

# Two for the Price of One: Lifting Separation Logic Assertions

Jacob Thamsborg and Lars Birkedal  
IT University of Copenhagen

Hongseok Yang  
Queen Mary University of London

## Abstract

Recently, data abstraction has been studied in the context of separation logic, with noticeable practical successes: the developed logics have enabled clean proofs of tricky challenging programs, such as subject-observer patterns, and they have become the basis of efficient verification tools for Java (*jStar*), C (*VeriFast*) and Hoare Type Theory (*Ynot*). In this paper, we give a new semantic analysis of such logic-based approaches using Reynolds’s relational parametricity. The core of the analysis is our lifting theorems, which give a sound and complete condition for when a true implication between assertions in the standard interpretation entails that the same implication holds in a relational interpretation. Using these theorems, we provide an algorithm for identifying abstraction-respecting client-side proofs; the proofs ensure that clients cannot distinguish two appropriately-related module implementations.

## 1 Introduction

Data abstraction is one of the key design principles for building computer software, and it has been the focus of active research from the early days of computer science. Recently, data abstraction has been studied in the context of separation logic [14, 4, 11, 15, 7], with noticeable practical successes: the developed logics have enabled clean proofs of tricky challenging programs, such as the subject-observer pattern, and they have become the basis of efficient verification tools for Java (*jStar* [8]), C (*VeriFast* [10]) and Hoare Type Theory (*Ynot* [12]).

In this paper, we give a new semantic analysis of these logic-based approaches using Reynolds’s relational parametricity. Our techniques can be used to prove *representation independence*, i.e., that clients cannot distinguish between related module implementations, a consequence that we would expect from using data abstraction, but (as we shall see) a consequence that only holds for certain good clients.

---

## INTERFACE SPECIFICATION

$$\begin{array}{l} \{1 \leftrightarrow \_ \} \mathbf{init}\{a\} \quad \{a\} \mathbf{nxt}\{b\} \quad \{b\} \mathbf{fin}\{1 \leftrightarrow \_ \} \\ \{a\} \mathbf{inc}\{a\} \quad \{b\} \mathbf{dec}\{b\} \end{array}$$

## TWO IMPLEMENTATIONS OF A COUNTER

$$\begin{array}{l} \mathbf{init}_1 \stackrel{\text{def}}{=} [1] := 0 \quad \mathbf{nxt}_1 \stackrel{\text{def}}{=} \mathbf{skip} \quad \mathbf{fin}_1 \stackrel{\text{def}}{=} \mathbf{skip} \\ \mathbf{inc}_1 \stackrel{\text{def}}{=} [1] := [1] + 1 \quad \mathbf{dec}_1 \stackrel{\text{def}}{=} [1] := [1] - 1 \\ \mathbf{init}_2 \stackrel{\text{def}}{=} [1] := 0 \quad \mathbf{nxt}_2 \stackrel{\text{def}}{=} [1] := -[1] \quad \mathbf{fin}_2 \stackrel{\text{def}}{=} [1] := -[1] \\ \mathbf{inc}_2 \stackrel{\text{def}}{=} [1] := [1] + 1 \quad \mathbf{dec}_2 \stackrel{\text{def}}{=} [1] := [1] + 1 \end{array}$$

## DEFINITIONS OF ABSTRACT PREDICATES

$$a_1 \stackrel{\text{def}}{=} 1 \leftrightarrow \_, \quad b_1 \stackrel{\text{def}}{=} 1 \leftrightarrow \_, \quad a_2 \stackrel{\text{def}}{=} 1 \leftrightarrow \_, \quad b_2 \stackrel{\text{def}}{=} 1 \leftrightarrow \_$$

## CLIENT-SIDE PROOF ATTEMPTS

$$\begin{array}{ll} \{1 \leftrightarrow \_ \} \mathbf{init}; \{a\} & \{1 \leftrightarrow \_ \} \mathbf{init}; \{a\} \\ \mathbf{inc}; \{a\} & \mathbf{inc}; \{a\} \\ \mathbf{nxt}; \{b\} & \mathbf{nxt}; \{b\} \\ \mathbf{dec}; \{b\} & [1] := [1] - 1; \{???\} \\ \mathbf{fin} \{1 \leftrightarrow \_ \} & \end{array}$$


---

Figure 1. Two-stage Counter

**Logic-based Data Abstraction** The basic idea of the logic-based approaches is that the private states of modules are exposed to clients only abstractly using *assertion variables* [4], also known as *abstract predicates* [14]. For concreteness, we consider a two-stage counter module and client programs in Figure 1. The module realizes a counter with increment and decrement operations, called *inc* and *dec*. An interesting feature is that the counter goes through two stages in its lifetime; in the first stage, it can perform only the increment operation, but in the second, it can only run the decrement. The *interface specification* in the figure formalizes this intended behavior of the counter using assertion variables *a* and *b*, where *a* means that the counter is in the first stage and *b* that the counter is in the second. The triple for *init* says that the initialization can turn the assertion  $1 \leftrightarrow \_$ , denoting heaps with cell 1, to the assertion variable *a*, which describes

an abstract state where we can only call `inc` or `nxt` (since  $a$  is the precondition of only those operations). The abstract state  $a$  can be changed to  $b$  by calling `nxt`, says the triple for the `nxt` operation. In  $b$  we are allowed to run `dec` but not `inc`. Finally, `fin` can turn the abstract state  $b$  back to  $1 \hookrightarrow \_$ . Note that by using  $a$  and  $b$ , the interface specification does not expose the private state of the module to the client. It reveals only *partial information* about the private state of the module; here it is whether the private state is in the first or the second stage. The flexibility afforded by revealing partial information is very useful in applications; see the examples mentioned in the references above.

In these logic-based approaches, proof attempts for clients of a module can succeed only when they are given with respect to the abstract interface specification, without making any further assumptions on assertion variables. For instance, the proof attempt on the bottom left of Figure 1 is successful, whereas the bottom right one is not, because the latter assumes that the assertion variable  $b$  entails the allocatedness of cell 1. This is so, even when the entailment holds for an actual definition of  $b$ .

**Representation Independence** In this paper, we analyze the condition for successful client-side proofs, using *representation independence*: given two implementations of a module, *if* there are coupling relations between the definitions of assertion variables in the two modules and, furthermore, the relations are preserved by the corresponding operations of the implementations, *then* a *proved* client will give the *same* result no matter whether it is executed with one or the other module implementation. For instance, Figure 1 describes two implementations of the counter, where cell 1 is used in both to represent their private states, but in different ways — the first stores the current value of the counter, but the second stores the current value or its negative version, depending on whether it is in the first stage or the second. In both cases, the assertion variables  $a$  and  $b$  are defined to be the same  $1 \hookrightarrow \_$ . The operations of the two implementations preserve the coupling relations  $r_a$  and  $r_b$ :

$$r_a \stackrel{\text{def}}{=} \{(h_1, h_2) \mid 1 \in \text{dom}(h_1) \cap \text{dom}(h_2) \wedge h_1(1) = h_2(1)\}$$

$$r_b \stackrel{\text{def}}{=} \{(h_1, h_2) \mid 1 \in \text{dom}(h_1) \cap \text{dom}(h_2) \wedge h_1(1) = -h_2(1)\}$$

Thus, the representation independence result says that all proved clients, such as the one on the left bottom of Figure 1, should behave the same for both implementations.

In earlier work [6] we were able to prove such a representation independence result for a more restricted form of logical data abstraction, namely one given

---

#### INTERFACE SPECIFICATION

$$\{1 \hookrightarrow \_ \} \text{init} \{1 \hookrightarrow \_ \wedge a * b\} \quad \{a\} \text{fin} \{1 \hookrightarrow \_ \}$$

$$\{1 \hookrightarrow \_ * a \vee 1 \hookrightarrow \_ * b\} \text{badfin} \{1 \hookrightarrow \_ \}$$

#### TWO IMPLEMENTATIONS

$$\text{init}_1 \stackrel{\text{def}}{=} \text{skip} \quad \text{fin}_1 \stackrel{\text{def}}{=} \text{skip} \quad \text{badfin}_1 \stackrel{\text{def}}{=} [1] := 1$$

$$\text{init}_2 \stackrel{\text{def}}{=} \text{skip} \quad \text{fin}_2 \stackrel{\text{def}}{=} \text{skip} \quad \text{badfin}_2 \stackrel{\text{def}}{=} [1] := 2$$

#### DEFINITIONS OF ABSTRACT PREDICATES

$$a_1 \stackrel{\text{def}}{=} 1 \hookrightarrow \_, \quad b_1 \stackrel{\text{def}}{=} \text{true}, \quad a_2 \stackrel{\text{def}}{=} 1 \hookrightarrow \_, \quad b_2 \stackrel{\text{def}}{=} \text{true}$$

#### TWO CLIENT-SIDE PROOFS

$\{1 \hookrightarrow \_ \}$	$\{1 \hookrightarrow \_ \}$
<code>init;</code>	<code>init;</code>
$\{1 \hookrightarrow \_ \wedge a * b\}$	$\{1 \hookrightarrow \_ \wedge a * b\}$
$\{a\}$	$\{1 \hookrightarrow \_ * a \vee 1 \hookrightarrow \_ * b\}$
<code>fin</code>	<code>badfin</code>
$\{1 \hookrightarrow \_ \}$	$\{1 \hookrightarrow \_ \}$

---

**Figure 2. Good or Bad Client-side Proofs**

by frame rules rather than general assertion variables. Roughly speaking, frame rules use a restricted form of assertion variables that are not exposed to clients at all, as can be seen from some models of separation logic in which frame rules are modelled via quantification over semantic assertions [5]. This means that the rules do not allow the exposure of even partial information about module internals. (On the other hand, frame rules implement information hiding, because they completely relieve clients of tracking the private state of a module, even in an abstracted form.) Our model in [6] exploited this restricted use of assertion variables, and gave relational meanings to Hoare triple specifications, which led to representation independence.

Removing this restriction and allowing assertion variables in client proofs turned out to be very challenging. The main problem was that it was unclear under which conditions a standard unary implication between assertions in the rule of consequence could be lifted to a *relational* one. “Always” is not an answer, because only some, not all, implications can be lifted. In this paper, we provide a sound and, in a certain sense, complete answer to when the lifting can be done.

For instance, consider the example in Figure 2. Our answer lets us conclude that the client on the left is *good* but the one on the right is *bad*. The client on the left calls `init` and ends with the post-condition  $(1 \hookrightarrow \_ \wedge a * b)$ . Since  $(1 \hookrightarrow \_ \wedge a * b) \implies a$  is true in the standard interpretation (in an intuitionistic setup<sup>1</sup>),

<sup>1</sup>In an intuitionistic setup,  $\varphi * \psi \implies \varphi$  holds for all  $\varphi, \psi$ .

the rule of consequence can be applied to yield the precondition of `fin`, which can be called, ending up with the postcondition  $(1 \hookrightarrow \_)$ . The key point here is the implication used in the rule of consequence. Our results imply that this implication can indeed be lifted to an implication between relational meanings of assertions  $(1 \hookrightarrow \_ \wedge a * b)$  and  $a$  (Theorem 8 in Section 4). They also entail that this lifting implies the representation independence theorem. We can confirm the claim of the theorem, simply noting that both `init1`; `fin1` and `init2`; `fin2` are the same skip command.

The client on the right also first calls `init` and then uses the rule of consequence. But this time our results say that a true implication  $(1 \hookrightarrow \_ \wedge a * b) \implies (1 \hookrightarrow \_ * a \vee 1 \hookrightarrow \_ * b)$  in the rule of consequence does *not* lift to an implication between relational meanings of the assertions (Theorem 15 in Section 5 or just Example 12.) Because of this failure, the proof of the client does not ensure representation independence. In fact, the client can indeed distinguish between the two module implementations — when the client is executed with the first module implementation, the final heap maps address 1 to 1, but when the client is executed with the second, the final heap maps address 1 to 2.

Note that we phrase the lifting only in terms of semantically true implications, without referring to how they are proved. By doing so, we make our results relevant to automatic tools that use the semantic model of separation logic to prove implications, such as the ones based on shallow embeddings of the assertion logic [12, 9].

To sum up, the question of whether representation independence holds for a client comes down to whether the implications used in the rule of consequence in a client-side proof can be lifted to a relational interpretation. In this paper, we give a sound and complete characterization of when that is the case.

## 2 Semantic Domain

In the following section we will define the meaning of an assertion to be an  $n$ -ary relation on heaps. To formalize this relational meaning, we need a semantic domain  $\mathbf{IRel}_n$  of relations, which we define and explain in this section.

Let  $\mathbf{Heap}$  be the set of finite partial functions from positive integers to integers (i.e.,  $\mathbf{Heap} \stackrel{\text{def}}{=} \mathbf{PosInt} \rightarrow_{\text{fin}} \mathbf{Int}$ ), ranged over by  $f, g, h$ . This is a commonly used set for modelling heaps in separation logic, and it has a partial commutative monoid structure  $(\square, \cdot)$ , where  $\square$  is the empty heap and the  $\cdot$  operator combines disjoint heaps:

$$\square \stackrel{\text{def}}{=} \emptyset, \quad f \cdot g \stackrel{\text{def}}{=} \begin{cases} f \cup g & \text{if } \text{dom}(f) \cap \text{dom}(g) = \emptyset \\ \text{undefined} & \text{otherwise} \end{cases}$$

The operator  $\cdot$  induces a partial order  $\sqsubseteq$  on  $\mathbf{Heap}$ , modelling heap extension, by  $f \sqsubseteq g$  iff  $g = f \cdot h$  for some  $h$ .

