakest precondition”; we conjecture that the development in Sect. 8 could alternatively be carried out using weakest precondition instead of strongest postcondition.

We now explain two of the cases in Fig. 5. In an assignment, $x := E$, the postcondition carries over all independences $[y \times w]$ in the precondition if $y \neq x$; these independences are unaffected by the assignment to x . Suppose that w does not occur in context G . Then x is not control dependent on w . Moreover, if all variables referenced in E are independent of w , then $[x \times w]$ will be in the postcondition of the assignment.

The case for **while** is best explained by means of an example.

Example 7.1 Consider the program

$$C = n := 0 ; \mathbf{while} \ y > n \ \mathbf{do} \ l := x ; x := y ; y := h ; n := n + 1.$$

Let $T_0^\# \dots T_{12}^\#$ be given by the following table, where V denotes the set $\{h, l, n, x, y\}$. For reasons of space, we will represent a non-empty set as the concatenation of its elements in the table – thus $\{h, l, n, x\}$ is represented as $hlnx$. For example, the entry in the column for $T_6^\#$ and in the row for x shows

⁵ For example, let $C = l := h - h$ and $T^\# = \{[l \times h]\}$. Then $[l \times h]$ is in the strongest semantic postcondition, since for all T and all s_1, s_2 we have $\llbracket C \rrbracket T \stackrel{l}{=} \llbracket C \rrbracket T \ s_2$ and therefore $\llbracket C \rrbracket T \models [l \times h]$, but not in the strongest provable postcondition.

that $[x \times h] \in T_6^\#$ and $[x \times l] \in T_6^\#$ and $[x \times n] \in T_6^\#$.

	$T_0^\#$	$T_1^\#$	$T_2^\#$	$T_3^\#$	$T_4^\#$	$T_5^\#$	$T_6^\#$	$T_7^\#$	$T_8^\#$	$T_9^\#$	$T_{10}^\#$	$T_{11}^\#$	$T_{12}^\#$
$h \times$	lnxy	lnxy	lnxy	lnxy	lnxy	lnxy	lnxy	lnxy	lnxy	lnxy	lnxy	lnxy	lnxy
$n \times$	hlxy	V	V	V	V	hlxn	hlxn	hlxn	hlxn	lnx	lnx	lnx	lnx
$l \times$	hnxy	hnxy	hln	hln	hln	hln	hn	ln	ln	ln	n	ln	ln
$x \times$	hlxy	hlxy	hlxy	hlxn	hlxn	hlxn	hln	hln	lnx	lnx	ln	ln	lnx
$y \times$	hlxn	hlxn	hlxn	hlxn	lnx	lnx	lnx	lnx	lnx	lnx	lnx	lnx	lnx
$\perp \times$	V	V	V	V	V	V	hlxn	hlxn	hlxn	hlxn	lnx	lnx	lnx

Our goal is to compute $sp(\emptyset, C, T_0^\#)$ and doing so involves the fixpoint computation sketched below.

	Iteration		
	first	second	third
while $y > n$ do	$T_1^\#$	$T_6^\#$	$T_{10}^\#$
$G_0 :$	$\{y\}$	$\{h, y\}$	$\{h, y\}$
$l := x$	$T_2^\#$	$T_7^\#$	$T_{11}^\#$
$x := y$	$T_3^\#$	$T_8^\#$	$T_{12}^\#$
$y := h$	$T_4^\#$	$T_8^\#$	$T_{12}^\#$
$n := n + 1$	$T_5^\#$	$T_9^\#$	$T_{12}^\#$
$\perp \cap T_1^\# \setminus [\perp \times G_0]$	$T_6^\#$	$T_{10}^\#$	$T_{10}^\#$

For example, the entry $T_9^\#$ in the column marked “second” and in the row marked “ $n := n + 1$ ”, denotes that $sp(\{h, y\}, n := n + 1, T_8^\#) = T_9^\#$.

Note that at the end of the first iteration, the independence set is $T_6^\#$ and $[l \times h]$ is still present; it takes a second iteration – the independence set is then $T_{10}^\#$ – to filter $[l \times h]$ out and thus detect insecurity. The third iteration affirms that $T_{10}^\#$ is indeed a fixpoint (of the functional $\mathcal{H}_{\text{while}}^{T_1^\#, \emptyset}$ defined in Fig. 5).

Theorem 7.3 states the correctness of the function sp , that it indeed computes a postcondition. Then, Theorem 7.4 states that the postcondition computed

$$\begin{aligned}
sp(G, x := E, T_0^\#) &= \\
&\quad \{[y \times w] \mid (y \neq x \vee y = \perp) \wedge [y \times w] \in T_0^\#\} \\
&\quad \cup \{[x \times w] \mid w \notin G \wedge \forall y \in \text{fv}(E) \bullet [y \times w] \in T_0^\#\} \\
sp(G, C_1 ; C_2, T_0^\#) &= sp(G, C_2, sp(G, C_1, T_0^\#)) \\
sp(G, \text{if } E \text{ then } C_1 \text{ else } C_2, T_0^\#) &= \\
\text{let } G_0 = G \cup \{w \mid \exists x \in \text{fv}(E) \bullet [x \times w] \notin T_0^\#\} & \\
T_1^\# = sp(G_0, C_1, T_0^\#) & \\
T_2^\# = sp(G_0, C_2, T_0^\#) & \\
\text{in } T_1^\# \cap T_2^\# & \\
sp(G, \text{while } E \text{ do } C_0, T_0^\#) &= \\
\text{let } \mathcal{H}_C^{T_0^\#, G} : \mathbf{Independ} \rightarrow \mathbf{Independ} \text{ be given by } (C = \text{while } E \text{ do } C_0) & \\
\mathcal{H}_C^{T_0^\#, G}(T^\#) = & \\
\text{let } G_0 = G \cup \{w \mid \exists x \in \text{fv}(E) \bullet [x \times w] \notin T^\#\} & \\
\text{in } (sp(G_0, C_0, T^\#) \cap T_0^\#) \setminus \{[\perp \times w] \mid w \in G_0\} & \\
\text{in } \text{lfp}(\mathcal{H}_C^{T_0^\#, G}) &
\end{aligned}$$

Fig. 5. Strongest Postcondition.

by sp is the strongest postcondition. We shall rely on the following property:

Lemma 7.2 (Monotonicity) *For all commands C , the following holds (for all $G, G_1, T^\#, T_1^\#$):*

- (1) $sp(G, C, T^\#)$ is well-defined;
- (2) $\mathcal{H}_C^{T^\#, G}$ (when C is a **while** loop) is a monotone function;
- (3) if $G \subseteq G_1$ then $sp(G, C, T^\#) \preceq sp(G_1, C, T^\#)$;
- (4) if $T^\# \preceq T_1^\#$ then $sp(G, C, T^\#) \preceq sp(G, C, T_1^\#)$.

PROOF. Induction in C , where the four parts of the lemma are proved simultaneously. We do a case analysis on C ; the only non-trivial case is where C is of the form **while** E **do** C_0 .

Using the induction hypothesis on C_0 , we infer that for all $T_0^\#, G$ it holds that $\mathcal{H}_C^{T_0^\#, G}$ is a monotone function on the complete lattice **Independ**. Hence $\text{lfp}(\mathcal{H}_C^{T_0^\#, G})$, and thus $sp(G, C, T_0^\#)$, is indeed well-defined.

