Galois connecting call-by-value and call-by-name

Dylan McDermott
Reykjavik University, Iceland

Alan Mycroft
University of Cambridge, UK

Abstract

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

In languages with side-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 side-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 these properties. The key ingredient is Levy's call-by-push-value calculus, which we use as a framework for reasoning about evaluation orders. We construct maps between the call-by-value and call-by-name interpretations of types. We then identify properties of side-effects that imply these maps form a Galois connection. These properties hold for some side-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 side-effects including divergence and nondeterminism.

2012 ACM Subject Classification Theory of computation → Semantics and reasoning; Theory of computation → Denotational semantics; Theory of computation → Categorical semantics

Keywords and phrases computational effect, evaluation order, call-by-push-value, categorical semantics

Funding *Dylan McDermott*: Supported by an EPSRC studentship, and by Icelandic Research Fund project grant no. 196323-053

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

- If there are no side-effects at all (in particular, programs are strongly normalizing), the choice of evaluation order is irrelevant: M and M' terminate with the same result.
- 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 side-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) [13]. Levy describes how to translate (possibly open) expressions e into CBPV terms $(\!|e|\!)^{\mathrm{v}}$ and $(\!|e|\!)^{\mathrm{n}}$, which respectively correspond to call-by-value and call-by-name. We study the relationship between the behaviour of $(\!|e|\!)^{\mathrm{v}}$ and the behaviour of $(\!|e|\!)^{\mathrm{n}}$ in a given program context.

The main obstacle is that $(\!|e|\!)^{\mathrm{v}}$ and $(\!|e|\!)^{\mathrm{n}}$ have different types, and hence cannot be directly compared. Our solution to this is based on Reynold's work relating direct and continuation semantics of the λ-calculus [24]: we identify maps between the call-by-value and call-by-name interpretations, and compose these with the translations of expressions to arrive at two terms that *can* be compared directly. We show that, under certain conditions (satisfied only for some side-effects, such as our examples), the maps between call-by-value and call-by-name form a *Galois connection* (Theorem 22). This fact gives rise to a general *reasoning principle* (Theorem 25) that we use to compare call-by-value with call-by-name. Given any preorder $\preccurlyeq$ that captures the property we wish to show about programs, our reasoning principle gives sufficiency conditions for showing $M \preccurlyeq M'$, 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 $\preccurlyeq$; each of these relations indicates the extent to which changing evaluation order affects the behaviour of the program. In the divergence example $N \preccurlyeq N'$ is defined to mean termination of N implies termination of N' with the same result; in the other examples $\preccurlyeq$ similarly mirrors the properties described informally above.

Rather than just considering some fixed collection of (allowable) side-effects, we work abstractly and identify properties of side-effects that enable us to relate call-by-value and call-by-name. An advantage of our approach is that the properties can be derived by looking at the structure of the two maps between evaluation orders.

Our reasoning principle relies on the existence of some denotational model of the side-effects. We construct the Galois connections and relate the call-by-value and call-by-name translations inside the model itself. Crucially, we use *order-enriched* models, which order 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 maps between the call-by-value and call-by-name translations (Section 5), and show that they form a Galois connection for side-effects satisfying certain conditions (Theorem 22).
- We use the Galois connection to prove a novel reasoning principle (Theorem 25) that relates the call-by-value and call-by-name translations of expressions (Section 6).

Throughout, we consider three different examples: no side-effects, divergence, and non-determinism. We apply our reasoning principle to each, proving all of the above properties. 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).

2 Call-by-push-value, call-by-value, and call-by-name

Levy [13, 15] 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.

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

$$
\begin{array}{rrcl}
\text{value types} & A, B & ::= & \mathbf{unit} \mid A_1 \times A_2 \mid \mathbf{bool} \mid \mathbf{U}\,\underline{C} \\
\text{computation types} & \underline{C}, \underline{D} & ::= & \underline{C}_1 \mathbin{\underline{\times}} \underline{C}_2 \mid A \to \underline{C} \mid \mathbf{F}\,A \\
\text{values} & V, W & ::= & x \mid () \mid (V_1, V_2) \mid \mathbf{true} \mid \mathbf{false} \mid \mathbf{thunk}\,M \\
\text{computations} & M, N & ::= & \lambda\{1.\,M_1, 2.\,M_2\} \mid 1\text{`}M \mid 2\text{`}M \\
 & & & \mid \lambda x\!:\!A.\,M \mid V\text{`}M \mid \mathbf{return}\,V \mid M \;\mathbf{to}\; x.\,N \\
 & & & \mid \mathbf{match}\; V \;\mathbf{with}\; (x_1, x_2).\,M \\
 & & & \mid \mathbf{if}\; V \;\mathbf{then}\; M_1 \;\mathbf{else}\; M_2 \mid \mathbf{force}\,V
\end{array}
$$

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

The value type $\mathbf{U}\,\underline{C}$ is the type of *thunks* of computations of type $\underline{C}$. Elements of $\mathbf{U}\,\underline{C}$ are introduced using **thunk**: the value $\mathbf{thunk}\,M$ is the suspension of the computation term M. The corresponding eliminator is **force**, which is the inverse of **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 $1\text{`}M$ and $2\text{`}M$. Computation types also include function types (where functions send values to computations). Function application is written $V\text{`}M$, where V is the argument and M is the function to apply. The *returner type* $\mathbf{F}\,A$ has as elements computations that return elements of the value type A; these computations may have side-effects. Elements of $\mathbf{F}\,A$ are introduced by **return**; the computation $\mathbf{return}\,V$ immediately returns the value V (with no side-effects). Computations can be sequenced using M **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 >>= \x -> N` in Haskell.) The syntax we give here does not include any method of introducing effects; we extend CBPV with divergence (via recursion) and with nondeterminism 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 **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.

CBPV has two typing judgments: $\Gamma \vdash V : A$ for values and $\Gamma \vdash_c M : \underline{C}$ for computations. *Typing contexts* Γ 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 $\diamond$ for the empty typing context, $V : A$ as an abbreviation for $\diamond \vdash V : A$, and $M : \underline{C}$ as an abbreviation for $\diamond \vdash_c M : \underline{C}$.

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 *terminal computations*, which are the subset of computations with an introduction form on the outside:

$$R ::= \lambda\{1.\,M_1, 2.\,M_2\} \mid \lambda x\!:\!A.\,M \mid \mathbf{return}\,V$$

$\boxed{\Gamma \vdash V : A}$

$$\frac{}{\Gamma \vdash x : A}\ \text{if } (x : A) \in \Gamma \qquad \frac{}{\Gamma \vdash () : \mathbf{unit}} \qquad \frac{\Gamma \vdash V_1 : A_1 \qquad \Gamma \vdash V_2 : A_2}{\Gamma \vdash (V_1, V_2) : A_1 \times A_2}$$

$$\frac{}{\Gamma \vdash \mathbf{true} : \mathbf{bool}} \qquad \frac{}{\Gamma \vdash \mathbf{false} : \mathbf{bool}} \qquad \frac{\Gamma \vdash_c M : \underline{C}}{\Gamma \vdash \mathbf{thunk}\, M : \mathbf{U}\, \underline{C}}$$

$\boxed{\Gamma \vdash_c M : \underline{C}}$

$$\frac{\Gamma \vdash_c M_1 : \underline{C}_1 \qquad \Gamma \vdash_c M_2 : \underline{C}_2}{\Gamma \vdash_c \lambda\{1.\, M_1, 2.\, M_2\} : \underline{C}_1 \mathbin{\underline{\times}} \underline{C}_2} \qquad \frac{\Gamma \vdash_c M : \underline{C}_1 \mathbin{\underline{\times}} \underline{C}_2}{\Gamma \vdash_c 1\text{‘}M : \underline{C}_1} \qquad \frac{\Gamma \vdash_c M : \underline{C}_1 \mathbin{\underline{\times}} \underline{C}_2}{\Gamma \vdash_c 2\text{‘}M : \underline{C}_2}$$

$$\frac{\Gamma, x : A \vdash_c M : \underline{C}}{\Gamma \vdash_c \lambda x{:}A.\, M : A \to \underline{C}} \qquad \frac{\Gamma \vdash V : A \qquad \Gamma \vdash_c M : A \to \underline{C}}{\Gamma \vdash_c V\text{‘}M : \underline{C}}$$

$$\frac{\Gamma \vdash V : A}{\Gamma \vdash_c \mathbf{return}\, V : \mathbf{F}\, A} \qquad \frac{\Gamma \vdash_c M : \mathbf{F}\, A \quad \Gamma, x : A \vdash_c N : \underline{C}}{\Gamma \vdash_c M\ \mathbf{to}\ x.\, N : \underline{C}}$$

$$\frac{\Gamma \vdash V : A_1 \times A_2 \quad \Gamma, x_1 : A_1, x_2 : A_2 \vdash_c M : \underline{C}}{\Gamma \vdash_c \mathbf{match}\ V\ \mathbf{with}\ (x_1, x_2).\, M : \underline{C}} \qquad \frac{\Gamma \vdash V : \mathbf{bool} \quad \Gamma \vdash_c M_1 : \underline{C} \quad \Gamma \vdash_c M_2 : \underline{C}}{\Gamma \vdash_c \mathbf{if}\ V\ \mathbf{then}\ M_1\ \mathbf{else}\ M_2 : \underline{C}}$$

$$\frac{\Gamma \vdash V : \mathbf{U}\, \underline{C}}{\Gamma \vdash_c \mathbf{force}\, V : \underline{C}}$$