We also consider the  $+$  operator for combining possibly-overlapping but consistent heaps, and the  $-$  operator for subtracting one heap from another:

$$f + g \stackrel{\text{def}}{=} \begin{cases} f \cup g & \text{if } \forall l \in \text{dom}(f) \cap \text{dom}(g). f(l) = g(l) \\ \text{undefined} & \text{otherwise} \end{cases}$$

$$(f - g)(l) \stackrel{\text{def}}{=} \begin{cases} f(l) & \text{if } l \in \text{dom}(f) \setminus \text{dom}(g) \\ \text{undefined} & \text{otherwise} \end{cases}$$

We call an  $n$ -ary relation  $r \subseteq \mathbf{Heap}^n$  *upward closed* iff  $(f_1, \dots, f_n) \in r \wedge (\forall i. f_i \sqsubseteq g_i) \implies (g_1, \dots, g_n) \in r$ .

**Definition 1**  $\mathbf{IRel}_n$  is the family of upward closed  $n$ -ary relations on heaps.

Note that  $\mathbf{IRel}_1$  consists of upward closed *sets* of heaps, which are frequently used to interpret assertions in separation logic for garbage-collected languages. We call elements of  $\mathbf{IRel}_1$  *predicates* and denote them by  $p, q$ .

For every  $n \geq 1$ , domain  $\mathbf{IRel}_n$  has a complete lattice structure: join and meet are given by union and intersection, bottom is the empty relation, and top is  $\mathbf{Heap}^n$ . The domain also has a semantic separating conjunction connective defined by

$$(f_1, \dots, f_n) \in r * s \stackrel{\text{def}}{\iff} \exists (g_1, \dots, g_n) \in r. \exists (h_1, \dots, h_n) \in s. \\ (f_1, \dots, f_n) \cdot (g_1, \dots, g_n) = (h_1, \dots, h_n).$$

Here we use the component-wise extension of  $\cdot$  for tuples. Intuitively, a tuple is related by  $r * s$  when it can be splitted into two disjoint tuples, one related by  $r$  and the other by  $s$ .

The domain  $\mathbf{IRel}_1$  of predicates is related to  $\mathbf{IRel}_n$  for every  $n$ , by the map  $\Delta_n \stackrel{\text{def}}{=} \lambda p. \{(f, \dots, f) \mid f \in p\}^\uparrow$ , where  $\uparrow$  is the upward closure on relations. Note that each predicate is turned into an  $n$ -ary identity relation on  $p$  modulo the upward closure. This map behaves well with respect to the structures discussed on  $\mathbf{IRel}_1$  and  $\mathbf{IRel}_n$ , as expressed by the lemma below:

**Lemma 2** Function  $\Delta_n$  preserves the complete lattice structure and the  $*$  operator.

## 3 Assertions and Relational Semantics

Let  $\mathbf{Var}$  and  $\mathbf{AVar}$  be disjoint sets of normal variables  $x, y, \dots$  and assertion variables  $a, b, \dots$ , respectively. Our assertions  $\varphi$  are from higher-order separation logic, and they conform to the following grammar:

$$E ::= x \mid 0 \mid 1 \mid E + E \mid \dots \quad P ::= E \hookrightarrow E \mid \dots \\ \varphi ::= P \mid a \mid \varphi * \varphi \mid \mathbf{true} \mid \varphi \wedge \varphi \mid \mathbf{false} \mid \varphi \vee \varphi \\ \quad \mid \forall x. \varphi \mid \exists x. \varphi$$

$\llbracket P \rrbracket_{\eta, \rho}^n \stackrel{\text{def}}{=} \Delta_n(\llbracket P \rrbracket_{\eta})$	$\llbracket \varphi * \psi \rrbracket_{\eta, \rho}^n \stackrel{\text{def}}{=} \llbracket \varphi \rrbracket_{\eta, \rho}^n * \llbracket \psi \rrbracket_{\eta, \rho}^n$
$\llbracket a \rrbracket_{\eta, \rho}^n \stackrel{\text{def}}{=} \rho(a)$	$\llbracket \varphi \wedge \psi \rrbracket_{\eta, \rho}^n \stackrel{\text{def}}{=} \llbracket \varphi \rrbracket_{\eta, \rho}^n \cap \llbracket \psi \rrbracket_{\eta, \rho}^n$
$\llbracket \text{true} \rrbracket_{\eta, \rho}^n \stackrel{\text{def}}{=} \text{Heap}^n$	$\llbracket \varphi \vee \psi \rrbracket_{\eta, \rho}^n \stackrel{\text{def}}{=} \llbracket \varphi \rrbracket_{\eta, \rho}^n \cup \llbracket \psi \rrbracket_{\eta, \rho}^n$
$\llbracket \text{false} \rrbracket_{\eta, \rho}^n \stackrel{\text{def}}{=} \emptyset$	$\llbracket \forall x. \varphi \rrbracket_{\eta, \rho}^n \stackrel{\text{def}}{=} \bigcap_{v \in \text{Int}} \llbracket \varphi \rrbracket_{\eta[x \mapsto v], \rho}^n$
	$\llbracket \exists x. \varphi \rrbracket_{\eta, \rho}^n \stackrel{\text{def}}{=} \bigcup_{v \in \text{Int}} \llbracket \varphi \rrbracket_{\eta[x \mapsto v], \rho}^n$

where  $\llbracket P \rrbracket_{\eta}$  is the standard semantics of  $P$  as an upward closed set of heaps, which satisfies:

$$\llbracket E \hookrightarrow F \rrbracket_{\eta} = \{f \mid \llbracket E \rrbracket_{\eta} \in \text{dom}(f) \wedge f(\llbracket E \rrbracket_{\eta}) = \llbracket F \rrbracket_{\eta}\}.$$

**Figure 3. Interpretation of Assertions**

In the grammar,  $E$  is a heap-independent expression, and  $P$  is a primitive predicate, which in the standard interpretation denotes an upward closed set of heaps. For instance,  $E \hookrightarrow E'$  means heaps containing cell  $E$  with contents  $E'$ . Formulas  $\varphi$  are assertions in higher-order separation logic that do not include negation nor implication. Note that these assertions can include assertion variables, thus called higher-order. The dots in the grammar indicate possible extensions of cases, such as multiplication for  $E$  and inductive predicates for  $P$ . We will use the abbreviation  $E \hookrightarrow_{\cdot}$  for  $\exists y. E \hookrightarrow y$ .

An assertion  $\varphi$  is given a meaning  $\llbracket \varphi \rrbracket_{\eta, \rho}^n \in \mathbf{IRel}_n$  as an  $n$ -ary relation on heaps, where  $n$  is a parameter of the semantics. Here environment  $\eta$  maps normal variables in  $\varphi$  to integers, and  $\rho$  maps assertion variables in  $\varphi$  to  $n$ -ary relations in  $\mathbf{IRel}_n$ . When  $\varphi$  does not contain any assertion variables, we often omit  $\rho$  and write  $\llbracket \varphi \rrbracket_{\eta}^n$ , because the meaning of  $\varphi$  does not depend on  $\rho$ .

We define the semantics of  $\varphi$ , using the complete lattice structure and the  $*$  operator of the domain  $\mathbf{IRel}_n$ ; see Figure 3. Note that the relational semantics of primitive predicates is defined by embedding their standard meanings via  $\Delta_n$ . In fact, this embedding relationship holds for all assertions, because  $\Delta_n$  preserves the semantic structures of the domains (Lemma 2).

**Lemma 3** *For all  $\varphi$  and  $\eta, \rho, \rho'$ , if  $\Delta_n(\rho(a)) = \rho'(a)$  for every  $a \in \text{AVar}$ , we have that  $\Delta_n(\llbracket \varphi \rrbracket_{\eta, \rho}^1) = \llbracket \varphi \rrbracket_{\eta, \rho'}^1$ .*

We write  $\varphi \models^n \psi$  to mean that  $\llbracket \varphi \rrbracket_{\eta, \rho}^n \subseteq \llbracket \psi \rrbracket_{\eta, \rho}^n$  holds for all environments  $\eta, \rho$ . If  $n=1$ , this reduces to the standard semantics of assertions in separation logic. We will use the phrase “ $\varphi \implies \psi$  is  $n$ -ary valid” to mean that  $\varphi \models^n \psi$  holds. In addition, we write  $\varphi \models_{\eta}^n \psi$  for a fixed  $\eta$  to mean that  $\llbracket \varphi \rrbracket_{\eta, \rho}^n \subseteq \llbracket \psi \rrbracket_{\eta, \rho}^n$  holds for all environments  $\rho$ ; we say that “ $\varphi \implies \psi$  is  $n$ -ary  $\eta$ -valid” if this is true.

## 4 Lifting Theorems and Completeness

We call an assertion  $\varphi$  *simple* if it is of the form  $(\bigvee_{i=1}^I \bigwedge_{j=1}^J \varphi_{(i,j)} * \mathbf{a}_{(i,j)})$ , where  $\mathbf{a}_{(i,j)}$  is a vector of assertion variables and  $\varphi_{(i,j)}$  is an assertion not containing any assertion variables. We will consider the question of lifting an implication between simple assertions  $\varphi, \psi$  to a binary relational interpretation: when does  $\varphi \models^1 \psi$  imply that  $\varphi \models^2 \psi$ ?

The starting point is to realize that it is sufficient to study implications of the form:

$$\bigwedge_{i=1}^M \varphi_i * a_{i,1} * \dots * a_{i, M_i} \implies \bigvee_{j=1}^N \psi_j * b_{j,1} * \dots * b_{j, N_j} \quad (1)$$

where  $\varphi_i$ 's and  $\psi_j$ 's do not contain assertion variables, and no assertion variables occur only on the right hand side of the implication.

**Lemma 4** *There is an algorithm taking simple assertions  $\varphi, \psi$  and returning finitely many implications  $\{\varphi^l \implies \psi^l\}_{l \in L}$ , such that (a)  $\varphi^l \implies \psi^l$  has the form (1) and (b) for any  $n \in \{1, 2\}$ , we have that  $\varphi \models^n \psi$  holds iff  $\varphi^l \models^n \psi^l$  holds for all  $l \in L$ .*

The algorithm in the lemma is given in Appendix B.

Thus, in this section, we will focus on lifting implications of the form (1). Specifically, we will give a *complete* answer to the following question: Given one such implication that is  $\eta$ -valid in the unary interpretation for some environment  $\eta$ , can we decide if the implication is  $\eta$ -valid in the binary interpretation *merely* by inspection of the layout of the assertion variables? The answer will come in two parts. The first part, in Section 4.2, provides three lifting theorems, each of which has a criterion on the variable layout that, if met, implies that  $\eta$ -validity may be lifted regardless of the  $\varphi_i$ 's and  $\psi_j$ 's. The second part, in Section 4.3, is a completeness theorem; it states that if the variables fail the criteria of all three lifting theorems then there are choices of  $\varphi_i$ 's and  $\psi_j$ 's with no variables such that we have unary but not binary validity.

This approach has ups and downs. Assume that we have an implication of the aforementioned form that is valid in the unary interpretation, and we would like to know if it is valid in the binary interpretation too. Trying out the layout of the variables against the criteria of the three lifting theorems is an easily decidable and purely syntactical process – and if it succeeds then we have binary validity. If it fails, however, we are at a loss; we know that there are  $\varphi_i$ 's and  $\psi_j$ 's with the same variable layout such that lifting fails but we do not learn anything about our concrete implication. There is, however, an alternate use of the theory below if the

lifting criteria fail: We can sidestep the general lifting theorems and try to verify directly the Parametricity Condition from Definition 6 for all environments  $\eta$ . It is a semantic condition and probably undecidable in general, but it involves no assertion variables and only unary semantics – and if it holds then, by Proposition 7, we have binary validity. Indeed, the lifting theorems with direct verification of the Parametricity Condition as backup can be thought of as an algorithm. It is complete in some cases: Theorem 15 expresses that binary  $\eta$ -validity and the Parametricity Condition imply each other if  $M_i \leq 2$  for all  $1 \leq i \leq M$ .

#### 4.1 Notation

We need some notation that will accompany us throughout this section. Consider an implication of the form (1). Let  $V = \bigcup_{i=1}^M \{a_{i,1}, \dots, a_{i,M_i}\}$  be the set of all left hand side assertion variables, these include the right hand side assertion variables too by assumption. Define  $\Pi : \{1, \dots, M\} \rightarrow \text{Nat}^V$  and  $\Omega : \{1, \dots, N\} \rightarrow \text{Nat}^V$  by the following:

$$\Pi(i)(c) \stackrel{\text{def}}{=} |\{k \mid a_{i,k} \equiv c\}|, \quad \Omega(j)(c) \stackrel{\text{def}}{=} |\{k \mid b_{j,k} \equiv c\}|.$$

These functions give vectors of assertion variable counts for each conjunct and disjunct. For  $1 \leq i \leq M$  and  $1 \leq j \leq N$  we write  $\Pi(i) \geq \Omega(j)$  if we have  $\Pi(i)(c) \geq \Omega(j)(c)$  for each variable  $c \in V$ , i.e., if conjunct  $i$  has the same or a greater number of occurrences of all variables than disjunct  $j$ . We write  $\Pi(i) \not\geq \Omega(j)$  if this fails, i.e., if there is  $c \in V$  such that  $\Pi(i)(c) < \Omega(j)(c)$ . If a conjunct, say conjunct  $i$ , has no variables, i.e., if  $\Pi(i)(c) = 0$  holds for all  $c \in V$ , then we say it *empty*; the same goes for the disjuncts.