Next assume that $T^\# \preceq T_1^\#$ and that $G \subseteq G_1$. Then clearly $\mathcal{H}_C^{T^\#, G} \preceq \mathcal{H}_C^{T_1^\#, G_1}$ (by the pointwise ordering) and therefore $\text{lfp}(\mathcal{H}_C^{T^\#, G}) \preceq \text{lfp}(\mathcal{H}_C^{T_1^\#, G_1})$ which amounts to the desired relation $sp(G, C, T^\#) \preceq sp(G_1, C, T_1^\#)$.

Theorem 7.3 For all $C, G, T_0^\#$, it holds that $G \vdash \{T_0^\#\} C \{sp(G, C, T_0^\#)\}$.

PROOF. Go by structural induction on C ; we perform a case analysis.

$C = x := E$. Let $T^\# = sp(G, C, T_0^\#)$, and assume $[z \times w] \in T^\#$. There are two cases:

- if $z \neq x$ then $[z \times w] \in T_0^\#$.
- if $z = x$ then $w \notin G$, and $\forall y \in \text{fv}(E) \bullet [y \times w] \in T_0^\#$.

This establishes $G \vdash \{T_0^\#\} C \{T^\#\}$.

$C = C_1 ; C_2$. Assume that $sp(G, C_1 ; C_2, T_0^\#) = T^\#$ because $T^\# = sp(G, C_2, T_1^\#)$ where $T_1^\# = sp(G, C_1, T_0^\#)$. By the induction hypothesis on C_1 and on C_2 , we have

$$G \vdash \{T_0^\#\} C_1 \{T_1^\#\} \text{ and } G \vdash \{T_1^\#\} C_2 \{T^\#\}$$

from which we infer the desired relation $G \vdash \{T_0^\#\} C_1 ; C_2 \{T^\#\}$.

$C = \text{if } E \text{ then } C_1 \text{ else } C_2$. Assume that $sp(G, \text{if } E \text{ then } C_1 \text{ else } C_2, T_0^\#) = T^\#$ because $G_0 = G \cup \{w \mid \exists x \in \text{fv}(E) \bullet [x \times w] \notin T_0^\#\}$, $T_1^\# = sp(G_0, C_1, T_0^\#)$ and $T_2^\# = sp(G_0, C_2, T_0^\#)$ and $T^\# = T_1^\# \cap T_2^\#$. Inductively, we have

$$G_0 \vdash \{T_0^\#\} C_1 \{T_1^\#\} \text{ and } G_0 \vdash \{T_0^\#\} C_2 \{T_2^\#\}.$$

As $T^\# \subseteq T_1^\#$ and $T^\# \subseteq T_2^\#$ we have $T_1^\# \preceq T^\#$ and $T_2^\# \preceq T^\#$; by [Sub], this implies

$$G_0 \vdash \{T_0^\#\} C_1 \{T^\#\} \text{ and } G_0 \vdash \{T_0^\#\} C_2 \{T^\#\}.$$

This establishes the desired $G \vdash \{T_0^\#\} \text{if } E \text{ then } C_1 \text{ else } C_2 \{T^\#\}$, since $G \subseteq G_0$ and $w \notin G_0$ implies $\forall x \in \text{fv}(E) \bullet [x \times w] \in T_0^\#$.

$C = \text{while } E \text{ do } C_0$. Assume that $sp(G, C, T_0^\#) = T^\#$ so we want to prove $G \vdash \{T_0^\#\} C \{T^\#\}$. We have $T^\# = \text{lfp}(\mathcal{H}_C^{T_0^\#, G})$. By definition of a fixpoint, $T^\# = \mathcal{H}_C^{T_0^\#, G}(T^\#)$. Thus

$$T^\# = (sp(G_0, C_0, T^\#) \cap T_0^\#) \setminus \{[\perp \times w] \mid w \in G_0\}$$

where $G_0 = G \cup \{w \mid \exists x \in \text{fv}(E) \bullet [x \times w] \notin T^\#\}$. Hence $sp(G_0, C_0, T^\#) \preceq T^\#$ and $T_0^\# \preceq T^\#$. We claim $G \vdash \{T^\#\} C \{T^\#\}$, which by [Sub] implies the desired $G \vdash \{T_0^\#\} C \{T^\#\}$.

It remains to prove the claim, $G \vdash \{T^\#\} C \{T^\#\}$. By the induction hypothesis on C_0 we have $G_0 \vdash \{T^\#\} C_0 \{sp(G_0, C_0, T^\#)\}$ and by [Sub] therefore

$$G_0 \vdash \{T^\#\} C_0 \{T^\#\}.$$

Now we get $G \vdash \{T^\#\} C \{T^\#\}$ by an application of [While] because $G \subseteq G_0$, because $w \notin G_0$ implies $\forall x \in \text{fv}(E) \bullet [x \times w] \in T^\#$, and because if $w \in G_0$ then $[\perp \times w] \notin T^\#$.

Theorem 7.4 *For all judgements $G \vdash \{T_0^\#\} C \{T^\#\}$, $sp(G, C, T_0^\#) \preceq T^\#$.*

PROOF. We perform induction in the derivation of $G \vdash \{T_0^\#\} C \{T^\#\}$, and do a case analysis on the last rule applied:

[Sub]. Assume that $G \vdash \{T_0^\#\} C \{T^\#\}$ because with $G \subseteq G_1$ and $T_0^\# \preceq T_2^\#$ and $T_3^\# \preceq T^\#$ we have $G_1 \vdash \{T_2^\#\} C \{T_3^\#\}$. Applying the induction hypothesis on that derivation, we get

$$sp(G_1, C, T_2^\#) \preceq T_3^\#$$

and by Lemma 7.2 we get $sp(G, C, T_0^\#) \preceq sp(G_1, C, T_2^\#)$. This yields the desired relation

$$sp(G, C, T_0^\#) \preceq T_3^\# \preceq T^\#.$$

[Assign], with $C = x := E$. Assume that $G \vdash \{T_0^\#\} C \{T^\#\}$, and let $T_1^\# = sp(G, C, T_0^\#)$. We want $T_1^\# \preceq T^\#$. Accordingly, assume $[y \times w] \in T^\#$ to show $[y \times w] \in T_1^\#$. We have two cases:

- $x \neq y$. Then $[y \times w] \in T_0^\#$; hence $[y \times w] \in T_1^\#$ by the definition of sp .
- $x = y$. Then $w \notin G$ and $\forall z \in \text{fv}(E) \bullet [z \times w] \in T_0^\#$; hence $[y \times w] \in T_1^\#$ by the definition of sp .

[Seq], with $C = C_1 ; C_2$. Assume $G \vdash \{T_0^\#\} C \{T^\#\}$ because $G \vdash \{T_0^\#\} C_1 \{T_2^\#\}$ and $G \vdash \{T_2^\#\} C_2 \{T^\#\}$. By the induction hypothesis on these derivations,

$$sp(G, C_1, T_0^\#) \preceq T_2^\# \text{ and } sp(G, C_2, T_2^\#) \preceq T^\#$$

which by Lemma 7.2 enables us to infer that

$$sp(G, C_2, sp(G, C_1, T_0^\#)) \preceq T^\#.$$

This is as desired, since $sp(G, C_1 ; C_2, T_0^\#) = sp(G, C_2, sp(G, C_1, T_0^\#))$.

