GALOIS CONNECTING CALL-BY-VALUE AND CALL-BY-NAME

DYLAN MCDERMOTT \textsuperscript{a} AND ALAN MYCROFT \textsuperscript{b}

\textsuperscript{a} Reykjavik University, Iceland  
\textit{e-mail address}: dylan@dylanm.org

\textsuperscript{b} University of Cambridge, UK  
\textit{e-mail address}: Alan.Mycroft@cl.cam.ac.uk

Abstract. We establish a general framework for reasoning about the relationship between call-by-value and call-by-name.

In languages with computational effects, call-by-value and call-by-name executions of programs often have different, but related, observable behaviours. For example, if a program might diverge but otherwise has no effects, then whenever it terminates under call-by-value, it terminates with the same result under call-by-name. We propose a technique for stating and proving properties like these. The key ingredient is Levy’s call-by-push-value calculus, which we use as a framework for reasoning about evaluation orders. We show that the call-by-value and call-by-name translations of expressions into call-by-push-value have related observable behaviour under certain conditions on computational effects, which we identify. We then use this fact to construct maps between the call-by-value and call-by-name interpretations of types, and identify further properties of effects that imply these maps form a Galois connection. These properties hold for some computational effects (such as divergence), but not others (such as mutable state). This gives rise to a general reasoning principle that relates call-by-value and call-by-name. We apply the reasoning principle to example computational effects including divergence and nondeterminism.

1. Introduction

Suppose that we have a language in which terms can be statically tagged either as using call-by-value evaluation or as using call-by-name evaluation. Each program in this language would therefore use a mix of call-by-value and call-by-name at runtime. Given any such program \( M \), we can construct a new program \( M' \) by changing call-by-value to call-by-name for some subterm. The question we consider in this paper is: what is the relationship between the observable behaviour of \( M \) and the observable behaviour of \( M' \)?

For a language with computational effects (such as divergence), changing the evaluation order in this way will in general change the behaviour of the program, but for some effects we can often say something about how we expect the behaviour to change:

- If there are no effects at all (in particular, programs are normalizing), the choice of evaluation order is irrelevant: \( M \) and \( M' \) terminate with the same result.

\textit{Key words and phrases}: computational effect, evaluation order, call-by-push-value, categorical semantics.  
\textsuperscript{*} This article is an extended version of [MM22].
• If there are diverging terms (for instance, via recursion), then the behaviour may change:
a program might diverge under call-by-value and return a result under call-by-name.
However, we can say something about how the behaviour changes: if $M$ terminates with
some result, then $M'$ terminates with the same result.
• If nondeterminism is the only effect, every result of $M$ is a possible result of $M'$.

These three instances of the problem are intuitively obvious, and each can be proved
separately. We develop a general technique for proving these properties.

The idea is to use a calculus that captures both call-by-value and call-by-name, as
a setting in which we can reason about both evaluation orders (this is where $M$ and $M'$
live). The calculus we use is Levy’s call-by-push-value (CBPV) [Lev99]. Levy describes
how to translate (possibly open) expressions $e$ into CBPV terms $V\langle e \rangle$ and $N\langle e \rangle$, which
respectively correspond to call-by-value and call-by-name. We study the relationship between
the behaviour of $V\langle e \rangle$ and the behaviour of $N\langle e \rangle$ in a given program context.

The main obstacle is that $V\langle e \rangle$ and $N\langle e \rangle$ have different types. The former has a “call-
by-value type” $F(V\langle \tau \rangle)$ and the latter a “call-by-name type” $N\langle \tau \rangle$, defined in Section 2.1.
They hence cannot be directly compared. Our solution is inspired by Reynolds’s work
relating direct and continuation semantics of the $\lambda$-calculus [Rey74].

The first step is to define a family of (set-theoretic) relations (in the style of a logical
relation) that compares the observable behaviour of a term of call-by-value type with
observable behaviour of a term of call-by-name type. We can then ask whether $V\langle e \rangle$ is
related in this sense to $N\langle e \rangle$. This is not the case in general. In the presence of arbitrary
computational effects, we cannot expect to say anything useful about how the behaviour of
$V\langle e \rangle$ relates to the behaviour of $N\langle e \rangle$. However, under certain conditions satisfied only for
certain effects, $V\langle e \rangle$ is related to $N\langle e \rangle$. These conditions say roughly that we can discard,
duplicate, and reorder effects. The main result of the first step is a theorem relating the two
translations of $e$ when these conditions hold (Theorem 4.7). This does not quite say what
happens if we were to replace call-by-value with call-by-name within some program; that is
the goal of the second step.

The second step is to identify maps between the call-by-value and call-by-name interpre-
tations, forming Galois connections (one for each source-language type) between the two
interpretations. We compose these maps with the translations of expressions, to arrive at
two terms that can be compared directly. For this step we assume a stronger condition on
computational effects than in the first, saying informally that effects can be thunked. Under
this condition we show that the maps between call-by-value and call-by-name represent the
relations from the first step. By combining this fact with Theorem 4.7 we prove a result that
directly relates the two terms we construct by composition with the Galois connections.

We therefore arrive at a general reasoning principle (Theorem 6.2) that we use to
compare call-by-value with call-by-name. Given any preorder $\preceq$ that captures the property
we wish to show about programs, our reasoning principle gives conditions that imply $M \preceq M'$,
where $M'$ is constructed as above by replacing call-by-value with call-by-name. We apply
our reasoning principle to examples by choosing different relations $\preceq$; each of these relations
indicates the extent to which changing evaluation order affects the behaviour of the program.
In the divergence example $N \preceq N'$ is defined to mean termination of $N$ implies termination
of $N'$ with the same result; in the other examples $\preceq$ similarly mirrors the properties described
informally above.
Rather than just considering some fixed collection of (allowable) effects, we work abstractly and identify properties of computational effects that enable us to relate call-by-value and call-by-name.

Our reasoning principle relies on the existence of some denotational model $M$ of the computational effects, and we reason primarily inside $M$. In the first step we in fact relate the behaviour of the call-by-value and call-by-name translations of terms within the given model $M$. On the other hand, in the second step we are able to prove a result (our reasoning principle Theorem 6.2) in which the conclusion is independent of $M$ (though Theorem 6.2 does assume the existence of a suitable $M$, since the proof relies on the first step).

Crucially, we use order-enriched models, which come with a partial order on the denotations of terms. The ordering on denotations is necessary to obtain a general reasoning principle. (Our example properties cannot be proved by showing that denotations are equal, because they are not symmetric.) Working inside the semantics rather than using syntactic logical relations makes it easier to prove and to use our reasoning principle, especially for the divergence example.

In Section 2 we summarize the call-by-push-value calculus (CBPV) and the call-by-value and call-by-name translations. We then make the following contributions:

- We describe an order-enriched categorical semantics for CBPV (Section 3).
- We define a family of relations for comparing the observable behaviours of a term of call-by-value type with a term of call-by-name type (Section 4). We prove that, for effects satisfying certain conditions, the call-by-value and call-by-name translations of expressions are related by these (Theorem 4.7). As a corollary, we directly relate the call-by-value and call-by-name translations of closed expressions of type $\text{bool}$ (Corollary 4.8).
- We define the Galois connections between the call-by-value and call-by-name translations (Section 5), and show that they represent the relations from the first step (Lemma 5.8).
- We use the Galois connections to prove a novel reasoning principle (Theorem 6.2) that relates the call-by-value and call-by-name translations of expressions (Section 6).

We apply our reasoning principle to three different examples: no effects, divergence, and nondeterminism. In this way we establish all of three facts listed at the beginning of this introduction. Our motivation is partly to demonstrate the Galois connection technique as a way of reasoning about different semantics of a given language. Call-by-value and call-by-name is one example of this (and Reynolds’s original application to direct and continuation semantics is another).

This paper is a revised and extended version of [MM22]. The primary difference is the addition of Section 4, containing the first step outlined above. The conference version [MM22] skips this step and goes directly to the Galois connections. The first step in particular enables us to prove a statement about closed terms of type $\text{bool}$ (Corollary 4.8) under weaker assumptions than in the conference version [MM22, Corollary 22]. We also add an extra example (immutable state), add products to the source language, and include more detailed proofs than in the conference version.
2. Call-by-push-value, call-by-value, and call-by-name

Levy [Lev99, Lev06] introduced call-by-push-value (CBPV) as a calculus that captures both call-by-value and call-by-name. We reason about the relationship between call-by-value and call-by-name evaluation inside CBPV.\(^1\)

The syntax of CBPV terms is stratified into two kinds: values \(V, W\) do not reduce, computations \(M, N\) might reduce (possibly with computational effects). The syntax of types is similarly stratified into value types \(A, B\) and computation types \(C, D\).

<table>
<thead>
<tr>
<th>value types</th>
<th>(A, B) ::= unit</th>
<th>(A_1 \times A_2)</th>
<th>bool</th>
<th>(U_C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>computation types</td>
<td>(C, D) ::= (C_1 \times C_2)</td>
<td>(A \rightarrow C)</td>
<td>(FA)</td>
<td></td>
</tr>
<tr>
<td>values</td>
<td>(V, W) ::= (x)</td>
<td>()</td>
<td>((V_1, V_2))</td>
<td>(\text{true})</td>
</tr>
<tr>
<td>computations</td>
<td>(M, N) ::= (\lambda{1. M_1, 2. M_2})</td>
<td>(1'M)</td>
<td>(2'M)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(\lambda x. A. M)</td>
<td>(V'M)</td>
<td>(\text{return}(V))</td>
<td>(M \text{ to } x. N)</td>
</tr>
<tr>
<td></td>
<td>(\text{match}\ V \text{ with } (x_1, x_2). M)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(\text{if } V \text{ then } M_1 \text{ else } M_2)</td>
<td>(\text{force}\ V)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We restrict to only the subset of CBPV required for this paper.

The value type \(U_C\) is the type of \textit{thunks} of computations of type \(C\). Elements of \(U_C\) are introduced using \textit{thunk}: the value \textit{thunk} \(M\) is the suspension of the computation term \(M\). The corresponding eliminator is \textit{force}, which behaves as the inverse of \textit{thunk}. Computation types include binary products; the pairing of two computations \(M_1\) and \(M_2\) is written \(\lambda\{1. M_1, 2. M_2\}\), and the first and second projections are \(1'M\) and \(2'M\). Computation types also include function types (where functions send values to computations). Function application is written \(V'M\), where \(V\) is the argument and \(M\) is the function to apply. The \textit{returner type} \(FA\) has as elements computations that return elements of the value type \(A\); these computations may have effects. Elements of \(FA\) are introduced by \textit{return}; the computation \textit{return} \(V\) immediately returns the value \(V\) (with no effects). Computations can be sequenced using \(M \text{ to } x. N\). This first evaluates \(M\) (which is required to have returner type), and then evaluates \(N\) with \(x\) bound to the result of \(M\). (It is similar to \(M \gg \lambda x \rightarrow N\) in Haskell.) The syntax we give here does not include any method of introducing effects; we extend CBPV with some example computational effects in Section 2.2.

The evaluation order in CBPV is fixed for each program. The only primitive that causes the evaluation of two separate computations is \textit{to}, which implements eager sequencing. Thunks give us more control over the evaluation order: they can be arbitrarily duplicated and discarded, and can be forced in any order chosen by the program. This is how CBPV captures both call-by-value and call-by-name (see Section 2.1 below).

CBPV has two typing judgments: \(\Gamma \vdash V : A\) for values and \(\Gamma \vdash_c M : C\) for computations. Typing contexts \(\Gamma\) are ordered lists of (variable, value type) pairs. We require that no variable appears more than once in any typing context. Figure 1 gives the typing rules. Rules that add a new variable to a typing context implicitly require that the variable is fresh. We write \(\emptyset\) for the empty typing context, \(V : A\) as an abbreviation for \(\emptyset \vdash V : A\), and \(M : C\) as an abbreviation for \(\emptyset \vdash_c M : C\).

---

\(^1\)Our use of CBPV should not be regarded as essential. We use it here because it is known to capture call-by-value and call-by-name in a strong sense (see [Lev99]). It may be possible to replace CBPV with some other language that captures call-by-value and call-by-name, and obtain similar results to ours.
We give an operational semantics for CBPV. This consists of a big-step evaluation relation \( M \Downarrow R \), which means the computation \( M \) evaluates to \( R \). Here \( R \) ranges over

\[
\begin{align*}
\lambda\{1, M_1, 2, M_2\} \Downarrow \lambda\{1, M_1, 2, M_2\} \\
\lambda x: A. M \Downarrow \lambda x: A. M \\
\text{return} V \Downarrow \text{return} V \\
\text{if true then} M_1 \text{ else } M_2 \Downarrow R \\
\text{match} (V_1, V_2) \text{ with } (x_1, x_2). M \Downarrow R \\
\end{align*}
\]

\[
\begin{align*}
M \Downarrow \lambda\{1, N_1, 2, N_2\} \\
N_i \Downarrow R \\
M \Downarrow \lambda x: A. N \\
N[x \mapsto V] \Downarrow R \\
M \Downarrow \text{return} V \\
N[x \mapsto V] \Downarrow R \\
M \Downarrow \text{force} (\text{thunk} M) \\
\end{align*}
\]
terminal computations, which are the subset of computations with an introduction form on the outside:

\[ R ::= \lambda \{1, M_1, 2, M_2\} | \lambda x : A. M | \text{return } V \]

We only evaluate closed, well-typed computations, so when we write \( M \Downarrow R \) we assume \( M : C \) for some \( C \) (this implies \( R : C \)). Reduction therefore cannot get stuck. The rules defining \( \Downarrow \) are given in Figure 2. All terminal computations evaluate to themselves. Products of computations are lazy: to evaluate a projection \( i \cdot M \), only the \( i \)th component of the pair \( M \) is evaluated. Since we have not yet included any way of forming impure computations, the semantics is deterministic and normalizing: given any \( M : C \), there is exactly one terminal computation \( R \) such that \( M \Downarrow R \). Section 2.2 extends the semantics in ways that violate these properties. We are primarily interested in evaluating computations of returner type.

A CBPV program is a closed computation \( M : \text{F} \text{bool} \). The reasoning principle we give for call-by-value and call-by-name relates open terms in program contexts. A program relation consists of a preorder\(^2\) \( \preceq \) on programs. For example, we could use

\[ M \preceq M' \text{ if and only if } \forall V : \text{bool}. (M \Downarrow \text{return } V) \Rightarrow (M' \Downarrow \text{return } V) \]

We could also use, for example, the total relation for \( \preceq \) (and in this case apply our reasoning principle for call-by-value and call-by-name even if we include e.g. mutable state as a side effect -- but then of course the conclusion of our reasoning principle would be trivial). Given any program relation \( \preceq \), we define a contextual preorder \( M \preceq_{\text{ctx}} M' \) on arbitrary well-typed computations (in typing context \( \Gamma \)) by considering the behaviour of \( M \) and \( M' \) in programs as follows. A computation context \( E \) is a computation term, with a single hole \( \square \) where a computation term is expected. We write \( E[M] \) for the computation that results from replacing \( \square \) with \( M \) (which may capture some of the free variables of \( M \)). For example, if \( E = \) the computation context \( N \text{ to } x. \square \) then \( E[\text{return } x] \) is the computation \( N \text{ to } x. \text{return } x \), where \( x \) is captured. We use computation contexts to define \( \preceq_{\text{ctx}} \).

**Definition 2.1** (Contextual preorder). Suppose that \( \preceq \) is a program relation, and that \( \Gamma \vdash_c M : C \) and \( \Gamma \vdash_c M' : C \) are two computations of the same type. We write \( M \preceq_{\text{ctx}} M' \) if, for all computation contexts \( E \) such that \( E[M], E[M'] : \text{F} \text{bool} \), we have \( E[M] \preceq E[M'] \). We write \( M \equiv_{\text{ctx}} M' \), and say that \( M \) and \( M' \) are contextually equivalent, when both \( M \preceq_{\text{ctx}} M' \) and \( M' \preceq_{\text{ctx}} M \) hold.

We sometimes omit \( \Gamma \), and write just \( M \preceq_{\text{ctx}} M' \) or \( M \equiv_{\text{ctx}} M' \).

### 2.1. Call-by-value and call-by-name.

We use CBPV (instead of e.g. Moggi’s monadic metalanguage [Mog91]) because it captures both call-by-value and call-by-name in a strong sense (see the introduction of [Lev99] for a detailed discussion of this). Levy [Lev99] gives two compositional translations from a source language into CBPV: one for call-by-value and one for call-by-name. We recall both translations in this section; our goal is to reason about the relationship between them.

For the source language, we use the following syntax of types \( \tau \) and expressions \( e \):

\[
\begin{align*}
\tau & ::= \text{unit} \mid \text{bool} \mid \tau_1 \times \tau_2 \mid \tau \rightarrow \tau' \\
e & ::= x \mid () \mid (e_1, e_2) \mid \text{fst } e \mid \text{snd } e \mid \text{true} \mid \text{false} \mid \text{if } e_0 \text{ then } e_1 \text{ else } e_2 \mid \lambda x : \tau. e \mid ee'
\end{align*}
\]

\(^2\)We do not actually need to assume that \( \preceq \) is reflexive or transitive at any point, but because of constraints we add later (such as existence of an adequate model), we do not expect there to be any interesting examples in which \( \preceq \) is not a preorder.
We include two base types \texttt{unit} and \texttt{bool} to be used in examples.\footnote{Unlike in Levy [Lev99], we do not include general sum types, only \texttt{bool}. We expect that including arbitrary sum types would complicate Section 4, because it is difficult to extend logical relations of varying arity with sums. The difficulty, and techniques for dealing with it, are discussed e.g. in [AS19, FS99, Kat08].} The source language has a typing judgement of the form $\Gamma \vdash e : \tau$, defined by the usual rules.