Figure 1 CBPV typing rules

$$\frac{}{\lambda\{1.\,M_1, 2.\,M_2\} \Downarrow \lambda\{1.\,M_1, 2.\,M_2\}} \qquad \frac{M \Downarrow \lambda\{1.\,N_1, 2.\,N_2\} \quad N_i \Downarrow R}{i`M \Downarrow R}\ i \in \{1,2\}$$

$$\frac{}{\lambda x{:}A.\,M \Downarrow \lambda x{:}A.\,M} \qquad \frac{M \Downarrow \lambda x{:}A.\,N \quad N[x \mapsto V] \Downarrow R}{V`M \Downarrow R}$$

$$\frac{}{\mathbf{return}\,V \Downarrow \mathbf{return}\,V} \qquad \frac{M \Downarrow \mathbf{return}\,V \quad N[x \mapsto V] \Downarrow R}{M\ \mathbf{to}\ x.\,N \Downarrow R}$$

$$\frac{M_1 \Downarrow R}{\mathbf{if\ true\ then}\ M_1\ \mathbf{else}\ M_2 \Downarrow R} \qquad \frac{M_2 \Downarrow R}{\mathbf{if\ false\ then}\ M_1\ \mathbf{else}\ M_2 \Downarrow R}$$

$$\frac{M[x_1 \mapsto V_1, x_2 \mapsto V_2] \Downarrow R}{\mathbf{match}\ (V_1, V_2)\ \mathbf{with}\ (x_1, x_2).\,M \Downarrow R}$$

$$\frac{M \Downarrow R}{\mathbf{force}\,(\mathbf{thunk}\,M) \Downarrow R}$$

Figure 2 Big-step operational semantics of CBPV

We only evaluate closed, well-typed computations, so when we write $M \Downarrow R$ we assume $M : \underline{C}$ for some $\underline{C}$ (this implies $R : \underline{C}$). 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 $i`M$, only the ith component of the pair M is evaluated. Since we have not yet included any way of forming impure computations, the semantics is deterministic and strongly normalizing: given any $M : \underline{C}$, 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 *ground type* G is a value type that does not contain thunks:

$$G \ ::= \ \mathbf{unit} \mid G_1 \times G_2 \mid \mathbf{empty} \mid G_1 + G_2$$

A CBPV *program* is a closed computation $M : \mathbf{F}\,G$, where G is a ground type. 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[1] $\preccurlyeq$ on closed computations of type G, for each ground type G. (We leave G implicit when writing $\preccurlyeq$.) For example, we could use

$$M \preccurlyeq M' \quad \text{if and only if} \quad \forall V{:}\mathbf{bool}.\ (M \Downarrow \mathbf{return}\,V) \ \Rightarrow\ (M' \Downarrow \mathbf{return}\,V)$$

We could also use, for example, the indiscrete relation for $\preccurlyeq$ (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 reasoning principle would be trivial).

[1] We do not actually need to assume that $\preccurlyeq$ is reflexive or transitive at any point, but because of constraints we add later (such as existence of an adequate model), it is unlikely that there are any interesting examples in which $\preccurlyeq$ is not a preorder.

$$\begin{array}{rcl@{\qquad}rcl}
1`\lambda\{1.\,M_1, 2.\,M_2\} & \equiv & M_1 & \textbf{if true then } M_1 \textbf{ else } M_2 & \equiv & M_1 \\
2`\lambda\{1.\,M_1, 2.\,M_2\} & \equiv & M_2 & \textbf{if false then } M_1 \textbf{ else } M_2 & \equiv & M_2 \\
V`\lambda x{:}A.\,M & \equiv & M[x \mapsto V] & \textbf{force}\,\textbf{thunk}\,M & \equiv & M \\
\textbf{match } (V_1, V_2) \textbf{ with } (x_1, x_2).\,M & \equiv & M[x_1 \mapsto V_1, x_2 \mapsto V_2] & & &
\end{array}$$

$$\begin{array}{rcl@{\qquad}rcl}
V & \equiv & () & V & \equiv & \textbf{thunk}\,\textbf{force}\,V \\
M[x \mapsto V] & \equiv & \textbf{match } V \textbf{ with } (x_1, x_2).\,M[x \mapsto (x_1, x_2)] & M & \equiv & \lambda\{1.\,1`M, 2.\,2`M\} \\
M[x \mapsto V] & \equiv & \textbf{if } V \textbf{ then } M[x \mapsto \textbf{true}] \textbf{ else } M[x \mapsto \textbf{false}] & M & \equiv & \lambda x{:}A.\,x`M
\end{array}$$

$$\begin{array}{l}
\textbf{return}\,V \textbf{ to } x.\,M \;\equiv\; M[x \mapsto V] \qquad\qquad M \;\equiv\; M \textbf{ to } x.\,\textbf{return}\,x \\
(M_1 \textbf{ to } x.\,M_2) \textbf{ to } y.\,M_3 \;\equiv\; M_1 \textbf{ to } x.\,(M_2 \textbf{ to } y.\,M_3) \\
\lambda\{1.\,M \textbf{ to } x.\,N_1, 2.\,M \textbf{ to } x.\,N_2\} \;\equiv\; M \textbf{ to } x.\,\lambda\{1.\,N_1, 2.\,N_2\} \\
\lambda y{:}A.\,M \textbf{ to } x.\,N \;\equiv\; M \textbf{ to } x.\,\lambda y{:}A.\,N
\end{array}$$

Figure 3 (Typed) equations between CBPV terms

Given any program relation $\preccurlyeq$, we define a *contextual preorder* $M \preccurlyeq^{\Gamma}_{\mathrm{ctx}} M'$ on arbitrary well-typed computations (in typing context Γ) by considering the behaviour of M and M' in programs as follows. A *computation context* $\mathcal{E}$ is a computation term, with a single hole $\square$ where a computation term is expected. We write $\mathcal{E}[M]$ for the computation that results from replacing $\square$ with M (which may capture some of the free variables of M). For example, if $\mathcal{E}$ is the computation context N **to** $x.\,\square$ then $\mathcal{E}[\textbf{return}\,x]$ is the computation N **to** $x.\,\textbf{return}\,x$, where x is captured. We use computation contexts to define $\preccurlyeq^{\Gamma}_{\mathrm{ctx}}$.

▶ **Definition 1** (Contextual preorder). *Suppose that $\preccurlyeq$ is a program relation, and that $\Gamma \vdash_c M : \underline{C}$ and $\Gamma \vdash_c M' : \underline{C}$ are two computations of the same type. We write $M \preccurlyeq^{\Gamma}_{\mathrm{ctx}} M'$ if, for all ground types G and computation contexts $\mathcal{E}$ such that $\mathcal{E}[M], \mathcal{E}[M'] : \mathbf{F}\,G$, we have $\mathcal{E}[M] \preccurlyeq \mathcal{E}[M']$.*

We sometimes omit Γ, and write just $M \preccurlyeq_{\mathrm{ctx}} M'$.

In parts of this paper, we use an equational theory on terms. We write $\equiv$ for the smallest equivalence relation on terms of the same type that is closed under the axioms in Figure 3 and under the syntax of CBPV terms (for example, $M \equiv N$ implies $\textbf{thunk}\,M \equiv \textbf{thunk}\,N$ and $V \equiv W$ implies $\textbf{return}\,V \equiv \textbf{return}\,W$).[2] All of the axioms should be read as subject to suitable typing constraints. In particular, $V \equiv ()$, which is the η-law for **unit**, can only be applied if V has type $()$. The group of axioms at the top of Figure 3 contains the β-laws for all of the type formers except $\mathbf{F}$. The second group contains η-laws. The bottom group contains axioms governing the behaviour of sequencing of computations: there is a left-unit axiom, a right-unit axiom, an associativity axiom, and axioms for commuting sequencing with the introduction forms for pairs of computations and for functions. If two computations are related by $\equiv$ then they behave the same under evaluation. In particular, at returner types:

[2] This is not exactly Levy's equational theory for CBPV, because we do not include *complex values*.