[If], with $C = \text{if } E \text{ then } C_1 \text{ else } C_2$. Assume that $G \vdash \{\mathsf{T}_0^\#\} C \{\mathsf{T}^\#\}$ because $G_1 \vdash \{\mathsf{T}_0^\#\} C_1 \{\mathsf{T}^\#\}$ and $G_1 \vdash \{\mathsf{T}_0^\#\} C_2 \{\mathsf{T}^\#\}$ where $G \subseteq G_1$ and where $w \notin G_1$ implies that $\forall x \in \text{fv}(E) \bullet [x \times w] \in \mathsf{T}_0^\#$. Inductively, via the judgements for C_1 and C_2 , we obtain

$$sp(G_1, C_1, \mathsf{T}_0^\#) \preceq \mathsf{T}^\# \text{ and } sp(G_1, C_2, \mathsf{T}_0^\#) \preceq \mathsf{T}^\#.$$

Let $G_0 = G \cup \{w \mid \exists x \in \text{fv}(E) \bullet [x \times w] \notin \mathsf{T}_0^\#\}$. Note that $G_0 \subseteq G_1$. Thus by Lemma 7.2 we get

$$sp(G_0, C_1, \mathsf{T}_0^\#) \preceq \mathsf{T}^\# \text{ and } sp(G_0, C_2, \mathsf{T}_0^\#) \preceq \mathsf{T}^\#.$$

This yields the claim since $sp(G, C, \mathsf{T}_0^\#) = sp(G_0, C_1, \mathsf{T}_0^\#) \cap sp(G_0, C_2, \mathsf{T}_0^\#)$.

[While], with $C = \text{while } E \text{ do } C_0$. Assume that $G \vdash \{\mathsf{T}^\#\} C \{\mathsf{T}^\#\}$ because $G_1 \vdash \{\mathsf{T}^\#\} C_0 \{\mathsf{T}^\#\}$ where $G \subseteq G_1$, and where $w \notin G_1$ implies that $\forall x \in \text{fv}(E) \bullet [x \times w] \in \mathsf{T}^\#$, and where $[\perp \times w] \in \mathsf{T}^\#$ implies $w \notin G_1$. Assume $\mathsf{T}_1^\# = sp(G, C, \mathsf{T}^\#)$ to show $\mathsf{T}_1^\# \preceq \mathsf{T}^\#$. By the definition of sp , $\mathsf{T}_1^\# = \text{Ifp}(\mathcal{H}_C^{\mathsf{T}^\#, G})$. Let $G_0 = G \cup \{w \mid \exists x \in \text{fv}(E) \bullet [x \times w] \notin \mathsf{T}^\#\}$; then

$$\begin{aligned} \mathcal{H}_C^{\mathsf{T}^\#, G}(\mathsf{T}^\#) &= (sp(G_0, C_0, \mathsf{T}^\#) \cap \mathsf{T}^\#) \setminus \{[\perp \times w] \mid w \in G_0\} \\ &\supseteq (sp(G_1, C_0, \mathsf{T}^\#) \cap \mathsf{T}^\#) \setminus \{[\perp \times w] \mid w \in G_1\} = \mathsf{T}^\# \end{aligned}$$

Here the set inclusion follows from the observation that $G_0 \subseteq G_1$ and an application of Lemma 7.2; the last equality follows since $w \in G_1$ implies $[\perp \times w] \notin \mathsf{T}^\#$, and since the induction hypothesis tells us that $sp(G_1, C_0, \mathsf{T}^\#) \preceq \mathsf{T}^\#$.

This shows that $\mathcal{H}_C^{\mathsf{T}^\#, G}$ is reductive at $\mathsf{T}^\#$, that is $\mathcal{H}_C^{\mathsf{T}^\#, G}(\mathsf{T}^\#) \preceq \mathsf{T}^\#$, so by Tarski's theorem we infer the desired relation $\mathsf{T}_1^\# = \text{Ifp}(\mathcal{H}_C^{\mathsf{T}^\#, G}) \preceq \mathsf{T}^\#$.

The following result, where we clearly cannot allow $y = \perp$, is useful for the developments in Sections 8:

Lemma 7.5 *Given C , and $y \in \mathbf{Var}$ with $y \notin \text{modified}(C)$. Then for all $\mathsf{T}_0^\#, G, w$:*

$$[y \times w] \in \mathsf{T}_0^\# \text{ implies } [y \times w] \in sp(G, C, \mathsf{T}_0^\#).$$

PROOF. Go by structural induction on C ; we perform a case analysis. In each case, our assumption is that $y \notin \text{modified}(C)$ and that $[y \times w] \in \mathsf{T}_0^\#$; we must show $[y \times w] \in sp(G, C, \mathsf{T}_0^\#)$.

$C = x := E$. Our assumptions imply that $y \neq x$, from which the result trivially follows.

$C = C_1 ; C_2$. Since $\mathbf{y} \notin \text{modified}(C_1)$ we can apply the induction hypothesis on C_1 , giving us $[\mathbf{y} \times \mathbf{w}] \in \text{sp}(G, C_1, T_0^\#)$. Since $\mathbf{y} \notin \text{modified}(C_2)$ we can then apply the induction hypothesis on C_2 , giving us $[\mathbf{y} \times \mathbf{w}] \in \text{sp}(G, C_2, \text{sp}(G, C_1, T_0^\#))$. This yields the claim.

$C = \text{if } E \text{ then } C_1 \text{ else } C_2$. Since $\mathbf{y} \notin \text{modified}(C_1)$ and $\mathbf{y} \notin \text{modified}(C_2)$, we can apply the induction hypothesis twice, yielding (using the terminology in Fig. 5)

$$[\mathbf{y} \times \mathbf{w}] \in T_1^\# \text{ and } [\mathbf{y} \times \mathbf{w}] \in T_2^\#.$$

Since $\text{sp}(G, C, T_0^\#) = T_1^\# \cap T_2^\#$, this yields the claim.

$C = \text{while } E \text{ do } C_0$. Let $T^\# = \text{sp}(G, C, T_0^\#)$, then $T^\# = \text{lfp}(\mathcal{H}_C^{T_0^\#, G})$. Define

$$T_1^\# = T^\# \cup \{[\mathbf{y} \times \mathbf{w}]\}$$

and let G_0 be as in Fig. 5. By applying the induction hypothesis on C_0 (possible since $\mathbf{y} \notin \text{modified}(C_0)$) we get $[\mathbf{y} \times \mathbf{w}] \in \text{sp}(G_0, C_0, T_1^\#)$; since $[\mathbf{y} \times \mathbf{w}] \in T_0^\#$ and $\mathbf{y} \neq \perp$ this implies

$$[\mathbf{y} \times \mathbf{w}] \in \mathcal{H}_C^{T_0^\#, G}(T_1^\#). \quad (1)$$

Since $\mathcal{H}_C^{T_0^\#, G}$ is a monotone function (by Lemma 7.2), we from $T_1^\# \preceq T^\#$ infer that $\mathcal{H}_C^{T_0^\#, G}(T_1^\#) \preceq \mathcal{H}_C^{T_0^\#, G}(T^\#) = T^\#$ and thus

$$T^\# \subseteq \mathcal{H}_C^{T_0^\#, G}(T_1^\#). \quad (2)$$

Combining (1) and (2), we infer $T_1^\# \subseteq \mathcal{H}_C^{T_0^\#, G}(T_1^\#)$, that is $\mathcal{H}_C^{T_0^\#, G}(T_1^\#) \preceq T_1^\#$. This shows that $\mathcal{H}_C^{T_0^\#, G}$ is reductive at $T_1^\#$, so by Tarski's theorem we infer $T^\# = \text{lfp}(\mathcal{H}_C^{T_0^\#, G}) \preceq T_1^\#$, that is $T_1^\# \subseteq T^\#$. This demonstrates that $[\mathbf{y} \times \mathbf{w}] \in T^\#$, as desired.