The two translations from the source language to CBPV are defined in Figure 3. For call-by-value, each source language type $\tau$ is mapped to a CBPV value type $V(\tau)$ that contains the results of call-by-value computations. For call-by-name, $\tau$ is translated to a computation type $N(\tau)$, which contains the computations themselves. Products in call-by-value use the value-type products of CBPV (which means they are necessarily \textit{strict}: both components of a pair are always evaluated). For call-by-name we give a lazy interpretation of binary products, using products of CBPV computation types. (Though note that we do not interpret \texttt{unit} as a nullary product of computation types. We instead treat \texttt{unit} as a base type, so that effects can happen at type \texttt{unit}, which matches typical functional languages.) Functions under the call-by-value translation accept values of type $V(\tau)$ as arguments; arguments are evaluated before being passed to the function. Under the call-by-name translation, functions accept thunks of computations as arguments; instead of evaluating them, arguments are thunked before passing them to call-by-name functions. Source-language typing contexts $\Gamma$ are translated to CBPV typing contexts $V(\Gamma)$ and $N(\Gamma)$. In call-by-value they contain values, in call-by-name they contain thunks of computations. Source-language expressions $e$ are mapped to CBPV computations $V(e)$ and $N(e)$. The translation uses some auxiliary program variables, which are assumed fresh.

For call-by-value we arbitrarily choose left-to-right evaluation for both pairing and function application. Under the call-by-name translation, computational effects occur only at the base types \texttt{unit} and \texttt{bool} (since this is where the returner types appear).

Of course, we have to justify that these translations actually capture call-by-value and call-by-name. There are two semantics of interest for the source language: a call-by-value semantics (that evaluates left-to-right), and a call-by-name semantics (with lazy products). Since we consider the observable behaviour of CBPV terms, the properties we want are that if the call-by-value translations $V(e)$ and $V(e')$ have the same observable behaviour then $e$ and $e'$ have the same observable behaviour with respect to the call-by-value semantics, and similarly for call-by-name. Levy [Lev99] proves both of these properties (though without products in the source language). We take this as the required justification, and do not give the details.

\section*{2.2. Examples.} We consider three collections of (allowable) effects as examples throughout the paper.

\textbf{Example 2.2} (No effects). We include the simplest possible example: the case where there are no computational effects at all. For this example, call-by-value and call-by-name turn out to have identical behaviour. We define the program relation $\approx_{\text{pure}}$ for closed computations $M, M' : F\texttt{bool}$ as:

$$M \approx_{\text{pure}} M' \text{ if and only if } \exists V : \texttt{bool}. (M \Downarrow \texttt{return} V) \land (M' \Downarrow \texttt{return} V)$$

In other words, $M$ and $M'$ both evaluate to the same result $V$. Since evaluation is deterministic, $V$ is necessarily unique. The contextual preorder $\approx_{\text{ctx}}$ $M' \text{ means if we construct two programs by wrapping } M \text{ and } M' \text{ in the same computation context, then}$
Figure 3. Translations from the source language into CBPV

<table>
<thead>
<tr>
<th>type τ → value type ‖τ‖</th>
<th>typing context Γ → typing context ‖Γ‖</th>
</tr>
</thead>
<tbody>
<tr>
<td>unit unit</td>
<td>⊙ ⊙</td>
</tr>
<tr>
<td>‖τ1‖ × ‖τ2‖</td>
<td>Γ, x : τ → ‖Γ‖, x : ‖τ‖</td>
</tr>
<tr>
<td>bool bool</td>
<td>Γ, x : τ → ‖Γ‖, x : ‖τ‖</td>
</tr>
<tr>
<td>τ → τ′</td>
<td>Γ, x : τ → ‖Γ‖, x : ‖τ‖</td>
</tr>
</tbody>
</table>

expression Γ ⊢ e : τ → computation ‖Γ‖ ⊢ e : F ‖τ‖

<table>
<thead>
<tr>
<th>expression Γ ⊢ e : τ → computation ‖Γ‖ ⊢ e : N ‖τ‖</th>
</tr>
</thead>
<tbody>
<tr>
<td>x → return x</td>
</tr>
<tr>
<td>() → return ()</td>
</tr>
<tr>
<td>(e1, e2) → N ‖e1‖ to z to N ‖e2‖ to z2 return z1</td>
</tr>
<tr>
<td>fst e → N ‖e‖ to z match z with (z1, z2) return z1</td>
</tr>
<tr>
<td>snd e → N ‖e‖ to z match z with (z1, z2) return z2</td>
</tr>
<tr>
<td>true → return true</td>
</tr>
<tr>
<td>false → return false</td>
</tr>
<tr>
<td>if e0 then e1 else e2 → N ‖e0‖ to z if z then N ‖e1‖ else N ‖e2‖</td>
</tr>
<tr>
<td>λx : τ. e → return thunk λx : N ‖τ‖. N ‖e‖</td>
</tr>
<tr>
<td>e e′ → N ‖e‖ to y N ‖e′‖ to z match z with (z1, z2) return y</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>expression Γ ⊢ e : τ → computation ‖Γ‖ ⊢ e : N ‖τ‖</th>
</tr>
</thead>
<tbody>
<tr>
<td>x → force x</td>
</tr>
<tr>
<td>() → return ()</td>
</tr>
<tr>
<td>(e1, e2) → λ{1. N ‖e1‖, 2. N ‖e2‖}</td>
</tr>
<tr>
<td>fst e → 1: N ‖e‖</td>
</tr>
<tr>
<td>snd e → 2: N ‖e‖</td>
</tr>
<tr>
<td>true → return true</td>
</tr>
<tr>
<td>false → return false</td>
</tr>
<tr>
<td>if e0 then e1 else e2 → N ‖e0‖ to z if z then N ‖e1‖ else N ‖e2‖</td>
</tr>
<tr>
<td>λx : τ. e → λx : U(N ‖τ‖), N ‖e‖</td>
</tr>
<tr>
<td>e e′ → (thunk N ‖e′‖) N ‖e‖</td>
</tr>
</tbody>
</table>

(A) Call-by-value translation V(−)

(b) Call-by-name translation N(−)
these two programs evaluate to the same result. This relation is symmetric. Our other examples use non-symmetric relations.

**Example 2.3** (Divergence). For our second example, the only effect is divergence (via recursion). In this case, call-by-value and call-by-name do not have identical behaviour (they are not related by $\preceq_{\text{ctx}}$ as it is defined in our no-effects example). We instead show that replacing call-by-value with call-by-name does not change a terminating program into a diverging one.

We extend our two languages with recursion. For CBPV we extend the syntax of computations with fixed points $\text{rec } x : U C. M$, and correspondingly extend the type system and operational semantics with the following rules:

$\Gamma, x : U C \vdash_c M : C \quad M[x \mapsto \text{thunk}(\text{rec } x : U C. M)] \Downarrow R$

$\begin{array}{c}
\text{rec } x : U C. M \Downarrow R
\end{array}$

The variable $x$ is bound to a thunk of the recursive computation, so recursion is done by forcing $x$. (This is not the only way to add recursion to CBPV [DCL18], but is the most convenient for our purposes.) Of course, by adding recursion we lose normalization (but the semantics is still deterministic). We extend the source language, and the two translations into CBPV, with recursive functions:

$e ::= \ldots \mid \text{rec } f : \tau \to \tau', \lambda x. e \quad \Gamma, f : \tau \to \tau', x : \tau \vdash e : \tau'$

$\Gamma \vdash f : \tau \to \tau', \lambda x. e : \tau \to \tau$

$\mathcal{V}(\text{rec } f : \tau \to \tau', \lambda x. e) = \text{return thunk}(\text{rec } f : U (\mathcal{V}(\tau)) \to F (\mathcal{V}(\tau')), \lambda x : \mathcal{V}(\tau). \mathcal{V}(e))$

$\mathcal{N}(\text{rec } f : \tau \to \tau', \lambda x. e) = \text{rec } f : U (\mathcal{N}(\tau)) \to F (\mathcal{N}(\tau')), \lambda x : U (\mathcal{N}(\tau)). \mathcal{N}(e)$

Again, the translations are the same as those given by Levy [Lev99], except that Levy has general fixed points for call-by-name, rather than just recursive functions. The expression $\Omega_\tau = (\text{rec } f : \text{bool} \to \tau. \lambda x. f x) \text{false}) : \tau$ enables us to distinguish between call-by-value and call-by-name: $(\lambda x : \tau. \text{true}) \Omega_\tau$ diverges in call-by-value but not in call-by-name. In particular, we have $\mathcal{N}(\lambda x : \tau. \text{true}) \Omega_\tau) \Downarrow \text{return true}$, but there is no $R$ such that $\mathcal{V}(\lambda x : \tau. \text{true}) \Omega_\tau) \Downarrow R$.

For this example, we define the program relation $\preceq_{\text{div}}$ by

$M \preceq_{\text{div}} M' \text{ if and only if } \forall V : \text{bool}. (M \Downarrow \text{return } V) \Rightarrow (M' \Downarrow \text{return } V)$

so that $M \preceq_{\text{ctx}} M'$ informally means if a program containing $M$ terminates with some result then the same program with $M'$ instead of $M$ terminates with the same result.

**Example 2.4** (Nondeterminism). For our third example, we consider finite nondeterminism. Again call-by-value and call-by-name have different behaviour, but any result of a call-by-value execution is also a result of a call-by-name execution (if suitable nondeterministic choices are made).

We consider CBPV without recursion, but augmented with computations $\text{fail}_C$ for nullary nondeterministic choice and $M \text{ or } N$ for binary nondeterministic choice between computations; the typing and evaluation rules are standard:

$\begin{array}{c}
\Gamma \vdash_c \text{fail}_C : C \\
\Gamma \vdash_c M : C \\
\Gamma \vdash_c N : C \\
M \Downarrow R \\
N \Downarrow R \\
\text{or } M \Downarrow R \\
\text{or } N \Downarrow R
\end{array}$
(There is no \( R \) such that \( \text{fail}_C \downarrow R \).) The computation \( \text{fail}_C \) is the unit for \( \text{or} \), so \( \text{fail}_C \text{ or } M \) and \( M \text{ or } \text{fail}_C \) have the same behaviour as \( M \). For each closed computation \( M : F \mathcal{A} \) there might be zero, one or several values \( V : A \) such that \( M \downarrow \text{return } V \).

We similarly include nullary and binary nondeterminism in the source language, and extend the call-by-value and call-by-name translations:

\[
\begin{align*}
\Gamma \vdash e : \tau & \quad \Gamma \vdash e' : \tau \\
\Gamma \vdash \text{fail} : \tau &
\end{align*}
\]

\[
\begin{align*}
\mathcal{V}(\text{fail}) &= \text{fail} \mathcal{F}(\mathcal{V}(\tau)) \\
\mathcal{V}(e \text{ or } e') &= \mathcal{V}(e) \text{ or } \mathcal{V}(e') \\
\mathcal{N}(\text{fail}) &= \text{fail} \mathcal{N}(\tau) \\
\mathcal{N}(e \text{ or } e') &= \mathcal{N}(e) \text{ or } \mathcal{N}(e')
\end{align*}
\]

As an example, evaluating the expression \( e = (\lambda x. \text{if } x \text{ then } x \text{ else true})(\text{true or false}) \) under call-by-value necessarily results in \( \text{true} \), but under call-by-name we can also get \( \text{false} \). (We have \( \mathcal{V}(e) \downarrow \text{return false} \) but \( \mathcal{N}(e) \downarrow \text{return false} \).)

For nondeterminism, we define \( \preceq_{\text{nd}} \) in the same way as our divergence example:

\[
M \preceq_{\text{nd}} M' \quad \text{if and only if} \quad \forall V : \text{bool} \quad (M \downarrow \text{return } V) \Rightarrow (M' \downarrow \text{return } V)
\]

This captures the property that any result that arises from an execution of \( M \) (which may involve call-by-value) might arise from an execution of \( M' \) (which may involve call-by-name).

**Example 2.5** (Immutable state). Finally, we consider the basic languages enriched with an extra construct for getting the value of an immutable state whose value is either \( \text{true} \) or \( \text{false} \). Once again we do not expect there to be any difference between call-by-value and call-by-name, and it is indeed the case that if \( e \) is a closed expression of type \( \text{bool} \), then call-by-value and call-by-name evaluations of \( e \) have the same behaviour (this is an instance of Corollary 4.8). Notably however, the model we use for this example fails to satisfy the assumptions of our main theorem (Theorem 6.2).

We augment CBPV with a computation \( \text{get} \). This gets the value of the state, producing either \( \text{true} \) or \( \text{false} \).

\[
\Gamma \vdash \text{get} : \text{F bool}
\]

Big-step evaluation has a slightly different form in this case. We write \( M \downarrow b R \) to mean \( M \) evaluates to \( R \) when the state is \( b \in \{ \text{true}, \text{false} \} \). The rules are those of Figure 2 (with the subscript \( b \) added), plus

\[
\text{get } \downarrow_{\text{true}} \text{return true} \quad \text{get } \downarrow_{\text{false}} \text{return false}
\]

Again we extend the source language, and also the call-by-value and call-by-name translations:

\[
\begin{align*}
e & := \ldots | \text{get} \\
\Gamma \vdash \text{get} : \text{bool} \\
\mathcal{V}(\text{get}) &= \mathcal{N}(\text{get}) = \text{get}
\end{align*}
\]

We define the program relation \( \preceq_{\text{get}} \) as follows:

\[
M \preceq_{\text{get}} M' \quad \text{if and only if} \quad \forall b \in \{ \text{true}, \text{false} \} . \exists V : \text{bool} \quad (M \downarrow b \text{return } V) \land (M' \downarrow b \text{return } V)
\]
3. Order-enriched denotational semantics

We give a denotational semantics for CBPV, which we use to prove instances of ≼_{ctx}. Since ≼_{ctx} is not in general symmetric, we use order-enriched models, which come with partial orders ⊆ between denotations. In an adequate model, [M] ⊆ [N] implies M ≼_{ctx} N. Our semantics is based on Levy’s algebra models [Lev06] for CBPV, in which each computation type is interpreted as a monad algebra. (We restrict to algebra models for simplicity. Other forms of model, such as adjunction models [Lev03] can be used for the same purpose.)

3.1. Order-enriched categories and strong monads. We define the basic categorical notions we need for the rest of the paper. We assume no knowledge of enriched category theory; instead we give the relevant order-enriched (specifically Poset-enriched) definitions here. (We do however assume some basic ordinary category theory.)

Definition 3.1. A Poset-category C is an ordinary category, together with a partial order ⊆ on each hom-set C(X, Y), such that composition is monotone.

If C is a Poset-category, we refer to the ordinary category as the underlying ordinary category, and write |C| for the class of objects.

Example 3.2. We use the following three Poset-categories.

| Poset-category C | Objects X ∈ |C| | Morphisms f : X → Y | Order f ⊆ f' |
|------------------|-------------|------------------|------------------|
| Set              | sets        | functions        | equality         |
| Poset            | posets      | monotone functions | pointwise         |
| ωCpo             | ωcpos       | ω-continuous functions | pointwise         |

In each case, composition and identities are defined in the usual way. For Set, since the hom-posets Set(X, Y) are discrete, all of the Poset-enriched definitions coincide with the ordinary (unenriched) definitions. The objects of ωCpo are posets (X, ⊆) for which ⊆ is ω-complete, i.e. for which every ω-chain x_0 ⊆ x_1 ⊆ ⋯ has a least upper bound ⋃ x. Morphisms are ω-continuous functions, i.e. monotone functions that preserve least upper bounds of ω-chains.

Let C be a Poset-category. We say that C is cartesian when its underlying category has a terminal object 1 and binary products X_1 × X_2, such that the pairing functions ⟨−, −⟩ : C(W, X_1) × C(W, X_2) → C(W, X_1 × X_2) are monotone. We write π_i : X_1 × X_2 → X_i for the projections from a product, and write ⟨⟩_X : X → 1 for the unique map into the terminal object. In every cartesian category, there are canonical associativity isomorphisms assoc_{X_1,X_2,X_3} : (X_1 × X_2) × X_3 → X_1 × (X_2 × X_3). We say that C is cartesian closed when it is cartesian and its underlying category has exponentials X ⇒ Y for which the currying functions Λ : C(W × X, Y) → C(W, X ⇒ Y) are monotone. We write ev_{X,Y} for the evaluation morphism Λ^{-1}id : (X ⇒ Y) × X → Y.

Binary coproducts in C are just binary coproducts in the underlying ordinary category, except that the copairing functions [−, −] : C(X_1, W) × C(X_2, W) → C(X_1 + X_2, W) are required to be monotone. We write inl : X_1 → X_1 + X_2 and inr : X_2 → X_1 + X_2 for the coprojections. The Poset-categories Set, Poset, and ωCpo are all cartesian closed, and have binary coproducts given by disjoint union.
Above we ask for monotonicity of the bijections  
\[ (-, -) : C(W, X_1) \times C(W, X_2) \to C(W, X_1 \times X_2) \]
\[ \Lambda : C(W \times X, Y) \to C(W, X \Rightarrow Y) \]
\[ [-, -] : C(X_1, W) \times C(X_2, W) \to C(X_1 + X_2, W) \]
We do not need to require monotonicity of their inverses explicitly, because this holds automatically. In particular, the uncurrying functions \( \Lambda^{-1} : C(W, X \Rightarrow Y) \to C(W \times X, Y) \) are monotone because \( \Lambda^{-1} f = ev_{X,Y} \circ (f \times id_X) \), and \( \circ \) and \( \times \) are both monotone.