▶ **Lemma 2.** *If $M, N : \mathbf{F}\,A$ are closed computations that satisfy $M \equiv N$, and $M \Downarrow \mathbf{return}\,V$, then there is some $W : A$ such that $N \Downarrow \mathbf{return}\,W$ and $V \equiv W$.*

The proof of this uses a bisimulation argument (see Pitts [21]).

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

We use CBPV (instead of e.g. the monadic metalanguage [20]) because it captures both call-by-value and call-by-name evaluation. Levy [13] 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 τ and expressions e:

$$\begin{aligned}
\tau &::= \mathbf{unit} \mid \tau_1 \times \tau_2 \mid \tau \to \tau' \\
e &::= x \mid () \mid (e_1, e_2) \mid \mathbf{fst}\,e \mid \mathbf{snd}\,e \mid \mathbf{true} \mid \mathbf{false} \mid \mathbf{if}\ e_0\ \mathbf{then}\ e_1\ \mathbf{else}\ e_2 \mid \lambda x\!:\!\tau.\, e \mid e\,e'
\end{aligned}$$

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 4. For call-by-value, each source language type τ maps to a CBPV value type $(\!|\tau|\!)^{\mathrm{v}}$ that contains the results of call-by-value computations. For call-by-name, τ is translated to a computation type $(\!|\tau|\!)^{\mathrm{n}}$, which contains the computations themselves. Products in call-by-value use the value-type products of CBPV (which means they are necessarily *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 **unit** as a nullary product of computation types.) Functions under the call-by-value translation accept values of type $(\!|\tau|\!)^{\mathrm{v}}$ 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 Γ are translated into CBPV typing contexts $(\!|\Gamma|\!)^{\mathrm{v}}$ and $(\!|\Gamma|\!)^{\mathrm{n}}$. 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 $(\!|e|\!)^{\mathrm{v}}$ and $(\!|e|\!)^{\mathrm{n}}$. 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, side-effects occur only at the base types **unit** and **bool** (since this is where the returner types appear). As a consequence, η-expansion of source-language expressions is semantics-preserving: $(\!|e|\!)^{\mathrm{n}} \equiv (\!|\lambda x : A.\, e\, x|\!)^{\mathrm{n}}$ for functions and $(\!|e|\!)^{\mathrm{n}} \equiv (\!|(\mathbf{fst}\,e, \mathbf{snd}\,e)|\!)^{\mathrm{n}}$ for binary products. This is not the case for call-by-value.

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 $(\!|e|\!)^{\mathrm{v}}$ and $(\!|e'|\!)^{\mathrm{v}}$ 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 [13] 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.

2.2 Examples

We consider three collections of (allowable) side-effects as examples throughout the paper.

$$\boxed{\text{types } \tau \;\mapsto\; \text{value types } (\!|\tau|\!)^{\mathrm{v}}}$$

$$\begin{array}{rcl}
\mathbf{unit} & \mapsto & \mathbf{unit} \\
\tau_1 \times \tau_2 & \mapsto & (\!|\tau_1|\!)^{\mathrm{v}} \times (\!|\tau_2|\!)^{\mathrm{v}} \\
\mathbf{bool} & \mapsto & \mathbf{bool} \\
\tau \to \tau' & \mapsto & \mathbf{U}\,((\!|\tau|\!)^{\mathrm{v}} \to \mathbf{F}\,(\!|\tau'|\!)^{\mathrm{v}})
\end{array}$$

$$\boxed{\text{typing contexts } \Gamma \;\mapsto\; \text{typing contexts } (\!|\Gamma|\!)^{\mathrm{v}}}$$

$$\begin{array}{rcl}
\diamond & \mapsto & \diamond \\
\Gamma, x : \tau & \mapsto & (\!|\Gamma|\!)^{\mathrm{v}}, x : (\!|\tau|\!)^{\mathrm{v}}
\end{array}$$

$$\boxed{\text{expressions } \Gamma \vdash e : \tau \;\mapsto\; \text{computation } (\!|\Gamma|\!)^{\mathrm{v}} \vdash_c (\!|e|\!)^{\mathrm{v}} : \mathbf{F}\,(\!|\tau|\!)^{\mathrm{v}}}$$

$$\begin{array}{rcl}
x & \mapsto & \mathbf{return}\,x \\
() & \mapsto & \mathbf{return}\,() \\
(e_1, e_2) & \mapsto & (\!|e_1|\!)^{\mathrm{v}}\ \mathbf{to}\ z_1.\,(\!|e_2|\!)^{\mathrm{v}}\ \mathbf{to}\ z_2.\,\mathbf{return}\,(z_1, z_2) \\
\mathsf{fst}\,e & \mapsto & (\!|e|\!)^{\mathrm{v}}\ \mathbf{to}\ z.\,\mathbf{match}\ z\ \mathbf{with}\ (z_1, z_2).\,\mathbf{return}\,z_1 \\
\mathsf{snd}\,e & \mapsto & (\!|e|\!)^{\mathrm{v}}\ \mathbf{to}\ z.\,\mathbf{match}\ z\ \mathbf{with}\ (z_1, z_2).\,\mathbf{return}\,z_2 \\
\mathsf{true} & \mapsto & \mathbf{return}\,\mathbf{true} \\
\mathsf{false} & \mapsto & \mathbf{return}\,\mathbf{false} \\
\mathsf{if}\ e_0\ \mathsf{then}\ e_1\ \mathsf{else}\ e_2 & \mapsto & (\!|e_0|\!)^{\mathrm{v}}\ \mathbf{to}\ z.\,\mathbf{if}\ z\ \mathbf{then}\ (\!|e_1|\!)^{\mathrm{v}}\ \mathbf{else}\ (\!|e_2|\!)^{\mathrm{v}} \\
\lambda x{:}\tau.\,e & \mapsto & \mathbf{return}\,\mathbf{thunk}\,\lambda x{:}(\!|\tau|\!)^{\mathrm{v}}.\,(\!|e|\!)^{\mathrm{v}} \\
e\,e' & \mapsto & (\!|e|\!)^{\mathrm{v}}\ \mathbf{to}\ y.\,(\!|e'|\!)^{\mathrm{v}}\ \mathbf{to}\ z.\,z\,\text{‘}\,\mathbf{force}\,y
\end{array}$$

(a) Call-by-value translation $(\!|-|\!)^{\mathrm{v}}$

$$\boxed{\text{types } \tau \;\mapsto\; \text{computation types } (\!|\tau|\!)^{\mathrm{n}}}$$

$$\begin{array}{rcl}
\mathbf{unit} & \mapsto & \mathbf{F}\,\mathbf{unit} \\
\tau_1 \times \tau_2 & \mapsto & (\!|\tau_1|\!)^{\mathrm{n}} \underline{\times} (\!|\tau_2|\!)^{\mathrm{n}} \\
\mathbf{bool} & \mapsto & \mathbf{F}\,\mathbf{bool} \\
\tau \to \tau' & \mapsto & (\mathbf{U}\,(\!|\tau|\!)^{\mathrm{n}}) \to (\!|\tau'|\!)^{\mathrm{n}}
\end{array}$$

$$\boxed{\text{typing contexts } \Gamma \;\mapsto\; \text{typing contexts } (\!|\Gamma|\!)^{\mathrm{n}}}$$

$$\begin{array}{rcl}
\diamond & \mapsto & \diamond \\
\Gamma, x : \tau & \mapsto & (\!|\Gamma|\!)^{\mathrm{n}}, x : \mathbf{U}\,(\!|\tau|\!)^{\mathrm{n}}
\end{array}$$

$$\boxed{\text{expressions } \Gamma \vdash e : \tau \;\mapsto\; \text{computations } (\!|\Gamma|\!)^{\mathrm{n}} \vdash_c (\!|e|\!)^{\mathrm{n}} : (\!|\tau|\!)^{\mathrm{n}}}$$

$$\begin{array}{rcl}
x & \mapsto & \mathbf{force}\,x \\
() & \mapsto & \mathbf{return}\,() \\
(e_1, e_2) & \mapsto & \lambda\{1.\,(\!|e_1|\!)^{\mathrm{n}}, 2.\,(\!|e_2|\!)^{\mathrm{n}}\} \\
\mathsf{fst}\,e & \mapsto & 1\,\text{‘}\,(\!|e|\!)^{\mathrm{n}} \\
\mathsf{snd}\,e & \mapsto & 2\,\text{‘}\,(\!|e|\!)^{\mathrm{n}} \\
\mathsf{true} & \mapsto & \mathbf{return}\,\mathbf{true} \\
\mathsf{false} & \mapsto & \mathbf{return}\,\mathbf{false} \\
\mathsf{if}\ e_0\ \mathsf{then}\ e_1\ \mathsf{else}\ e_2 & \mapsto & (\!|e_0|\!)^{\mathrm{n}}\ \mathbf{to}\ z.\,\mathbf{if}\ z\ \mathbf{then}\ (\!|e_1|\!)^{\mathrm{n}}\ \mathbf{else}\ (\!|e_2|\!)^{\mathrm{n}} \\
\lambda x{:}\tau.\,e & \mapsto & \lambda x{:}\mathbf{U}\,(\!|\tau|\!)^{\mathrm{n}}.\,(\!|e|\!)^{\mathrm{n}} \\
e\,e' & \mapsto & (\mathbf{thunk}\,(\!|e'|\!)^{\mathrm{n}})\,\text{‘}\,(\!|e|\!)^{\mathrm{n}}
\end{array}$$