8 Modular Reasoning and the Frame Rule

Although we have not emphasized it in this article, it is possible to state the logic in Fig. 3 in a way such that the judgements for a particular command mention only the variables relevant to the command. In this manner, one may obtain “small specifications” (O’Hearn et al., 2001) for every command. For instance, the rule for assignment could be rewritten as:

$$\text{[Assign]} \quad G \vdash \{\mathsf{T}_0^\#\} x := E \{\mathsf{T}^\#\} \quad \text{if } \forall [\mathbf{y} \times \mathbf{w}] \in \mathsf{T}^\# \bullet (\mathbf{y} = x \wedge \mathbf{w} \notin G \wedge \forall z \in \text{fv}(E) \bullet [z \times \mathbf{w}] \in \mathsf{T}_0^\#)$$

An independence $[\mathbf{y} \times \mathbf{w}]$ with $\mathbf{y} \neq x$ is invariant with respect to $x := E$, since \mathbf{y} is not modified, and hence there is no need to mention it. If it is required in a larger context, as in a sequential composition, it can be retrieved using a *frame rule*, as in separation logic (Ishtiaq and O’Hearn, 2001; Reynolds, 2002). The frame rule is motivated by the desire for local reasoning: if C_1 and C_2 modify disjoint regions of a heap, reasoning about C_1 can be performed independently of the reasoning about C_2 . In our setting, with $\text{lhs}(\mathsf{T}^\#) = \{\mathbf{y} \mid [\mathbf{y} \times \mathbf{w}] \in \mathsf{T}^\#\}$, we can state the frame rule as

$$\text{[Frame]} \quad \frac{G \vdash \{\mathsf{T}_0^\#\} C \{\mathsf{T}^\#\}}{G \vdash \{\mathsf{T}_0^\# \cup \mathsf{T}_1^\#\} C \{\mathsf{T}^\# \cup \mathsf{T}_1^\#\}} \quad \text{if } \text{modified}(C) \cap \text{lhs}(\mathsf{T}_1^\#) = \emptyset \text{ and } \perp \notin \text{lhs}(\mathsf{T}_1^\#)$$

In the rest of this section, we will show results about modular reasoning in the context of the calculation of strongest postconditions; this entails (Corollary 8.3) that [Frame] as stated above is indeed admissible. For a development where the frame rule plays a key rôle, we refer the interested reader to our more recent work (Amtoft et al., 2006).

Theorem 8.1 (Frame rule (I)) *Given $\mathsf{T}_0^\#$ and C . Then for all $\mathsf{T}^\#, G$:*

- (1) *If $\text{lhs}(\mathsf{T}_0^\#) \subseteq \mathbf{Var}$ and $\text{lhs}(\mathsf{T}_0^\#) \cap \text{modified}(C) = \emptyset$ then $\text{sp}(G, C, \mathsf{T}^\# \cup \mathsf{T}_0^\#) \supseteq \text{sp}(G, C, \mathsf{T}^\#) \cup \mathsf{T}_0^\#$.*
- (2) *If $\text{lhs}(\mathsf{T}_0^\#) \subseteq \mathbf{Var}$ and $\text{lhs}(\mathsf{T}_0^\#) \cap \text{fv}(C) = \emptyset$ then $\text{sp}(G, C, \mathsf{T}^\# \cup \mathsf{T}_0^\#) = \text{sp}(G, C, \mathsf{T}^\#) \cup \mathsf{T}_0^\#$.*

Note that the weaker premise in 1 does not imply the stronger consequence in 2, since (with $[z \times \mathbf{w}]$ playing the role of $\mathsf{T}_0^\#$)

$$\begin{aligned} \text{sp}(\emptyset, x := \mathbf{y} + z, \{[\mathbf{y} \times \mathbf{w}]\} \cup \{[z \times \mathbf{w}]\}) &= \{[\mathbf{y} \times \mathbf{w}], [z \times \mathbf{w}], [x \times \mathbf{w}]\} \\ \text{sp}(\emptyset, x := \mathbf{y} + z, \{[\mathbf{y} \times \mathbf{w}]\}) \cup \{[z \times \mathbf{w}]\} &= \{[\mathbf{y} \times \mathbf{w}], [z \times \mathbf{w}]\}. \end{aligned}$$

Theorem 8.1 is proved by observing that part (1) follows from Lemmas 7.5 and 7.2; then part (2) follows using the following result:

Lemma 8.2 *Let $\mathsf{T}_0^\#$ and C be given, with $\text{lhs}(\mathsf{T}_0^\#) \subseteq \mathbf{Var}$ and $\text{lhs}(\mathsf{T}_0^\#) \cap \text{fv}(C) = \emptyset$. Then for all $\mathsf{T}^\#$ and G , $\text{sp}(G, C, \mathsf{T}^\# \cup \mathsf{T}_0^\#) \subseteq \text{sp}(G, C, \mathsf{T}^\#) \cup \mathsf{T}_0^\#$.*

PROOF. Go by structural induction on C ; we perform a case analysis.

$C = x := E$. The claim follows from the following calculation, using that $lhs(T_0^\#) \cap fv(E) = \emptyset$ and that $x \notin lhs(T_0^\#)$.

$$\begin{aligned}
& sp(G, C, T^\# \cup T_0^\#) \\
&= \{[y \times w] \mid y \neq x \wedge [y \times w] \in T^\# \cup T_0^\#\} \cup \\
&\quad \{[x \times w] \mid w \notin G \wedge \forall y \in fv(E) \bullet [y \times w] \in T^\# \cup T_0^\#\} \\
&= \{[y \times w] \mid y \neq x \wedge [y \times w] \in T^\# \cup T_0^\#\} \cup \\
&\quad \{[x \times w] \mid w \notin G \wedge \forall y \in fv(E) \bullet [y \times w] \in T^\#\} \\
&= T_0^\# \cup \{[y \times w] \mid y \neq x \wedge [y \times w] \in T^\#\} \cup \\
&\quad \{[x \times w] \mid w \notin G \wedge \forall y \in fv(E) \bullet [y \times w] \in T^\#\} \\
&= T_0^\# \cup sp(G, C, T^\#)
\end{aligned}$$

$C = C_1 ; C_2$. Using our induction hypothesis and Lemma 7.2, we get

$$\begin{aligned}
& sp(G, C, T^\# \cup T_0^\#) = sp(G, C_2, sp(G, C_1, T^\# \cup T_0^\#)) \\
&\subseteq sp(G, C_2, sp(G, C_1, T^\#) \cup T_0^\#) \subseteq sp(G, C_2, sp(G, C_1, T^\#)) \cup T_0^\# \\
&= sp(G, C, T^\#) \cup T_0^\#
\end{aligned}$$

$C = \text{if } E \text{ then } C_1 \text{ else } C_2$. Let

$$G_0 = G \cup \{w \mid \exists x \in fv(E) \bullet [x \times w] \notin T^\# \cup T_0^\#\}$$

where our assumptions imply that $lhs(T_0^\#) \cap fv(E) = \emptyset$ and therefore also

$$G_0 = G \cup \{w \mid \exists x \in fv(E) \bullet [x \times w] \notin T^\#\}.$$

Using the induction hypothesis, we now get

$$\begin{aligned}
sp(G, C, T^\# \cup T_0^\#) &= sp(G_0, C_1, T^\# \cup T_0^\#) \cap sp(G_0, C_2, T^\# \cup T_0^\#) \\
&\subseteq (sp(G_0, C_1, T^\#) \cup T_0^\#) \cap (sp(G_0, C_2, T^\#) \cup T_0^\#) \\
&= (sp(G_0, C_1, T^\#) \cap sp(G_0, C_2, T^\#)) \cup T_0^\# \\
&= sp(G, C, T^\#) \cup T_0^\#
\end{aligned}$$