We shall write  $-$  to denote  $\exists n, m. n \multimap m$ , meaning heaps with at least one cell. On the semantic side, we write  $[m]$  for  $m \in \text{PosInt}$  to denote the heap that stores 0 at location  $m$  and nothing else. For  $m_0, \dots, m_n \in \text{PosInt}$  different we write  $[m_0, \dots, m_n]$  for  $[m_0] \cdot \dots \cdot [m_n]$ .

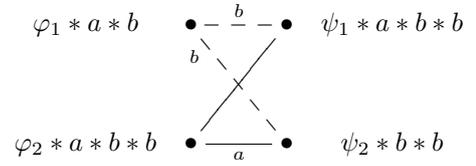
Finally we introduce a piece of sanity-preserving graphical notation. We depict an implication of the form (1) as a complete bipartite graph with the conjuncts lined up on the left hand side and the disjuncts on the right hand side. For any  $1 \leq i \leq M$  and any  $1 \leq j \leq N$  we draw a solid line from conjunct  $i$  to disjunct  $j$  if  $\Pi(i) \geq \Omega(j)$ . We label that line with all the  $c \in V$  such that  $\Pi(i)(c) > \Omega(j)(c)$  if indeed there are any such. If, on the other hand,  $\Pi(i) \not\geq \Omega(j)$  then we draw a dashed line instead and label it with all the  $c \in V$  such that  $\Pi(i)(c) < \Omega(j)(c)$ . Note that the drawing of edges depend solely on the layout of the variables; the  $\varphi_i$ 's and  $\psi_j$ 's have no say in the matter. As examples we refer the reader to Examples 12 and 13 for graphical

representations of  $1 \multimap \_ \wedge a * b \implies 1 \multimap \_ * a \vee 1 \multimap \_ * b$  and  $\_ * a * b \wedge a * a \implies \_ * a * a \vee \_ * \_ * b$  respectively.

#### 4.2 Layouts that Lift

The following is a first example of a layout of variables that ensure that for any choice of  $\varphi_i$ 's and  $\psi_j$ 's we get that unary  $\eta$ -validity of the implication yields binary  $\eta$ -validity. That it holds is a consequence of Theorem 8 but we have spelled out a concrete proof that will serve as a guide to the further development.

**Example 5 (Shadow-Lift)** For any four assertions  $\varphi_1, \varphi_2, \psi_1, \psi_2$  with no assertion variables and any appropriate environment  $\eta$  we have that unary  $\eta$ -validity of the following implication implies binary  $\eta$ -validity:



Assume that we have unary  $\eta$ -validity. Before we go on to consider the binary case we derive a simple unary consequence that does not involve assertion variables: For any  $h \in \text{Heap}$  with subheaps  $h_1 \sqsubseteq h$  and  $h_2 \sqsubseteq h$  such that  $h_1 \in \llbracket \varphi_1 \rrbracket_\eta^1$  and  $h_2 \in \llbracket \varphi_2 \rrbracket_\eta^1$  we get that  $h_2 \in \llbracket \psi_1 \rrbracket_\eta^1$  or that  $h_2 \in \llbracket \psi_2 \rrbracket_\eta^1$ .

To prove this, let  $h, h_1$  and  $h_2$  be as assumed. We build  $\rho : \{a, b\} \rightarrow \text{IRel}_1$  by letting  $\rho(a) = \text{Heap}$  and letting  $\rho(b)$  be the following union of sets of heaps:

$$\{(h - h_1) \cdot [n, n + 1]\}^\uparrow \cup \{(h - h_2) \cdot [n]\}^\uparrow \cup \{[n + 1]\}^\uparrow$$

where  $n = \max(\text{dom}(h) \cup \{0\}) + 1$ . It is now immediate that  $h \cdot [n, n + 1]$  lies in the interpretation of both conjuncts and by our assumption on the original implication it must lie in the interpretation on one of the disjuncts too. Suppose that we have

$$h \cdot [n, n + 1] \in \llbracket \psi_1 * a * b * b \rrbracket_{\eta, \rho}^1 = \llbracket \psi_1 \rrbracket_\eta^1 * \rho(b) * \rho(b),$$

where the equality holds because  $\rho(a) = \text{Heap}$  is the unit for  $*$ . We then write  $h \cdot [n, n + 1] = g_1 \cdot g_2 \cdot g_3$  for  $g_1 \in \llbracket \psi_1 \rrbracket_\eta^1$  and  $g_2, g_3 \in \rho(b)$ . But as  $g_2$  and  $g_3$  have disjoint domains we must have  $(h - h_2) \cdot [n] \sqsubseteq g_2$  and  $[n + 1] \sqsubseteq g_3$  or the version with  $g_2$  and  $g_3$  swapped. In any case we have that

$$\begin{aligned} \text{dom}(g_1) &= \text{dom}(h \cdot [n, n + 1]) \setminus (\text{dom}(g_2 \cdot g_3)) \\ &\subseteq \text{dom}(h \cdot [n, n + 1]) \setminus (\text{dom}(h - h_2) \cup \{n, n + 1\}) \\ &= \text{dom}(h_2). \end{aligned}$$

But then we have  $g_1 \sqsubseteq h_2$  and since  $g_1 \in \llbracket \psi_1 \rrbracket_\eta^1$  we get  $h_2 \in \llbracket \psi_1 \rrbracket_\eta^1$  too. If we have  $h \cdot [n, n + 1] \in \llbracket \psi_2 * b * b \rrbracket_{\eta, \rho}^1$  we proceed similarly.

The above short proof is the crux of the example. It implies unary  $\eta$ -validity – this we knew already – but also the binary  $\eta$ -validity. To see this, we pick an arbitrary environment  $\rho : \{a, b\} \rightarrow \mathbf{IRel}_2$ , we take arbitrary  $(h_1, h_2) \in \llbracket \varphi_1 * a * b \wedge \varphi_2 * a * b * b \rrbracket_{\eta, \rho}^2$  and we aim to prove that  $(h_1, h_2) \in \llbracket \psi_1 * a * b * b \vee \psi_2 * b * b \rrbracket_{\eta, \rho}^2$  too. We split  $(h_1, h_2)$  according to the conjuncts. Because of Lemma 3 and the upward closedness condition of  $\mathbf{IRel}_2$ , we can write

$$(h_1, h_2) = (g^1, g^1) \cdot (g_1^2, g_2^2) \cdot (g_1^3, g_2^3)$$

for  $g^1 \in \llbracket \varphi_1 \rrbracket_{\eta}^1$ ,  $(g_1^2, g_2^2) \in \rho(a)$  and  $(g_1^3, g_2^3) \in \rho(b)$ . Also we can write

$$(h_1, h_2) = (f^1, f^1) \cdot (f_1^2, f_2^2) \cdot (f_1^3, f_2^3) \cdot (f_1^4, f_2^4)$$

for  $f^1 \in \llbracket \varphi_2 \rrbracket_{\eta}^1$ ,  $(f_1^2, f_2^2) \in \rho(a)$  and  $(f_1^3, f_2^3), (f_1^4, f_2^4) \in \rho(b)$ . But now  $g^1 + f^1$  with subheaps  $g^1$  and  $f^1$  fulfills the above properties and so we get  $f^1 \in \llbracket \psi_1 \rrbracket_{\eta}^1$  or  $f^1 \in \llbracket \psi_2 \rrbracket_{\eta}^1$  and the second splitting of  $(h_1, h_2)$  shows that  $(h_1, h_2)$  lie in the binary interpretation of the first or second disjunct, respectively. Notice that neither  $g^1 \in \llbracket \psi_1 \rrbracket_{\eta}^1$  nor  $g^1 \in \llbracket \psi_2 \rrbracket_{\eta}^1$  would have worked since the first conjunct has too few variables, i.e.,  $\Pi(1) \not\geq \Omega(1)$  and  $\Pi(1) \not\geq \Omega(2)$  ■

Generalizing the unary consequence that served as the crucial stepping stone in the above example we arrive at the following condition on our implications:

**Definition 6 (Parametricity Condition)** *Assume that we have an implication of the form (1) and an appropriate environment  $\eta$ . For all  $h, h_1, \dots, h_M \in \mathbf{Heap}$  with  $h_i \sqsubseteq h$  and  $h_i \in \llbracket \varphi_i \rrbracket_{\eta}^1$  for all  $1 \leq i \leq M$  we must have one or two of these options:*

1. *There are  $1 \leq i \leq M$  and  $1 \leq j \leq N$  such that  $h_i \in \llbracket \psi_j \rrbracket_{\eta}^1$ ,  $\Pi(i) \geq \Omega(j)$  and the  $j$ -th disjunct is not empty.*
2. *There is  $1 \leq j \leq N$  such that  $h \in \llbracket \psi_j \rrbracket_{\eta}^1$  and the  $j$ -th disjunct is empty.*

Note that specializing the Parametricity Condition, henceforth just the PC, to an implication of the form treated in the above example yields the stated unary consequence because no disjuncts are empty. The second option in the PC will be motivated later.

We emphasize that the PC may hold or may fail for any given combination of an implication and environment  $\eta$ . But if it holds then we have binary  $\eta$ -validity; the proof in case of the first option of the PC is an easy generalization of the latter half of the above example:

**Proposition 7** *The PC implies binary  $\eta$ -validity.*

We arrive now at the first lifting theorem. It is a generalization of the former half of Example 5; the proof of the theorem has a lot more details to it than the example but the overall idea is the same. The theorem states a criterion on the layout of the variables that, if met, means that unary  $\eta$ -validity implies the PC and hence also binary  $\eta$ -validity. The criterion is, loosely, that we can remove all variables that occur as labels of solid lines without introducing new solid lines and without emptying any disjuncts:

**Theorem 8 (Shadow-Lift)** *Unary  $\eta$ -validity of an implication implies the PC if each dashed line has a label that is not a label on a solid line and each disjunct has an occurrence of a variable that is not a label on a solid line. Spelling it out in symbols, we require, with  $L = \{(i, j) \mid 1 \leq i \leq M \wedge 1 \leq j \leq N\}$ , that*

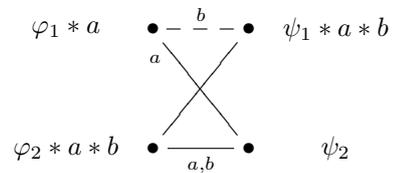
$$\begin{aligned} \forall (i, j) \in L. \Pi(i) \not\geq \Omega(j) &\implies \\ \exists c \in V. \Pi(i)(c) < \Omega(j)(c) \wedge \\ (\forall (k, l) \in L. \Pi(k) \geq \Omega(l)) &\implies \Pi(k)(c) = \Omega(l)(c) \end{aligned}$$

and

$$\begin{aligned} \forall 1 \leq j \leq N. \exists c \in V. \Omega(j)(c) > 0 \wedge \\ (\forall (k, l) \in L. \Pi(k) \geq \Omega(l)) &\implies \Pi(k)(c) = \Omega(l)(c). \end{aligned}$$

As motivation for the next lifting theorem, we note that the variable layout criterion of the above theorem fails if one or more disjuncts are empty. Correspondingly, we never touch upon the second option of the PC. But there are variable layouts with empty disjuncts that ensure lifting:

**Example 9 (Balloon-Lift)** For any four assertions  $\varphi_1, \varphi_2, \psi_1, \psi_2$  with no assertion variables and any appropriate environment  $\eta$  we have that unary  $\eta$ -validity of the following implication implies binary  $\eta$ -validity:



Assume unary  $\eta$ -validity. As in Example 5 we derive a unary consequence as intermediate result: For any  $h \in \mathbf{Heap}$  with subheaps  $h_1 \sqsubseteq h$  and  $h_2 \sqsubseteq h$  such that  $h_1 \in \llbracket \varphi_1 \rrbracket_{\eta}^1$  and  $h_2 \in \llbracket \varphi_2 \rrbracket_{\eta}^1$  we have that either  $h_2 \in \llbracket \psi_1 \rrbracket_{\eta}^1$  or  $h \in \llbracket \psi_2 \rrbracket_{\eta}^1$ .

To prove this, let  $h, h_1$  and  $h_2$  be as assumed. We construct  $\rho : \{a, b\} \rightarrow \mathbf{IRel}_1$  by letting  $\rho(a) = \mathbf{Heap}$  and

$\rho(b) = \{h - h_2\}^\uparrow$ . We get that

$$h \sqsupseteq h_1 \in \llbracket \varphi_1 \rrbracket_\eta^1 = \llbracket \varphi_1 \rrbracket_\eta^1 * \mathbf{Heap} = \llbracket \varphi_1 * a \rrbracket_{\eta, \rho}^1,$$

and

$$\begin{aligned} h &= h_2 \cdot (h - h_2) \\ &\in \llbracket \varphi_2 \rrbracket_\eta^1 * \rho(b) = \llbracket \varphi_2 \rrbracket_\eta^1 * \mathbf{Heap} * \rho(b) = \llbracket \varphi_2 * a * b \rrbracket_{\eta, \rho}^1. \end{aligned}$$

This means that  $h$  must lie in the interpretation of one of the disjuncts. If it is the first, we inspect the interpretation and get that

$$h = g_1 \cdot g_2 \cdot g_3$$

for  $g_1 \in \llbracket \psi_1 \rrbracket_\eta^1$ ,  $g_2 \in \mathbf{Heap}$  and  $g_3 \sqsupseteq h - h_2$ . It means that

$$\begin{aligned} \text{dom}(g_1) &= \text{dom}(h) \setminus \text{dom}(g_2 \cdot g_3) \subseteq \text{dom}(h) \setminus \text{dom}(g_2) \\ &\subseteq \text{dom}(h) \setminus \text{dom}(h - h_2) = \text{dom}(h_2) \end{aligned}$$

which implies that  $g_1 \sqsubseteq h_2$  and so  $h_2 \in \llbracket \psi_1 \rrbracket_\eta^1$ . If, on the other hand,  $h$  lies in the interpretation of the second disjunct then we are done immediately.