(b) Call-by-name translation $(\!|-|\!)^{\mathrm{n}}$

Figure 4 Translations from the source language into CBPV

▶ **Example 3** (No side-effects)**.** We include the simplest possible example: the case where there are no side-effects at all. For this example, call-by-value and call-by-name turn out to have identical behaviour. We define the program relation $M \preccurlyeq M'$ (for closed computations $M, M' : \mathbf{F}\,\mathbf{bool}$) as:

$$M \preccurlyeq M' \quad \text{if and only if} \quad \exists V\!:\!\mathbf{bool}.\ (M \Downarrow \mathbf{return}\,V) \ \wedge\ (M' \Downarrow \mathbf{return}\,V)$$

In other words, M and M' both evaluate to the same result V. Since evaluation is deterministic, V is necessarily unique. The contextual preorder $M \preccurlyeq^{\Gamma}_{\text{ctx}} M'$ means if we construct two programs by wrapping M and M' in the same computation context, then these two programs evaluate to the same result. This relation is symmetric. Our other examples use non-symmetric relations.

▶ **Example 4** (Divergence)**.** For our second example, the only side-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 $\preccurlyeq_{\text{ctx}}$ as it is defined in our no-side-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 $\mathbf{rec}\,x\!:\!\mathbf{U}\underline{C}.\,M$, and correspondingly extend the type system and operational semantics with the following rules:

$$\frac{\Gamma, x : \mathbf{U}\underline{C} \vdash_c M \ :\ \underline{C}}{\Gamma \vdash_c \mathbf{rec}\,x\!:\!\mathbf{U}\underline{C}.\,M \ :\ \underline{C}} \qquad \frac{M[x \mapsto \mathbf{thunk}\,(\mathbf{rec}\,x\!:\!\mathbf{U}\underline{C}.\,M)] \Downarrow R}{\mathbf{rec}\,x\!:\!\mathbf{U}\underline{C}.\,M \Downarrow R}$$

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 [3], but is the most convenient for our purposes.) Of course, by adding recursion we lose strong normalization (but the semantics is still deterministic). We extend the source language, and the two translations into CBPV, with recursive functions:

$$e \ ::=\ \ldots \mid \mathbf{rec}\,f\!:\!\tau \to \tau'.\,\lambda x.\,e \qquad \frac{\Gamma, f : \tau \to \tau', x : \tau \vdash e : \tau'}{\Gamma \vdash \mathbf{rec}\,f\!:\!\tau \to \tau'.\,\lambda x.\,e : \tau \to \tau'}$$

$$\begin{aligned}
(\!|\mathbf{rec}\,f\!:\!\tau \to \tau'.\,\lambda x.\,e|\!)^{\mathrm{v}} &= \mathbf{return}\,\mathbf{thunk}\,(\mathbf{rec}\,f\!:\!\mathbf{U}\,((\!|\tau|\!)^{\mathrm{v}} \to \mathbf{F}\,(\!|\tau'|\!)^{\mathrm{v}}).\,\lambda x\!:\!(\!|\tau|\!)^{\mathrm{v}}.\,(\!|e|\!)^{\mathrm{v}}) \\
(\!|\mathbf{rec}\,f\!:\!\tau \to \tau'.\,\lambda x.\,e|\!)^{\mathrm{n}} &= \mathbf{rec}\,f\!:\!\mathbf{U}\,(\mathbf{U}\,(\!|\tau|\!)^{\mathrm{n}} \to (\!|\tau'|\!)^{\mathrm{n}}).\,\lambda x\!:\!\mathbf{U}\,(\!|\tau|\!)^{\mathrm{n}}.\,(\!|e|\!)^{\mathrm{n}}
\end{aligned}$$

Again, the translations are the same as those given by Levy [13], except that Levy has general fixed points for call-by-name, rather than just recursive functions. The expression $\Omega_\tau = ((\mathbf{rec}\,f : \mathbf{bool} \to \tau.\,\lambda x.\,f\,x)\,\mathbf{false}) : \tau$ enables us to distinguish between call-by-value and call-by-name: $(\lambda x\!:\!\tau.\,\mathbf{false})\,\Omega_\tau$ diverges in call-by-value but not in call-by-name. In particular, we have $(\!|(\lambda x : \tau.\,\mathbf{true})\,\Omega_\tau|\!)^{\mathrm{n}} \Downarrow \mathbf{return}\,\mathbf{true}$, but there is no R such that $(\!|(\lambda x\!:\!\tau.\,\mathbf{true})\,\Omega_\tau|\!)^{\mathrm{v}} \Downarrow R$.

For this example, we define the program relation $\preccurlyeq$ by

$$M \preccurlyeq M' \quad \text{if and only if} \quad \forall V\!:\!\mathbf{bool}.\ (M \Downarrow \mathbf{return}\,V) \ \Rightarrow\ (M' \Downarrow \mathbf{return}\,V)$$

so that $M \preccurlyeq^{\Gamma}_{\text{ctx}} M'$ 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 5** (Nondeterminism)**.** Finally, 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 $\mathbf{fail}_{\underline{C}}$ for nullary nondeterministic choice and $M \,\mathbf{or}\, N$ for binary nondeterministic choice between computations; the typing and evaluation rules are standard:

$$\frac{}{\Gamma \vdash_c \mathbf{fail}_{\underline{C}} : \underline{C}} \qquad \frac{\Gamma \vdash_c M : \underline{C} \qquad \Gamma \vdash_c N : \underline{C}}{\Gamma \vdash_c M \,\mathbf{or}\, N : \underline{C}} \qquad \frac{M \Downarrow R}{M \,\mathbf{or}\, N \Downarrow R} \qquad \frac{N \Downarrow R}{M \,\mathbf{or}\, N \Downarrow R}$$

(There is no R such that $\mathbf{fail}_{\underline{C}} \Downarrow R$.) The computation $\mathbf{fail}_{\underline{C}}$ is the unit for $\mathbf{or}$, so $\mathbf{fail}_{\underline{C}} \,\mathbf{or}\, M$ and $M \,\mathbf{or}\, \mathbf{fail}_{\underline{C}}$ have the same behaviour as M. For each closed computation $M : \mathbf{F}\,A$ there might be zero, one or several values $V : A$ such that $M \Downarrow \mathbf{return}\, V$.

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

$$\frac{}{\Gamma \vdash \mathbf{fail}_\tau : \tau} \qquad \frac{\Gamma \vdash e : \tau \qquad \Gamma \vdash e' : \tau}{\Gamma \vdash e \,\mathbf{or}\, e' : \tau} \qquad \begin{array}{l} (\!|\mathbf{fail}_\tau|\!)^{\mathrm{v}} = \mathbf{fail}_{\mathbf{F}\,(\!|\tau|\!)^{\mathrm{v}}} \\ (\!|e \,\mathbf{or}\, e'|\!)^{\mathrm{v}} = (\!|e|\!)^{\mathrm{v}} \,\mathbf{or}\, (\!|e'|\!)^{\mathrm{v}} \end{array} \qquad \begin{array}{l} (\!|\mathbf{fail}_\tau|\!)^{\mathrm{n}} = \mathbf{fail}_{(\!|\tau|\!)^{\mathrm{n}}} \\ (\!|e \,\mathbf{or}\, e'|\!)^{\mathrm{n}} = (\!|e|\!)^{\mathrm{n}} \,\mathbf{or}\, (\!|e'|\!)^{\mathrm{n}} \end{array}$$

As an example, evaluating the expression $e = (\lambda x.\,\mathbf{if}\ x\ \mathbf{then}\ x\ \mathbf{else}\ \mathbf{true})(\mathbf{true}\,\mathbf{or}\,\mathbf{false})$ under call-by-value necessarily results in **true**, but under call-by-name we can also get **false**. (We have $(\!|e|\!)^{\mathrm{v}} \not\Downarrow \mathbf{return}\,\mathbf{false}$ but $(\!|e|\!)^{\mathrm{n}} \Downarrow \mathbf{return}\,\mathbf{false}$.)