We interpret computation types as (Eilenberg–Moore) algebras for an order-enriched monad \( T \), which we need to be strong (just as models of Moggi’s monadic metalanguage [Mog91] use a strong monad). The definitions of strong Poset-monad and of \( T \)-algebra we give are slightly non-standard, but are equivalent to the standard ones (see for example [MU22]). In particular, it is more convenient for us to bake the strength into the (Kleisli) extension of the monad instead of having a separate strength.

**Definition 3.3** (Strong Poset-monad). Let \( C \) be a cartesian Poset-category. A strong Poset-monad \( T \) on \( C \) consists of:
- an object \( TX \in |C| \) for each \( X \in |C| \);
- a morphism \( \eta_X : X \to TX \) for each \( X \in |C| \) (the unit);
- a monotone function \( (-)^{W \times \square} : C(W \times TX, TY) \to C(W \times TX, TY) \) (Kleisli extension) for each \( W, X, Y \in |C| \).

These are required to satisfy the following four laws.\(^4\)

- **Naturality of extension in \( W \):**
  \[ f^{(W \times \square)} \circ (w \times id_{TX}) = \left( f \circ (w \times id_X) \right)^{W \times \square} \]
  for all \( f : W \times X \to TY \) and \( w : W' \to W \).
- **Left unit:**
  \[ f^{W \times \square} \circ (id_W \times \eta_X) = f \]
  for all \( f : W \times X \to TY \).
- **Right unit:**
  \[ (\eta_X \circ \pi_2)^{1 \times \square} = \pi_2 \]
  for all \( X \in |C| \).
- **Associativity:**
  \[ (g^{W' \times \square} \circ (id_{W'} \circ f) \circ assoc)^{(W' \times W) \times \square} = g^{W' \times \square} \circ (id_{W'} \times f^{W \times \square}) \circ assoc \]
  for all \( f : W \times X \to TY \) and \( g : W' \times Y \to TZ \).

Specializing the Kleisli extension of \( T \) to \( W = 1 \) produces a (non-strong) extension operator \( (-)^\dagger : C(X, TY) \to C(TX, TY) \), satisfying the usual monad laws:
\[ f^\dagger \circ \eta_X = f \quad \eta_X^{\dagger} = id_X \quad (g^\dagger \circ f)^\dagger = g^\dagger \circ f^\dagger \]

We use this to define, for every \( f : X \to Y \), a morphism \( Tf : TX \to TY \) by \( Tf = (\eta_Y \circ f)^\dagger \). The latter definition makes \( T \) into a Poset-functor: the mapping \( f \mapsto Tf \) is monotone, and

\(^4\)The conference version [MM22] of this paper incorrectly omits naturality in \( W \) from the definition of strong Poset-monad and from the definition of Eilenberg–Moore algebra. (Naturality in \( W \) is required in [MU22, Definition 4.1].)
preserves identities and composition. The definition of $T$ on morphisms also ensures that the unit and Kleisli extension of $T$ satisfy the following naturality laws:

$T f \circ \eta_X = \eta_Y \circ f \quad (T g \circ f) |^{W \times \Box} = T g \circ f |^{W \times \Box} \quad (f \circ (id_W \times g)) |^{W \times \Box} = f |^{W \times \Box} \circ (id_W \times T g)$

In the notation $f |^{\Box \times W} : W \times TX \rightarrow TY$, the square $\Box$ indicates the position of $T$ in the domain. Since products are symmetric, choosing to put $T$ to the right of $W$ is arbitrary. We construct a Kleisli extension operator with the square to the left as follows:

$f |^{\Box \times W} = (f \circ (\pi_2, \pi_1)) |^{\Box \times W} \circ (\pi_2, \pi_1) : TX \times W \rightarrow TY$ (where $f : X \times W \rightarrow TY$)

We also define two natural transformations for sequencing of computations: $seq^L$ for left-to-right and $seq^R$ for right-to-left, as follows.

$seq^L_{X_1, X_2} = (\eta_{X_1 \times X_2} |^{\Box \times X_2}) |^{\Box \times TX_2} : TX_1 \times TX_2 \rightarrow T(X_1 \times X_2)$

$seq^R_{X_1, X_2} = (\eta_{X_1 \times X_2} |^{\Box \times X_2}) |^{\Box \times TX_1} : TX_1 \times TX_2 \rightarrow T(X_1 \times X_2)$

We further define an effectful pairing operation $\langle\langle - , - \rangle\rangle$:

$\langle\langle f_1, f_2 \rangle\rangle = seq^L_{X_1, X_2} \circ (f_1, f_2) : W \rightarrow T(X_1 \times X_2)$ (where $f_i : W \rightarrow TX_i$)

This evaluates from left to right; we do not need the right-to-left version.

**Definition 3.4 (Eilenberg–Moore algebra).** Let $T$ be a strong Poset-monad on a cartesian Poset-category $C$. A $T$-algebra $Z = (Z, (-)^T)$ is a pair of:

- an object $Z \in |C|$ (the carrier);
- a monotone function $(-)^T |^{W \times \Box} : C(W \times X, Z) \rightarrow C(W \times TX, Z)$ (the extension operator) for each $W, X \in |C|$.

These are required to satisfy the following three laws.

- **Naturality in $W$:**

  $f |^{W \times \Box} \circ (w \times id_{TX}) = (f \circ (w \times id_X)) |^{W' \times \Box}$

  for all $f : W \times X \rightarrow Z$ and $w : W' \rightarrow W$.

- **Left unit:**

  $f |^{W \times \Box} \circ (id_W \times \eta_X) = f$

  for all $f : W \times X \rightarrow Z$.

- **Associativity:**

  $g |^{W' \times \Box} \circ (id_W \times f) \circ assoc |^{\Box \times (W' \times W)} = g |^{W' \times \Box} \circ (id_W \times f |^{W \times \Box}) \circ assoc$

  for all $f : W \times X \rightarrow TY$ and $g : W' \times Y \rightarrow Z$.

  For each $T$-algebra $Z$, we write $U_T Z$ for the carrier $Z \in |C|$.

Just as for the extension operator of a strong Poset-monad, we specialize the extension operator of a $T$-algebra to $W = 1$ and obtain a (non-strong) extension operator $(-)^T : C(X, Z) \rightarrow C(TX, Z)$. We also have an extension operator with reversed products, written $(-)^{\Box \times W} : C(X \times W, Z) \rightarrow C(TX \times W, Z)$.

The following constructions of algebras are standard.
Definition 3.5. Let $T$ be a strong Poset-monad on a cartesian closed Poset-category $C$.

- The free $T$-algebra $FTX$ on an object $X \in |C|$ has carrier $TX$; the extension operator is Kleisli extension $(-)^T$.
- If $Z_1$ and $Z_2$ are $T$-algebras, then their product $Z_1 \times Z_2$ is the $T$-algebra with carrier $Z_1 \times Z_2$, and extension operator
  \[ f_{Z_1 \times Z_2} = \langle \pi_1 \circ f \rangle_{Z_1 \times Z_2}, \langle \pi_2 \circ f \rangle_{Z_1 \times Z_2} \]
- If $Y \in |C|$ and $Z$ is a $T$-algebra, then their power $Y \Rightarrow Z$ is the $T$-algebra with carrier $Y \Rightarrow Z$ and extension operator
  \[ f_{Y \Rightarrow Z} = \Lambda((\Lambda^{-1} f \circ \beta_{W, Y, X})_{Y \times X}) \circ \beta_{W, TX, Y} \]
where $\beta_{X_1, X_2, X_3} = \langle \langle \pi_1 \circ \pi_1, \pi_2 \circ \pi_1 \rangle : (X_1 \times X_2) \times X_3 \rightarrow (X_1 \times X_3) \times X_2 \rangle$.

We use these constructions to interpret CBPV computation types: returner types are interpreted as free $T$-algebras, product types $C_1 \times C_2$ are interpreted as product $T$-algebras, and function types $A \rightarrow C$ are interpreted as power $T$-algebras.

3.2. Models of CBPV. We define the notion of (order-enriched, algebra) model as follows.

Definition 3.6. A model $M = (C, T)$ of CBPV consists of

- a cartesian closed Poset-category $C$ that admits the coproduct $2 = 1 + 1$;
- a strong Poset-monad $T$ on $C$.

Given a model $M = (C, T)$, the interpretation $[-]$ of CBPV is defined in Figure 4. Value types $A$ are interpreted as objects $[A] \in |C|$, while computation types $C$ are interpreted as $T$-algebras. The value type $UC$ is interpreted as the carrier $UT[C]$ of the $T$-algebra $[C]$. Typing contexts $\Gamma$ are interpreted as objects $[\Gamma] \in C$ using the cartesian structure of $C$; if $(x : A) \in \Gamma$ then we write $\pi_x$ for the corresponding projection $[\Gamma] \rightarrow [A]$. Values $\Gamma \vdash V : A$ (respectively computations $\Gamma \vdash c : M : C$) are interpreted as morphisms $[\Gamma] \rightarrow [A]$ (resp. $[\Gamma] \rightarrow [M] : [C]$) in $C$; we often omit the typing context and type when writing these. Programs $\circ \vdash_{\mathsf{ctx}} M : \mathsf{bool}$ are therefore interpreted as morphisms $[M] : 1 \rightarrow T2$. To interpret if, we use the fact that, since $C$ is cartesian closed, products distribute over the coproduct $2 = 1 + 1$. This means that for every $W \in |C|$, the coproduct $W + W$ also exists in $C$, and the canonical morphism
  \[ W + W \rightarrow \frac{\left([id_W, inl]{\omega}, [id_W, inr]{\omega}\right)}{\omega} \rightarrow W \times 2 \]
has an inverse $\text{dist}_{W} : W \times 2 \rightarrow W + W$.

By composing the semantics of CBPV with the two translations of the source language, we obtain a call-by-value semantics $\mathcal{V}[-] = [\mathcal{V}(-)]$ and a call-by-name semantics $\mathcal{N}[-] = [\mathcal{N}(-)]$ of the source language. For convenience, we spell out these composed semantics in Figure 5.

We use the denotational semantics as a tool for proving instances of contextual preorders; for this we need adequacy.

Definition 3.7. A model of CBPV is adequate with respect to a given program relation $\preceq$ if for all computations $\Gamma \vdash_{\mathsf{ctx}} M : C$ and $\Gamma \vdash_{\mathsf{ctx}} M' : C$ we have
  \[ \Gamma \vdash M : C \subseteq \Gamma \vdash M' : C \quad \Rightarrow \quad M \preceq_{\mathsf{ctx}} M' \]
The interpretation of the \( \pi \) operator is completely determined once the carrier is fixed; it is analogous to the Kleisli extension.

\[
\text{Example 3.9.} \quad \text{For divergence, we use } C = \omega \text{Cpo. The strong Poset-monad } T \text{ freely adjoins a least element } \bot \text{ to each } \omega \text{Cpo. The unit } \eta_X \text{ is the inclusion } X \hookrightarrow TX, \text{ while Kleisli extension is given by}
\]

\[
f^T(w, x) = \begin{cases} \bot & \text{if } x = \bot \\ f(w, x) & \text{otherwise} \end{cases}
\]

A \( T \)-algebra \( Z \) is equivalently an \( \omega \text{Cpo} \) \( Z \) with a least element \( \bot \in Z \). The extension operator is completely determined once the carrier is fixed; it is analogous to \( (-)^1 \). In this case, the product \( Z_1 \times Z_2 \) is the set of pairs ordered pointwise, and the exponential \( Y \Rightarrow Z \) is the set of set of \( \omega \)-continuous functions ordered pointwise. Hence \( Z_1 \times Z_2 \) has a least element \( (\bot, \bot) \) (so forms a \( T \)-algebra) whenever \( Z_1 \) and \( Z_2 \) have least
Figure 5. Denotational semantics of call-by-value and call-by-name
elements, and \( Y \Rightarrow Z \) has a least element (the commonly-\( \bot \) function) whenever \( Z \) has a least element.

If \( Z \) is a \( T \)-algebra, then every \( \omega \)-continuous function \( f : Z \rightarrow Z \) has a least fixed point \( \text{fix } f = \bigsqcup_{n \in \mathbb{N}} f^n \bot \in Z \). These enable us to interpret recursive computations, by defining \( \text{fix } x. U C, M \rho = \text{fix}(\lambda x. [M](\rho, x)) \). The interpretation \( [M] \) of a closed computation \( M : F\text{bool} \) is either \( \bot \) (signifying divergence), or one of the two elements of \( 2 \).

**Example 3.10.** For finite nondeterminism, we use \( C = \text{Poset} \). The strong \( \text{Poset} \)-monad \( T \) freely adds finite joins to each poset. It is defined by

\[
\text{TX} = (\{ \downarrow S' | S' \in \mathcal{P}_{\text{fin}}X \}, \subseteq) \quad \eta_X x = \downarrow \{ x \} \quad f^{1 \times \Box}(w, S) = \bigcup_{x \in S} f(w, x)
\]

where \( \mathcal{P}_{\text{fin}}X \) is the set of finite subsets of \( X \), and \( \downarrow S' = \{ x \in X | \exists x' \in S'. x \subseteq x' \} \) is the **downwards-closure** of \( S' \subseteq X \). Each \( T \)-algebra is again completely determined by its carrier; a \( T \)-algebra \( Z \) is equivalently a poset \( Z \) that has finite joins. The extension operator is necessarily given by \( f^{1 \times \Box}(w, S) = \bigsqcup_{x \in S} f(w, x) \). (The latter join exists because \( S \) is the downwards-closure of a finite set, even though \( S \) itself might not be finite.) The product \( Z_1 \times Z_2 \) is the set of pairs ordered componentwise, with joins given by \( \bigsqcup_i(z_i, z'_i) = (\bigsqcup_i z_i, \bigsqcup_i z'_i) \).

The power \( Y \Rightarrow Z \) is the set of monotone functions ordered pointwise, with joins given by \( (\bigsqcup_i f_i)x = \bigsqcup_i (f_ix) \).

We interpret nondeterministic computations using nullary and binary joins:

\[
[[\text{fail}_C]] \rho = \bot \quad [[M \text{ or } N]] \rho = [[M]] \rho \uplus [[N]] \rho
\]

The interpretation \( [[M]] \) of a closed computation \( M : F\text{bool} \) is one of the four subsets of \( 2 \).

**Example 3.11.** For immutable state, we use \( C = \text{Set} \), with the **reader** monad

\[
\text{TX} = (2 \Rightarrow X) \quad \eta_X x = \lambda b. x \quad f^{1 \times \Box}(w, t) = \lambda b. f(w, t b) \ b
\]

where \( 2 = \{ \text{true}, \text{false} \} \). The CBPV computation \( \text{get} \) is interpreted as

\[
[[\text{get}]] \rho = \lambda b. b
\]
that relate elements $f^v$ of $T(\mathcal{V}[\tau])$ with elements $f^n$ of $U_T(\mathcal{N}[\tau])$. Here by element we mean generalized element, so $f^v$ and $f^n$ are morphisms

$$f^v : W \to T(\mathcal{V}[\tau]) \quad f^n : W \to U_T(\mathcal{N}[\tau])$$

from some $W$. The definition of $\mathcal{R}[\tau]_W$ is in the style of a logical relation, by induction on the type $\tau$. The cases are listed in Figure 6. Informally, we have the following.

- For $\tau = \text{unit}$ and $\tau = \text{bool}$, we compare $f^v$ and $f^n$ directly using the order relation $\sqsubseteq$ on morphisms in $\mathcal{C}$. (We can do this because $T(\mathcal{V}[\tau]) = U_T(\mathcal{N}[\tau])$.)
- For a product type $\tau_1 \times \tau_2$, we compare the first components and compare the second components. We get these components by composing with the call-by-value and call-by-name interpretations of the projections $\text{fst}$ and $\text{snd}$.
- For a function type $\tau \to \tau'$, we relate $f^v$ to $f^n$ when these give related results when applied to related arguments. Here we use the call-by-value and call-by-name interpretations of application.

Note that in the function case we quantify over morphisms $w : W' \to W$ to permit varying arities $W$ (cf. the Kripke logical relations of varying arity of [JT93]). (Precisely, this ensures that $\mathcal{R}[\tau]$ is closed under precomposition with morphisms $w : W' \to W$, as in Lemma 4.5(2) below.)

We define $\times$ in terms of $\mathcal{R}[-]$. To state the definition, we need some more notation. Let $\Gamma' = x_1 : A_1, \ldots, x_k : A_k$ be a CBPV context. Given a morphism $f_i : W \to [A_i]$ for each $i \leq k$, we obtain a morphism $\langle f_i \rangle_i : W \to [\Gamma]$, by iterated pairing. Given instead a morphism $f_i : W \to T[A_i]$ for each $i$, we obtain a morphism $\langle\langle f_i \rangle_i : W \to T[\Gamma]$ by iterated pairing and left-to-right evaluation:

$$\langle f_i \rangle_{i \leq k} = \begin{cases} \langle \rangle W & \text{if } k = 0 \\ \langle\langle f_i \rangle_{i \leq (k-1)}, f_k \rangle & \text{if } k > 0 \end{cases} \quad \langle\langle f_i \rangle_i \rangle_{i \leq k} = \begin{cases} \eta_1 \circ \langle \rangle W & \text{if } k = 0 \\ \langle\langle f_i \rangle_{i \leq (k-1)}, f_k \rangle & \text{if } k > 0 \end{cases}$$

**Definition 4.1.** Let $\mathcal{M}$ be a CBPV model, and let

$$g^v : \mathcal{V}[\Gamma] \to T(\mathcal{V}[\tau']) \quad g^n : \mathcal{N}[\Gamma] \to U_T(\mathcal{N}[\tau'])$$

be morphisms, where $\Gamma = x_1 : \tau_1, \ldots, x_k : \tau_k$. We write

$$g^v \prec g^n$$

when, for all objects $W \in |\mathcal{C}|$ and families of morphisms

$$(f^v_i : W \to T(\mathcal{V}[\tau_i]));_i \quad (f^n_i : W \to U_T(\mathcal{N}[\tau_i]));_i$$
we have
\[ f^n_1 R^n W f^n_1 \land \cdots \land f^n_k R^n W f^n_k \implies (g^{n+} \circ \langle f^n \rangle_i) \theta^n W (g^{n} \circ \langle f^n \rangle_i) \]

As we mention above, our goal is to relate the observable behaviour of \( V[e] \) to the observable behaviour of \( N[e] \). Precisely, we want to prove \( V[e] \preceq N[e] \). By considering what this means for specific expressions \( e \), we can see that this is not true in general, for three reasons:

- Consider the expression
  \[ e = (\lambda x: \text{bool}.()) : \text{bool} \to \text{unit} \]
  If we apply this to an argument, then in call-by-value we evaluate the argument but in call-by-name we do not.