$C = \mathbf{while\ E\ do\ } C_0$. Let $H = \mathcal{H}_C^{\top^\#, G}$ and let $H_0 = \mathcal{H}_C^{\top^\# \cup \top_0^\#, G}$, we must show that $lfp(H_0) \subseteq lfp(H) \cup \top_0^\#$. Define $\top_1^\# = lfp(H_0)$, and let

$$G_0 = G \cup \{\mathbf{w} \mid \exists x \in \text{fv}(E) \bullet [x \times \mathbf{w}] \notin \top_1^\# \setminus \top_0^\#\}$$

where our assumptions imply that $lhs(\top_0^\#) \cap \text{fv}(E) = \emptyset$ and therefore also

$$G_0 = G \cup \{\mathbf{w} \mid \exists x \in \text{fv}(E) \bullet [x \times \mathbf{w}] \notin \top_1^\#\}.$$

Using Lemma 7.2 and our induction hypothesis on C_0 we get (where $\top_\perp^\#$ denotes $\{[\perp \times \mathbf{w}] \mid \mathbf{w} \in G_0\}$)

$$\begin{aligned} \top_1^\# \setminus \top_0^\# &= H_0(\top_1^\#) \setminus \top_0^\# \\ &= ((sp(G_0, C_0, \top_1^\#) \cap (\top^\# \cup \top_0^\#)) \setminus \top_\perp^\#) \setminus \top_0^\# \\ &= (sp(G_0, C_0, \top_1^\#) \cap \top^\#) \setminus \top_0^\# \setminus \top_\perp^\# \\ &\subseteq (sp(G_0, C_0, (\top_1^\# \setminus \top_0^\#) \cup \top_0^\#) \cap \top^\#) \setminus \top_0^\# \setminus \top_\perp^\# \\ &\subseteq ((sp(G_0, C_0, (\top_1^\# \setminus \top_0^\#)) \cup \top_0^\#) \cap \top^\#) \setminus \top_0^\# \setminus \top_\perp^\# \\ &\subseteq (sp(G_0, C_0, (\top_1^\# \setminus \top_0^\#)) \cap \top^\#) \setminus \top_\perp^\# \\ &= H(\top_1^\# \setminus \top_0^\#) \end{aligned}$$

We have proved $H(\top_1^\# \setminus \top_0^\#) \preceq \top_1^\# \setminus \top_0^\#$. This shows that H is reductive at $\top_1^\# \setminus \top_0^\#$, so by Tarski's theorem we infer $lfp(H) \preceq \top_1^\# \setminus \top_0^\#$. This implies the desired relation

$$lfp(H_0) = \top_1^\# \subseteq (\top_1^\# \setminus \top_0^\#) \cup \top_0^\# \subseteq lfp(H) \cup \top_0^\#.$$

As a consequence of Theorem 8.1, we get the following result:

Corollary 8.3 (Frame rule (II)) *Assume that $G \vdash \{\top_1^\#\} C \{\top_2^\#\}$ and that $lhs(\top_0^\#) \cap (\text{modified}(C) \cup \{\perp\}) = \emptyset$. Then $G \vdash \{\top_1^\# \cup \top_0^\#\} C \{\top_2^\# \cup \top_0^\#\}$.*

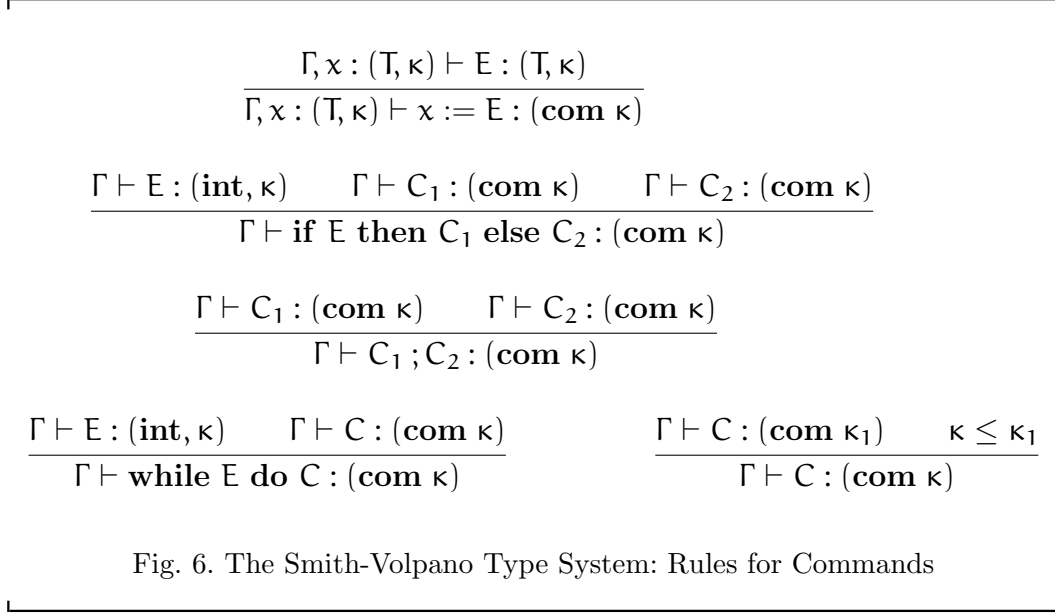
PROOF. Using Theorems 8.1 and 7.4 we get

$$sp(G, C, \top_1^\# \cup \top_0^\#) \supseteq sp(G, C, \top_1^\#) \cup \top_0^\# \supseteq \top_2^\# \cup \top_0^\#.$$

Since by Theorem 7.3 we have $G \vdash \{\top_1^\# \cup \top_0^\#\} C \{sp(G, C, \top_1^\# \cup \top_0^\#)\}$, the result follows by [Sub].

Example 8.4 Assume that

$$G \vdash \{\top_1^\#\} C_1 \{\top_3^\#\} \text{ and } G \vdash \{\top_2^\#\} C_2 \{\top_4^\#\}.$$



Further assume that $lhs(T_2^\#) \cap (modified(C_1) \cup \{\perp\}) = \emptyset$ and that $lhs(T_3^\#) \cap (modified(C_2) \cup \{\perp\}) = \emptyset$. Then Corollary 8.3 yields

$$G \vdash \{T_1^\# \cup T_2^\#\} C_1 \{T_3^\# \cup T_2^\#\} \text{ and } G \vdash \{T_3^\# \cup T_2^\#\} C_2 \{T_3^\# \cup T_4^\#\}$$

and therefore $G \vdash \{T_1^\# \cup T_2^\#\} C_1 ; C_2 \{T_3^\# \cup T_4^\#\}$.

A traditional view of modularity in the security literature is the “hook-up property” (McCullough, 1987): if two programs are secure then their composition is secure as well. Our logic satisfies the hook-up property for sequential composition; in our context, a secure program is one which has $[l \times h]$ as an invariant (if $[l \times h]$ is in the precondition, it is also in the strongest postcondition). With this interpretation, Sabelfeld and Sands’s hook-up theorem holds (Sabelfeld and Sands, 2001, Theorem 5).

9 The Smith-Volpano Security Type System

In the Smith-Volpano type system (Volpano and Smith, 1997), variables are labeled by security types; for example, $x : (T, \kappa)$ means that x has type T and security level κ . The security typing rules are given in Fig. 6. To handle implicit flows due to conditionals, the technical development requires commands to be typed $(\mathbf{com} \ \kappa)$ with the intention that all variables assigned to in such commands have level at least κ . The judgement $\Gamma \vdash C : (\mathbf{com} \ \kappa)$ says that in the security type context Γ that binds free variables in C to security types, command C has type $(\mathbf{com} \ \kappa)$.