Now we prove the claim of binary  $\eta$ -validity. We pick an arbitrary environment  $\rho : \{a, b\} \rightarrow \mathbf{IRel}_2$ , we take arbitrary  $(h_1, h_2) \in \llbracket \varphi_1 * a \wedge \varphi_2 * a * b \rrbracket_{\eta, \rho}^2$  and we must prove that  $(h_1, h_2) \in \llbracket \psi_1 * a * b \vee \psi_2 \rrbracket_{\eta, \rho}^2$  too. We write

$$(h_1, h_2) = (g^1, g^1) \cdot (g_1^2, g_2^2)$$

for  $g^1 \in \llbracket \varphi_1 \rrbracket_\eta^1$  and  $(g_1^2, g_2^2) \in \rho(a)$ , and

$$(h_1, h_2) = (f^1, f^1) \cdot (f_1^2, f_2^2) \cdot (f_1^3, f_2^3)$$

for  $f^1 \in \llbracket \varphi_2 \rrbracket_\eta^1$ ,  $(f_1^2, f_2^2) \in \rho(a)$  and  $(f_1^3, f_2^3) \in \rho(b)$ . But now  $g^1 + f^1$  with subheaps  $g^1$  and  $f^1$  satisfies the above properties and so we get  $f^1 \in \llbracket \psi_1 \rrbracket_\eta^1$  or  $g^1 + f^1 \in \llbracket \psi_2 \rrbracket_\eta^1$ . If  $f^1 \in \llbracket \psi_1 \rrbracket_\eta^1$  holds then the second splitting shows that  $(h_1, h_2)$  is in the interpretation of the first disjunct. If  $g^1 + f^1 \in \llbracket \psi_2 \rrbracket_\eta^1$ , we are done too, since we may write  $(h_1, h_2) = (g^1 + f^1, g^1 + f^1) \cdot (e_1, e_2)$  for some  $(e_1, e_2) \in \mathbf{Heap}^2$  and so  $(h_1, h_2)$  lies in the interpretation  $\llbracket \psi_2 \rrbracket_\eta^2 = \Delta(\llbracket \psi_2 \rrbracket_\eta^1)$  of the second conjunct. ■

Just as we did for Example 5 we may generalize the former half of this example, this yields Theorem 10 below. The latter half of the example, on the other hand, constitutes an example of the approach of the proof of Proposition 7 in case we run into the second option of the PC.

**Theorem 10 (Balloon-Lift)** *Unary  $\eta$ -validity of an implication implies the PC if there is a subset  $B \subseteq V$  with the following three properties. First, each conjunct has at most one occurrence of a variable from  $B$ , i.e.,*

$$\forall 1 \leq i \leq M. \sum_{c \in B} \Pi(i)(c) \leq 1.$$

*Second, each disjunct is empty or has exactly one occurrence of a variable from  $B$ , i.e.,*

$$\forall 1 \leq j \leq N. \sum_{c \in V} \Omega(j)(c) = 0 \vee \sum_{c \in B} \Omega(j)(c) = 1.$$

*Third, each dashed line must have a label from  $B$ . That is, when  $L = \{(i, j) \mid 1 \leq i \leq M \wedge 1 \leq j \leq N\}$ ,*

$$\forall (i, j) \in L. \Pi(i) \not\leq \Omega(j) \implies \exists c \in B. \Pi(i)(c) < \Omega(j)(c).$$

One thing to note about the theorem is that if we have no empty disjuncts, none of the variables in the subset  $B \subseteq V$  can be labels of a solid line. In particular, the conditions of Theorem 8 are met, so the above theorem is really only useful if one or more disjuncts are empty.

The final lifting theorem captures the oddities of the special case of just one conjunct:

**Theorem 11 (Lonely-Lift)** *Unary  $\eta$ -validity of an implication implies the PC if there is just one conjunct, i.e.,  $M=1$ , and all lines are solid, i.e.,  $\Pi(1) \geq \Omega(j)$  for all  $1 \leq j \leq N$ .*

### 4.3 Completeness

It is now time for examples of implications that do not lift, i.e., that are valid in the unary interpretation but not in the binary. The first is based on the following observation: If  $h \in \llbracket 1 \leftrightarrow \_ \rrbracket_\eta^1$  and  $h \in p * q$  for  $h \in \mathbf{Heap}$  and  $p, q \in \mathbf{IRel}_1$  then we have  $h \in \llbracket 1 \leftrightarrow \_ \rrbracket_\eta^1 * p$  or  $h \in \llbracket 1 \leftrightarrow \_ \rrbracket_\eta^1 * q$ . This is because we must have  $[1 \mapsto n] \sqsubseteq h$  for some  $n \in \mathbf{Int}$  and so writing  $h = h_1 \cdot h_2$  with  $h_1 \in p$  and  $h_2 \in q$  gives us  $[1 \mapsto n] \sqsubseteq h_1$  or  $[1 \mapsto n] \sqsubseteq h_2$ . But this line of argument breaks down if we change to binary reading. We have, e.g.,  $([1], [1]) \in \llbracket 1 \leftrightarrow \_ \rrbracket_\eta^2$  and  $([1], [1]) \in \{([1], [])\}^\uparrow * \{([], [1])\}^\uparrow$  but both  $\llbracket 1 \leftrightarrow \_ \rrbracket_\eta^2 * \{([1], [])\}^\uparrow$  and  $\llbracket 1 \leftrightarrow \_ \rrbracket_\eta^2 * \{([], [1])\}^\uparrow$  are empty. We can recast this as an implication that cannot be lifted:

**Example 12 (Fan-Counter)** This implication is valid on the unary but not on the binary level:

$$\begin{array}{ccc} 1 \leftrightarrow \_ & \bullet \text{---} \frac{a}{\text{---}} \text{---} \bullet & 1 \leftrightarrow \_ * a \\ & \text{b} \text{---} \diagdown & \\ & \text{b} \text{---} \diagup & \\ a * b & \bullet \text{---} \frac{a}{\text{---}} \text{---} \bullet & 1 \leftrightarrow \_ * b \end{array}$$

First we argue that the implication holds on the unary level. Let  $\rho : \{a, b\} \rightarrow \mathbf{IRel}_1$  be an arbitrary environment of upwards closed sets of heaps to  $a$  and  $b$ . Let  $h \in \mathbf{Heap}$  be arbitrary and assume that

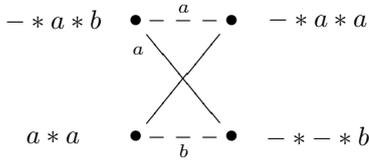
$$h \in \llbracket 1 \leftrightarrow \_ \wedge (a * b) \rrbracket_{\eta, \rho}^1 = \llbracket 1 \leftrightarrow \_ \rrbracket_\rho^1 \cap (\rho(a) * \rho(b)).$$

By the above observation we get either  $h \in \llbracket 1 \leftrightarrow \_ \rrbracket_{\eta}^1 * \rho(a)$  or  $h \in \llbracket 1 \leftrightarrow \_ \rrbracket_{\eta}^1 * \rho(b)$  which matches the right hand side of the implication.

Now we move on to prove that the implication fails on the binary level. Define an environment  $\rho : \{a, b\} \rightarrow \text{IRel}_2$  by  $\rho(a) = \{([1], [])\}^\uparrow$  and  $\rho(b) = \{([], [1])\}^\uparrow$ . Then,  $\llbracket 1 \leftrightarrow \_ \wedge a * b \rrbracket_{\eta, \rho}^2 = \llbracket 1 \leftrightarrow \_ \rrbracket_{\eta, \rho}^2 \cap (\rho(a) * \rho(b))$ , which contains the pair  $([1], [1])$ . But, as observed, both disjuncts have empty binary interpretations. ■

An observation of similar nature is that for  $p \in \text{IRel}_1$  we have either  $p = \text{Heap}$  or  $p \subseteq \llbracket - \rrbracket_{\eta}^1 = \{[m \mapsto n] \mid m \in \text{PosInt}, n \in \text{Int}\}^\uparrow$  because if  $h \neq \text{Heap}$  then it cannot contain the empty heap. On the binary level, however, we have  $\text{Heap}^2 \neq \{([1], [])\}^\uparrow \not\subseteq \llbracket - \rrbracket_{\eta}^2 = \{([m \mapsto n], [m \mapsto n]) \mid m \in \text{PosInt}, n \in \text{Int}\}^\uparrow$ . One consequence is this:

**Example 13 (Bridge-Counter)** This implication is valid on the unary but not on the binary level:



First we argue that the implication holds on the unary level. Let  $\rho : \{a, b\} \rightarrow \text{IRel}_1$  be an arbitrary environment that assigns upwards closed sets of heaps to each of the two variables. We branch on the value of  $\rho(a)$ . If  $\rho(a) \neq \text{Heap}$  then we have  $\rho(a) \subseteq \llbracket - \rrbracket_{\eta}^1$  which again means that the first conjunct directly implies the second disjunct. If  $\rho(a) = \text{Heap}$  holds, we get that

$$\begin{aligned} \llbracket - * a * b \rrbracket_{\eta, \rho}^1 &= \llbracket - \rrbracket_{\eta, \rho}^1 * \text{Heap} * \rho(b) = \llbracket - \rrbracket_{\eta, \rho}^1 * \rho(b) \\ &\subseteq \llbracket - \rrbracket_{\eta, \rho}^1 = \llbracket - \rrbracket_{\eta, \rho}^1 * \text{Heap} * \text{Heap} = \llbracket - * a * a \rrbracket_{\eta, \rho}^1 \end{aligned}$$

because  $\text{Heap}$  is the unit for  $*$ . Hence we get that the first conjunct implies the first disjunct and we have proved that the implication holds unarily.

Now we prove that the implication fails on the binary level. Define an environment  $\rho : \{a, b\} \rightarrow \text{IRel}_2$  by  $\rho(a) = \{([1], [])\}^\uparrow \cup \{([2], [2])\}^\uparrow$  and  $\rho(b) = \text{Heap}^2$ . Observe now that  $([1, 2], [2]) = ([2], [2]) \cdot ([1], []) \cdot ([], [])$ , which implies that  $([1, 2], [2]) \in \llbracket - * a * b \rrbracket_{\eta, \rho}^2$ . From the rewriting  $([1, 2], [2]) = ([1], []) \cdot ([2], [2])$ , we get  $([1, 2], [2]) \in \llbracket a * a \rrbracket_{\eta, \rho}^2$  too and so this pair of heaps lies in the interpretation of the left hand side. But it does not belong to the interpretation of either disjunct. An easy – if somewhat indirect – way of realizing this is to note that any pair of heaps in either  $\llbracket - \rrbracket_{\eta, \rho}^2$  or in  $\llbracket a * a \rrbracket_{\eta, \rho}^2$  must have a second component with nonempty domain. But then any pair of heaps in

the interpretation of either disjunct must have a second component with a domain of at least two elements. In particular, neither can contain the pair  $([1, 2], [2])$ . ■

In principle, the above two observations are all that we need to prove completeness. Or, phrased differently, assume that we have a layout of variables that fail the criteria of all three lifting theorems; by applying one of the two observations, we can then build a concrete implication with that variable layout and with unary but not binary validity.

Having said that, the territory to cover is huge; the full completeness proof is a lengthy and rather technical journey, the details of which do not provide much insight. We supply it as a series of lemmas in Appendix D; these include generalizations of Example 12 and Example 13 above. If one verifies the lemmas in the order listed and apply them as sketched then it is feasible, if not exactly easy, to prove the following:

**Theorem 14 (Completeness)** *If a variable layout meets none of the criteria in Theorems 8, 10 and 11, then there are choices of  $\varphi_i$ 's and  $\psi_j$ 's with no variables such we have unary but not binary validity.*

## 5 Higher Arities and Parametricity

We saw in Proposition 7 that the PC implies binary  $\eta$ -validity of an implication. It is easy to show that the PC also implies unary  $\eta$ -validity, either directly or by observing that binary implies unary. A natural question to ask is whether we can reverse this? Example 12 shows that unary validity does not entail the PC, because the latter fails for that concrete implication. But as binary validity fails too, we could hope that binary validity would enforce the PC. Unfortunately, this is not the case, as demonstrated by the implication

$$1 \leftrightarrow \_ \wedge a * a * b \implies 1 \leftrightarrow \_ * a \vee 1 \leftrightarrow \_ * b.$$

Here the PC is the same as for Example 12 and hence still is not true, but we do have binary validity. We do not, however, have ternary validity but the example could easily be scaled: having  $n$  occurrences of  $a$  in the second conjunct means  $n$ -ary but not  $n+1$ -ary validity for any  $n \geq 1$ . In summary, we have seen that for any  $n \geq 1$  we can have  $n$ -ary validity whilst the PC fails.

What does hold, however, is the following:

**Theorem 15** *For an implication of the form (1) and an appropriate environment  $\eta$  we have that  $n$ -ary  $\eta$ -validity implies the PC if  $n \geq \max\{M_1, \dots, M_M\}$ .*

Notice how this fits nicely with the above example: with  $n$  occurrences of  $a$  we have  $n$ -ary validity but we

need  $(n+1)$ -ary validity to prove the PC since there is also a single  $b$ . The proof is in Appendix E, and reuses techniques from the proofs of Theorems 8 and 10.