- Consider the expression
  \[ e = (\lambda x: \text{bool}. \text{if } x \text{ then } x \text{ else } x) : \text{bool} \to \text{bool} \]
  If we apply this, then in call-by-value the argument is evaluated once, but in call-by-name the argument is evaluated twice.

- Consider the expression
  \[ e = (\lambda x: \text{bool}. \lambda y: \text{bool}. \text{if } y \text{ then } x \text{ else } x) : \text{bool} \to \text{bool} \to \text{bool} \]
  In call-by-value the argument to the outer function is evaluated first, and the argument to the inner function is evaluated second. In call-by-name the arguments are evaluated in the opposite order.

This suggests we should assume that computations can be discarded, copied, and reordered with respect to other computations. Precisely, we want the following properties.

**Definition 4.2.** Let \( T \) be a strong Poset-monad. A morphism \( f : X \to TY \) is:

- **lax discardable** when
  \[ T\langle \rangle_Y \circ f \sqsubseteq \eta_1 \circ \langle \rangle_X : X \to T1 \]

- **lax copyable** when
  \[ T\langle id_Y, id_Y \rangle \circ f \sqsubseteq \text{seq}_{Y,Y}^L \circ \langle f, f \rangle : X \to T(Y \times Y) \]
  equivalently, when
  \[ T\langle id_Y, id_Y \rangle \circ f \sqsubseteq \text{seq}_{Y,Y}^R \circ \langle f, f \rangle : X \to T(Y \times Y) \]

- **lax central** when
  \[ \text{seq}_{Y,W}^L \circ (f \times id_{TW}) \sqsubseteq \text{seq}_{Y,W}^R \circ (f \times id_{TW}) : X \times TW \to T(Y \times W) \]
  for all \( W \in |C| \).

The non-lax versions of these properties were first defined by Führmann [Füh99].

**Example 4.3.** For each of our examples from Section 3.3, every morphism \( f : X \to TY \) is lax discardable, lax copyable, and lax central.

Here we define these three properties for morphisms in the model \( M \), but there are similar notions for CBPV computations, as the following lemma shows.

**Lemma 4.4.** Let \( \Gamma \vdash c \mathcal{M} : F \mathcal{A} \) be a CBPV computation. The following hold for every CBPV model that is adequate with respect to a program relation \( \preceq \).
following closure properties.

Lemma 4.5. Let \( R \) be a CBPV model. For each \( \tau \), the family of relations \( R[\tau] \) has the following closure properties.

1. For all \( f^v, g^v, f^n, g^n \), we have
   \[
   g^v \sqsubseteq f^v \land f^v \rhd[\tau]_W f^n \land f^n \sqsubseteq g^n \quad \Rightarrow \quad g^v \rhd[\tau]_W g^n
   \]

2. For all \( w : W' \rightarrow W \), and \( f^v, f^n \), we have
   \[
   f^v \rhd[\tau]_W f^n \quad \Rightarrow \quad (f^v \circ w) \rhd[\tau]_{W'} (f^n \circ w)
   \]

3. For all \( f^v, f^n \), we have
   \[
   f^v \rhd[\tau]_{W \times X} f^n \quad \Rightarrow \quad (f^v)^{+\times \Box} \rhd[\tau]_{W \times T X} (f^n)^{+\times \Box}
   \]

4. For all \( W_1, W_2 \) such that the coproduct \( W_1 + W_2 \) exists, and all \( f_1^v, f_2^v, f_1^n, f_2^n \), we have
   \[
   f_1^v \rhd[\tau]_{W_1} f_1^n \land f_2^v \rhd[\tau]_{W_2} f_2^n \quad \Rightarrow \quad [f_1^v, f_2^v] \rhd[\tau]_{W_1 + W_2} [f_1^n, f_2^n]
   \]

Proof. The proof of each property is by induction on the type \( \tau \).

(1) The unit and \textbf{bool} cases are trivial, while the cases for product and function types follow from the inductive hypothesis by monotonicity of composition and pairing.
(2) The unit and bool cases follow from monotonicity of composition. The case for product types follows from the inductive hypothesis. For a function type $\tau \to \tau'$ we need to show, for every $w' : W'' \to W'$ and $g^v, g^n$, that $g^v \mathcal{R}[\tau]_{W''} g^n$ implies

$$(ev^\dag \circ \langle f^v \circ w \circ w', g^v \rangle) \mathcal{R}[\tau']_{W'} (ev \circ \langle f^n \circ w \circ w', g^n \rangle)$$

This follows immediately from the assumption $f^v \mathcal{R}[\tau \to \tau']_W f^n$, instantiated with the morphism $w \circ w' : W'' \to W$.

(3) The unit and bool cases follow from monotonicity of extension operators. The case for product types follows from the inductive hypothesis, by naturality of extension operators and the definition of the product $T$-algebra:

$$(T \pi_i \circ (f^v)^{W \times \square}) = ((T \pi_i \circ f^v)^{W \times \square}) \mathcal{R}[\tau_i]_{W \times TX} ((\pi_i \circ f^n)^{W \times \square}) = (\pi_i \circ (f^n)^{W \times \square})$$

For a function type $\tau \to \tau'$ we show, for every $w : W' \to W \times TX$ and $g^v, g^n$ satisfying $g^v \mathcal{R}[\tau]_{W'} g^n$, that

$$(ev^\dag \circ \langle ((f^v)^{W \times \square} \circ w, g^v) \rangle) \mathcal{R}[\tau']_{W'} (ev \circ \langle (f^n)^{W \times \square} \circ w, g^n \rangle)$$

By property (2) above, we have

$$(g^n \circ \pi_2) \mathcal{R}[\tau]_{(W \times X) \times W'} (g^n \circ \pi_2)$$

so that $f^v \mathcal{R}[\tau \to \tau']_{W \times TX} f^n$ implies

$$(ev^\dag \circ \langle f^v \circ \pi_1, g^v \circ \pi_2 \rangle) \mathcal{R}[\tau]_{(W \times X) \times W'} (ev \circ \langle f^n \circ \pi_1, g^n \circ \pi_2 \rangle)$$

Hence, by applying (2) and the inductive hypothesis for $\tau'$, we have $h^v \mathcal{R}[\tau']_{W'} h^n$, where we define

$$h^v = (ev^\dag \circ \langle f^v \circ \pi_1, g^v \circ \pi_2 \rangle) \beta_{W', TX, X}^{(W \times W') \times \square} \circ \beta_{W, TX, W'} \circ \langle w, id_{W'} \rangle$$

$$h^n = (ev \circ \langle f^n \circ \pi_1, g^n \circ \pi_2 \rangle) \beta_{W, TX, W'}^{(W \times W') \times \square} \circ \beta_{W, TX, W'} \circ \langle w, id_{W'} \rangle$$

with $\beta$ as in Definition 3.5. It then remains to show that $h^v$ and $h^n$ are the two sides of the required instance of $\mathcal{R}[\tau']_{W'}$, which we prove as follows. To prove we have the correct left-hand side, we use the associativity law, naturality of Kleisli extension, and the associativity law again, as follows.

$$h^v = ev^\dag \circ \langle (f^v \circ \pi_1, g^v \circ \pi_2) \rangle \beta_{W', TX, X}^{(W \times W') \times \square} \circ \beta_{W, TX, W'} \circ \langle w, id_{W'} \rangle$$

$$= ev^\dag \circ \langle seq^L \circ (f^v \times id) \beta_{W, T(V[\tau]), X}^{(W \times (W \times V[\tau]) \times \square)} \circ \beta_{W, TX, T(V[\tau])} \langle w, g^n \rangle$$

$$= ev^\dag \circ \langle seq^L \circ ((f^v)^{W \times \square} \times id_{T(V[\tau])}) \rangle \langle w, g^n \rangle$$

$$= ev^\dag \circ \langle (f^v)^{W \times \square} \circ w, g^n \rangle$$

To prove we have the correct right-hand side, we use naturality of extension, and the definition of power $T$-algebras, as follows.

$$h^n = (ev \circ \langle f^n \times id_{U_T(V[\tau])} \rangle \beta_{W, U_T(V[\tau]), TX}^{(W \times W') \times \square} \circ \beta_{W, TX, U_T(V[\tau])} \circ \langle w, g^n \rangle$$

$$= ev \circ \langle (f^n)^{W \times \square} \times id_{U_T(V[\tau])} \rangle \langle w, g^n \rangle$$

$$= ev \circ \langle (f^n)^{W \times \square} \circ w, g^n \rangle$$
(4) The **unit** and **bool** cases are immediate from monotonicity of the copairing operator \([-,-]\). For product types, it is enough to note that \(T\pi_i \circ [f^1_i, f^2_i] = [T\pi_i \circ f^1_i, T\pi_i \circ f^2_i]\)
and \(\pi_i \circ [f^1_i, f^2_i] = [\pi_i \circ f^1_i, \pi_i \circ f^2_i]\), and then apply the inductive hypothesis. For a function type \(\tau \to \tau'\), we show that \(g^\nu \mathcal{R}[\tau]_{W'} g^n\) implies
\[(ev^\dagger \circ \langle \langle f^1_i, f^2_i \rangle \circ w, g^n \rangle) \mathcal{R}_{W'} (ev \circ \langle \langle f^1_i, f^2_i \rangle \circ w, g^n \rangle)\]
where \(w : W' \to W_1 + W_2\). To do this, consider the following objects \(W'_i, W'_2\) and morphisms \(w'_1, w'_2\).
\[W'_i = W' \times i \quad w'_i = \pi_2 : W'_i \to W_i \quad (i \in \{1, 2\})\]
Property (2) implies
\[(g^\nu \circ \pi_1) \mathcal{R}_{W'_i} (g^n \circ \pi_1)\]
so from the assumption
\[f^i_\nu \mathcal{R}_{W'} \tau \to \tau' \quad f^1_i\]
we obtain \(k^i_\nu \mathcal{R}_{W'_i} k^n_i\), where
\[k^i_\nu = ev^\dagger \circ \langle \langle f^i_\nu \circ w'_i, g^n \circ \pi_1 \rangle \rangle \quad k^n_i = ev \circ \langle f^n_i \circ w'_i, g^n \circ \pi_1 \rangle\]
Coproducts are distributive (because we assume cartesian closure), so the coproduct \(W'_1 + W'_2\) exists and is isomorphic to \(W' \times (W_1 + W_2)\). We can therefore apply the inductive hypothesis for \(\tau\), and then (2), to obtain
\[(\langle [k^i_\nu, k^n_i] \circ h \rangle) \mathcal{R}_{W'} (\langle [k^n_1, k^n_2] \circ h \rangle)\]
where
\[h : W' \xrightarrow{(id_{W'}, w)} W' \times (W_1 + W_2) \xrightarrow{proj} W'_1 + W'_2\]
The result follows because
\[\langle [k^i_\nu, k^n_i] \circ h \rangle = ev^\dagger \circ \langle \langle f^i_\nu \circ w'_i, g^n \rangle \rangle \quad \langle [k^n_1, k^n_2] \circ h \rangle = ev \circ \langle [f^n_1, f^n_2] \circ w, g^n \rangle\]

The second lemma consists of some technical consequences of lax discardability, lax copability and lax centrality; we state them here for use in the proof of Theorem 4.7 below. For convenience, we render each of the inequalities in the statement of the lemma in the syntax of CBPV (we will not need the syntactic inequalities in the following, so we omit the precise statements and proof).

**Lemma 4.6.** Let \(T\) be a strong **Poset**-monad.

1. Let \(f_1 : W \to TX_1\) and \(f_2 : W \to TX_2\) be morphisms. For each \(i \in \{1, 2\}\), if \(f_j\) is lax discardable for \(j \neq i\), then
\[T\pi_i \circ \langle f_1, f_2 \rangle \sqsubseteq f_i : W \to TX_i\]
2. Let \(f_1 : W_1 \to TX_1, f_2 : W_2 \to TX_2\) and \(g : X_1 \times X_2 \to TY\) be morphisms. If \(f_1\) is lax central, then
\[(g^\uparrow_{TX_1 \times X_2} \circ f_1 \times f_2) \sqsubseteq (g^\uparrow_{TX_1 \times X_2} \circ f_1 \times f_2) : W_1 \times W_2 \to TY\]
If \(f_2\) is lax central, then
\[(g^\uparrow_{X_1 \times X_2} \circ f_1 \times f_2) \sqsubseteq (g^\uparrow_{X_1 \times X_2} \circ f_1 \times f_2) : W_1 \times W_2 \to TY\]
\[ M_2 \text{ to } x_2. M_1 \text{ to } x_1. N \preceq_{\text{ctx}} M_1 \text{ to } x_1. M_2 \text{ to } x_2. N \quad (f_i = [M_i], g = [N]) \]

(3) Let \( f : W \to TX \), \( g : X \to TY \) and \( h : X \times Y \to TZ \) be morphisms. If \( f \) is lax copyable and lax central, then

\[
(h^{\times Y} \circ (\text{id}_X, g)) \circ f \sqsubseteq (h^{\times Y} \circ (f \times \text{id}_Y)) \circ (\text{id}_W, g) \circ f : W \to TZ
\]

\[ M \text{ to } x. N \text{ to } y. N' \preceq_{\text{ctx}} (M \text{ to } x. N) \text{ to } y. M \text{ to } x. N' \quad (f = [M], g = [N], h = [N']) \]

(4) Let \( f : W \to TX \), \( g_1 : X \to TY_1 \) and \( g_2 : X \to TY_2 \) be morphisms. If \( f \) is lax copyable and lax central, then

\[
\langle\langle g_1, g_2 \rangle\rangle \circ f \sqsubseteq \langle\langle g_1 \circ f, g_2 \circ f \rangle\rangle : W \to T(Y_1 \times Y_2)
\]

Proof. (1) The following proves the statement for \( i = 1 \); the proof for \( i = 2 \) is similar.

\[
T\pi_1 \circ \langle\langle f_1, f_2 \rangle\rangle = T\pi_1 \circ T(\text{id}_{X_1} \times (\cdot)_{X_2}) \circ \langle\langle f_1, f_2 \rangle\rangle
\]
\[
= T\pi_1 \circ \langle\langle f_1, T(\cdot)_{X_2} \circ f_2 \rangle\rangle
\]
\[
\sqsubseteq T\pi_1 \circ \langle\langle f_1, \eta_{X_2} \circ f_2 \rangle\rangle
\]
\[
= T\pi_1 \circ (\eta_{X_1 \times 1})^{\square \times 1} \circ \langle\langle f_1, \cdot \rangle\rangle
\]
\[
= f_1
\]

(2) Lax centrality of \( f_1 \) implies the first inequality of (2) as follows.

\[
(g^{\square \times T X_2} \circ (\cdot)_{X_2}) \circ (f_1 \times f_2)
\]
\[
= g \circ \text{seq}_{X_1, X_2}^{\text{R}} \circ (f_1 \times \text{id}_{T X_2}) \circ (\text{id}_{W_1} \times f_2)
\]
\[
\sqsubseteq g \circ \text{seq}_{X_1, X_2}^{\text{R}} \circ (f_1 \times \text{id}_{T X_2}) \circ (\text{id}_{W_1} \times f_2)
\]
\[
= (g^{\square \times X_2})^{T X_1 \times \square} \circ (f_1 \times f_2)
\]

For the other inequality, precomposing with the isomorphisms \( \langle\pi_2, \pi_1 \rangle \) swaps the roles of \( f_1 \) and \( f_2 \), so that we can reuse the first inequality.
(3) The following proves the result.

\[(h^{X \times Y} \circ (id_X, g))^\dagger \circ f\]
\[= (h^{X \times Y} \circ (id_X \times g))^\dagger \circ T(id_X, id_X) \circ f\]  
(naturality of extension)
\[\sqsubseteq (h^{X \times Y} \circ (id_X \times g))^\dagger \circ seq_{X,Y} \circ (f, f)\]  
(lax copyability)
\[= ((h^{X \times Y} \circ (id_X \times g))^\dagger \circ T \circ X \times Y)^{W \times X} \circ (f \times g) \circ (id_W, f)\]  
(associativity, left unit)
\[\sqsubseteq ((h^{X \times Y} \circ (f \times id_Y))^\dagger \circ (id_W \times g))^\dagger \circ (id_W, f)\]  
(naturality of extension)
\[= (h^{X \times Y} \circ (f \times id_Y))^\dagger \circ (id_W \times g)^\dagger \circ (id_W, f)\]  
(associativity law)
\[= (h^{X \times Y} \circ (f \times id_Y))^\dagger \circ (id_W, g^\dagger \circ f)\]  

(4) The required inequality is equivalent, by postcomposing with the isomorphism \(T(\pi_2, \pi_1)\), to

\[(seq_{Y_2,Y_1}^R \circ (g_2, g_1))^\dagger \circ f \sqsubseteq seq_{Y_2,Y_1}^R \circ (g^\dagger \circ f, g_1^\dagger \circ f)\]

We prove this as follows, using (3) with \(h = \eta_{Y_2,Y_1}^{X \times Y_1} \circ (g_2 \times id_{Y_1})\).

\[= ((\eta^{X \times Y_1} \circ (g_2 \times id_{Y_1}))^\dagger \circ (id_X, g_1)^\dagger \circ f\]  
(naturality of extension)
\[\sqsubseteq ((\eta^{X \times Y_1} \circ (f \times id_{Y_1}))^\dagger \circ (id_W, g_1^\dagger \circ f)\]  
(Lemma 4.6(3))
\[= (\eta^{X \times Y_1} \circ (g_2^\dagger \times id_{Y_1}) \circ (f \times id_{Y_1}))^\dagger \circ (id_W, g_1^\dagger \circ f)\]  
(associativity law)
\[= seq_{Y_2,Y_1}^R \circ (g_2^\dagger \circ f, g_1^\dagger \circ f)\]  
(naturality of extension)

We are now ready to prove the main result of this section.