We now show a conservative extension: if a program is well-typed in the Smith-Volpano system, then for any two precludes, the current values of low variables are independent of the initial values of high variables. For simplicity, we consider a program with only two variables, h with level H and l with level L .

Theorem 9.1 *Assume that C can be given a security type wrt. the environment $h : (-, H), l : (-, L)$. Then for all $T^\#$, if $[l \times h] \in T^\#$ then $[l \times h] \in sp(\emptyset, C, T^\#)$.*

The upshot of the theorem is that a well-typed program has $[l \times h]$ as *invariant*: if $[l \times h]$ appears in the precondition, then it also appears in the strongest postcondition.

The theorem is a straightforward consequence of the following lemma which facilitates a proof by induction. For L commands, the assumption $h \notin G$ in the lemma says that L commands cannot be control dependent on H guards.

Lemma 9.2 (1) *Suppose $h : (-, H), l : (-, L) \vdash C : (\mathbf{com} H)$. Then for all $G, T^\#$, if $[l \times h] \in T^\#$ then $[l \times h] \in sp(G, C, T^\#)$.*

(2) *Suppose $h : (-, H), l : (-, L) \vdash C : (\mathbf{com} L)$. Then for all $G, T^\#$, if $[l \times h] \in T^\#$ and $h \notin G$ then $[l \times h] \in sp(G, C, T^\#)$.*

PROOF. We prove the two parts of the lemma in turn. In both cases we go by induction on the derivation of C (in the system of Fig. 6), with cases on the last rule used.

(Part 1.)

$C = z := E$. Clearly, $z \neq l$ as otherwise the assignment cannot be typed. By definition,

$$sp(G, z := E, T^\#) \supseteq \{[u \times w] \mid u \neq z \wedge [u \times w] \in T^\#\} \ni [l \times h].$$

$C = C_1 ; C_2$. We have $h : (-, H), l : (-, L) \vdash C_1 : (\mathbf{com} H)$ and $h : (-, H), l : (-, L) \vdash C_2 : (\mathbf{com} H)$. Inductively, we get

$$[l \times h] \in sp(G, C_1, T^\#) \text{ and then } [l \times h] \in sp(G, C_2, sp(G, C_1, T^\#)).$$

Thus $[l \times h] \in sp(G, C_1 ; C_2, T^\#)$.

$C = \mathbf{if} E \mathbf{then} C_1 \mathbf{else} C_2$. Then $h : (-, H), l : (-, L) \vdash C_1 : (\mathbf{com} H)$ and $h : (-, H), l : (-, L) \vdash C_2 : (\mathbf{com} H)$. Assume $[l \times h] \in T^\#$. Inductively,

$$[l \times h] \in sp(G_0, C_1, T^\#) \text{ and } [l \times h] \in sp(G_0, C_2, T^\#)$$

where $G_0 = G \cup \{\mathbf{w} \mid \exists x \in \text{fv}(E) \bullet [x \times \mathbf{w}] \notin T^\#\}$. Thus $[l \times h] \in \text{sp}(G_0, C_1, T^\#) \cap \text{sp}(G_0, C_2, T^\#)$, and we are done.

$C = \mathbf{while} E \mathbf{do} C_0$. Then $h : (-, H), l : (-, L) \vdash C_0 : (\mathbf{com} H)$. Let $T_0^\# = \text{sp}(G, C, T^\#)$. Then $T_0^\# = \text{lf}p(\mathcal{H}_C^{T^\#, G})$. Hence $T_0^\# = \mathcal{H}_C^{T^\#, G}(T_0^\#)$.

Let $T_1^\# = T_0^\# \cup [l \times h]$. Now

$$\mathcal{H}_C^{T^\#, G}(T_1^\#) = (\text{sp}(G_0, C_0, T_1^\#) \cap T^\#) \setminus \{[\perp \times \mathbf{w}] \mid \mathbf{w} \in G_0\}$$

where $G_0 = G \cup \{\mathbf{w} \mid \exists x \in \text{fv}(E) \bullet [x \times \mathbf{w}] \notin T_1^\#\}$. Inductively, as $[l \times h] \in T_1^\#$, we get $[l \times h] \in \text{sp}(G_0, C_0, T_1^\#)$. Thus,

$$[l \times h] \in \mathcal{H}_C^{T^\#, G}(T_1^\#). \quad (1)$$

Because $\mathcal{H}_C^{T^\#, G}$ is a monotone function, from $T_1^\# \preceq T_0^\#$ we get $\mathcal{H}_C^{T^\#, G}(T_1^\#) \preceq \mathcal{H}_C^{T^\#, G}(T_0^\#) = T_0^\#$. Thus

$$T_0^\# \subseteq \mathcal{H}_C^{T^\#, G}(T_1^\#). \quad (2)$$

Combining (1) and (2), we get $T_1^\# \subseteq \mathcal{H}_C^{T^\#, G}(T_1^\#)$, i.e., $\mathcal{H}_C^{T^\#, G}(T_1^\#) \preceq T_1^\#$. This shows that $\mathcal{H}_C^{T^\#, G}$ is reductive at $T_1^\#$, so by Tarski's theorem, $T_0^\# = \text{lf}p(\mathcal{H}_C^{T^\#, G}) \preceq T_1^\#$, that is, $T_1^\# \subseteq T_0^\#$. Hence $[l \times h] \in T_0^\#$.

Subtyping. For the subtyping rule, the result trivially follows by induction on the smaller derivation tree for C .

This completes Part 1 of the proof.

(Part 2.)

Subtyping. Assume $[l \times h] \in T^\#$ and $h \notin G$. By the typing rule, $h : (-, H), l : (-, L) \vdash C : (\mathbf{com} L)$ follows because $h : (-, H), l : (-, L) \vdash C : (\mathbf{com} \kappa_1)$ for some κ_1 . If $\kappa_1 = L$, the result follows inductively. If $\kappa_1 = H$, then we can apply Part 1 of the theorem to conclude that $[l \times h] \in \text{sp}(G, C, T^\#)$ holds for any G ; in particular, it must therefore hold for the G where $h \notin G$.

$C = z := E$. Assume $[l \times h] \in T^\#$ and $h \notin G$. By typing, $z : (-, L)$ and $h \notin \text{fv}(E)$, as otherwise the assignment cannot be typed. Hence $z = l$. By definition⁶,

⁶ In case there are several low variables, we would use the argument:

$$\text{sp}(G, z := E, T^\#) \supseteq \{[u \times \mathbf{w}] \mid u \neq z \wedge [u \times \mathbf{w}] \in T^\#\} \ni [l \times h]$$

to deal with low variables not assigned to.

$$\begin{aligned} sp(\mathbf{G}, z := E, \mathbf{T}^\#) \supseteq \{[z \times w] \mid w \notin \mathbf{G} \wedge \forall u \in \text{fv}(E) \bullet [u \times w] \in \mathbf{T}^\#\} \\ \ni [l \times h]. \end{aligned}$$

$C = C_1 ; C_2$. Easy induction.