By an easy generalization of Proposition 7 we have the following corollary to the above theorem:

**Corollary 16** *The PC holds iff we have  $n$ -ary  $\eta$ -validity for all  $n \geq 1$ .*

This corollary can be read, loosely, as a coincidence between parametric polymorphism as introduced by Strachey [19] and relational parametricity as proposed by Reynolds [18]: The PC corresponds to Strachey parametricity in the loose sense that if it holds, then there is an approach, parametric in the assertion variables, that produce right hand side proofs of heap membership from the left hand side ones: Take a heap, split it along the conjuncts, apply the PC to the parts in the interpretations of the  $\varphi$ 's and you are done, possibly after discarding some variables. This involves no branching or other intrinsic operations on the assertion variables, which we are free to discard by our intuitionistic setup. If, on the other hand, the implication is  $\eta$ -valid for arbitrary arity, then it is fair to call it relationally parametric. Note also that the Examples 12 and 13 branch on assertion variable values.

This result is analogous to the conjecture of coincidence between Strachey parametricity and  $n$ -ary relational parametricity for traditional type-based parametricity [17, Page 2].

Finally we note that as a consequence of the above corollary we have that the lifting theorems in the previous really show that unary validity can be lifted to validity of arbitrary arity. In some sense, they are stronger than required for representation independence, for which binary validity suffices. The authors are unaware of any practical applications of this fact.

## 6 Representation Independence

In this section, we relate our lifting theorems to representation independence. We consider separation logic with assertion variables where the rule of consequence is restricted according to our lifting theorems, and we define a relational semantics of the logic, which gives a representation independence theorem: all proved clients cannot distinguish between appropriately related module implementations. For space reasons, our presentation will be minimalistic, covering only a part of the logic; the missing parts, in particular allocation of new cells, we believe, can be handled using FM sets, following [6].

We consider commands  $C$  given by the grammar:

$$C ::= k \mid [E]:=E \mid \mathbf{let} \ y=[E] \ \mathbf{in} \ C \mid C; C \mid \mathbf{if} \ B \ C \ C$$

$$\begin{array}{c} \hline \text{CHK}(\varphi', \varphi) \quad \varphi' \models^1 \varphi \quad \{\varphi\}C\{\psi\} \quad \psi \models^1 \psi' \quad \text{CHK}(\psi, \psi') \\ \hline \{\varphi'\}C\{\psi'\} \\ \hline \frac{\{\varphi\}C\{\varphi'\}}{\{\varphi * \psi\}C\{\varphi' * \psi\}} \quad \frac{\{\varphi\}C\{\psi\}}{\{\exists x. \varphi\}C\{\exists x. \psi\}} \quad x \notin \text{FV}(C) \\ \hline \frac{\Gamma, \{\varphi\}k\{\psi\} \vdash \{\varphi\}k\{\psi\}}{\Gamma, \{\varphi * E \leftrightarrow x\} \mathbf{let} \ x=[E] \ \mathbf{in} \ C\{\psi\}} \quad \frac{\{E \leftrightarrow \_ \}[E] := F\{E \leftrightarrow F\}}{\{\exists x. \varphi * E \leftrightarrow x\} \mathbf{let} \ x=[E] \ \mathbf{in} \ C\{\psi\}} \quad x \notin \text{FV}(\psi) \\ \hline \frac{\{\varphi\}C\{\varphi'\} \quad \{\varphi'\}C'\{\psi\}}{\{\varphi\}C; C'\{\psi\}} \quad \frac{\{\varphi \wedge B\}C\{\psi\} \quad \{\varphi \wedge \neg B\}C'\{\psi\}}{\{\varphi\} \mathbf{if} \ B \ C \ C'\{\psi\}} \\ \hline \end{array}$$

**Figure 4. Proof Rules**

Here  $B$  is a heap-independent boolean expression, such as  $x=0$ . Commands  $C$  are from the loop-free simple imperative language. They can call module operations  $k$ , and manipulate heap cells; command  $[x]:=E$  assigns  $E$  to the heap cell  $x$ , and this assigned value is read by  $\mathbf{let} \ y=[x] \ \mathbf{in} \ C$ , which also binds  $y$  to the read value and runs  $C$  under this binding.

Properties of commands  $C$  are specified using Hoare triples  $\Gamma \vdash \{\varphi\}C\{\psi\}$ , where the context  $\Gamma$  is a set of triples for module operations. Figure 4 shows rules for proving these properties. In the figure, we omit contexts, if the same context  $\Gamma$  is used for all the triples.

The rule of consequence deserves attention. Note that the rule uses semantic implications  $\models^1$  in the standard unary interpretation, thus allowing the use of existing theorem provers for (higher-order) separation logic. The rule does not allow all semantic implications, but only those that pass our algorithm CHK, so as to ensure that the implications can lift to the relational level. Our algorithm  $\text{CHK}(\varphi, \psi)$  performs two checks, and returns true only when both succeed. The first check is whether  $\varphi$  and  $\psi$  can be transformed to simple assertions  $\varphi'$  and  $\psi'$ , using only the distribution of  $*$  over  $\exists x$  and  $\vee$  and distributive lattice laws for  $\vee$  and  $\wedge$ . If this check succeeds and gives  $\varphi'$  and  $\psi'$ , the algorithm transforms  $\varphi \models^1 \psi'$  to a set of implications of the form (1) in Section 4 (Lemma 4). Then, it uses the method outlined just before Section 4.1, and does the second check on whether all the implications in the resulting set lift to the relational level.

Commands  $C$  are interpreted in a standard way, as functions of the type:  $\llbracket C \rrbracket_{\eta, u} \in \text{Heap} \rightarrow (\text{Heap} \cup \{\text{err}\})$ . Here  $\text{err}$  denotes a memory error, and  $\eta$  and  $u$  are environments that provide the meanings of, respectively, free ordinary variables and module operations. For instance,  $\llbracket k \rrbracket_{\eta, u}$  is  $u(k)$ .

Our semantics of triples, on the other hand, is not standard, and uses the binary interpretation of asser-

tions:  $(\eta, \rho, \mathbf{u}) \models^2 \{\varphi\}C\{\psi\}$  iff

$$\forall r \in \text{IRel}_2. \forall f, g \in \text{Heap}. (f, g) \in \llbracket \varphi \rrbracket_{\eta, \rho}^2 * r \implies (\llbracket C \rrbracket_{\eta, u_1}(f), \llbracket C \rrbracket_{\eta, u_2}(g)) \in \llbracket \psi \rrbracket_{\eta, \rho}^2 * r.$$

The environment  $\rho$  provides the meanings of assertion variables, and the 2-dimensional vector  $\mathbf{u}$  gives the two meanings for module operations; intuitively, each  $u_i$  corresponds to the  $i$ -th module implementation. The interpretation means that if two module implementations  $\mathbf{u}$  are used by the same client  $C$ , then these combinations should result in the same computation, in the sense that they map  $\varphi$ -related input heaps to  $\psi$ -related outputs. The satisfaction of triples can be extended to  $(\eta, \rho, \mathbf{u}) \models^2 \Gamma$ , by asking that all triples in  $\Gamma$  should hold wrt.  $(\eta, \rho, \mathbf{u})$ . Using these satisfaction relations on triples and contexts, we define the notion of 2-validity of judgements:  $\Gamma \vdash \{\varphi\}C\{\psi\}$  is 2-valid iff

$$\forall (\eta, \rho, \mathbf{u}). (\eta, \rho, \mathbf{u}) \models^2 \Gamma \implies (\eta, \rho, \mathbf{u}) \models^2 \{\varphi\}C\{\psi\}.$$

**Theorem 17** *Every derivable  $\Gamma \vdash \{\varphi\}C\{\psi\}$  is 2-valid.*

It is this theorem that we use to derive the representation independence results mentioned in the introduction. Consider again the example in Figure 1. Since the proof of the left hand side client  $C$  is derivable using the above rules, with  $\Gamma$  being the interface specification for the operations of the counter, we get 2-validity of  $\Gamma \vdash \{1 \leftrightarrow \_ \} C \{1 \leftrightarrow \_ \}$ . Therefore, when we run the client  $C$  with the related module implementations in the introduction, we find that  $C$  maps  $\llbracket 1 \leftrightarrow \_ \rrbracket^2$ -related heaps (i.e, heaps with the same value at cell 1) to  $\llbracket 1 \leftrightarrow \_ \rrbracket^2$ -related heaps again.

## 7 Conclusion and Discussion

In this paper, we have given a sound and complete characterization of when *semantic* implications between assertions in higher-order separation logic can be lifted to a relational interpretation. This characterization has, then, been used to identify proofs of clients that respect the abstraction of module internals, specified by means of assertion variables, and to show representation independence for clients with such proofs. We hope that our results provide a solid semantic basis for recent logic-based approaches to data abstraction.

In earlier work, Banerjee and Naumann [2] studied relational parametricity for dynamically allocated heap objects in a Java-like language. Banerjee and Naumann made use of a non-trivial semantic notion of confinement to describe internal resources of a module; here instead we use higher-order separation logic to describe which resources are internal to the module.

Relational interpretations have also been used to give models of programming languages with local state, which can validate representation independence results [13, 16, 3, 1]. These results typically rely on the module allocating the private state, whereas we use the power of separation logic and allow the ownership transfer of states from client to module. For instance, in the two-stage counter in the introduction, the ownership of the cell 1 is transferred from the client to the module upon calling `init`. Even with this ownership transfer, representation independence is guaranteed, because we consider only those clients having (good) proofs in separation logic. This contrasts with representation independence results in local state models, which consider not some but all well-typed clients.

## References

- [1] A. Ahmed, D. Dreyer, and A. Rossberg. State-dependent representation independence. In *POPL*, pages 340–353, 2009.
- [2] A. Banerjee and D. Naumann. Ownership confinement ensures representation independence for object-oriented programs. *JACM*, 52(6), 2005.
- [3] N. Benton and B. Leperchey. Relational reasoning in a nominal semantics for storage. In *TLCA*, pages 86–101, 2005.
- [4] B. Biering, L. Birkedal, and N. Torp-Smith. BI hyperdoctrines and higher-order separation logic. In *ESOP*, 2005.
- [5] L. Birkedal, N. Torp-Smith, and H. Yang. Semantics of separation-logic typing and higher-order frame rules. In *LICS*, pages 260–269, 2005.
- [6] L. Birkedal and H. Yang. Relational parametricity and separation logic. In *FOSSACS*, pages 93–107, 2007.
- [7] W.-N. Chin, C. David, H. H. Nguyen, and S. Qin. Enhancing modular oo verification with separation logic. In *POPL*, 2008.
- [8] D. Distefano and M. Parkinson. jStar: towards practical verification for java. In *OOPSLA*, pages 213–226, 2008.
- [9] R. Dockins, A. Hobor, and A. Appel. A fresh look at separation algebras and share accounting. In *APLAS*, pages 161–177, 2009.
- [10] B. Jacobs and F. Piessens. The VeriFast program verifier. Technical Report CW-520, Katholieke Universiteit Leuven, 2008.
- [11] A. Nanevski, A. Ahmed, G. Morrisett, and L. Birkedal. Abstract predicates and mutable ADTs in Hoare type theory. In *ESOP*, 2007.
- [12] A. Nanevski, G. Morrisett, A. Shinnar, P. Govereau, and L. Birkedal. Ynot: dependent types for imperative programs. In *ICFP*, pages 229–240, 2008.
- [13] P. O’Hearn and R. Tennent. Parametricity and local variables. *JACM*, 42(3):658–709, May 1995.
- [14] M. Parkinson and G. Bierman. Separation logic and abstraction. In *POPL*, pages 247–258, 2005.
- [15] M. Parkinson and G. Bierman. Separation logic, abstraction and inheritance. In *POPL*, pages 75–86, 2008.
- [16] A. Pitts and I. Stark. Observable properties of higher order functions that dynamically create local names, or: What’s new? In *MFCS*, pages 122–141, 1993.
- [17] G. Plotkin and M. Abadi. A logic for parametric polymorphism. In *TLCA*, pages 361–375, 1993.
- [18] J. C. Reynolds. Types, abstraction, and parametric polymorphism. In *Information Processing 83*, pages 513–523, 1983.
- [19] C. Strachey. Fundamental concepts in programming languages. In *Proceedings of the 1967 International Summer School in Computer Programming, Copenhagen, Denmark.*, 1967.

## A Proofs of Lemma 2 and Lemma 3

**Lemma 2** *Function  $\Delta_n$  preserves the complete lattice structure and the  $*$  operator.*

**Proof:** From the definition, it is immediate that  $\Delta_n(\text{Heap}) = \text{Heap}^n$  and  $\Delta_n(\emptyset) = \emptyset$ . Now consider a non-empty family  $\{p_i\}_{i \in I}$  of predicates in  $\text{IRel}_1$ . In order to show the preservation of the complete lattice structure, we need to prove that

$$\Delta_n\left(\bigcap_{i \in I} p_i\right) = \bigcap_{i \in I} \Delta_n(p_i) \quad \wedge \quad \bigcup_{i \in I} \Delta_n(p_i) = \Delta_n\left(\bigcup_{i \in I} p_i\right).$$

The  $\subseteq$  direction in both cases is easy; it follows from the monotonicity of  $\Delta_n$ .