**Theorem 4.7.** Let \(\mathcal{M} = (\mathcal{C}, \mathcal{T})\) be a CBPV model. If every morphism \(f : X \rightarrow TY\) is lax discardable, lax copyable, and lax central, then for every expression \(\Gamma \vdash e : \tau\) we have

\[\mathcal{V}[e] \times \mathcal{N}[e]\]

**Proof.** Explicitly, we are required to show, for each expression \(e\), that for all families of morphisms

\[(f^u_i : W \rightarrow T(\mathcal{V}[\tau_i]))_i \quad (f^n_i : W \rightarrow U_T(\mathcal{N}[\tau_i]))_i\]

we have

\[f^n_i \mathcal{R}[\tau_i]_W f^u_i \wedge \cdots \wedge f^n_k \mathcal{R}[\tau_k]_W f^u_k \Rightarrow (\mathcal{V}[e]^\dagger \circ (f^u_i)_i) \mathcal{R}[\tau']_W (\mathcal{N}[e] \circ (f^n_i)_i)\]

We prove this by induction on \(e\).
For a variable $x_j$, we are required to prove
\[(T \pi_j \circ \langle f_i^v \rangle_i) \mathcal{R}[\tau_j]_W f_j^v\]
A simple induction on $j$, using Lemma 4.6(1), tells us that
\[T \pi_j \circ \langle f_i^v \rangle_i \sqsubseteq f_j^v\]
The result then follows from the assumption $f_j^v \mathcal{R}[\tau_j]_W f_j^v$ via Lemma 4.5(1).

For the expression $(),$ we are required to show
\[(T \langle \nu[\eta] \rangle \circ \langle f_i^v \rangle_i) \mathcal{R}[\text{unit}]_W (\eta \circ \langle \rangle_W)\]
By definition of $\mathcal{R}[\text{unit}]$, this is the same as
\[(T \langle \nu[\eta] \rangle) \circ \langle f_i^v \rangle_i \sqsubseteq (\eta \circ \langle \rangle_W)\]
which is immediate from lax discardability.

For a pair $(e_1, e_2)$, the inductive hypothesis tells us that
\[(\nu[e_j]^\dagger \circ \langle f_i^v \rangle_i) \mathcal{R}[\tau_j^\dagger]_W (\nu[e_j] \circ \langle f_i^n \rangle_i)\]
for each $j \in \{1, 2\}$. We also have
\[T \pi_j \circ \langle \nu[e_1], \nu[e_2] \rangle^\dagger \circ \langle f_i^v \rangle_i\]
\[\sqsubseteq T \pi_j \circ \langle \langle \nu[e_1] \rangle^\dagger \circ \langle f_i^v \rangle_i \rangle_1, \langle \nu[e_2] \rangle^\dagger \circ \langle f_i^v \rangle_i \rangle_2 \rangle_1 \quad \text{(Lemma 4.6(4))}\]
\[\sqsubseteq \nu[e_j]^\dagger \circ \langle f_i^v \rangle_i \quad \text{(Lemma 4.6(1))}\]
so that Lemma 4.5 implies
\[(T \pi_j \circ \langle \nu[e_1], \nu[e_2] \rangle^\dagger \circ \langle f_i^v \rangle_i) \mathcal{R}[\tau_j^\dagger \times \tau_j^\dagger]_W (\nu[\tau_j^\dagger]_W \circ \langle f_i^n \rangle_i)\]
Hence
\[(\langle \nu[e_1], \nu[e_2] \rangle^\dagger \circ \langle f_i^v \rangle_i) \mathcal{R}[\tau_1^\dagger \times \tau_2^\dagger]_W (\langle \nu[e_1], \nu[e_2] \rangle \circ \langle f_i^n \rangle_i)\]
as required.

For $\text{fst} e$, we need to show
\[(T \pi_1 \circ \langle \nu[e] \rangle^\dagger \circ \langle f_i^v \rangle_i) \mathcal{R}[\tau_1^\dagger]_W (\pi_1 \circ \langle \nu[e] \rangle \circ \langle f_i^n \rangle_i)\]
which is immediate from the inductive hypothesis and the definition of $\mathcal{R}[\tau_1^\dagger \times \tau_2^\dagger]$.

The $\text{snd}$ case is similar to the $\text{fst}$ case.

For the expression $\text{true}$, we need to show
\[(T (\text{inl} \circ \langle \nu[\eta] \rangle) \circ \langle f_i^v \rangle_i) \mathcal{R}[\text{bool}]_W (\eta_2 \circ \text{inl} \circ \langle \rangle_W)\]
which means
\[(T (\text{inl} \circ \langle \nu[\eta] \rangle) \circ \langle f_i^v \rangle_i) \sqsubseteq (\eta_2 \circ \text{inl} \circ \langle \rangle_W)\]
This follows from lax discardability, and naturality of $\eta$.

The $\text{false}$ case is similar to the $\text{true}$ case.

For $\text{if } e_0 \text{ then } e_1 \text{ else } e_2$, the inductive hypothesis gives us
\[g_0^v \sqsubseteq g_0^v \mathcal{R}[\tau']_W g_1^v \quad g_0^v \mathcal{R}[\tau']_W g_2^v\]
where we define
\[g_j^v = \nu[e_j]^\dagger \circ \langle f_i^v \rangle_i \quad g_j^n = \nu[e_j] \circ \langle f_i^n \rangle_i \quad (j \in \{0, 1, 2\})\]
By applying all of the closure properties of Lemma 4.5, we therefore have
\[(([g_1^v, g_2^v] \circ \text{dist})^{W \times \square} \circ (\text{id}_W \times g_0^v) \circ (\text{id}, \text{id})) \mathcal{R}[\tau']_W ([g_1^v, g_2^v] \circ \text{dist})^{W \times \square} \circ (\text{id}_W \times g_0^v) \circ (\text{id}, \text{id}))\]
This is not quite what we need, but it does imply the result via another use of Lemma 4.5(1), as follows. We have

\[
([\mathcal{V}[e_1], \mathcal{V}[e_2]] \circ \text{dist})^{\Box \times 2} = [\mathcal{V}[e_1]] \circ \text{dist}
\]

by precomposing with \(\text{dist}^{-1}\) and using the universal property of the coproduct \(T(\mathcal{V}[\Gamma]) + T(\mathcal{V}[\Gamma])\). It follows that

\[
\mathcal{V}[\text{if } e_0 \text{ then } e_1 \text{ else } e_2] \circ \langle f_i^n \rangle
\]

\[
= ([\mathcal{V}[e_1], \mathcal{V}[e_2]] \circ \text{dist})^{\Box \times 2} \circ (\text{id, } \mathcal{V}[e_0])) \circ \langle f_i^n \rangle_i
\]

\[
\subseteq ([\mathcal{V}[e_1], \mathcal{V}[e_2]] \circ \text{dist})^{\Box \times 2} \circ (\langle f_i^n \rangle_i \times \text{id}_2))^{\Box \times 2} \circ (\text{id}_W, g_0^n) \quad \text{(Lemma 4.6(3))}
\]

\[
= ([\mathcal{V}[e_1] \uparrow, \mathcal{V}[e_2] \uparrow] \circ \text{dist} \circ (\langle f_i^n \rangle_i \times \text{id}_2))^{\Box \times 2} \circ (\text{id}_W, g_0^n)
\]

\[
= ([g_1^n, g_2^n] \circ \text{dist})^{\Box \times 2} \circ (\text{id}_W \times g_0^n) \circ (\text{id}, \text{id})
\]

We also have

\[
\mathcal{N}[\text{if } e_0 \text{ then } e_1 \text{ else } e_2] \circ (f_i^n)_i = ([g_1^n, g_2^n] \circ \text{dist})^{\Box \times 2} \circ (\text{id}_W \times g_0^n) \circ (\text{id}, \text{id})
\]

so that Lemma 4.5(1) implies

\[
(\mathcal{V}[\text{if } e_0 \text{ then } e_1 \text{ else } e_2] \uparrow) \circ \langle f_i^n \rangle_i \circ (\mathcal{N}[\text{if } e_0 \text{ then } e_1 \text{ else } e_2] \circ (f_i^n)_i)
\]

as required.

- For a lambda abstraction \(\lambda x : \tau. e\) of type \(\tau \rightarrow \tau'\), consider arbitrary \(w : W' \rightarrow W\) and \(g^n, g^n\) such that

\[
g^n \rightarrow [\tau]_{W'} g^n
\]

We need to show

\[
(\text{ev} \circ ([\mathcal{V}[\lambda x : \tau. e] \uparrow] \circ \langle f_i^n \rangle_i \circ w), g^n)) \circ \mathcal{R}[\tau']_{W'} (\text{ev} \circ (\mathcal{N}[\lambda x : \tau. e] \circ (f_i^n)_i \circ w, g^n))
\]

Applying Lemma 4.5(3) to the assumptions yields

\[
(f_i^n \circ w) \circ \mathcal{R}[\tau]_{W'} (f_i^n \circ w)
\]

so that, by the inductive hypothesis, we have

\[
(\mathcal{V}[e] \uparrow \circ \langle f_i^n \circ w \rangle_i, g^n)) \circ \mathcal{R}[\tau']_{W'} (\mathcal{N}[e] \circ \langle f_i^n \circ w \rangle_i, g^n))
\]

We show that this is the required instance of \(\mathcal{R}[\tau']_{W'}\) by rewriting both sides as follows. For the left-hand side, we have

\[
\text{ev} \circ ([\mathcal{V}[\lambda x : \tau. e] \uparrow] \circ \langle f_i^n \rangle_i \circ w, g^n))
\]

\[
= \text{ev} \circ \langle T(\Lambda(\mathcal{V}[e])) \circ (f_i^n \circ w)_i, g^n\rangle
\]

\[
= (\text{ev} \circ (\Lambda(\mathcal{V}[e]) \times \text{id}_{\mathcal{V}[\tau]}))^\dagger \circ \langle (f_i^n \circ w)_i, g^n\rangle \quad \text{(naturality of extension)}
\]

\[
= \mathcal{V}[e] \uparrow \circ \langle f_i^n \circ w \rangle_i, g^n\rangle
\]

For the right-hand side:

\[
\text{ev} \circ (\mathcal{N}[\lambda x : \tau. e] \circ (f_i^n)_i \circ w, g^n) = \text{ev} \circ (\Lambda(\mathcal{N}[e]) \circ (f_i^n)_i \circ w, g^n)
\]

\[
= \mathcal{N}[e] \circ \langle f_i^n \circ w \rangle_i, g^n\rangle
\]
• For an application \( e e' \), where \( e \) has type \( \tau \rightarrow \tau' \), the inductive hypothesis for \( e' \) gives us

\[
(V[e']^\dagger \circ \langle f_i^n \rangle_i) \circ (N[e'] \circ \langle f_i^n \rangle_i)
\]

so that, by the inductive hypothesis for \( e \) with \( w = id_W \), we have

\[
(ev^\dagger \circ \langle V[e']^\dagger \circ \langle f_i^n \rangle_i \rangle, (V[e']^\dagger \circ \langle f_i^n \rangle_i)) \circ (N[e] \circ \langle f_i^n \rangle_i))
\]

We rewrite both sides as follows. For the left we have

\[
V[e']^\dagger \circ \langle f_i^n \rangle_i = (ev^{TV[\tau \rightarrow \tau'] \times V[\tau]}) \circ (V[e] \circ \langle f_i^n \rangle_i) \quad \text{(naturality of extension)}
\]

\[
= ev^\dagger \circ \langle V[e], V[e'] \rangle \circ \langle f_i^n \rangle_i \quad \text{(left unit, associativity)}
\]

\[
\subseteq ev^\dagger \circ \langle (V[e']^\dagger \circ \langle f_i^n \rangle_i), (V[e']^\dagger \circ \langle f_i^n \rangle_i) \rangle
\]

and for the right,

\[
N[e] \circ \langle f_i^n \rangle_i = ev \circ (N[e] \circ \langle f_i^n \rangle_i).
\]

We therefore have

\[
(V[e e']^\dagger \circ \langle f_i^n \rangle_i) \circ (N[e e'] \circ \langle f_i^n \rangle_i)
\]

as required.

As a corollary, we can directly compare the call-by-value and call-by-name translations of source-language programs (closed expressions of type \( \text{bool} \)).

**Corollary 4.8.** Let \( M = (C, T) \) be a CBPV model that is adequate with respect to a program relation \( \ll \). If every morphism \( f : X \rightarrow TY \) is lax discardable, lax copyable, and lax central, then for every closed expression \( e : \text{bool} \), we have

\[
V[e] \ll N[e].
\]

**Proof.** By Theorem 4.7 we have \( V[e] \propto N[e] \), so in particular \( V[e] \circ \text{bool} \subseteq N[e] \). By definition, the latter means \( V[e] \subseteq N[e] \), which implies the result by adequacy.

The conclusion of this corollary, namely \( V[e] \ll N[e] \) is independent of the choice of model \( M \). In contrast, the conclusion of Theorem 4.7 is not independent of \( M \). Theorem 4.7 should therefore be viewed as a result about the denotations \( V[e] \) and \( N[e] \), rather than about the translations \( V[e] \) and \( N[e] \). We rectify this in Section 6 below, where the conclusion of our main result Theorem 6.2 relates \( V[e] \) with \( N[e] \) via the contextual preorder, which is independent of \( M \).

### 4.1. Examples

To conclude this section, we discuss the consequences of the results above for each of our examples.

We first note that we can in fact simplify the definition of \( \prec \) for each of these examples, by using the fact that, in each of the three Poset-categories \( \text{Set}, \text{Poset}, \omega\text{Cpo} \), morphisms are in particular functions, and are ordered pointwise. (We treat a set as a discrete poset here.) It follows that instead of considering generalized elements \( f : W \rightarrow X \), it is enough to consider ordinary elements \( t \in X \) (which we can identify with morphisms \( t : 1 \rightarrow X \)). The simplification of \( R[-] \) we obtain is defined as follows.
The precise relationship between $\mathcal{R}[-]$ and $\mathcal{R}'[-]$ is as follows.

**Lemma 4.9.** For each of our four example models, we have 

$$f^v \mathcal{R}[\tau]_W f^n \iff \forall w \in W. f^v w \mathcal{R}'[\tau] f^n w$$

for every $f^v : W \rightarrow T(\mathcal{V}[\tau])$ and $f^n : W \rightarrow U_\Gamma(\mathcal{N}[\tau])$. Hence, for each $g^v : \mathcal{V}[\Gamma] \rightarrow T(\mathcal{V}[\tau])$ and $g^n : \mathcal{N}[\Gamma] \rightarrow U_\Gamma(\mathcal{N}[\tau])$, we have 

$$g^v \prec g^n$$

exactly when, for all $(a^v_i \in T(\mathcal{V}[\tau_i]))_i$ and $(a^n_i \in U_\Gamma(\mathcal{N}[\tau_i]))_i$,

$$a^v_i \mathcal{R}'[\tau_i] a^n_i \land \cdots \land a^n_k \mathcal{R}'[\tau_k] a^n_k \Rightarrow (g^v)^{\langle a_i^v \rangle}_i) \mathcal{R}'[\tau'] (g^n)^{\langle a_i^n \rangle}_i$$

where $\Gamma = x_1 : \tau_1, \ldots, x_k : \tau_k$.

**Proof.** By induction on $\tau$. This is trivial for $\text{unit}$, $\text{bool}$ and product types. For a function type $\tau \rightarrow \tau'$, the ($\Leftarrow$) direction is again trivial. The ($\Rightarrow$) direction follows from the fact that, identifying an element $w \in W$ with a morphism $w : 1 \rightarrow W$, we have $\mathcal{R}'[\tau] = \mathcal{R}[\tau]_1$ and $\mathcal{R}'[\tau'] = \mathcal{R}[\tau']_1$ by the inductive hypothesis.

We now consider each of our examples in turn. Note that, since the proof of Theorem 4.7 is by induction on the expression $e$, we need to extend the proof with cases for the extra syntax we add in these examples.

**Corollary 4.10.** For our no effects example, we have

$$\mathcal{V}[e] \times \mathcal{N}[e]$$

for every $\Gamma \vdash e : \tau$, and in particular,

$$\exists V : \text{bool}. (\mathcal{V}[e'] \downarrow \text{return } V) \land (\mathcal{N}[e'] \downarrow \text{return } V)$$

for every closed expression $e' : \text{bool}$.