$C = \text{if } E \text{ then } C_1 \text{ else } C_2$. By typing, $h : (-, H), l : (-, L) \vdash E : (\mathbf{int}, L)$. Hence $h \notin \text{fv}(E)$ and thus $h \notin \mathbf{G}_0$, where $\mathbf{G}_0 = \mathbf{G} \cup \{w \mid \exists x \in \text{fv}(E) \bullet [x \times w] \notin \mathbf{T}^\#\}$. Then inductively, we obtain $[l \times h] \in sp(\mathbf{G}_0, C_1, \mathbf{T}^\#)$ and $[l \times h] \in sp(\mathbf{G}_0, C_2, \mathbf{T}^\#)$. Hence $[l \times h] \in sp(\mathbf{G}_0, C_1, \mathbf{T}^\#) \cap sp(\mathbf{G}_0, C_2, \mathbf{T}^\#) = sp(\mathbf{G}, C, \mathbf{T}^\#)$.

$C = \text{while } E \text{ do } C_0$. By typing, $h : (-, H), l : (-, L) \vdash E : (\mathbf{int}, L)$. Hence $h \notin \text{fv}(E)$. Now the proof proceeds similarly to the corresponding case in Part 1 and is omitted, except that we note that to use the induction hypothesis on the derivation of C_0 , we need to show $h \notin \mathbf{G}_0$, where $\mathbf{G}_0 = \mathbf{G} \cup \{w \mid \exists x \in \text{fv}(E) \bullet [x \times w] \notin \mathbf{T}_1^\#\}$. Since $h \notin \mathbf{G}$, it is sufficient to show that for all $x \in \text{fv}(E)$ it holds that $[x \times h] \in \mathbf{T}_1^\#$. But this follows since $h \notin \text{fv}(E)$ and $[l \times h] \in \mathbf{T}_1^\#$.

10 Discussion

Perspective. This article specifies an information flow analysis for confidentiality using a Hoare-like logic and shows an application of the analysis to a program transformation, namely, slicing. The concrete pre- and post-conditions of a program are abstracted by independences. Independences can be statically checked against the logic and can be statically computed using strongest post-conditions. We also show how the notion of independences underlies a classic type-based information flow analysis due to Smith and Volpano (1997).

Giacobazzi and Mastroeni (2004) consider attackers as abstract interpretations and generalize the notion of noninterference by parameterizing it wrt. what an attacker can analyze about the input/output information flow. For instance, assume an attacker can only analyze the *parity* (odd/even) of values. Then

while h **do** $l := l + 2 ; h := h - 1$

is secure, although it contains an update of a low variable under a high guard. We might try to model this approach in our framework by parameterizing Definition 3.1 wrt. parity, but it is not clear how to alter the proof rules accordingly. Instead, we envision our logic to be put on top of abstract interpretations. In the parity example, the above program would be abstracted

to

```
while h do h := h - 1
```

which our logic already deems secure for an attacker not able to observe non-termination.

Related Work. The most closely related work is that of Clark, Hankin, and Hunt (2002), who consider a language similar to ours and then extend it to Idealized Algol, requiring distinguishing between identifiers and locations. The analysis for Idealized Algol is split in two stages: the first stage does a control-flow analysis, specified using a flow logic (Nielson et al., 1999). The second stage specifies what is an acceptable information flow analysis with respect to the control-flow analysis. The precision of the control-flow analysis influences the precision of the information flow analysis. Flow logics usually do not come with a frame rule so it is unclear what modularity properties their analysis satisfies. For each statement S in the program, they compute the set of dependences introduced by S ; a pair (x, y) is in that set if different values for y prior to execution of S may result in different values for x after execution of S . For a complete program, they thus, as expected, compute essentially the same information as we do, but the information computed *locally* is different from ours: we estimate if different *initial* values of y , i.e., values of y prior to execution of *the whole program*, may result in different values for x after execution of S .

Joshi and Leino (2000) provide an elegant semantic characterization of non-interference that allows handling both termination-sensitive and termination-insensitive noninterference. Their notion of security for a command C is equationally characterized by $C ; HH = HH ; C ; HH$, where HH means that an arbitrary value is assigned to a high variable. They show how to express their notion of security in Dijkstra’s weakest precondition calculus. Although they do not consider synthesizing loop invariants, this can certainly be done via a fixpoint computation with weakest preconditions. However, their work is not concerned with computing dependences.

Darvas, Hähnle and Sands (2004) use dynamic logic to express secure information flow in JavaCard. They discuss several ways that noninterference can be expressed in a program logic, one of which is as follows: consider a program with variables l and h . Consider another copy of the program with l , h relabeled to fresh variables l' , h' respectively. Then, noninterference holds in the following situation: running the original program and the copy sequentially such that the initial state satisfies $l = l'$ should yield a final state satisfying $l = l'$. They are also interested in showing insecurity, by exhibiting distinct initial values for high variables that give distinct current values of low variables;

to achieve this accuracy, they need the power of a general purpose theorem prover, which is also helpful in that they can express declassification, as well as treat exceptions (which most approaches based on static analysis cannot easily be extended to deal with).

Barthe, D’Argenio and Rezk (2004) use the same idea of self-composition (i.e., composing a program with a copy of itself) as Darvas et alii and investigate “abstract” noninterference (Giacobazzi and Mastroeni, 2004) for several languages. By parameterizing noninterference with a property, they are able to handle more general information flow policies, including a form of declassification known as delimited information release (Sabelfeld and Myers, 2004). They show how self-composition can be formulated in logics describing these languages, namely, Hoare logic, separation logic, linear temporal logic, etc. They also discuss how to use their results for model checking programs with finite state spaces to check satisfaction of their generalized definition of noninterference.

The first work that used a Hoare-style semantics to reason about information flow was by Andrews and Reitman (1980). Their assertions keep track of the security level of variables, and are able to deal even with parallel programs. However, no formal correctness result is stated.

Differences from the Conference Version. Apart from the removal of some infelicities of notation, we have made three additional contributions not present in the conference version of this article (Amtoft and Banerjee, 2004b).

- We consider nontermination sensitive noninterference here, compared to nontermination insensitive noninterference in the conference version. For that purpose, we needed to add a new kind of independence, which in turn necessitated changes in the logic – specifically, in the rule for loops.
- To prove the resulting logic semantically correct, the semantics had to be modified, since the one given in (Amtoft and Banerjee, 2004b) had no explicit notion of non-termination. As an extra benefit, the resulting denotational semantics is significantly simpler than the previous one.
- We show an application of the logic to forward slicing. Although the connection between information flow analysis and slicing was explored by Abadi et. al. (1999), that paper did not provide a means to compute forward slices, which we present here.

On the other hand, the conference version of this article contained a section on counterexample generation, which we have chosen to omit here. We feel that the results might merit a separate paper after some strengthening.

Conclusion. The work reported in this article was inspired in part by presentations by Roberto Giacobazzi and Reiner Hähnle at the Dagstuhl Seminar on Language-based Security (October 2003). The reported work is only the first step in our goal to formulate more general definitions of noninterference in terms of program independence, such that the definitions support modular reasoning. We are in the process of extending the framework in this article to handle a richer language, with methods, pointers, objects and dynamic memory allocation. An obvious goal is interprocedural reasoning about variable and field independences, perhaps using a higher-order version of the frame rule (O’Hearn et al., 2004). We would also like to explore, via abstract interpretation and perhaps following Schmidt’s development (Schmidt, 2002), whether our Hoare-like logic is the “best” possible one with respect to the underlying abstract interpretation.

Note Added in Print. In work subsequent to this article (Amtoft et al., 2006), we have extended the framework to handle modular reasoning about information flow in an object-oriented language. Thus we have addressed one of the future goals mentioned in the conclusion of this article. Also, in a forthcoming paper, Hunt and Sands (2006) discuss a flow-sensitive security type system for a simple imperative language. They show that their flow-sensitive security type system is equivalent to the logic of independences discussed in the conference version of this paper Amtoft and Banerjee (2004b).