We start with the  $\supseteq$  direction for the meet operator. Pick  $(h_1, \dots, h_n)$  from  $\bigcap_{i \in I} \Delta_n(p_i)$ . Then,

$$\forall i \in I. (h_1, \dots, h_n) \in \Delta_n(p_i).$$

By the definition of  $\Delta_n$ , this means that

$$\forall i \in I. \exists f_i \in p_i. f_i \sqsubseteq h_1 \wedge \dots \wedge f_i \sqsubseteq h_n. \quad (2)$$

Let  $f = \sum_{i \in I} f_i$ . The sum here is well-defined, because (a) there are only finitely many  $f$ 's such that  $f \sqsubseteq h_k$  for all  $1 \leq k \leq n$ , and (b) any two such  $f$  and  $g$  should have the same value for every location in  $\text{dom}(f) \cap \text{dom}(g)$ . Since all  $f_i$ 's satisfy (2), their sum  $f$  also satisfies

$$f \sqsubseteq h_1 \wedge \dots \wedge f \sqsubseteq h_n.$$

Furthermore,  $f \in \bigcap_{i \in I} p_i$ , because  $p_i$ 's are upward closed and  $f$  is an extension of  $f_i$  in  $p_i$ . Hence,  $\Delta_n(\bigcap_{i \in I} p_i) \subseteq \bigcap_{i \in I} \Delta_n(p_i)$ .

Next we prove the  $\supseteq$  direction for the join operator. Pick  $(h_1, \dots, h_n)$  from  $\Delta_n(\bigcup_{i \in I} p_i)$ . Then,

$$\exists i \in I. \exists f \in p_i. f \sqsubseteq h_1 \wedge \dots \wedge f \sqsubseteq h_n.$$

Hence, by the definition of  $\Delta_n$ ,

$$(h_1, \dots, h_n) \in \Delta_n(p_i) \subseteq \bigcup_{i \in I} \Delta_n(p_i),$$

as desired.

Finally, it remains to show that  $\Delta_n$  preserves the  $*$  operator. Consider predicates  $p, q \in \text{IRel}_1$ . We need to prove that

$$\Delta_n(p * q) = \Delta_n(p) * \Delta_n(q).$$

Choose an arbitrary  $(h_1, \dots, h_n)$  from  $\Delta_n(p * q)$ . By the definition of  $\Delta_n(p * q)$ , it follows that

$$\begin{aligned} \exists f \in p. \exists g \in q. (\text{dom}(f) \cap \text{dom}(g) = \emptyset) \wedge \\ f \sqsubseteq h_1 \wedge \dots \wedge f \sqsubseteq h_n \wedge g \sqsubseteq h_1 \wedge \dots \wedge g \sqsubseteq h_n. \end{aligned}$$

Now, define  $f_i = f$  and  $g_i = h_i - f$  for  $i \in \{1, \dots, n\}$ . Then,

$$\begin{aligned} (\forall i \in \{1, \dots, n\}. f_i \cdot g_i = h_i) \\ \wedge (f_1, \dots, f_n) \in \Delta_n(p) \wedge (g_1, \dots, g_n) \in \Delta_n(q). \end{aligned}$$

Hence,  $(h_1, \dots, h_n) \in \Delta_n(p) * \Delta_n(q)$ . This shows that  $\Delta_n(p * q) \subseteq \Delta_n(p) * \Delta_n(q)$ . For the other inclusion, suppose that

$$(h_1, \dots, h_n) \in \Delta_n(p) * \Delta_n(q).$$

Then, by the definition of  $*$ ,

$$\begin{aligned} \exists (f_1, \dots, f_n) \in \Delta_n(p). \exists (g_1, \dots, g_n) \in \Delta_n(q). \\ (\forall i \in \{1, \dots, n\}. f_i \cdot g_i = h_i). \end{aligned}$$

Since  $(f_1, \dots, f_n) \in \Delta_n(p)$  and  $(g_1, \dots, g_n) \in \Delta_n(q)$ , there are  $f \in p$  and  $g \in q$  such that

$$f \sqsubseteq f_1 \wedge \dots \wedge f \sqsubseteq f_n \wedge g \sqsubseteq g_1 \wedge \dots \wedge g \sqsubseteq g_n.$$

Furthermore, since  $f_1$  and  $g_1$  have disjoint domains, their subheaps  $f$  and  $g$  must have disjoint domains as well. Consequently,  $f \cdot g$  is well defined, and it satisfies

$$f \cdot g \in p * q \wedge (\forall i \in \{1, \dots, n\}. f \cdot g \sqsubseteq f_i \cdot g_i \sqsubseteq h_i).$$

This implies that  $(h_1, \dots, h_n) \in \Delta_n(p * q)$ , as desired. ■

**Lemma 3** *For all  $\varphi$  and  $\eta, \rho, \rho'$ , if  $\Delta_n(\rho(a)) = \rho'(a)$  for every  $a \in \text{AVar}$ , we have that  $\Delta_n(\llbracket \varphi \rrbracket_{\eta, \rho}^1) = \llbracket \varphi \rrbracket_{\eta, \rho'}^n$ .*

**Proof:** We prove by induction on the structure of  $\varphi$ . All the inductive cases and the cases of true and false follow from the preservation result of Lemma 2. Thus, it is sufficient to show the lemma when  $\varphi \equiv a$  or  $\varphi \equiv P$ . When  $\varphi \equiv a$ , the assumption of the lemma implies that

$$\Delta_n(\llbracket a \rrbracket_{\eta, \rho}^1) = \Delta_n(\rho(a)) = \rho'(a) = \llbracket a \rrbracket_{\eta, \rho'}^n.$$

When  $\varphi \equiv P$ , we note that  $\Delta_n \circ \Delta_1 = \Delta_n$ , and conclude that

$$\Delta_n(\llbracket P \rrbracket_{\eta, \rho}^1) = \Delta_n(\Delta_1(\llbracket P \rrbracket_{\eta})) = \Delta_n(\llbracket P \rrbracket_{\eta}) = \llbracket P \rrbracket_{\eta, \rho'}^n.$$

■

## B Proof of Lemma 4

**Lemma 4** *There is an algorithm taking simple assertions  $\varphi, \psi$  and returning finitely many implications  $\{\varphi^l \implies \psi^l\}_{l \in L}$ , such that (a)  $\varphi^l \implies \psi^l$  has the form (1) and (b) for any  $n \in \{1, 2\}$ , we have that  $\varphi \models^n \psi$  holds iff  $\varphi^l \models^n \psi^l$  holds for all  $l \in L$ .*

**Proof:** The algorithm first transforms  $\psi$  in the conjunctive normal form, using proof rules in classical logic, which hold in all the  $n$ -ary semantics. This gives an implication of the form:

$$\bigvee_{i=1}^I \bigwedge_{j=1}^J \varphi_{(i,j)} * \mathbf{a}_{(i,j)} \implies \bigwedge_{k=1}^K \bigvee_{l=1}^L \psi_{(k,l)} * \mathbf{b}_{(k,l)}.$$

Then, the algorithm constructs the below set:

$$\left\{ \bigwedge_{j=1}^J \varphi_{(i,j)} * \mathbf{a}_{(i,j)} \implies \bigvee_{l=1}^L \psi_{(k,l)} * \mathbf{b}_{(k,l)} \right\}_{1 \leq i \leq I, 1 \leq k \leq K}$$

Finally, it removes, in each implication, all the disjuncts that include assertion variables not appearing on the LHS of the implication. The outcome of this removal becomes the result of the algorithm. ■

## C Layouts that Lift

**Lemma 18 (Segregation)** *For any  $I, J \geq 1$  there are non-empty, finite segregating subsets  $\mathbf{S}_{i,j}^{I,J} \subseteq \text{PosInt}$  for all  $1 \leq i \leq I$  and  $1 \leq j \leq J$  with these properties:*

1.  $\forall 1 \leq i_1, i_2 \leq I. \bigcup_{1 \leq j \leq J} \mathbf{S}_{i_1,j}^{I,J} = \bigcup_{1 \leq j \leq J} \mathbf{S}_{i_2,j}^{I,J}$ .
2.  $\forall 1 \leq i \leq I. \forall 1 \leq j_1 \neq j_2 \leq J. \mathbf{S}_{i,j_1}^{I,J} \cap \mathbf{S}_{i,j_2}^{I,J} = \emptyset$ .
3.  $\forall 1 \leq i_1 \neq i_2 \leq I. \forall 1 \leq j_1, j_2 \leq J. \mathbf{S}_{i_1,j_1}^{I,J} \cap \mathbf{S}_{i_2,j_2}^{I,J} \neq \emptyset$ .

By 1 we define  $\mathbf{S}^{I,J} = \bigcup_{1 \leq j \leq J} \mathbf{S}_{i,j}^{I,J}$  for any  $1 \leq i \leq I$ .

**Theorem 8 (Shadow-Lift)** *Unary  $\eta$ -validity of an implication implies the PC if each dashed line has a label that is not a label on a solid line and each disjunct has an occurrence of a variable that is not a label on a solid line. Spelling it out in symbols, we require, with  $L = \{(i, j) \mid 1 \leq i \leq M \wedge 1 \leq j \leq N\}$ , that*

$$\begin{aligned} \forall (i, j) \in L. \Pi(i) \not\geq \Omega(j) &\implies \\ \exists c \in V. \Pi(i)(c) < \Omega(j)(c) \wedge & \\ (\forall (k, l) \in L. \Pi(k) \geq \Omega(l) \implies \Pi(k)(c) = \Omega(l)(c)) & \end{aligned}$$

and

$$\begin{aligned} \forall 1 \leq j \leq N. \exists c \in V. \Omega(j)(c) > 0 \wedge & \\ (\forall (k, l) \in L. \Pi(k) \geq \Omega(l) \implies \Pi(k)(c) = \Omega(l)(c)). & \end{aligned}$$

**Proof:** Assume that we have an implication of the form (1) in Section 4 and an appropriate environment  $\eta$ , that the stated criterion on the variable layout holds and that we have unary  $\eta$ -validity. We must show that the PC holds.

According to Definition 6 we assume that we have heaps  $h, h_1, \dots, h_M \in \text{Heap}$  with  $h_i \sqsubseteq h$  and  $h_i \in \llbracket \varphi_i \rrbracket_\eta^1$  for all  $1 \leq i \leq M$ . The core of the proof is the construction of a particular environment  $\rho : V \rightarrow \text{IRel}_1$ . For that purpose we need some notation. For a subset  $M \subseteq \text{PosInt}$  we denote by  $[M]$  the heap that has domain  $M$  and stores some fixed value, say 0, at all these locations. Let  $C \subseteq V$  be the set of assertion variables that do not occur as labels on solid edges, i.e., for a  $c \in V$  we have that  $c \in C$  iff

$$\begin{aligned} \forall 1 \leq i \leq M. \forall 1 \leq j \leq N. \\ \Pi(i) \geq \Omega(j) \implies \Pi(i)(c) = \Omega(j)(c). \end{aligned}$$

For each  $1 \leq i \leq M$  we let  $K_i$  be the set of second indices of all variables in conjunct  $i$  that lie in  $C$ , i.e., we set  $K_i = \{1 \leq k \leq M_i \mid a_{i,k} \in C\}$ . If non-empty, we let  $k_i = \min(K_i)$ .

We now define  $\rho(c) = \text{Heap}$  for  $c \in V \setminus C$ . For a variable  $c \in C$  we let  $\rho(c)$  be the union of

$$\bigcup_{\substack{1 \leq i \leq M, \\ K_i \neq \emptyset, \\ a_{i,k_i} \equiv c}} (h - h_i) \cdot [\mathbf{S}_{i,k_i}^{M,K} + L] \cdot \prod_{\substack{1 \leq k \leq K, \\ k \notin K_i}} [\mathbf{S}_{i,k}^{M,K} + L]$$

and

$$\bigcup_{1 \leq i \leq M, K_i \neq \emptyset, k \in K_i \setminus \{k_i\}, a_{i,k} \equiv c} [\mathbf{S}_{i,k}^{M,K} + L],$$

where we have used  $K = \max\{M_1, \dots, M_M\}$  and  $L = \max(\text{dom}(h) \cup \{0\})$ . For each  $1 \leq i \leq M$  we can write  $h \cdot [\mathbf{S}^{M,K} + L]$  as the following product

$$h_i \cdot (h - h_i) \cdot \prod_{k \in K_i} [\mathbf{S}_{i,k}^{M,K} + L] \cdot \prod_{1 \leq k \leq K, k \notin K_i} [\mathbf{S}_{i,k}^{M,K} + L],$$

which implies that we have  $h \cdot [\mathbf{S}^{M,K} + L]$  a member of  $\llbracket \varphi_i * a_{i,1} * \dots * a_{i,M_i} \rrbracket_{\eta, \rho}^1$ . In summary, we have shown that  $h \cdot [\mathbf{S}^{M,K} + L]$  lies in the unary interpretation of the left hand side in the environments  $\eta$  and  $\rho$ . By assumption, the same must hold for the right hand side and from this we aim to derive the PC.