**Proof.** This is immediate from Theorem 4.7 and Corollary 4.8, with $(\approx) = (\approx_{\text{pure}})$.

**Corollary 4.11.** For our divergence example, we have

$$\mathcal{V}[e] \times \mathcal{N}[e]$$

for every $\Gamma \vdash e : \tau$, and in particular,

$$\forall V : \text{bool}. (\mathcal{V}[e'] \downarrow \text{return } V) \Rightarrow (\mathcal{N}[e'] \downarrow \text{return } V)$$

for every closed expression $e' : \text{bool}$.
Proof. We first extend the inductive proof of Theorem 4.7 with a case for recursive functions \( \text{rec} f : \tau \rightarrow \tau', \lambda x. e \). In light of Lemma 4.9 above, it suffices for this to show that, if 
\[
\begin{align*}
  h^v : V[\tau \rightarrow \tau'] \times V[\tau] \rightarrow T(V[\tau]) & \quad h^n : U_T(N[\tau \rightarrow \tau']) \times U_T(N[\tau]) \rightarrow U_T(N[\tau])
\end{align*}
\]
are \( \omega \)-continuous functions that satisfy 
\[
\begin{align*}
  t^v \mathcal{R}'[\tau \rightarrow \tau'] t^n \wedge a^v \mathcal{R}'[\tau] a^n & \Rightarrow h^v(t^v, a^v) \mathcal{R}' h^n(t^n, a^n)
\end{align*}
\]
when \( t^v, a^v \neq \bot \), then we have 
\[
\begin{align*}
  \text{fix} \left( \lambda g^v. h^v(g^v, \cdot) \right) \mathcal{R}' \text{fix} \left( \lambda g^n. h^n(g^n, \cdot) \right)
\end{align*}
\]
This follows from the fact that each \( \mathcal{R}'[\tau'''] \) relates \( \bot \) to \( \bot \) and is closed under least upper bounds of \( \omega \)-chains, which can be proved by a simple induction on \( \tau''' \).

The second part follows from Corollary 4.8, by adequacy of the model with respect to the program relation defined in Example 2.3.

**Corollary 4.12.** For our nondeterminism example, we have 
\[
\begin{align*}
  V[e] \times N[e]
\end{align*}
\]
for every \( \Gamma \vdash e : \tau \), and in particular, 
\[
\begin{align*}
  \forall V : \text{bool}. \ (V[e'] \downarrow \text{return } V) & \Rightarrow (N[e'] \downarrow \text{return } V)
\end{align*}
\]
for every closed expression \( e' : \text{bool} \).

**Proof.** We first extend the inductive proof of Theorem 4.7 with two extra cases: one for \textit{fail} and one for \textit{or}. Following Lemma 4.9 above, to prove both of these cases, it suffices to show that \( \mathcal{R}'[\tau] \) is closed under finite joins for each \( \tau' \), i.e. that 
\[
\begin{align*}
  \bot \mathcal{R}'[\tau'] \bot t^v_1 \mathcal{R}'[\tau] t^n_1 \wedge t^v_2 \mathcal{R}'[\tau] t^n_2 & \Rightarrow (t^v_1 \sqcup t^v_2) \mathcal{R}' (t^n_1 \sqcup t^n_2)
\end{align*}
\]
This is a simple induction on \( \tau' \). The result follows from Theorem 4.7 and Corollary 4.8, by adequacy of the model with respect to the program relation defined in Example 2.4.

**Corollary 4.13.** For our immutable state example, we have 
\[
\begin{align*}
  V[e] \times N[e]
\end{align*}
\]
for every \( \Gamma \vdash e : \tau \), and in particular, 
\[
\begin{align*}
  \forall b \in \{\text{true, false}\}. \exists V : \text{bool}. \ (V[e'] \downarrow b \text{ return } V) \wedge (N[e'] \downarrow b \text{ return } V)
\end{align*}
\]
for every closed expression \( e' : \text{bool} \).

**Proof.** Once again, we need to add the extra case to Theorem 4.7. It is enough to show that 
\[
\begin{align*}
  (\text{get} \circ () \downarrow (f^v_1)) \mathcal{R}'[\text{bool}] (\text{get} \circ () \downarrow w),
\end{align*}
\]
which follows from lax discardability. We can then apply Theorem 4.7 and Corollary 4.8 to obtain the result, using adequacy with respect to the program relation defined in Example 2.5.
5. A Galois connection between call-by-value and call-by-name

We improve on the results of the previous section by showing how to directly relate the call-by-value semantics $\mathcal{V}[e]$ of an expression to a morphism derived from the call-by-name semantics $\mathcal{N}[e]$. Under a further condition on the model (which again restricts the allowable effects), we prove a statement (Theorem 5.10) of the form

$$\mathcal{V}[e] \subseteq \psi_\tau \circ \mathcal{N}[e] \circ \hat{\phi}_\Gamma$$

Involving morphisms $\phi_\tau$ and $\psi_\tau$ for mapping between the call-by-value and call-by-name semantics:

$$T(\mathcal{V}[\tau]) \xleftarrow{\phi_\tau} U_\tau(\mathcal{N}[\tau]) \xrightarrow{\psi_\tau}$$

(In the inequality above, $\hat{\phi}_\Gamma : \mathcal{V}[\Gamma] \to \mathcal{N}[\Gamma]$ is constructed by extending the morphisms $(\phi_\tau \circ \eta) : \mathcal{V}[\tau] \to U_\tau(\mathcal{N}[\tau])$ from types to contexts.)

We do not want just any maps between call-by-value and call-by-name. We show (Corollary 5.9) that the maps we define (precisely, the monotone functions $\circ (-)$ and $\psi_\tau \circ (-)$) form Galois connections [MSS86]. This is the crucial property that allows us to prove the inequality above.\(^5\)

**Definition 5.1.** A Galois connection consists of two posets $X$, $Y$ and two monotone functions $\phi : X \to Y$, $\psi : Y \to X$, such that $x \sqsubseteq \psi(\phi x)$ for all $x \in X$ and $\phi(\psi y) \sqsubseteq y$ for all $y \in Y$.

The results of the previous section are helpful here. We have relations $\mathcal{R}[\tau]$ that in some sense capture the relationship between call-by-value and call-by-name. This suggests we should look for morphisms $\phi_\tau$ and $\psi_\tau$ that represent the relations $\mathcal{R}[\tau]$, i.e. that satisfy the following equivalences.

$$\phi_\tau \circ f^v \sqsubseteq f^n \iff f^v \mathcal{R}[\tau]f^n \iff f^v \sqsubseteq \psi_\tau \circ f^n$$

These equivalences uniquely determine $\psi_\tau$ and $\phi_\tau$, and guarantee that we have Galois connections $(\psi_\tau \circ (-), \phi_\tau \circ (-))$. Furthermore, these equivalences enable us to prove

$$\mathcal{V}[e] \subseteq \psi_\tau \circ \mathcal{N}[e] \circ \hat{\phi}_\Gamma$$

as a corollary of $\mathcal{V}[e] \times \mathcal{N}[e]$. That is, the main result of this section (Theorem 5.10) is a corollary of the main result of the previous section (Theorem 4.7).

Given a CBPV model $\mathcal{M} = (\mathcal{C}, \mathcal{T})$, the morphisms $\phi_\tau$ and $\psi_\tau$ are defined in Figure 7.\(^6\)

As for the relations $\mathcal{R}[\tau]$, the definition is by induction on $\tau$. This is mutual induction: at contravariant positions, the definition of $\phi$ uses $\psi$, and vice-versa.

Of course we do not expect to be able to prove the properties outlined above for a general model $\mathcal{M}$. To see what conditions we should require $\mathcal{M}$ to satisfy, suppose that we do have Galois connections $(\phi_\tau \circ (-), \psi_\tau \circ (-))$, equivalently, that we have

$$\phi_\tau \circ \psi_\tau \sqsubseteq id_U(\mathcal{N}[\tau]) \quad id_T(\mathcal{V}[\tau]) \sqsubseteq \psi_\tau \circ \phi_\tau$$

\(^5\)The proof of Theorem 5.10 that we give here does not directly use the fact that the maps are Galois connections; instead, it uses Theorem 4.7. This is simply to avoid another induction on expressions. In the conference version of this paper, the corresponding fact [MM22, Lemma 20] was proved directly using the fact that the maps are Galois connections.

\(^6\)We present the definitions in a different way to [MM22], but the morphisms are in fact the same. The syntactic maps in Figure 8 below are similarly presented differently to [MM22].
The round-trip from call-by-value to call-by-name and back thunks the computational effects.
These suggest that lax discardability and lax copyability will be useful, and indeed we use
them,
thus
moving
from
left
to
right
discards
effects.
Similarly,
converting
a
strict
pair
\(M\) of type
\(F \langle V \{\text{unit} \times \text{unit}\} \rangle = F \langle \text{unit} \times \text{unit} \rangle\) to call-by-name and back duplicates the effects of
\(M\):

\[
[M] \subseteq (\psi_{\text{unit} \times \text{unit}} \circ \phi_{\text{unit} \times \text{unit}} \circ \psi_{\text{unit} \times \text{unit}})[N] \subseteq [N]
\]

These suggest that lax discardability and lax copyability will be useful, and indeed we use
both of these properties in the proof of Lemma 5.8 below.

For function types we need even more. Consider what happens when we convert a
CBPV computation \(M : F \langle V \{\text{unit} \rightarrow \text{unit}\} \rangle = F \langle U \{\text{unit} \rightarrow \text{Funit} \} \rangle\) to call-by-name and then back to call-by-value. By doing this we obtain the denotation of a computation that
immediately returns:

\[
[M] \subseteq (\psi_{\text{unit} \rightarrow \text{unit}} \circ \phi_{\text{unit} \rightarrow \text{unit}} \circ \psi_{\text{unit} \rightarrow \text{unit}})[M] = [\text{return thunk } \lambda x : \text{unit}. \ M \to z . \ x \ \text{force } z]
\]

The round-trip from call-by-value to call-by-name and back thunks the computational effects
of \(M\), suspending them until the function is applied. The property we ask for the model to
satisfy in order to make this a valid inequality is \textit{lax thunkability} of morphisms.

\[
\phi_{\text{unit}} = \text{id}_1 : T1 \rightarrow T1
\]
\[
\phi_{\text{bool} \times T_2} = \langle \phi_{\text{bool}} \circ T_1, \phi_{T_2} \circ T_1 \rangle : T(\langle V_1 \times V_2 \rangle) \rightarrow U_T(\langle N_1 \rangle) \times U_T(\langle N_2 \rangle)
\]
\[
\phi_\text{bool} = \text{id}_2 : T2 \rightarrow T2
\]
\[
\psi_{\text{unit}} = \text{id}_{T1} : T1 \rightarrow T1
\]
\[
\psi_{\text{bool} \times T_2} = \langle \psi_{\text{bool}} \circ T_1, \psi_{T_2} \circ T_1 \rangle : U_T(\langle N_1 \rangle) \times U_T(\langle N_2 \rangle) \rightarrow T(\langle V_1 \times V_2 \rangle)
\]
\[
\psi_\text{bool} = \text{id}_{T2} : T2 \rightarrow T2
\]
\[
\psi_{\text{unit} \times T_2} = \langle \psi_{\text{unit}} \circ T_1, \psi_{T_2} \circ T_1 \rangle : U_T(\langle N_1 \rangle) \times U_T(\langle N_2 \rangle) \rightarrow T(\langle V_1 \times V_2 \rangle)
\]
\[
\psi_\text{unit} = \text{id}_{T1} : T1 \rightarrow T1
\]
\[
\psi_{\text{bool} \times T_2} = \langle \psi_{\text{bool}} \circ T_1, \psi_{T_2} \circ T_1 \rangle : U_T(\langle N_1 \rangle) \times U_T(\langle N_2 \rangle) \rightarrow T(\langle V_1 \times V_2 \rangle)
\]
\[
\psi_\text{bool} = \text{id}_{T2} : T2 \rightarrow T2
\]
\[
\psi_{\text{unit} \times T_2} = \langle \psi_{\text{unit}} \circ T_1, \psi_{T_2} \circ T_1 \rangle : U_T(\langle N_1 \rangle) \times U_T(\langle N_2 \rangle) \rightarrow T(\langle V_1 \times V_2 \rangle)
\]
\[
\psi_\text{unit} = \text{id}_{T1} : T1 \rightarrow T1
\]
\[
\psi_{\text{bool} \times T_2} = \langle \psi_{\text{bool}} \circ T_1, \psi_{T_2} \circ T_1 \rangle : U_T(\langle N_1 \rangle) \times U_T(\langle N_2 \rangle) \rightarrow T(\langle V_1 \times V_2 \rangle)
\]
\[
\psi_\text{bool} = \text{id}_{T2} : T2 \rightarrow T2
\]
Definition 5.2. Let $T$ be a strong Poset-monad on a cartesian Poset-category $C$. A morphism $f : X \to TY$ is lax thunkable if $T\eta_Y \circ f \sqsubseteq \eta_{TY} \circ f$. If every such $f$ is lax thunkable (equivalently, if $T\eta_Y \sqsubseteq \eta_{TY}$ for every $Y \in |C|$), then we say that $T$ is lax idempotent.\footnote{Lax idempotent Poset-monads are a special case of lax idempotent 2-monads, which are well-known, and are often called Kock-Zöberlein monads [Koc95].}

Again this a lax version of a property defined by Führmann [Füh99]. For the corresponding property in the syntax of CBPV we have the following.

Lemma 5.3. Let $\Gamma \vdash_c M : FA$ be a computation. For every adequate CBPV model, if $J^\Gamma M$ is lax thunkable, then $\text{return thunk return } x \approx_{\text{ctx}} \text{return thunk } M$.

Proof. We have

$$[M \text{ to } x. \text{return thunk return } x] = T\eta[A] \circ [M]$$

$$\sqsubseteq \eta_{T[A]} \circ [M]$$

$$= [\text{return thunk } M]$$

which implies the result by adequacy. \qed

Example 5.4. For three of our examples the strong Poset-monad $T$ is lax idempotent. For no effects, we use the identity monad, which is trivially lax idempotent because $T\eta_Y = id_Y = \eta_{TY}$. For divergence, the monad (Example 3.9) is lax idempotent because the left hand side of $T\eta_Y t \sqsubseteq \eta_{TY} t$ is $\bot$ when $t = \bot$ (intuitively, we can thunk diverging computations), and otherwise the two sides are equal. For nondeterminism the monad (Example 3.9) is lax idempotent because, since $\downarrow\{y\} \subseteq S$ for every $y \in S \in TY$, we have

$$T\eta_Y S = \downarrow\{\downarrow\{y\} \mid y \in S\} \subseteq \downarrow\{S\} = \eta_{TY} S$$

(intuitively, we can postpone nondeterministic choices).

On the other hand, the reader monad we use for immutable state (Example 3.11) is not lax idempotent. Indeed, a morphism $f : X \to 2 \Rightarrow Y$ is lax thunkable exactly when it satisfies $f x \text{ true } = f x \text{ false }$ for all $x \in X$. In particular, $[\text{get}]$ is not lax thunkable. As a consequence, we cannot apply the results of this section to this model. (In fact, Proposition 5.11 below implies that the conclusion of Theorem 5.10 is false in this case.)

This does not mean that our reasoning principle (Theorem 6.2) does not apply to immutable state, only that this model is not good enough to instantiate it. Indeed, it is known that this model of immutable state fails to be fully abstract, i.e. that it distinguishes between computations that are contextually equivalent [KKS22]. A different model, such as the identity monad on Set $\times$ Set, which is lax idempotent, may enable us to apply Theorem 6.2.

Lax thunkability is difficult to use directly in proofs, so we establish the following characterizations of lax thunkable morphisms.

Lemma 5.5. Let $T$ be a strong Poset-monad and $f : X \to TY$ be a morphism. The following are equivalent:

(1) $f$ is lax thunkable;

(2) the implication

$$g_1 \sqsubseteq (g_2 \circ (id_W \times \eta_Y))^{W \times \Box} \Rightarrow g_1 \circ (id_W \times f) \sqsubseteq g_2 \circ (id_W \times f)$$

holds for all $T$-algebras $Z$ and morphisms $g_1, g_2 : W \times TY \to Z$;
(3) the implication
\[
g_1 \sqsubseteq (g_2 \circ (\eta_Y \times id_W))^\uplus_{\square \times W} \Rightarrow g_1 \circ (f \times id_W) \sqsubseteq g_2 \circ (f \times id_W)
\]
holds for all \(T\)-algebras \(Z\) and morphisms \(g_1, g_2 : TY \times W \to Z\);

(4) the implication
\[
g_1 \sqsubseteq (g_2 \circ \eta_Y)^\uplus_{W \times \square} \Rightarrow g_1 \circ f \sqsubseteq g_2 \circ f
\]
holds for all \(T\)-algebras \(Z\) and morphisms \(g_1, g_2 : TY \to Z\).

Proof. (1) \(\Rightarrow\) (2): If \(g_1 \sqsubseteq (g_2 \circ (id_W \times \eta_Y))^\uplus_{W \times \square}\) then
\[
g_1 \circ (id_W \times f) \sqsubseteq (g_2 \circ (id_W \times \eta_Y))^\uplus_{W \times \square} \circ (id_W \times f) \quad \text{(assumption)}
\]
\[
= g_2^\uplus_{W \times \square} \circ (id_W \times T\eta_Y) \circ (id_W \times f) \quad \text{(naturality of extension)}
\]
\[
\sqsubseteq g_2^\uplus_{W \times \square} \circ (id_W \times \eta_{TY}) \circ (id_W \times f) \quad \text{(lax thunkability of } f)\]
\[
= g_2 \circ (id_W \times f) \quad \text{(left unit law)}
\]

(2) \(\Rightarrow\) (4): Specializing (2) to \(W = 1\) yields (4).