For nondeterminism, we define $\preccurlyeq$ in the same way as our divergence example:

$$M \preccurlyeq M' \quad \text{if and only if} \quad \forall V\!:\!\mathbf{bool}.\ (M \Downarrow \mathbf{return}\, V) \ \Rightarrow\ (M' \Downarrow \mathbf{return}\, V)$$

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).

3 Order-enriched denotational semantics

We give a denotational semantics for CBPV, which we use to prove instances of $\preccurlyeq_{\mathrm{ctx}}$. Since $\preccurlyeq_{\mathrm{ctx}}$ is not in general symmetric, we use *order-enriched* models, which come with partial orders $\sqsubseteq$ between denotations. In an *adequate* model, $[\![M]\!] \sqsubseteq [\![N]\!]$ implies $M \preccurlyeq_{\mathrm{ctx}} N$. Our semantics is based on Levy's *algebra models* [15] 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* [14] can be used for the same purpose.)

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 6.** *A* **Poset***-category* **C** *is an ordinary category, together with a partial order $\sqsubseteq$ on each hom-set* $\mathbf{C}(X, Y)$*, such that composition is monotone.*

If **C** is a **Poset**-category, we refer to the ordinary category as the *underlying* ordinary category, and write $|\mathbf{C}|$ for the class of objects.

▶ **Example 7.** We use the following three **Poset**-categories.

Poset-category **C**	Objects $X \in \lvert\mathbf{C}\rvert$	Morphisms $f : X \to Y$	Order $f \sqsubseteq f'$
Set	sets	functions	equality
Poset	posets	monotone functions	pointwise
ω**Cpo**	ωcpos	ω-continuous functions	pointwise

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

Let $\mathbf{C}$ be a **Poset**-category. We say that $\mathbf{C}$ is *cartesian* when its underlying category has a terminal object 1 and binary products $X_1 \times X_2$, such that the pairing functions $\langle -, - \rangle : \mathbf{C}(W, X_1) \times \mathbf{C}(W, X_2) \to \mathbf{C}(W, X_1 \times X_2)$ are monotone. When this is the case, there are canonical isomorphisms $assoc_{X_1,X_2,X_3} : (X_1 \times X_2) \times X_3 \to X_1 \times (X_2 \times X_3)$. We say that $\mathbf{C}$ is *cartesian closed* when it is cartesian and its underlying category has exponentials $X \Rightarrow Y$ for which the currying functions $\Lambda : \mathbf{C}(W \times X, Y) \to \mathbf{C}(W, X \Rightarrow Y)$ are monotone. (It follows that the uncurrying functions $\Lambda^{-1} : \mathbf{C}(W, X \Rightarrow Y) \to \mathbf{C}(W \times X, Y)$ are also monotone.) We write $ev_{X,Y}$ for the canonical morphism $\Lambda^{-1} id : (X \Rightarrow Y) \times X \to Y$. Binary *coproducts* in $\mathbf{C}$ are just binary coproducts in the underlying ordinary category, except that the copairing functions $[-, -] : \mathbf{C}(X_1, W) \times \mathbf{C}(X_2, W) \to \mathbf{C}(X_1 + X_2, W)$ are required to be monotone. The **Poset**-categories **Set**, **Poset**, and $\omega\mathbf{Cpo}$ are all cartesian closed, and have binary coproducts.

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 [20] 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. (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 8** (Strong **Poset**-monad). *A strong* **Poset**-*monad* $\mathsf{T} = (T, \eta, (-)^\dagger)$ *on a cartesian* **Poset**-*category* $\mathbf{C}$ *consists of an object* $TX \in |\mathbf{C}|$ *and morphism* $\eta_X : X \to TX$ *for each* $X \in |\mathbf{C}|$, *and a monotone function (*Kleisli extension*)* $(-)^\dagger : \mathbf{C}(W \times X, TY) \to \mathbf{C}(W \times TX, TY)$ *for each* $W, X, Y \in |\mathbf{C}|$, *such that*

$$\begin{aligned} f^\dagger \circ (W \times \eta_X) &= f : W \times X \to TY && \textit{for all } f : W \times X \to TY \\ (\eta_X \circ \pi_2)^\dagger &= \pi_2 : 1 \times TX \to TX && \textit{for all } X \in |\mathbf{C}| \\ (g^\dagger \circ (W' \times f) \circ assoc)^\dagger &= g^\dagger \circ (W' \times f^\dagger) \circ assoc && \textit{for all } f : W \times X \to TY, \\ &\quad : (W' \times W) \times TX \to TZ && \qquad g : W' \times Y \to TZ \end{aligned}$$

Specializing the Kleisli extension of T to $W = 1$ produces a (non-strong) extension operator $(-)^\dagger : \mathbf{C}(X, TY) \to \mathbf{C}(TX, TY)$. 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*.)

▶ **Definition 9** (Eilenberg-Moore algebra). *Let* T *be a strong* **Poset**-*monad on a cartesian* **Poset**-*category* $\mathbf{C}$. *A* T-algebra $\mathsf{Z} = (Z, (-)^\ddagger)$ *is a pair of an object* $Z \in |\mathbf{C}|$ *(the* carrier*) and monotone function (*extension *operator)* $(-)^\ddagger : \mathbf{C}(W \times X, Z) \to \mathbf{C}(W \times TX, Z)$ *for each* $W, X \in |\mathbf{C}|$, *such that*

$$\begin{aligned} f^\ddagger \circ (W \times \eta_X) &= f : W \times X \to Z && \textit{for all } f : W \times X \to TY \\ (g^\ddagger \circ (W' \times f) \circ assoc)^\ddagger &= g^\ddagger \circ (W' \times f^\dagger) \circ assoc && \textit{for all } f : W \times X \to TY, \\ &\quad : (W' \times W) \times TX \to Z && \qquad g : W' \times Y \to Z \end{aligned}$$

For each $X \in |\mathbf{C}|$, *we write* $F_{\mathsf{T}} X$ *for the* free T-*algebra* $(TX, (-)^\dagger)$, *and for each* T-*algebra* Z, *we write* $U_{\mathsf{T}} \mathsf{Z}$ *for the carrier* $Z \in |\mathbf{C}|$.

Specializing the extension operator of a T-algebra Z to $W = 1$ produces a (non-strong) extension operator $(-)^{\ddagger} : \mathbf{C}(X, Z) \to \mathbf{C}(TX, Z)$.

Let T be a strong **Poset**-monad on a cartesian closed **Poset**-category $\mathbf{C}$. If $\mathsf{Z}_1, \mathsf{Z}_2$ are T-algebras, then there is a T-algebra $\mathsf{Z}_1 \times \mathsf{Z}_2$ with carrier $Z_1 \times Z_2$ and extension operator

$$f^{\ddagger} = \langle (\pi_1 \circ f)^{\ddagger}, (\pi_2 \circ f)^{\ddagger} \rangle : W \times TX \to Z_1 \times Z_2 \qquad \text{for } f : W \times X \to Z_1 \times Z_2$$

(These are binary products in the Eilenberg-Moore **Poset**-category of T.) If $Y \in |\mathbf{C}|$ and Z is a T-algebra, then there is a T-algebra $Y \Rightarrow \mathsf{Z}$ with carrier $Y \Rightarrow Z$ and extension operator

$$f^{\ddagger} = \Lambda((\Lambda^{-1} f \circ \beta_{W,Y,X})^{\ddagger} \circ \beta_{W,TX,Y}) : W \times TX \to Y \Rightarrow Z \qquad \text{for } f : W \times X \to Y \Rightarrow Z$$

where $\beta_{X_1,X_2,X_3} = \langle \langle \pi_1 \circ \pi_1, \pi_2 \rangle, \pi_2 \circ \pi_1 \rangle : (X_1 \times X_2) \times X_3 \to (X_1 \times X_3) \times X_2$. These provide the interpretations of binary products $\underline{C}_1 \mathbin{\underline{\times}} \underline{C}_2$ of computation types, and the interpretations of function types $A \to \underline{C}$.