Acknowledgments. We thank Reiner Hähnle, David Naumann, Peter O’Hearn, Tamara Rezk, David Sands, and Hongseok Yang, as well as the participants of the *Open Software Quality* meeting in Santa Cruz, May 2004, and the anonymous reviewers of both Static Analysis Symposium (SAS) 2004 as well as this special issue, for useful comments on a draft of this article. Thanks to Sruthi Bandhakavi for implementing the analysis reported in this article and for her help with Example 7.1. Thanks to David Schmidt for many discussions on abstract interpretation. Finally, thanks are due to Roberto Giacobazzi for organizing a very stimulating SAS 2004, and to the city of Verona for its warmth and its ice creams.

References

Abadi, M., Banerjee, A., Heintze, N., Riecke, J. G., 1999. A core calculus of dependency. In: ACM Symposium on Principles of Programming Languages (POPL). pp. 147–160.

- Amtoft, T., Bandhakavi, S., Banerjee, A., 2006. A logic for information flow in object-oriented programs. In: ACM Symposium on Principles of Programming Languages (POPL). To appear.
- Amtoft, T., Banerjee, A., Apr. 2004a. Information flow analysis in logical form. Tech. Rep. CIS TR 2004-3, Kansas State University.
URL <http://www.cis.ksu.edu/~ab/Publications/ifalftr.pdf>
- Amtoft, T., Banerjee, A., 2004b. Information flow analysis in logical form. In: Static Analysis Symposium (SAS). Vol. 3148 of Lecture Notes in Computer Science. Springer-Verlag, pp. 100–115.
- Andrews, G. R., Reitman, R. P., Jan. 1980. An axiomatic approach to information flow in programs. *ACM Trans. Prog. Lang. Syst.* 2 (1), 56–75.
- Banerjee, A., Naumann, D. A., 2005. Stack-based access control and secure information flow. *Journal of Functional Programming* 15 (2), 131–177, Special Issue on Language-based Security.
- Barthe, G., D’Argenio, P. R., Rezk, T., 2004. Secure information flow by self-composition. In: IEEE Computer Security Foundations Workshop (CSFW). IEEE Computer Society Press, pp. 100–114.
- Bell, D., LaPadula, L., 1973. Secure computer systems: Mathematical foundations. Tech. Rep. MTR-2547, MITRE Corp.
- Clark, D., Hankin, C., Hunt, S., 2002. Information flow for Algol-like languages. *Computer Languages* 28 (1), 3–28.
- Cohen, E. S., 1978. Information transmission in sequential programs. In: DeMillo, R. A., Dobkin, D. P., Jones, A. K., Lipton, R. J. (Eds.), *Foundations of Secure Computation*. Academic Press, pp. 297–335.
- Cousot, P., Cousot, R., 1977a. Abstract interpretation: a unified lattice model for static analysis of programs by construction or approximation of fixpoints. In: *ACM Symposium on Principles of Programming Languages (POPL)*. ACM Press, New York, NY, pp. 238–252.
- Cousot, P., Cousot, R., Aug. 1977b. Automatic synthesis of optimal invariant assertions: mathematical foundations. In: *Proceedings of the ACM Symposium on Artificial Intelligence and Programming Languages, SIGPLAN Notices*. Vol. 12. ACM Press, pp. 1–12.
- Darvas, Á., Hähnle, R., Sands, D., 2004. A theorem proving approach to analysis of secure information flow. Tech. Rep. 2004-01, Department of Computing Science, Chalmers University of Technology and Göteborg University, a fuller version of a paper appearing in *Workshop on Issues in the Theory of Security*, 2003.
- Denning, D., Denning, P., 1977. Certification of programs for secure information flow. *Communications of the ACM* 20 (7), 504–513.
- Giacobazzi, R., Mastroeni, I., 2004. Abstract non-interference: Parameterizing non-interference by abstract interpretation. In: *ACM Symposium on Principles of Programming Languages (POPL)*. pp. 186–197.
- Goguen, J., Meseguer, J., 1982. Security policies and security models. In: *Proc. IEEE Symp. on Security and Privacy*. pp. 11–20.
- Heintze, N., Riecke, J. G., 1998. The SLam calculus: programming with se-

- crecy and integrity. In: ACM Symposium on Principles of Programming Languages (POPL). pp. 365–377.
- Hunt, S., Sands, D., 1991. Binding time analysis: A new PERspective. In: Partial Evaluation and Semantics-Based Program Manipulation (PEPM). Vol. 26 (9) of SIGPLAN Notices. pp. 154–165.
- Hunt, S., Sands, D., 2006. On flow-sensitive security types. In: ACM Symposium on Principles of Programming Languages (POPL). ACM Press, to appear.
- Ishtiaq, S., O’Hearn, P. W., 2001. BI as an assertion language for mutable data structures. In: ACM Symposium on Principles of Programming Languages (POPL). pp. 14–26.
- Joshi, R., Leino, K. R. M., 2000. A semantic approach to secure information flow. *Sci. Comput. Programming* 37, 113–138.
- McCullough, D., 1987. Specifications for multi-level security and a hook-up. In: IEEE Symposium on Security and Privacy. pp. 161–166.
- Nielson, F., Nielson, H. R., Hankin, C., 1999. Principles of Program Analysis. Springer-Verlag, web page at www.imm.dtu.dk/~riis/PPA/ppa.html.
- O’Hearn, P., Reynolds, J., Yang, H., 2001. Local reasoning about programs that alter data structures. In: *Computer Science Logic*. Vol. 2142 of Lecture Notes in Computer Science. Springer-Verlag, pp. 1–19.
- O’Hearn, P., Yang, H., Reynolds, J., 2004. Separation and information hiding. In: ACM Symposium on Principles of Programming Languages (POPL). pp. 268–280.
- Ørbæk, P., Palsberg, J., Nov. 1997. Trust in the λ -calculus. *Journal of Functional Programming* 7 (6), 557–591.
- Pottier, F., Simonet, V., Jan. 2003. Information flow inference for ML. *ACM Trans. Prog. Lang. Syst.* 25 (1), 117–158.
- Reynolds, J. C., 2002. Separation logic: a logic for shared mutable data structures. In: IEEE Symposium on Logic in Computer Science (LICS). IEEE Computer Society Press, pp. 55–74.
- Sabelfeld, A., Myers, A., 2004. A model for delimited information release. In: Proceedings of the International Symposium on Software Security (ISSS). No. 3233 in Lecture Notes in Computer Science. Springer-Verlag, pp. 174–191.
- Sabelfeld, A., Myers, A. C., Jan. 2003. Language-based information-flow security. *IEEE J. Selected Areas in Communications* 21 (1), 5–19.
- Sabelfeld, A., Sands, D., 2001. A Per model of secure information flow in sequential programs. *Higher-order and Symbolic Computation* 14 (1), 59–91.
- Schmidt, D. A., 2002. Structure-preserving binary relations for program abstraction. In: *The Essence of Computation: Complexity, Analysis, Transformation – Essays dedicated to Neil D. Jones*. No. 2566 in Lecture Notes in Computer Science. Springer-Verlag, pp. 245–265.
- Volpano, D., Irvine, C., Smith, G., 1996. A sound type system for secure flow

analysis. *Journal of Computer Security* 4 (3), 167–188.

Volpano, D., Smith, G., 1997. A type-based approach to program security. In: *Proceedings of Theory and Practice of Software Development (TAPSOFT'97)*. No. 1214 in *Lecture Notes in Computer Science*. Springer-Verlag, pp. 607–621.