(4) \(\Rightarrow\) (1): Consider the \(T\)-algebra \(Z = F_T(TY)\) and morphisms \(g_1 = T\eta_Y \) and \(g_2 = \eta_{TY}\).
We have \(g_1 = T\eta_Y = (\eta_{TY} \circ \eta_Y)^\uplus = (g_2 \circ \eta_Y)^\uplus\) by the definition of \(T\) on morphisms, so (3) gives us the required inequality \(g_1 \circ f \sqsubseteq g_2 \circ f\).

(1) \(\Rightarrow\) (3): Similar to the proof that (1) implies (2).

(3) \(\Rightarrow\) (4): Similar to the proof that (2) implies (4).

Lax thunkability is a strong property. In particular, it implies all of the properties we assumed in the previous section. (The non-lax version of this fact is noted by Führmann in [Füh99].)

Lemma 5.6. Let \(T\) be a strong Poset-monad. If \(f : X \to TY\) is lax thunkable, then \(f\) is also lax discardable, lax copyable, and lax central.

Proof. Lax discardability: By the definition of \(T\) on morphisms, we have
\[
T\langle \rangle_Y = (\eta_1 \circ \langle \rangle_Y)^\uplus = ((\eta_1 \circ \langle \rangle_{TY}) \circ \eta_Y)^\uplus
\]
so Lemma 5.5(4) implies
\[
T\langle \rangle_Y \circ f \sqsubseteq (\eta_1 \circ \langle \rangle_{TY}) \circ f = \eta_1 \circ \langle \rangle_X
\]

Lax copyability: By the definition of \(T\) on morphisms, and the left unit law for \(T\), we have
\[
T\langle id_Y, id_Y \rangle = (\eta_{Y \times Y} \circ \langle id_Y, id_Y \rangle)^\uplus = (seq_{Y,Y}^1 \circ \langle \eta_Y, \eta_Y \rangle)^\uplus = ((seq_{Y,Y}^1 \circ \langle id_Y, id_Y \rangle) \circ \eta_Y)^\uplus
\]
so Lemma 5.5(4) implies
\[
T\langle id_Y, id_Y \rangle \circ f \sqsubseteq (seq_{Y,Y}^1 \circ \langle id_Y, id_Y \rangle) \circ f = seq_{Y,Y}^1 \circ \langle f, f \rangle
\]
Lax centrality: We have

\[ \text{seq}_Y^L = (\eta_Y \times W)^{\Box \times TW} (\text{definition of seq}^L) \]

\[ = (((\eta_Y \times W)^{\Box \times W} \circ (\eta_Y \times id_W))^Y \times \Box \times TW) (\text{left unit law}) \]

\[ = ((\eta_Y \times W)^{\Box \times W} \circ (\eta_Y \times id_{TW})) (\text{naturality of extension}) \]

\[ = (\text{seq}_Y^R \circ (\eta_Y \times id_{TW}))^{\Box \times TW} (\text{definition of seq}^R) \]

so Lemma 5.5(3) implies

\[ \text{seq}_Y^L \circ (f \times id_{TW}) \sqsubseteq \text{seq}_Y^R \circ (f \times id_{TW}) \]

Our aim is now to establish the relationship between \( R[\tau] \) and \( (\phi_\tau, \psi_\tau) \) outlined at the beginning of this section. For \( \tau = \text{unit} \) and \( \tau = \text{bool} \) this turns out to be easy. For product types, we use lax discardability and lax copyability, while for function types, we use lax thunkability. For the latter two cases the following lemma is useful.

**Lemma 5.7.** Let \( T \) be a strong Poset-monad.

1. If \( f : W \to T(X_1 \times X_2) \) is lax copyable, and both \( g_1 : W \to TX_1 \) and \( g_2 : W \to TX_2 \) are lax discardable, then

\[ f \sqsubseteq \langle \langle g_1, g_2 \rangle \rangle \iff T\pi_1 \circ f \sqsubseteq g_1 \land T\pi_2 \circ f \sqsubseteq g_2 \]

2. If \( f : W \to T(X \Rightarrow Z) \) is lax thunkable, where \( Z \) is a \( T \)-algebra, then for every \( g : W \times X \to Z \), we have

\[ f \sqsubseteq \eta_{X \Rightarrow Z} \circ \Lambda g \iff ev^{\Box \times X} \circ (f \times id_X) \sqsubseteq g \]

**Proof.** (1) For the (\( \Rightarrow \)) direction, we have

\[ T\pi_1 \circ f \sqsubseteq T\pi_1 \circ \langle \langle g_1, g_2 \rangle \rangle \quad \text{(assumption)} \]

\[ \sqsubseteq g_i \quad \text{(Lemma 4.6(1))} \]

for each \( i \in \{1, 2\} \). For the (\( \Leftarrow \)) direction, we have

\[ f = T(\pi_1 \times \pi_2) \circ T(id, id) \circ f \]

\[ \sqsubseteq T(\pi_1 \times \pi_2) \circ \text{seq}^L \circ \langle f, f \rangle \quad \text{(lax copyability)} \]

\[ = \langle \langle T\pi_1 \circ f, T\pi_2 \circ f \rangle \rangle \quad \text{(naturality of extension)} \]

\[ \sqsubseteq \langle \langle g_1, g_2 \rangle \rangle \quad \text{(assumption)} \]

(2) For the (\( \Rightarrow \)) direction, we have

\[ ev^{\Box \times X} \circ (f \times id_X) \sqsubseteq ev^{\Box \times X} \circ (\eta_{X \Rightarrow Z} \circ \Lambda g) \times id_X \quad \text{(assumption)} \]

\[ = ev \circ (\Lambda g \times id_X) \quad \text{(left unit law)} \]

\[ = g \]

For the (\( \Leftarrow \)) direction, we note that

\[ \eta_{X \Rightarrow Z} \circ \Lambda (ev^{\Box \times X}) \circ f = \eta_{X \Rightarrow Z} \circ \Lambda (ev^{\Box \times X} \circ (f \times id_X)) \]

\[ \sqsubseteq \eta_{X \Rightarrow Z} \circ \Lambda g \quad \text{(assumption)} \]
so it suffices to show \( f \sqsubseteq \eta_{X \Rightarrow Z} \circ \Lambda (ev^{\top} \circ \times X) \circ f \). Since \( f \) is lax thunkable, this is a consequence of Lemma 5.5(4) applied to the following.

\[
\begin{align*}
\text{id}_{T(X \Rightarrow Z)} &= \eta_{X \Rightarrow Z}^\dagger \\
&= (\eta_{X \Rightarrow Z} \circ \Lambda ev)^\dagger \\
&= (\eta_{X \Rightarrow Z} \circ \Lambda (ev^{\top} \circ (\eta_{X \Rightarrow Z} \times id_X)))^\dagger \\
&= (\eta_{X \Rightarrow Z} \circ \Lambda (ev^{\top} \circ X) \circ \eta_{X \Rightarrow Z})^\dagger
\end{align*}
\]

(right unit law)

We can now relate \( R[\tau] \) to the morphisms \( \phi_\tau \) and \( \psi_\tau \), as follows.

**Lemma 5.8.** Let \( \mathcal{M} = (C, T) \) be a CBPV model for which \( T \) is lax idempotent. For every type \( \tau \), object \( W \in |C| \), and pair of morphisms

\[
f^v : W \rightarrow T(\mathcal{V}[\tau]) \quad f^n : W \rightarrow U_T(\mathcal{N}[\tau])
\]

we have the following equivalences.

\[
\phi_\tau \circ f^v \sqsubseteq f^n \iff f^v \circ R[\tau]_W f^n \iff f^v \sqsubseteq \psi_\tau \circ f^n
\]

**Proof.** By induction on the type \( \tau \).

- **The unit** case is trivial.
- For a product type \( \tau_1 \times \tau_2 \), we have that \( f^v \circ R[\tau_1 \times \tau_2] f^n \) is equivalent, by expanding out the definition and applying the inductive hypothesis, to each of the following properties.

\[
\forall i \in \{1, 2\}. \phi_{\tau_i} \circ T_{\tau_i} \circ f^v \sqsubseteq \pi_i \circ f^n \quad \forall i \in \{1, 2\}. T_{\tau_i} \circ f^v \sqsubseteq \psi_{\tau_i} \circ \pi_i \circ f^n
\]

It therefore suffices to show that the left is equivalent to \( \phi_{\tau_1 \times \tau_2} \circ f^v \sqsubseteq f^n \), and that the right is equivalent to \( f^v \sqsubseteq \psi_{\tau_1 \times \tau_2} \circ f^n \). For the left, the inequality \( \phi_{\tau_1 \times \tau_2} \circ f^v \sqsubseteq f^n \) holds exactly when \( \pi_i \circ \phi_{\tau_1 \times \tau_2} \circ f^v \sqsubseteq \pi_i \circ f^n \) holds for all \( i \in \{1, 2\} \). The required equivalence therefore follows from

\[
\pi_i \circ \phi_{\tau_1 \times \tau_2} \circ f^v = T_{\tau_i} \circ \phi_{\tau_i} \circ f^v
\]

which is immediate from the definition of \( \phi_{\tau_1 \times \tau_2} \). For the right, it is enough to apply Lemma 5.7(1) to the inequality \( f^v \sqsubseteq \psi_{\tau_1 \times \tau_2} \circ f^n \). We can do this because \( T \) is lax idempotent, which implies lax discardability and lax copyability by Lemma 5.6.

- **The bool** case is trivial.
- For a function type \( \tau \rightarrow \tau' \), we consider the two equivalences separately.

For the equivalence on the left, we have

\[
\phi_{\tau \rightarrow \tau'} \circ f^v \sqsubseteq f^n \iff \Lambda^{-1}(\phi_{\tau \rightarrow \tau'} \circ f^v) \sqsubseteq \Lambda^{-1} f^n
\]

\[
\iff \phi_{\tau \rightarrow \tau'} \circ ev^\dagger \circ \langle \langle f^v \circ \pi_1, \psi_{\tau \rightarrow \tau'} \circ \pi_2 \rangle \rangle \sqsubseteq ev \circ (f^n \times id)
\]

\[
\iff (ev^\dagger \circ \langle \langle f^v \circ \pi_1, \psi_{\tau \rightarrow \tau'} \circ \pi_2 \rangle \rangle) \circ R[\tau']_W U_T(\mathcal{N}[\tau']) (ev \circ (f^n \times id))
\]

where the second step uses the definition of \( \phi_{\tau \rightarrow \tau'} \), and the final step uses the inductive hypothesis. Call the instance of \( R[\tau']_W U_T(\mathcal{N}[\tau']) \) above (*). To show that \( f^v \circ R[\tau \rightarrow \tau']_W f^n \) follows from (*), consider arbitrary \( w : W' \rightarrow W \) and \( g^v, g^n \) such that \( g^v \circ R[\tau']_W g^n \). By Lemma 4.5(2), we can precompose both sides of (*), with \( \langle w, g^n \rangle \) to obtain

\[
(ev^\dagger \circ \langle \langle f^v \circ w, \psi_{\tau \rightarrow \tau'} \circ g^n \rangle \rangle) \circ R[\tau']_W (ev \circ (f^n \circ w, g^n))
\]

The inductive hypothesis implies \( g^v \sqsubseteq \psi_{\tau \rightarrow \tau'} \circ g^n \), so by Lemma 4.5(1) it follows that

\[
(ev^\dagger \circ \langle \langle f^v \circ w, g^n \rangle \rangle) \circ R[\tau']_W (ev \circ (f^n \circ w, g^n))
\]
as required. Conversely, $f^v \mathcal{R}[\tau \to \tau']_W f^n$ implies $(\ast)$ by taking

$$W' = W \times U_T(\mathcal{N}[\tau]) \quad w = \pi_1 : W' \to W \quad g^v = \psi_\tau \circ \pi_2 \quad g^n = \pi_2$$

and noting that $\psi_\tau \circ \pi_2 \sqsubseteq \psi_\tau \circ \pi_2$ implies $(\psi_\tau \circ \pi_2) \mathcal{R}[\tau]_{W'} \pi_2$ by the inductive hypothesis.

For the remaining equivalence, we have

$$f^v \sqsubseteq \psi_{\tau \to \tau'} \circ f^n \iff ev^\Box \circ (f^v \times \text{id}) \sqsubseteq \psi_{\tau} \circ ev \circ (f^n \times (\phi_\tau \circ \eta))$$

$$\iff (ev^\Box \circ (f^v \times \text{id})) \mathcal{R}\mathcal{[\tau]}_{W \times \mathcal{V}[\tau]} (ev \circ (f^n \times (\phi_\tau \circ \eta)))$$

where the first step uses Lemma 5.7(2) and the second uses the inductive hypothesis. Call the instance of $\mathcal{R}[\tau]_{W \times \mathcal{V}[\tau]}$ above $(\ast \ast)$. To show that $(\ast \ast)$ implies $f^v \mathcal{R}[\tau \to \tau']_W f^n$, consider arbitrary $w : W' \to W$ and $g^v, g^n$ such that $g^v \mathcal{R}[\tau]_{W'} g^n$. By Lemma 4.5(2,3), we have

$$\begin{align*}
(ev^\Box \circ (f^v \times \text{id}))^{W \times \Box} \circ (w, g^v) \mathcal{R}\mathcal{[\tau]}_{W'} (ev \circ (f^n \times (\phi_\tau \circ \eta)))^{W \times \Box} \circ (w, g^n)
\end{align*}$$

Lax idempotence of $T$ implies lax centrality by Lemma 5.6, so we have

$$\begin{align*}
(ev^\Box \circ (f^v \times \text{id}))^{W \times \Box} \circ (w, g^n)
= (ev^\Box \circ (f^v \times \text{id}))^{T(\mathcal{V}[\tau] \to \mathcal{V}[\tau'])} \circ (f^v \circ w, g^n) \quad \text{(naturality of extension)}
= (ev^\Box \circ T(\mathcal{V}[\tau]))^{\Box \times \mathcal{V}[\tau]} \circ (f^v \circ w, g^n) \quad \text{(Lemma 4.6(2))}
= ev^\Box \circ \langle f^v \circ w, g^n \rangle \quad \text{(left unit, associativity)}
\end{align*}$$

Since $g^v \mathcal{R}[\tau]_{W'} g^n$, we also have

$$\begin{align*}
(ev \circ (f^n \times (\phi_\tau \circ \eta)))^{W \times \Box} \circ (w, g^v)
\subseteq (ev \circ (f^n \times \phi_\tau)) \circ (w, g^v) \quad \text{(Lemma 5.5(2))}
= ev \circ (f^n \circ w, \phi_\tau \circ g^n)
\subseteq ev \circ (f^n \circ w, g^n) \quad \text{(inductive hypothesis)}
\end{align*}$$

where we again use the fact that $T$ is lax idempotent. It follows by Lemma 4.5(1) that

$$\begin{align*}
(ev^\Box \circ \langle f^v \circ w, g^n \rangle) \mathcal{R}\mathcal{[\tau]}_{W'} (ev \circ (f^n \circ w, g^n))
\end{align*}$$

as required. Finally, to show that $f^v \mathcal{R}[\tau \to \tau']_W f^n$ implies $(\ast \ast)$, we take

$$W' = W \times \mathcal{V}[\tau] \quad w = \pi_1 \quad g^v = \eta \circ \pi_2 \quad g^n = \phi_\tau \circ \eta \circ \pi_2$$

noting that the inductive hypothesis implies $(\eta \circ \pi_2) \mathcal{R}[\tau]_{W'} (\phi_\tau \circ \eta \circ \pi_2)$. From this we obtain

$$\begin{align*}
(ev^\Box \circ \langle f^v \circ \pi_1, \eta \circ \pi_2 \rangle) \mathcal{R}\mathcal{[\tau]}_{W \times \mathcal{V}[\tau]} (ev \circ (f^n \times (\phi_\tau \circ \eta)))
\end{align*}$$

which simplifies to $(\ast \ast)$ by the left unit law.

An immediate corollary is that, as claimed above, the maps between call-by-value and call-by-name form Galois connections.

**Corollary 5.9.** Let $\mathcal{M} = (\mathcal{C}, T)$ be a CBPV model such that $T$ is lax idempotent. The monotone functions

$$\begin{align*}
\mathcal{C}(W, T(\mathcal{V}[\tau])) & \xrightarrow{\phi_\tau \circ (-)} \mathcal{C}(W, U_T(\mathcal{N}[\tau])) \\
\mathcal{C}(W, U_T(\mathcal{N}[\tau])) & \xrightarrow{\psi_\tau \circ (-)} \mathcal{C}(W, T(\mathcal{V}[\tau]))
\end{align*}$$

Vol. 20:1
form a Galois connection for every source-language type $\tau$ and object $W \in |C|$. 

**Proof.** Immediate from Lemma 5.8.

Another corollary of Lemma 5.8 is the following, which is the main result of this section. We use this result in the following section to establish our reasoning principle. To state it, let

$$ \text{Theorem 5.10.} \quad \text{Let} \quad \begin{array}{l}
\Gamma = x_1 : \tau_1, \ldots, x_k : \tau_k.
\end{array} $$

where $\Gamma = x_1 : \tau_1, \ldots, x_k : \tau_k$.