▶ **Definition 10.** *A* model *of CBPV consists of*

- *a cartesian closed* **Poset***-category* $\mathbf{C}$*;*
- *an object* 2 *and morphisms* $\mathit{inl}, \mathit{inr} : 1 \to 2$*, forming the coproduct* $1 + 1$*;*
- *a strong* **Poset***-monad* $\mathsf{T} = (T, \eta, (-)^{\dagger})$ *on* $\mathbf{C}$*.*

Given any model, the interpretation $[\![-]\!]$ of CBPV is defined in Figure 5. Value types A are interpreted as objects $[\![A]\!] \in |\mathbf{C}|$, while computation types $\underline{C}$ are interpreted as T-algebras. Typing contexts Γ are interpreted as objects $[\![\Gamma]\!] \in \mathbf{C}$ using the cartesian structure of $\mathbf{C}$; if $(x : A) \in \Gamma$ then we write π_x for the corresponding projection $[\![\Gamma]\!] \to [\![A]\!]$. Values $\Gamma \vdash V : A$ (respectively computations $\Gamma \vdash_c M : \underline{C}$) are interpreted as morphisms $[\![\Gamma \vdash V : A]\!]$ (resp. $[\![\Gamma \vdash_c M : \underline{C}]\!]$) in $\mathbf{C}$; we often omit the typing context and type when writing these. Programs $\diamond \vdash_c M : \mathbf{bool}$ are therefore interpreted as morphisms $[\![M]\!] : 1 \to F_{\mathsf{T}}2$. To interpret **if**, we use the fact that, since $\mathbf{C}$ is cartesian closed, products distribute over the coproduct $2 = 1 + 1$. This means that for every $W \in |\mathbf{C}|$, the coproduct $W + W$ also exists in $\mathbf{C}$, and the canonical morphism

$$W + W \xrightarrow{[\langle id_W, inl \circ \langle\rangle_W \rangle, \langle id_W, inr \circ \langle\rangle_W \rangle]} W \times 2$$

has an inverse $\mathit{dist}_W : W \times 2 \to W + W$.

By composing the semantics of CBPV with the two translations of the source language, we obtain a call-by-value semantics $[\![-]\!]^{\mathrm{v}} = [\![(\!|-|\!)^{\mathrm{v}}]\!]$ and a call-by-name semantics $[\![-]\!]^{\mathrm{n}} = [\![(\!|-|\!)^{\mathrm{n}}]\!]$ of the source language.

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

▶ **Definition 11.** *A model of CBPV is* adequate *(with respect to a given program relation $\preccurlyeq$) if for all computations $\Gamma \vdash_c M : \underline{C}$ and $\Gamma \vdash_c M' : \underline{C}$ we have*

$$[\![\Gamma \vdash_c M : \underline{C}]\!] \sqsubseteq [\![\Gamma \vdash_c M' : \underline{C}]\!] \quad \Rightarrow \quad M \preccurlyeq^{\Gamma}_{\mathrm{ctx}} M'$$

We give three different models, one for each of our three examples in Section 2.2. Each model is adequate with respect to the corresponding definition of $\preccurlyeq$; the proof in each case is a standard *logical relations* argument (e.g. [28]).

▶ **Example 12.** For CBPV with no side-effects, we use $\mathbf{C} = \mathbf{Set}$. The strong **Poset**-monad T is the identity on $\mathbf{Set}$. Each T-algebra Z is completely determined by its carrier Z; the extension operator $(-)^{\ddagger} : \mathbf{Set}(W \times X, Z) \to \mathbf{Set}(W \times X, Z)$ is necessarily the identity. The interpretation $[\![M]\!]$ of each program M is just an element of 2.

$$\boxed{\mathbf{C}\text{-object } \llbracket A \rrbracket}$$

$$\begin{aligned}
\llbracket \mathbf{unit} \rrbracket &= 1 \\
\llbracket A_1 \times A_2 \rrbracket &= \llbracket A_1 \rrbracket \times \llbracket A_2 \rrbracket \\
\llbracket \mathbf{bool} \rrbracket &= 2 \quad (= 1{+}1) \\
\llbracket \mathbf{U}\,\underline{C} \rrbracket &= U_{\mathsf{T}} \llbracket \underline{C} \rrbracket
\end{aligned}$$

$$\boxed{\mathsf{T}\text{-algebra } \llbracket \underline{C} \rrbracket}$$

$$\begin{aligned}
\llbracket \underline{C}_1 \mathbin{\underline{\times}} \underline{C}_2 \rrbracket &= \llbracket \underline{C}_1 \rrbracket \times \llbracket \underline{C}_2 \rrbracket \\
\llbracket A \to \underline{C} \rrbracket &= \llbracket A \rrbracket \Rightarrow \llbracket \underline{C} \rrbracket \\
\llbracket \mathbf{F}\, A \rrbracket &= F_{\mathsf{T}} \llbracket A \rrbracket
\end{aligned}$$

$$\boxed{\mathbf{C}\text{-object } \llbracket \Gamma \rrbracket}$$

$$\begin{aligned}
\llbracket \diamond \rrbracket &= 1 \\
\llbracket \Gamma, x : A \rrbracket &= \llbracket \Gamma \rrbracket \times \llbracket A \rrbracket
\end{aligned}$$

$$\boxed{\llbracket \Gamma \vdash V : A \rrbracket : \llbracket \Gamma \rrbracket \to \llbracket A \rrbracket}$$

$$\begin{aligned}
\llbracket x \rrbracket &= \pi_x \\
\llbracket () \rrbracket &= \langle\rangle_{\llbracket \Gamma \rrbracket} \\
\llbracket (V_1, V_2) \rrbracket &= \langle \llbracket V_1 \rrbracket, \llbracket V_2 \rrbracket \rangle \\
\llbracket \mathbf{true} \rrbracket &= \mathit{inl} \circ \langle\rangle_{\llbracket \Gamma \rrbracket} \\
\llbracket \mathbf{false} \rrbracket &= \mathit{inr} \circ \langle\rangle_{\llbracket \Gamma \rrbracket} \\
\llbracket \mathbf{thunk}\, M \rrbracket &= \llbracket M \rrbracket
\end{aligned}$$

$$\boxed{\llbracket \Gamma \vdash_c M : \underline{C} \rrbracket : \llbracket \Gamma \rrbracket \to U_{\mathsf{T}} \llbracket \underline{C} \rrbracket}$$

$$\begin{aligned}
\llbracket \lambda\{1.\, M_1, 2.\, M_2\} \rrbracket &= \langle \llbracket M_1 \rrbracket, \llbracket M_2 \rrbracket \rangle \\
\llbracket i`M \rrbracket &= \pi_i \circ \llbracket M \rrbracket \qquad\qquad (i \in \{1, 2\}) \\
\llbracket \lambda x \!:\! A.\, M \rrbracket &= \Lambda \llbracket M \rrbracket \\
\llbracket V`M \rrbracket &= \Lambda^{-1} \llbracket M \rrbracket \circ \langle \mathit{id}, \llbracket V \rrbracket \rangle \\
\llbracket \mathbf{return}\, V \rrbracket &= \eta \circ \llbracket V \rrbracket \\
\llbracket M \ \mathbf{to}\ x.\, N \rrbracket &= \llbracket N \rrbracket^{\ddagger} \circ \langle \mathit{id}, \llbracket M \rrbracket \rangle \\
\llbracket \mathbf{match}\, V\, \mathbf{with}\ (x, y).\, M \rrbracket &= \llbracket M \rrbracket \circ \mathit{assoc}^{-1} \circ \langle \mathit{id}, \llbracket V \rrbracket \rangle \\
\llbracket \mathbf{if}\ V\ \mathbf{then}\ M_1\ \mathbf{else}\ M_2 \rrbracket &= [\llbracket M_1 \rrbracket, \llbracket M_2 \rrbracket] \circ \mathit{dist} \circ \langle \mathit{id}, \llbracket V \rrbracket \rangle \\
\llbracket \mathbf{force}\, V \rrbracket &= \llbracket V \rrbracket
\end{aligned}$$

Figure 5 Denotational semantics of CBPV

▶ **Example 13.** For divergence, we use $\mathbf{C} = \omega\mathbf{Cpo}$. The strong **Poset**-monad T freely adjoins a least element $\bot$ to each $\omega\mathbf{Cpo}$. The unit η_X is the inclusion $X \hookrightarrow TX$, while Kleisli extension is given by

$$f^{\dagger}(w, x) = \begin{cases} \bot & \text{if } x = \bot \\ f(w, x) & \text{otherwise} \end{cases}$$

A T-algebra Z is equivalently an $\omega\mathbf{Cpo}$ Z with a least element $\bot \in Z$. The extension operator is completely determined once the carrier is fixed; it is analogous to $(-)^{\dagger}$.

In this case, the product $Z_1 \times Z_2$ is the set of pairs ordered componentwise, and the exponential $Y \Rightarrow Z$ is the set of set of ω-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 elements, and $Y \Rightarrow Z$ has a least element (the constantly-$\bot$ function) whenever Z has a least element.