We now know that  $h \cdot [\mathbf{S}^{M,K} + L]$  lies in the interpretation of some disjunct, say disjunct  $j$ . This means that

$$\begin{aligned} h \cdot [\mathbf{S}^{M,K} + L] \in \llbracket \psi_j * b_{j,1} * \dots * b_{j,N_j} \rrbracket_{\eta, \rho}^1 \\ = \llbracket \psi_j \rrbracket_\eta^1 * \prod_{k \in J} \rho(b_{j,k}), \end{aligned}$$

where  $J = \{1 \leq k \leq N_j \mid b_{j,k} \in C\}$  is the set of second indices of variables of disjunct  $j$  that are in  $C$ . By the

second assumption of the theorem we know that  $J \neq \emptyset$ . We write

$$h \cdot [\mathbf{S}^{M,K} + L] = g \cdot \prod_{k \in J} g_k$$

for  $g \in \llbracket \psi_j \rrbracket_\eta^1$  and  $g_k \in \rho(b_{j,k})$  for each  $k \in J$ . By the properties of segregating sets we get that there must be a common  $1 \leq i \leq M$  such that for all  $k \in J$  there is  $l_k \in K_i$  with

$$[\mathbf{S}_{i,l_k}^{M,K} + L] \sqsubseteq g_k,$$

i.e., the  $g_k$ 's are all 'from the same conjunct'. But this implies  $\Pi(i)(c) \geq \Omega(j)(c)$  for all  $c \in C$  as the segregating sets are non-empty. But then  $\Pi(i)(c) \geq \Omega(j)(c)$  must hold for  $c \in V \setminus C$  too by the first assumption of the lemma and so  $\Pi(i) \geq \Omega(j)$ . Also we must have  $\Pi(i)(c) = \Omega(j)(c)$  for each  $c \in C$  by definition of  $C$ . By construction we have

$$\text{dom} \left( \prod_{k \in J} g_k \right) \supseteq \text{dom}(h - h_i) \cup (\mathbf{S}^{M,K} + L)$$

But then  $\text{dom}(g) \subseteq h_i$  and so we have  $h_i \in \llbracket \psi_j \rrbracket_\eta^1$  too and we have proved the first option of the PC. ■

**Theorem 10 (Balloon-Lift)** *Unary  $\eta$ -validity of an implication implies the PC if there is a subset  $B \subseteq V$  with the following three properties. First, each conjunct has at most one occurrence of a variable from  $B$ , i.e.,*

$$\forall 1 \leq i \leq M. \sum_{c \in B} \Pi(i)(c) \leq 1.$$

*Second, each disjunct is empty or has exactly one occurrence of a variable from  $B$ , i.e.,*

$$\forall 1 \leq j \leq N. \sum_{c \in V} \Omega(j)(c) = 0 \vee \sum_{c \in B} \Omega(j)(c) = 1.$$

*Third, each dashed line must have a label from  $B$ . That is, when  $L = \{(i, j) \mid 1 \leq i \leq M \wedge 1 \leq j \leq N\}$ ,*

$$\forall (i, j) \in L. \Pi(i) \not\geq \Omega(j) \implies \exists c \in B. \Pi(i)(c) < \Omega(j)(c).$$

**Proof:** Assume that we have an implication of the form (1) in Section 4 and an appropriate environment  $\eta$ , that the stated criterion on the variable layout holds and that we have unary  $\eta$ -validity. We must show that the PC holds.

According to Definition 6 we assume that we have heaps  $h, h_1, \dots, h_M \in \mathbf{Heap}$  with  $h_i \sqsubseteq h$  and  $h_i \in \llbracket \varphi_i \rrbracket_\eta^1$  for all  $1 \leq i \leq M$ . The core of the proof is the construction of a particular environment  $\rho : V \rightarrow \mathbf{IRel}_1$ .

We define  $\rho(c) = \mathbf{Heap}$  for  $c \in V \setminus B$ . For a variable  $c \in B$  we let  $\rho(c)$  be the following union

$$\bigcup_{1 \leq i \leq M, 1 \leq k \leq M_i, a_{i,k} \equiv c} h - h_i.$$

For each  $1 \leq i \leq M$  we can write  $h = h_i \cdot (h - h_i)$  and so we have  $h$  in  $\llbracket \varphi_i * a_{i,1} * \dots * a_{i,M_i} \rrbracket_{\eta, \rho}^1$  by the first of the original assumption on the set  $B$ . In summary, we have shown that  $h$  lies in the unary interpretation of the left hand side in the environments  $\eta$  and  $\rho$ . By assumption, the same must hold for the right hand side and from this we aim to derive the PC.

We now know that  $h$  lies in the interpretation of some disjunct, say disjunct  $j$ . If this disjunct is empty we have proved the second option of the PC. Otherwise we know that there is exactly one  $1 \leq k \leq N_j$  such that  $b_{j,k} \in B$ . But then we have

$$h \in \llbracket \psi_j * b_{j,1} * \dots * b_{j,N_j} \rrbracket_{\eta, \rho}^1 = \llbracket \psi_j \rrbracket_\eta^1 * \rho(b_{j,k}).$$

We write

$$h = g \cdot g_k$$

for  $g \in \llbracket \psi_j \rrbracket_\eta^1$  and  $g_k \in \rho(b_{j,k})$ . There must be an  $1 \leq i \leq M$  such that  $g_k \supseteq h - h_i$  and such that  $\Pi(i)(b_{j,k}) = \Omega(j)(b_{j,k}) = 1$ . The first gives  $h_i \in \llbracket \psi_j \rrbracket_\eta^1$  and the second implies  $\Pi(i) \geq \Omega(j)$  by the third assumption on  $B$ . And we have arrived at the first option of the PC. ■

## D Completeness

**Lemma 19 (Fan-Counter)** *Suppose that the layout of variables is as follows. There are at least two conjuncts, i.e.,  $M \geq 2$ , and one conjunct has the property that each variable occurring in the conjunct also occurs as a label of a solid line leaving the conjunct and ending in a non-empty disjunct. In symbols the latter is*

$$\begin{aligned} \exists 1 \leq i \leq M. \forall c \in V. \Pi(i)(c) > 0 \implies \\ \exists 1 \leq j \leq N. \Pi(i) \geq \Omega(j) \wedge \Pi(i)(c) > \Omega(j)(c) \wedge \\ \exists d \in V. \Omega(j)(d) > 0. \end{aligned}$$

*Then there are choices of  $\varphi_i$ 's and  $\psi_j$ 's with no variables such that the implication holds on the unary level but not on the binary level.*

In the search for counterexamples we may without loss of generality assume the negation of the conditions of the above lemma. This means, provided at least two conjuncts, that for any non-empty set of solid lines leaving one common conjunct and ending in non-empty disjuncts there is a variable that occurs in the conjunct

but is not a label of either of the lines. If, loosely phrased, we invalidate that variable, then all the solid lines break down, i.e., become dashed.

**Lemma 20 (X-Counter)** *Suppose that the layout of variables is as follows. There are two distinct conjuncts  $i_0$  and  $i_1$  and two distinct non-empty disjuncts  $j_0$  and  $j_1$  such that  $\Pi(i_0) \not\geq \Omega(j_0)$  while  $\Pi(i_0) \geq \Omega(j_1)$ ,  $\Pi(i_1) \geq \Omega(j_0)$  and  $\Pi(i_1) \geq \Omega(j_1)$ . Then there are choices of  $\varphi_i$ 's and  $\psi_j$ 's that the implication holds on the unary level but not on the binary level.*

Again we may without loss of generality assume that the negation of this lemma holds when building counterexamples. Picture the graph of the implication without empty disjuncts and without dashed lines. The negation of the above means that we may arrive at all vertices in the connected component containing some vertex by paths from that vertex of length 2 or less. Also all connected components are complete, in particular no two vertices with a dashed line between them can belong to the same component.

**Lemma 21 (Bridge-Counter)** *Suppose that the layout of variables is as follows. There are at least two conjuncts, i.e.,  $M \geq 2$ , all disjuncts are non-empty and there is a dashed line with labels that all occur as labels on solid lines too. In symbols the last demand is*

$$\begin{aligned} \exists 1 \leq i \leq M. \exists 1 \leq j \leq N. \Pi(i) \not\geq \Omega(j) \wedge \\ \forall c \in V. \Pi(i)(c) < \Omega(j)(c) \implies \\ \exists 1 \leq k \leq M. \exists 1 \leq l \leq N. \\ \Pi(k) \geq \Omega(l) \wedge \Pi(k)(c) > \Omega(l)(c). \end{aligned}$$

*Then there are choices of  $\varphi_i$ 's and  $\psi_j$ 's with no variables such that the implication holds on the unary level but not on the binary level.*

This lemma deals with the case of a variable layout with at least two conjuncts and no empty disjuncts but where the first condition of Theorem 8 fails.

**Lemma 22 (All-Out-Counter)** *Suppose that the layout of variables is as follows. There are at least two conjuncts, i.e.,  $M \geq 2$ , at least one non-empty disjunct and for each variable one of the following two holds: Either the variable occurs as a label on a solid line ending in a non-empty disjunct. Or it occurs at least twice in a conjunct and we have an empty disjunct. In symbols the variable condition is*

$$\begin{aligned} \forall c \in V. (\exists 1 \leq i \leq M. \exists 1 \leq j \leq N. \Pi(i) \geq \Omega(j) \wedge \\ \Pi(i)(c) > \Omega(j)(c) \wedge \exists d \in V. \Omega(j)(d) > 0) \vee \\ (\exists 1 \leq i \leq M. \Pi(i)(c) \geq 2 \wedge \\ \exists 1 \leq j \leq N. \forall d \in V. \Omega(j)(d) = 0). \end{aligned}$$

*Then there are choices of  $\varphi_i$ 's and  $\psi_j$ 's with no variables such that the implication holds on the unary level but not on the binary level.*

This lemma deals with two cases. The first is the case of a variable layout with at least two conjuncts and no empty disjuncts but where the second condition of Theorem 8 fails while the first holds. The second is the case of a variable layout with at least two conjuncts, at least one empty disjunct and no dashed lines for which Theorem 10 fails.

## E Higher Arities and Parametricity

**Theorem 15** *For an implication of the form (1) and an appropriate environment  $\eta$  we have that  $n$ -ary  $\eta$ -validity implies the PC if  $n \geq \max\{M_1, \dots, M_M\}$ .*

**Proof:** Assume that we have an implication of the form (1) in Section 4 and an appropriate environment  $\eta$ , that  $n \geq \max\{M_1, \dots, M_M\}$  and that we have  $n$ -ary  $\eta$ -validity. We must show that the PC holds. We consider only the case  $n \geq 2$ , the case  $n = 1$  proceeds along the lines of the proof of Theorem 10.

According to Definition 6 we assume that we have heaps  $h, h_1, \dots, h_M \in \mathbf{Heap}$  with  $h_i \sqsubseteq h$  and  $h_i \in \llbracket \varphi_i \rrbracket_\eta^1$  for all  $1 \leq i \leq M$ . The core of the proof is the construction of a particular environment  $\rho : V \rightarrow \mathbf{IRel}_n$ . For that purpose we need some notation. Define, for each  $1 \leq k \leq n$ , a map  $\gamma_k : \mathbf{Heap} \rightarrow \mathbf{Heap}^n$  by letting

$$\gamma_k(h) = (\overbrace{\square, \dots, \square}^{k-1}, h, \overbrace{\square, \dots, \square}^{n-k})$$

for any  $h \in \mathbf{Heap}$ , i.e., it returns the  $n$ -tuple that has  $h$  as the  $k$ -th entry and the empty heap everywhere else. Similarly, we define  $\delta : \mathbf{Heap} \rightarrow \mathbf{Heap}^n$  by setting

$$\delta(h) = (\overbrace{h, \dots, h}^n)$$

for any  $h \in \mathbf{Heap}$ , i.e., it returns the  $n$ -tuple that has  $h$  as all entries. For a subset  $M \subseteq \mathbf{PosInt}$  we denote by  $[M]$  the heap that has domain  $M$  and stores some fixed value, say 0, at all these locations.

For a variable  $c \in V$  we now define  $\rho(c)$  to be the following union of relations in  $\mathbf{IRel}_n$ :

$$\bigcup_{1 \leq i \leq M, 1 \leq k \leq M_i, a_{(i,k)} \equiv c} \{\gamma_k(h - h_i) \cdot \gamma_1([\mathbf{S}_{i,k}^{M,K} + L])\}^\uparrow$$

where  $K = \max\{M_1, \dots, M_M\}$ ,  $L = \max(\text{dom}(h))$ . This is well-defined because of our assumption that  $n \geq \max\{M_1, \dots, M_M\}$ . For each  $1 \leq i \leq M$  we have that

$$\begin{aligned} \delta(h) &= \delta(h_i) \cdot \gamma_1(h - h_i) \cdots \gamma_n(h - h_i) \\ &\sqsupseteq \delta(h_i) \cdot \gamma_1(h - h_i) \cdots \gamma_{M_i}(h - h_i) \end{aligned}$$

where we use the extension order for heap tuples defined by pointwise extension in all entries. Also, we have that

$$\begin{aligned} [\mathbf{S}^{M,K} + L] &= [\mathbf{S}_{i,1}^{M,K} + L] \cdots \cdots [\mathbf{S}_{i,K}^{M,K} + L] \\ &\sqsubseteq [\mathbf{S}_{i,1}^{M,K} + L] \cdots \cdots [\mathbf{S}_{i,M_i}^{M,K} + L]. \end{aligned}$$

This gives us that  $\delta(h) \cdot \gamma_1([\mathbf{S}^{M,K} + L])$  extends the following  $n$ -tuple of heaps:

$$\delta(h_i) \cdot \prod_{1 \leq k \leq M_i} \gamma_k(h - h_i) \cdot \gamma_1([\mathbf{S}_{i,k}^{M,K} + L])$$

which again means that  $\delta(h) \cdot \gamma_1([\mathbf{S}^{M,K} + L])$  lies in

$$\llbracket \varphi_i * a_{i,1} * \cdots * a_{i,M_i} \rrbracket_{\eta,\rho}^n.$$

In summary, we have shown that  $\delta(h) \cdot \gamma_1([\mathbf{S}^{M,K} + L])$  lies in the  $n$ -ary interpretation of the left hand side in the environments  $\eta$  and  $\rho$ . By assumption, the same must hold for the right hand side and from this we aim to derive the PC.