**Theorem 5.10.** Let $\mathcal{M} = (C, T)$ be a CBPV model such that $T$ is lax idempotent. For all source-language expressions $\Gamma \vdash e : \tau$ we have

$$ \mathcal{V}[e] \sqsubseteq \psi_\tau \circ \mathcal{N}[e] \circ \hat{\phi}_\Gamma $$

**Proof.** The inequality we need to establish is equivalently

$$ \mathcal{V}[e] \sqsubseteq \psi_\tau \circ \mathcal{N}[e] \circ (\phi_{\tau_i} \circ \eta_{\mathcal{V}[\tau_i]} \circ \pi_{\tau_i})_i $$

where $\Gamma = x_1 : \tau_1, \ldots, x_k : \tau_k$. By Lemma 5.8 we have

$$ (\eta_{\mathcal{V}[\tau_i]} \circ \pi_{\tau_i}) \mathcal{R}[\tau_i] \mathcal{V}[\Gamma] (\phi_{\tau_i} \circ \eta_{\mathcal{V}[\tau_i]} \circ \pi_{\tau_i}) $$

for each $i$. We invoke Theorem 4.7 (using Lemma 5.6 to show lax discardability, lax copyability and lax centrality), to obtain

$$ (\mathcal{V}[e]^\dagger \circ (\eta_{\mathcal{V}[\tau_i]} \circ \pi_{\tau_i} )_i \mathcal{R}[\tau_i] \mathcal{V}[\Gamma] (\mathcal{N}[e] \circ (\phi_{\tau_i} \circ \eta_{\mathcal{V}[\tau_i]} \circ \pi_{\tau_i} )_i ) $$

The left-hand side is equal to $\mathcal{V}[e]$ by the left unit law of $T$, so Lemma 5.8 implies the required inequality.

This theorem has a partial converse, as follows.

**Proposition 5.11.** Let $\mathcal{M} = (C, T)$ be an arbitrary CBPV model. If $\mathcal{V}[e] \sqsubseteq \psi_\tau \circ \mathcal{N}[e] \circ \hat{\phi}_\Gamma$ for every $\Gamma \vdash e : \tau$, then for each object $X \in \{1, 2\}$ we have $T\eta_X \sqsubseteq \eta_{TX}$, and every morphism $Y \to TX$ is lax thunkable.

**Proof.** The first step is to show that $id \sqsubseteq \psi_\tau \circ \phi_\tau$ for every $\tau$, by applying the assumption to the expression $x : \text{unit} \vdash x() : \tau$. Indeed, we have

$$ ev = \mathcal{V}[x()] \sqsubseteq \psi_\tau \circ \mathcal{N}[x()] \circ \hat{\phi}_{\text{unit} \to \tau} = \psi_\tau \circ \phi_\tau \circ ev $$

which implies $id \sqsubseteq \psi_\tau \circ \phi_\tau$ because $ev : 1 \times (1 \Rightarrow T(\mathcal{V}[\tau])) \to T(\mathcal{V}[\tau])$ is an isomorphism. It follows for each $\tau' \in \{\text{unit, bool}\}$ that

$$ id_{T(1 \Rightarrow T(\mathcal{V}[\tau']))} \sqsubseteq \psi_{\text{unit} \to \tau'} \circ \phi_{\text{unit} \to \tau'} = \eta_{1 \Rightarrow T(\mathcal{V}[\tau'])} \circ id_{1 \Rightarrow T(\mathcal{V}[\tau'])}^\dagger $$

which implies

$$ id_{T(\mathcal{V}[\tau'])} \sqsubseteq \eta_{T(\mathcal{V}[\tau'])} \circ id_{T(\mathcal{V}[\tau'])}^\dagger $$

Hence, by naturality of Kleisli extension and the right unit law, we have

$$ T\eta_{\mathcal{V}[\tau']} = id_{T(T(\mathcal{V}[\tau']))} \circ T\eta_{\mathcal{V}[\tau']} \sqsubseteq \eta_{T(\mathcal{V}[\tau'])} \circ id_{T(\mathcal{V}[\tau'])}^\dagger \circ T\eta_{\mathcal{V}[\tau']} = \eta_{T(\mathcal{V}[\tau'])} $$

as required.
We now use the Galois connections defined in the previous section to relate the call-by-value and call-by-name computations. These are defined, again by induction on $\tau$, with $V$ convenient to work with however.

\[ \Phi_{\text{unit}} M = M \]
\[ \Phi_{\tau_1 \times \tau_2} M = \lambda \{ 1. \text{to} \ x. \ \text{match} \ x \ \text{with} \ (z_1, z_2). \ \Phi_{\tau_1} (\text{return} z_1), \]
\[ 2. \text{to} \ x. \ \text{match} \ x \ \text{with} \ (z_1, z_2). \ \Phi_{\tau_2} (\text{return} z_2) \} \]
\[ \Phi_{\text{bool}} M = M \]
\[ \Phi_{\tau \rightarrow \tau'} M = \lambda x: U(N(\tau)). \text{to} \ f. \ \Psi_{\tau} (\text{force} \ x) \ \text{to} \ y. \ \Phi_{\tau'} (y \ \text{force} \ f) \]

\[ \Gamma \vdash \psi \ L \ M : F(V(\tau)) \mapsto \Gamma \vdash \psi \ L \ N : N(V(\tau)) \]

\[ \Phi_{\text{unit}} N = N \]
\[ \Phi_{\tau_1 \times \tau_2} N = \Psi_{\tau_1} (1 : N) \ \text{to} \ z_1. \ \Psi_{\tau_2} (2 : N) \ \text{to} \ z_2. \ \text{return} \ (z_1, z_2) \]
\[ \Psi_{\text{bool}} N = N \]
\[ \Psi_{\tau \rightarrow \tau'} N = \text{return thunk} \lambda x : V(\tau). \ \Psi_{\tau'} ((\text{thunk} (\Phi_{\tau} (\text{return} \ x))) : N) \]

\textbf{Figure 8.} Syntactic maps $\Phi$ from call-by-value to call-by-name and $\Psi$ from call-by-name to call-by-value

In particular, it follows from this proposition that lax discardability, lax copyability, and lax centrality are not enough. Our immutable state example satisfies all three of those properties, but the morphism $[\text{get}] : 1 \rightarrow T2$ is not lax thunkable, so we do not have $V[e] \subseteq \psi \ L \ N[e] \circ \hat{\phi}_{\Gamma}$ for every $e$.

\section{The reasoning principle}

We now use the Galois connections defined in the previous section to relate the call-by-value and call-by-name translations of expressions, and arrive at our main reasoning principle.

Recall that the problem with comparing $V[e]$ with $N[e]$ directly is that they have different types. We render the Galois connections defined in the previous section in the syntax of CBPV, and then construct from $N[e]$ a computation that we can directly compare with $V[e]$:

\[ V[e] \simeq_{\text{ctx}} \Psi_{\tau} (N[e] [\hat{\phi}_{\Gamma}]) \]

More precisely, we render $\phi_{\tau}$ and $\psi_{\tau}$ in the syntax as maps $\Phi_{\tau}$ from call-by-value computations to call-by-name computations, and $\Psi_{\tau}$ from call-by-name to call-by-value.\(^8\) These are defined, again by induction on $\tau$, in Figure 8. (We use some auxiliary variables in the definition, which are assumed to be fresh.) We further define, for each source-language context $\Gamma = x_1 : \tau_1, \ldots, x_k : \tau_k$, a substitution\( \hat{\phi}_{\Gamma} = x_1 \mapsto \text{thunk} (\Phi_{\tau_1} (\text{return} \ x_1)), \ldots, x_k \mapsto \text{thunk} (\Phi_{\tau_k} (\text{return} \ x_k)) \)

\(^8\)We define $\Phi$ and $\Psi$ directly as maps from computations to computations, but we could instead have defined computations

\[ x : U(F(V(\tau))) \vdash_{c} \Phi'_{\tau} : N(V(\tau)) \quad x : U(N(V(\tau))) \vdash_{c} \Psi'_{\tau} : F(V(\tau)) \]

and then recovered $\Phi$ and $\Psi$ modulo $\beta\eta$-laws for thunks, by substitution. This definition is slightly less convenient to work with however.
for converting a call-by-value context into a call-by-name context. This has the following typing:

\[ \mathcal{N}(\Gamma) \vdash_c N : \mathcal{C} \rightsquigarrow \mathcal{V}(\Gamma) \vdash_c \hat{\Phi}_\Gamma N : \mathcal{C} \]

The maps \( \Phi, \Psi \) and \( \hat{\Phi} \) are syntactic renderings of \( \phi, \psi \) and \( \hat{\phi} \) in the following sense.

**Lemma 6.1.** Given any model of CBPV, we have:

1. \( [\Phi_\tau M] = \phi_\tau \circ [M] \) for all \( \Gamma \vdash_c M : \mathcal{F}(\mathcal{V}(\tau)) \);
2. \( [\Psi_\tau N] = \psi_\tau \circ [N] \) for all \( \Gamma \vdash_c N : \mathcal{N}(\tau) \);
3. \( [\mathcal{N}(\hat{\Phi}_\Gamma)] = [N] \circ \phi_\Gamma \) for all \( \mathcal{N}(\Gamma) \vdash_c N : \mathcal{C} \).

**Proof.** (1) and (2) are proved by mutual induction on the type \( \tau \), with each case being an easy calculation. (3) then follows immediately from (1) together with the evident substitution lemma for the denotational semantics of CBPV.

Given a source-language expression \( \Gamma \vdash e : \tau \), the computation we obtain by composing \( \mathcal{N}(\ell) \) with the maps between call-by-value and call-by-name has the same type as \( \mathcal{V}(\ell) \):

\[ \mathcal{V}(\Gamma) \vdash_c \Psi_\tau(\mathcal{N}(\ell)[\hat{\Phi}_\Gamma]) : \mathcal{F}(\mathcal{V}(\tau)) \]

We can therefore compare \( \mathcal{V}(\ell) \) with \( \Psi_\tau(\mathcal{N}(\ell)[\hat{\Phi}_\Gamma]) \) directly. In particular, it makes sense to replace \( \mathcal{V}(\ell) \) with \( \Psi_\tau(\mathcal{N}(\ell)[\hat{\Phi}_\Gamma]) \) within a CBPV computation, as outlined in the introduction. Using the results of the previous section, we establish the following result for reasoning about how replacing \( \mathcal{V}(\ell) \) in this way changes the behaviour of a computation.

Recall that a program relation \( \preceq \) is a preorder on CBPV programs, and that each program relation induces a contextual preorder \( \preceq_{\text{ctx}} \). Given any program relation \( \preceq \), to show that the call-by-value and call-by-name translations of source-language expressions are related by \( \preceq_{\text{ctx}} \), it is enough to find an adequate model involving a lax idempotent \( T \):

**Theorem 6.2** (Relationship between call-by-value and call-by-name). Let \( \mathcal{M} = (\mathcal{C}, T) \) be a CBPV model that is adequate with respect to a given program relation \( \preceq \). If \( T \) is lax idempotent, then for every source-language expression \( \Gamma \vdash e : \tau \) we have

\[ \mathcal{V}(\ell) \preceq_{\text{ctx}} \Psi_\tau(\mathcal{N}(\ell)[\hat{\Phi}_\Gamma]) \]

**Proof.** By adequacy it suffices to show \( [\mathcal{V}(\ell)] \subseteq [\Psi_\tau(\mathcal{N}(\ell)[\hat{\Phi}_\Gamma])] \), which, by Lemma 6.1, is equivalent to \( \Psi_\tau(\mathcal{N}(\ell)[\hat{\Phi}_\Gamma]) \). The result therefore follows from Theorem 5.10.

The generality of this theorem comes from two sources. First, we consider arbitrary program relations \( \preceq \). The only requirement on these is the existence of some adequate model in which morphisms are lax thunkable. Second, this theorem applies to terms that are open and have higher types, using the maps between the two evaluation orders (in contrast to Corollary 4.8 above).

For our first three examples (no effects, divergence, and nondeterminism), the model is adequate and has a lax idempotent \( T \). Thus in each case the assumptions of our reasoning principle are satisfied, establishing the claims stated informally at the beginning of the introduction.

**Remark 6.3.** Given an adequate model in which \( T \) is lax idempotent, it follows from Corollary 5.9 and Lemma 6.1 that the maps \( \Phi_\tau \) and \( \Psi_\tau \) on terms form a Galois connection (with respect to \( \preceq_{\text{ctx}} \)). In particular, we have

\[ M \preceq_{\text{ctx}} \Psi_\tau(\Phi_\tau M) \quad \Phi_\tau(\Psi_\tau N) \preceq_{\text{ctx}} N \]
Both of these inequalities are in general proper (they are not contextual equivalences). To see this, consider our divergence example, for which the above inequalities hold. For each $C$, let $\Omega_C$ be the diverging computation $\text{rec } x : U_C. \text{force } x$ (which has type $C$). Then if $\tau = \text{bool} \rightarrow \text{bool}$ and $M = \Omega_{F(\nu(\tau))}$, we do not have $M \succctx \Psi_{\tau}(\Phi_{\tau} M)$, because for $E = (\Box \text{ to } \text{return false})$ the computation $E[M]$ diverges but $E[\Psi_{\tau}(\Phi_{\tau} M)] \downarrow \text{return false}$. In this case we have $\Psi_{\tau}(\Phi_{\tau} M) \simeq_{\text{ctx}} \text{return thunk } \lambda x : \text{bool}. \Omega_{F\text{bool}}$. For a counterexample to $\Phi_{\tau}(\Psi_{\tau} N) \simeq_{\text{ctx}} N$, let $\tau = \text{bool} \rightarrow \text{bool}$ and $N = \lambda x : U(\text{F bool}). \text{return true}$. Then for $E' = ((\text{thunk } \Omega_{F\text{bool}}) \Box)$, the computation $E'[\Phi_{\tau}(\Psi_{\tau} N)]$ diverges but $E'[N] \downarrow \text{return true}$. Here we have $\Phi_{\tau}(\Psi_{\tau} N) \simeq_{\text{ctx}} \lambda x : U(\text{F bool}). \text{force } x \text{ to } y. \text{return true}.$

In particular, our maps between call-by-value and call-by-name are merely Galois connections, and not sections or retractions. This contrasts with Reynolds [Rey74], who obtains a retraction between direct and continuation semantics.

7. Related work

Comparing evaluation orders. Plotkin [Plo75] and many others (e.g. [IT16]) relate call-by-value and call-by-name. Crucially, they consider $\lambda$-calculi with no effects other than divergence. This makes a significant difference to the techniques that can be used, in particular because in this case the equational theory for call-by-name is strictly weaker than for call-by-value. This is not necessarily true for other effects. Other evaluation orders (such as call-by-need) have also been compared in similarly restricted settings [MOTW95, MM19, HH19]. We suspect our technique could also be adapted to these. Here we use CBPV as a calculus in which to reason about both call-by-value and call-by-name, but other calculi (e.g. the modal calculus of [ESPU22]) may be suitable for this purpose.

It might also be possible to recast some of our work in terms of the duality between call-by-value and call-by-name [Fil89, CH00, Wad03, Sel01], In particular, this may shed some light on our definitions of $\Phi$ and $\Psi$. It is not clear to us what the precise connection is however. While Selinger [Sel01] defines translations between call-by-value and call-by-name versions of Parigot’s $\lambda\mu$-calculus [Par92], these translations behave differently to ours, in particular, they are semantics-preserving.

Relating semantics of languages. The technique we use here to relate call-by-value and call-by-name is based on the idea used first by Reynolds [Rey74] to relate direct and continuation semantics of the $\lambda$-calculus, and later used by others (e.g. [MW85, Kuč98, CF94, Fil96]). Reynolds constructs a relation between the two semantics, and uses this to establish a retraction between direct and continuation semantics, just as we construct a relation between call-by-value and call-by-name and then use this to establish a Galois connection. A minor difference is that Reynolds relies on continuations with a large-enough domain of answers (e.g. a solution to a particular recursive domain equation). Our maps exist for any choice of model. We are the first to use this technique to relate call-by-value and call-by-name. There has been some work [SF92, LD93, SW96] on soundness and completeness properties of translations (similar to the translations into CBPV), in particular using Galois connections (and similar structures) for which the order is reduction of programs. Our results would fail if we used reduction of programs directly, so we consider only the observable behaviour of programs.
There are some similarities between our work and the work of New et al. [NL20, NLA21] on gradual typing. In particular, [NLA21] has embedding-projection pairs (a special case of Galois connections) for casting from a more dynamic type to a less dynamic type, and vice versa. Their application is quite different however. The double category perspective used in [NL20] may also be illuminating here.

8. Conclusions

In this paper, we give a general reasoning principle (Theorem 6.2) that relates the observable behaviour of terms under call-by-value and call-by-name. The reasoning principle works for various collections of computational effects, in particular, it enables us to obtain theorems about divergence and nondeterminism. It is about open expressions, and enables us to change evaluation order within programs.

The technique we use involves first relating the observable behaviour of the call-by-value and call-by-name translations of expressions via a logical relation (Theorem 4.7). We obtain a result about call-by-value and call-by-name evaluations of programs as a corollary (Corollary 4.8). Applying this to divergence, we show that if the call-by-value execution terminates with some result then the call-by-name execution terminates with the same result. For nondeterminism, we show that all possible results of call-by-value executions are possible results of call-by-name executions. There may be other collections of effects we can apply our technique to, including combinations of divergence and nondeterminism.

We expect that our technique can be applied to other evaluation orders. Two evaluation orders can be related by giving translations into some common language (here we use CBPV), constructing maps between the two translations, and showing that (for some models) these maps form Galois connections. A major advantage of the technique is that it allows us to identify axiomatic properties of computational effects (thunkable, etc.) that give rise to relationships between evaluation orders.

Acknowledgments

We thank the anonymous referees for helpful comments. The first author was supported by an EPSRC studentship, and by Icelandic Research Fund grants 196323-053 and 228684-052.

References