If Z is a T-algebra, then every ω-continuous function $f : Z \to Z$ has a least fixed point $\mathit{fix}\, f = \bigsqcup_{n\in\mathbb{N}} f^n \bot \in Z$. These enable us to interpret recursive computations, by defining $[\![\mathbf{rec}\, x : \mathbf{U}\underline{C}.\, M]\!]\rho = \mathit{fix}\big(x \mapsto [\![M]\!](\rho, x)\big)$. The interpretation $[\![M]\!]$ of a program $M : \mathbf{F\,bool}$ is either $\bot$ (signifying divergence), or one of the two elements of 2.

▶ **Example 14.** For finite nondeterminism, we use $\mathbf{C} = \mathbf{Poset}$. The strong **Poset**-monad T freely adds finite joins to each poset. It is defined by

$$TX = (\{{\downarrow}S' \mid S' \in \mathcal{P}_{\text{fin}} X\}, \subseteq) \qquad \eta_X\, x = {\downarrow}\{x\} \qquad f^{\dagger}(w, S) = \bigcup_{x\in S} f(w, x)$$

where $\mathcal{P}_{\text{fin}} X$ is the set of finite subsets of X, and ${\downarrow}S' = \{x \in X \mid \exists x' \in S'.\, x \sqsubseteq x'\}$ 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 $f^{\ddagger}(w, S) = \bigsqcup_{x\in S} f(w, x)$. (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 function space $Y \Rightarrow Z$ is the set of monotone functions ordered pointwise, with joins given by $(\bigsqcup_i f_i)x = \bigsqcup_i (f_i x)$.

We interpret nondeterministic computations using nullary and binary joins:

$$[\![\mathbf{fail}_{\underline{C}}]\!]\rho = \bot \qquad [\![M \mathbin{\mathbf{or}} N]\!]\rho = [\![M]\!]\rho \sqcup [\![N]\!]\rho$$

The interpretation $[\![M]\!]$ of a program $M : \mathbf{F\,bool}$ is one of the four subsets of 2.

4 Left-to-right or right-to-left?

Before we consider call-by-value and call-by-name, we consider a simpler case of a relationship between evaluation orders. We show that, under certain conditions (the assumptions of Theorem 15 below), left-to-right evaluation of pairs is equivalent to right-to-left evaluation in call-by-value. (We could also consider right-to-left evaluation of function application with no extra difficulty.)

The call-by-value translation $(\!|e|\!)^{\mathrm{v}}$ defined in Figure 4 evaluates pairs from left-to-right. We give another call-by-value translation $(\!|-|\!)^{\mathrm{r}}$, which has an identical definition to $(\!|-|\!)^{\mathrm{v}}$ except that pairs are evaluated right-to-left:

$$(\!|(e_1, e_2)|\!)^{\mathrm{r}} \;=\; (\!|e_2|\!)^{\mathrm{r}} \mathbin{\mathbf{to}} z_2.\, (\!|e_1|\!)^{\mathrm{r}} \mathbin{\mathbf{to}} z_1.\, \mathbf{return}\,(z_1, z_2)$$

The left-to-right and right-to-left translations have the same types, and therefore it makes sense to ask whether $(\!|e|\!)^{\mathrm{v}} \cong_{\mathrm{ctx}} (\!|e|\!)^{\mathrm{r}}$ holds, where $M \cong_{\mathrm{ctx}} M'$ means both $M \preccurlyeq_{\mathrm{ctx}} M'$ and $M' \preccurlyeq_{\mathrm{ctx}} M$ hold. (Relating call-by-value and call-by-name is more difficult because the two translations have different types.)

Since we reason using the denotational semantics, we first spell out the interpretations of these two translations of pairs given any CBPV model. We define two natural transformations for sequencing of computations: *seq* for left-to-right and $\mathit{seq}^{\mathrm{r}}$ for right-to-left. These are defined in terms of a monotone function $(-)^{\dagger^r} : \mathbf{C}(X \times W, Y) \to \mathbf{C}(TX \times W, Y)$ (which is Kleisli extension with the products reversed):

$$\begin{aligned} f^{\dagger^r} &= (f \circ \langle \pi_2, \pi_1 \rangle)^{\dagger} \circ \langle \pi_2, \pi_1 \rangle \\ \mathit{seq}_{X_1,X_2} &= ({\eta_{X_1 \times X_2}}^{\dagger})^{\dagger^r} \ : \ TX_1 \times TX_2 \to T(X_1 \times X_2) \\ \mathit{seq}^{\mathrm{r}}_{X_1,X_2} &= ({\eta_{X_1 \times X_2}}^{\dagger^r})^{\dagger} \ : \ TX_1 \times TX_2 \to T(X_1 \times X_2) \end{aligned}$$

In our three example models, the natural transformation *seq* is given by:

$$\begin{aligned} \mathit{seq}_{X_1,X_2}(x_1, x_2) &= (x_1, x_2) && \text{(no side-effects)} \\ \mathit{seq}_{X_1,X_2}(x_1, x_2) &= \begin{cases} \bot & \text{if } x_1 = \bot \vee x_2 = \bot \\ (x_1, x_2) & \text{otherwise} \end{cases} && \text{(divergence)} \\ \mathit{seq}_{X_1,X_2}(S_1, S_2) &= \{(x_1, x_2) \mid x_1 \in S_1 \wedge x_2 \in S_2\} && \text{(nondeterminism)} \end{aligned}$$

Writing $[\![-]\!]^{\mathrm{v}}$ for $[\![(\!|-|\!)^{\mathrm{v}}]\!]$ and $[\![-]\!]^{\mathrm{r}}$ for $[\![(\!|-|\!)^{\mathrm{r}}]\!]$, we have:

$$[\![(e_1, e_2)]\!]^{\mathrm{v}} = \mathit{seq} \circ \langle [\![e_1]\!]^{\mathrm{v}}, [\![e_2]\!]^{\mathrm{v}} \rangle \qquad [\![(e_1, e_2)]\!]^{\mathrm{r}} = \mathit{seq}^{\mathrm{r}} \circ \langle [\![e_1]\!]^{\mathrm{r}}, [\![e_2]\!]^{\mathrm{r}} \rangle$$

We do not have $(\!|e|\!)^{\mathrm{v}} \cong_{\mathrm{ctx}} (\!|e|\!)^{\mathrm{r}}$ in general, because side-effects might occur in (observably) different orders. However, we do have $(\!|e|\!)^{\mathrm{v}} \cong_{\mathrm{ctx}} (\!|e|\!)^{\mathrm{r}}$ for our three examples (no side-effects, divergence and nondeterminism). The crucial property is that in all three cases, the monad T is *commutative* [9], meaning that $\mathit{seq} = \mathit{seq}^{\mathrm{r}}$. Commutativity implies $[\![e]\!]^{\mathrm{v}} = [\![e]\!]^{\mathrm{r}}$, and then adequacy implies $(\!|e|\!)^{\mathrm{v}} \cong_{\mathrm{ctx}} (\!|e|\!)^{\mathrm{r}}$. Hence we arrive at a reasoning principle for relating left-to-right and right-to-left evaluation. Given a program relation $\preccurlyeq$, to show $(\!|e|\!)^{\mathrm{v}} \cong_{\mathrm{ctx}} (\!|e|\!)^{\mathrm{r}}$, it suffices to find an adequate model that satisfies commutativity:

▶ **Theorem 15** (Relationship between left-to-right and right-to-left)**.** *Suppose we are given a program relation $\preccurlyeq$, and a model of CBPV that is adequate with respect to $\preccurlyeq$ and satisfies $\mathit{seq} = \mathit{seq}^{\mathrm{r}}$. If $\Gamma \vdash e : \tau$ then $(\!|e|\!)^{\mathrm{v}} \cong_{\mathrm{ctx}} (\!|e|\!)^{\mathrm{r}}$.*

Proof. By induction on e. The only non-trivial case is pairing, for which we use $\mathit{seq} = \mathit{seq}^{\mathrm{r}}$. ◀

The key property required for the proof, which follows from commutativity, is *centrality*:

▶ **Definition 16** (Central [23])**.** *A computation $\Gamma \vdash_c M : \mathbf{F}\, A$ is* central *if for all computations $\Gamma \vdash_c N : \mathbf{F}\, B$ we have*

$$M \ \mathbf{to}\ x.\, N \ \mathbf{to}\ y.\, \mathbf{return}\,(x, y) \ \cong_{\mathrm{ctx}} \ N \ \mathbf{to}\ y.\, M \ \mathbf{to}\ x.\, \mathbf{return}\,(x, y)$$

A morphism $f : Z \to TX$ is central *if for all $g : Z \to TY$ we have*

$$\mathit{seq} \circ \langle f, g \rangle = \mathit{seq}^{\mathrm{r}} \circ \langle f, g \rangle$$

5 A Galois connection between call-by-value and call-by-name

We now return to the main contribution of this paper: relating call-by-value and call-by-name. This is more difficult than relating left-to-right and right-to-left evaluation, because the call-by-value and call-by-name translations of expressions have different types:

$$(\!|\Gamma|\!)^{\mathrm{v}} \vdash_c (\!|e|\!)^{\mathrm{v}} : \mathbf{F}\,(\!|\tau|\!)^{\mathrm{v}} \qquad (\!|\Gamma|\!)^{\mathrm{n}} \vdash_c (\!|e|\!)^{\mathrm{n}} : (\!|\tau|\!)^{\mathrm{n}}$$

We cannot ask whether $(\!|e|\!)^{\mathrm{v}}$ and $(\!|e|\!)^{\mathrm{n}}$ are related by $\preccurlyeq_{\mathrm{ctx}}$, because it only relates terms of the same type. It does not make sense to replace $(\!|e|\!)^{\mathrm{v}}$ with $(\!|e|\!)^{\mathrm{n}}$ inside a CBPV program, because the result would not be well-typed.