There is  $1 \leq j \leq N$  such that we have

$$\delta(h) \cdot \gamma_1([\mathbf{S}^{M,K} + L]) \in \llbracket \psi_j * b_{j,1} * \cdots * b_{j,N_j} \rrbracket_{\eta,\rho}^n.$$

Consider first the case of a non-empty disjunct, i.e., the case  $N_j > 0$ . We split along the disjunct and get

$$\delta(h) \cdot \gamma_1([\mathbf{S}^{M,K} + L]) = \delta(g) \cdot \mathbf{g}_1 \cdots \cdots \mathbf{g}_{N_j}$$

for  $g \in \llbracket \psi_j \rrbracket_{\eta}^1$  and  $\mathbf{g}_k \in \rho(b_{j,k})$  for all  $1 \leq k \leq N_j$ . By the properties of segregating sets we get that there must be a common  $1 \leq i \leq M$  such that for all  $1 \leq k \leq N_k$  there is  $1 \leq k_k \leq M_i$  with

$$\gamma_{k_k}(h - h_i) \cdot \gamma_1([\mathbf{S}_{i,k_k}^{M,K} + L]) \sqsubseteq \mathbf{g}_k,$$

i.e., the  $\mathbf{g}_k$ 's are all 'from the same conjunct'. But this implies  $\Pi(i) \geq \Omega(j)$  as the segregating sets are non-empty. Also the above equality enforces  $\text{dom}(g) \subseteq \text{dom}(h)$  by the definition of  $\gamma_1$ . Indeed we must have  $\text{dom}(g) \subseteq \text{dom}(h_i)$  since in particular we have

$$\gamma_{1_k}(h - h_i) \cdot \gamma_1([\mathbf{S}_{i,1_k}^{M,K} + L]) \sqsubseteq \mathbf{g}_1.$$

But then  $g \sqsubseteq h_i$  so we have  $h_i \in \llbracket \psi_j \rrbracket_{\eta}^1$  too and the first option of the PC holds.

We consider now the case of an empty disjunct, i.e., the case  $N_j = 0$ . As above we split along the disjunct and get

$$\delta(h) \cdot \gamma_1([\mathbf{S}^{M,K} + L]) = \delta(g) \cdot \mathbf{g}$$

for  $g \in \llbracket \psi_j \rrbracket_{\eta}^1$  and  $\mathbf{g} \in \text{Heap}^n$ . Again we must have  $\text{dom}(g) \subseteq \text{dom}(h)$  which implies  $g \sqsubseteq h$  and the second option of the PC holds. ■

## F Proof of Theorem 17

**Theorem 17** *Every derivable  $\Gamma \vdash \{\varphi\}C\{\psi\}$  is 2-valid.*

**Proof:** We will show that all the rules in Figure 4 are sound. This lets us prove the theorem by induction on the height of the derivation of a judgment, because using the soundness of the rules, we can handle all the base and inductive cases.

Let's start with the rule for the module operation  $k$ . Suppose that  $(\eta, \rho, \mathbf{u}) \models^2 (\Gamma, \{\varphi\}k\{\psi\})$ . Then, by the definition of  $\models^2$ , we should have  $(\eta, \rho, \mathbf{u}) \models^2 \{\varphi\}k\{\psi\}$  as well. From this follows the soundness of the rule.

Next, consider four rules: (a) the frame rule for adding  $- * \varphi$  to the pre and post-conditions, (b) the rule for adding  $\exists x$  to the pre and post-conditions, (c) the rule for sequencing, and (d) the rule for the conditional statement. All these rules are sound, because of the following four facts:

$$\begin{aligned} (\eta, \rho, \mathbf{u}) \models^2 \{\varphi\}C\{\varphi'\} \\ \implies (\eta, \rho, \mathbf{u}) \models^2 \{\varphi * \psi\}C\{\varphi' * \psi\} \end{aligned}$$

$$\begin{aligned} (x \notin \text{FV}(C)) \wedge (\eta, \rho, \mathbf{u}) \models^2 \{\varphi\}C\{\psi\} \\ \implies (\eta, \rho, \mathbf{u}) \models^2 \{\exists x.\varphi\}C\{\exists x.\psi\} \end{aligned}$$

$$\begin{aligned} (\eta, \rho, \mathbf{u}) \models^2 \{\varphi\}C\{\varphi'\} \wedge (\eta, \rho, \mathbf{u}) \models^2 \{\varphi'\}C'\{\psi\} \\ \implies (\eta, \rho, \mathbf{u}) \models^2 \{\varphi\}C; C'\{\psi\} \end{aligned}$$

$$\begin{aligned} (\eta, \rho, \mathbf{u}) \models^2 \{\varphi \wedge B\}C\{\psi\} \wedge \\ (\eta, \rho, \mathbf{u}) \models^2 \{\varphi \wedge \neg B\}C'\{\psi\} \\ \implies (\eta, \rho, \mathbf{u}) \models^2 \{\varphi\}\mathbf{if} B C C'\{\psi\} \end{aligned}$$

The first fact is an easy consequence of using the quantification over  $\text{IRel}_2$  in the semantics of triples. The second also follows easily from the semantics of triples and the fact that  $\llbracket C \rrbracket_{\eta, u_i} = \llbracket C \rrbracket_{\eta[x \mapsto v], u_i}$  for all  $v \in \text{Int}$ , as long as one remembers that the  $*$  operator distributes over union. The third and fourth are not different, and they follow from the semantics of triples and commands. Here we will go through the details of proving the fourth fact. Consider  $(\eta, \rho, \mathbf{u})$  satisfying the assumption of the fact. Pick  $r \in \text{IRel}_2$  and heaps  $f, g$  such that

$$(f, g) \in \llbracket \varphi \rrbracket_{\eta, \rho}^2 * r.$$

Now we do the case analysis on whether  $\llbracket B \rrbracket_{\eta}$  is true or not. If it is true, then

$$\begin{aligned} (f, g) \in \llbracket \varphi \wedge B \rrbracket_{\eta, \rho}^2 * r \\ \wedge \llbracket \mathbf{if} B C C' \rrbracket_{\eta, u_1}(f) = \llbracket C \rrbracket_{\eta, u_1}(f) \\ \wedge \llbracket \mathbf{if} B C C' \rrbracket_{\eta, u_2}(g) = \llbracket C \rrbracket_{\eta, u_2}(g). \end{aligned}$$

Hence, by assumption, we get that

$$\begin{aligned} & (\llbracket \mathbf{if} \ B \ C \ C' \rrbracket_{\eta, u_1}(f), \llbracket \mathbf{if} \ B \ C \ C' \rrbracket_{\eta, u_2}(g)) \\ &= (\llbracket C \rrbracket_{\eta, u_1}(f), \llbracket C \rrbracket_{\eta, u_2}(g)) \in \llbracket \psi \rrbracket_{\eta, \rho}^2 * r. \end{aligned}$$

If  $\llbracket B \rrbracket_{\eta}$  is not true, we reason similarly, but with  $C'$  instead of  $C$ , and get that

$$(\llbracket \mathbf{if} \ B \ C \ C' \rrbracket_{\eta, u_1}(f), \llbracket \mathbf{if} \ B \ C \ C' \rrbracket_{\eta, u_2}(g)) \in \llbracket \psi \rrbracket_{\eta, \rho}^2 * r.$$

We have just shown that in both cases, the outcomes of the conditional statements are related by  $\llbracket \psi \rrbracket_{\eta, \rho}^2 * r$ , as claimed by the fourth fact.

We move on to the rules for heap update and dereference. They are sound because of the below two facts:

$$(\eta, \rho, \mathbf{u}) \models^2 \{E \leftrightarrow \_ \} [E] := F \{E \leftrightarrow F\}$$

$$\begin{aligned} x \notin \text{FV}(\psi) \wedge (\eta, \rho, \mathbf{u}) \models^2 \{\varphi * E \leftrightarrow x\} \mathbf{let} \ x = [E] \ \mathbf{in} \ C \{\psi\} \\ \implies (\eta, \rho, \mathbf{u}) \models^2 \{\exists x. \varphi * E \leftrightarrow x\} \mathbf{let} \ x = [E] \ \mathbf{in} \ C \{\psi\} \end{aligned}$$

To prove the first, we pick  $(\eta, \rho, \mathbf{u})$ , a relation  $r \in \text{IRel}_2$  and heaps  $f, g$  such that

$$(f, g) \in \llbracket E \leftrightarrow \_ \rrbracket_{\eta, \rho}^2 * r.$$

Then, there exist heaps  $h, f_1, g_1$  such that

$$(f, g) = (h, h) \cdot (f_1, g_1) \wedge \llbracket E \rrbracket_{\eta} \in \text{dom}(h) \wedge (f_1, g_1) \in r.$$

Thus,

$$\begin{aligned} & (\llbracket [E] := F \rrbracket_{\eta, u_1}(f), \llbracket [E] := F \rrbracket_{\eta, u_2}(g)) \\ &= (h \llbracket [E]_{\eta} \mapsto [F]_{\eta} \rrbracket \cdot f_1, h \llbracket [E]_{\eta} \mapsto [F]_{\eta} \rrbracket \cdot g_1) \\ &\in \llbracket E \leftrightarrow F \rrbracket_{\eta, \rho}^2 * r, \end{aligned}$$

as desired by the first fact. For the proof of the second fact, suppose that the assumption of the second fact holds, and pick  $r \in \text{IRel}_2$  and heaps  $f, g$  such that

$$(f, g) \in \llbracket \exists x. \varphi * E \leftrightarrow x \rrbracket_{\eta, \rho}^2 * r.$$

Then, there exists an integer  $v$  and heaps  $h, f_1, g_1$  such that

$$\begin{aligned} (f, g) &= (h, h) \cdot (f_1, g_1) \wedge (f_1, g_1) \in r \\ &\wedge h \in \llbracket \varphi * E \leftrightarrow x \rrbracket_{\eta[x \mapsto v], \rho}^1. \end{aligned}$$

Thus,  $(f, g) \in \llbracket \varphi * E \leftrightarrow x \rrbracket_{\eta[x \mapsto v], \rho}^1 * r$ , and  $f(\llbracket E \rrbracket_{\eta}) = g(\llbracket E \rrbracket_{\eta}) = v$ . Using these and the assumed triple of the second fact, we derive the below:

$$\begin{aligned} & (\llbracket \mathbf{let} \ x = [E] \ \mathbf{in} \ C \rrbracket_{\eta, u_1}(f), \llbracket \mathbf{let} \ x = [E] \ \mathbf{in} \ C \rrbracket_{\eta, u_2}(g)) \\ &= (\llbracket C \rrbracket_{\eta[x \mapsto v], u_1}(f), \llbracket C \rrbracket_{\eta[x \mapsto v], u_2}(g)) \\ &\in \llbracket \psi \rrbracket_{\eta[x \mapsto v], \rho}^2 * r \\ &= \llbracket \psi \rrbracket_{\eta, \rho}^2 * r. \end{aligned}$$

The last equality holds, because  $x$  does not appear in  $\varphi$ . We have just proved that the output states of two dereferencing commands are  $(\llbracket \psi \rrbracket_{\eta, \rho}^2 * r)$ -related, as claimed by the second fact.

Finally, we prove that the rule of consequence is sound. It is sufficient to show that

$$\begin{aligned} & \text{CHK}(\varphi', \varphi) \wedge \varphi' \models^1 \varphi \wedge \\ & \text{CHK}(\psi, \psi') \wedge \psi \models^1 \psi' \wedge (\eta, \rho, \mathbf{u}) \models^2 \{\varphi\} C \{\psi\} \\ & \implies (\eta, \rho, \mathbf{u}) \models^2 \{\varphi'\} C \{\psi'\}. \end{aligned}$$

From the first four conjuncts of the assumption, it follows that

$$\varphi' \models^2 \varphi \wedge \psi \models^2 \psi'.$$

This is due to the correctness of CHK, which holds because all the transformations used in the first check of CHK are based on semantic equivalences holding in  $\text{IRel}_2$  and the second lifting check is sound because of our lifting theorems. In order to prove the conclusion of the above implication, pick  $r \in \text{IRel}_2$  and heaps  $f, g$  such that

$$(f, g) \in \llbracket \varphi' \rrbracket_{\eta, \rho}^2 * r.$$

Since the  $*$  operator is monotone and  $\varphi' \models^2 \varphi$ , we get that

$$(f, g) \in \llbracket \varphi \rrbracket_{\eta, \rho}^2 * r.$$

This relationship and the assumed triple  $\{\varphi\} C \{\psi\}$ , then, imply the below:

$$(\llbracket C \rrbracket_{\eta, u_1}(f), \llbracket C \rrbracket_{\eta, u_2}(g)) \in \llbracket \psi \rrbracket_{\eta, \rho}^2 * r.$$

Again, since  $\psi \models^2 \psi'$ , the monotonicity of the  $*$  operator implies that

$$(\llbracket C \rrbracket_{\eta, u_1}(f), \llbracket C \rrbracket_{\eta, u_2}(g)) \in \llbracket \psi' \rrbracket_{\eta, \rho}^2 * r.$$

Note that this is the conclusion that we are looking for.  $\blacksquare$