Reynolds [24] solves a similar problem when comparing direct and continuation semantics of the λ-calculus by defining maps between the two semantics, so that a denotation in the direct semantics can be viewed as a denotation in the continuation semantics and vice versa. We use the same idea here. Specifically, we define maps Φ from call-by-value computations to call-by-name computations, and Ψ from call-by-name to call-by-value:

$$\Gamma \vdash_c M : \mathbf{F}\,(\!|\tau|\!)^{\mathrm{v}} \;\mapsto\; \Gamma \vdash_c \Phi_\tau M : (\!|\tau|\!)^{\mathrm{n}} \qquad\qquad \Gamma \vdash_c N : (\!|\tau|\!)^{\mathrm{n}} \;\mapsto\; \Gamma \vdash_c \Psi_\tau N : \mathbf{F}\,(\!|\tau|\!)^{\mathrm{v}}$$

Then instead of replacing $(\!|e|\!)^{\mathrm{v}}$ with $(\!|e|\!)^{\mathrm{n}}$ directly, we use Φ and Ψ to convert $(\!|e|\!)^{\mathrm{n}}$ to a computation of the same type as $(\!|e|\!)^{\mathrm{v}}$ (we define this computation formally in Section 6):

$$(\!|\Gamma|\!)^{\mathrm{v}} \longrightarrow (\!|\Gamma|\!)^{\mathrm{n}} \xrightarrow{(\!|e|\!)^{\mathrm{n}}} (\!|\tau|\!)^{\mathrm{n}} \longrightarrow \mathbf{F}\,(\!|\tau|\!)^{\mathrm{v}}$$

This behaves like a call-by-name computation, but has the same type as a call-by-value computation. (We could instead have chosen to convert $(\!|e|\!)^{\mathrm{v}}$ into a computation of the same type as $(\!|e|\!)^{\mathrm{n}}$. As we show in Section 6.1, this choice is arbitrary.) We do not want just *any* maps between call-by-value and call-by-name. We show that, under certain conditions (the assumptions of Theorem 22 below) the maps we define form *Galois connections* [18] in the denotational semantics. This is crucial for the correctness of our reasoning principle. The conditions needed to show that we have Galois connections are where the choice of side-effects becomes important.

▶ **Definition 17.** *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 syntactic maps Φ_τ and Ψ_τ are defined in Figure 6. (We use some extra variables in the definition, which are assumed to be fresh.) The maps Φ_τ from call-by-value computations first evaluate the computation, and then map the result to call-by-name using $\hat{\Phi}_\tau$, which has the following typing:

$$\Gamma \vdash V : (\!|\tau|\!)^{\mathrm{v}} \quad\mapsto\quad \Gamma \vdash_c \hat{\Phi}_\tau V : (\!|\tau|\!)^{\mathrm{n}}$$

Given any model of CBPV, we correspondingly define morphisms $\phi_\tau : T[\![\tau]\!]^{\mathrm{v}} \to U_{\mathsf{T}}[\![\tau]\!]^{\mathrm{n}}$ and $\psi_\tau : U_{\mathsf{T}}[\![\tau]\!]^{\mathrm{n}} \to T[\![\tau]\!]^{\mathrm{v}}$ in Figure 7 (where $\hat{\phi}_\tau : [\![\tau]\!]^{\mathrm{v}} \to U_{\mathsf{T}}[\![\tau]\!]^{\mathrm{n}}$). These morphisms are the interpretations of Φ and Ψ in the following sense.

▶ **Lemma 18.** *Given any model of CBPV, if $\Gamma \vdash_c M : \mathbf{F}\,(\!|\tau|\!)^{\mathrm{v}}$ then $[\![\Phi_\tau M]\!] = \phi_\tau \circ [\![M]\!]$, and if $\Gamma \vdash_c N : (\!|\tau|\!)^{\mathrm{n}}$ then $[\![\Psi_\tau N]\!] = \psi_\tau \circ [\![N]\!]$.*

Proof sketch. By induction on the type τ. Each case is an easy calculation. ◀

$$
\begin{aligned}
\Phi_\tau M &= M \textbf{ to } x.\, \hat{\Phi}_\tau x \\[1ex]
\hat{\Phi}_{\textbf{unit}} V &= \textbf{return}\, V \\
\hat{\Phi}_{\tau_1 \times \tau_2} V &= \textbf{match } V \textbf{ with } (z_1, z_2).\ \ \lambda\{1.\, (\hat{\Phi}_{\tau_1} z_1),\ 2.\, (\hat{\Phi}_{\tau_2} z_2)\} \\
\hat{\Phi}_{\textbf{bool}} V &= \textbf{return}\, V \\
\hat{\Phi}_{\tau \to \tau'} V &= \lambda x : \mathbf{U}\, (\!|\tau|\!)^{\mathrm{n}}.\, \Psi_\tau(\textbf{force}\, x) \textbf{ to } y.\ (y \,\text{`}\, \textbf{force}\, V) \textbf{ to } z.\ \hat{\Phi}_{\tau'} z \\[1ex]
\Psi_{\textbf{unit}} N &= N \\
\Psi_{\tau_1 \times \tau_2} N &= \Psi_{\tau_1}(1\text{`}N) \textbf{ to } z_1.\, \Psi_{\tau_2}(2\text{`}N) \textbf{ to } z_2.\, \textbf{return}\, (z_1, z_2) \\
\Psi_{\textbf{bool}} N &= N \\
\Psi_{\tau \to \tau'} N &= \textbf{return}\, \textbf{thunk}\, \lambda x : (\!|\tau|\!)^{\mathrm{v}}.\, \Psi_{\tau'}\big((\textbf{thunk}\, (\hat{\Phi}_\tau x)) \,\text{`}\, N\big)
\end{aligned}
$$

Figure 6 Syntactic maps Φ from call-by-value to call-by-name and Ψ from call-by-name to call-by-value

$$\boxed{\phi_\tau : T[\![\tau]\!]^{\mathrm{v}} \to U_{\mathsf{T}}[\![\tau]\!]^{\mathrm{n}}}$$

$$\phi_\tau = \hat{\phi}_\tau^{\ddagger}$$

$$\boxed{\hat{\phi}_\tau : [\![\tau]\!]^{\mathrm{v}} \to U_{\mathsf{T}}[\![\tau]\!]^{\mathrm{n}}}$$

$$
\begin{aligned}
\hat{\phi}_{\textbf{unit}} &= \eta_1 \ :\ 1 \to T1 \\
\hat{\phi}_{\tau_1 \times \tau_2} &= \langle \hat{\phi}_{\tau_1} \circ \pi_1, \hat{\phi}_{\tau_2} \circ \pi_2 \rangle \ :\ [\![\tau_1]\!]^{\mathrm{v}} \times [\![\tau_2]\!]^{\mathrm{v}} \to U_{\mathsf{T}}[\![\tau_1]\!]^{\mathrm{n}} \times U_{\mathsf{T}}[\![\tau_2]\!]^{\mathrm{n}} \\
\hat{\phi}_{\textbf{bool}} &= \eta_2 \ :\ 2 \to T2 \\
\hat{\phi}_{\tau \to \tau'} &= \Lambda((\phi_{\tau'} \circ ev)^{\ddagger} \circ ([\![\tau \to \tau']\!]^{\mathrm{v}} \times \psi_\tau)) \ :\ [\![\tau]\!]^{\mathrm{v}} \Rightarrow T[\![\tau']\!]^{\mathrm{v}} \to U_{\mathsf{T}}[\![\tau]\!]^{\mathrm{n}} \Rightarrow U_{\mathsf{T}}[\![\tau']\!]^{\mathrm{n}}
\end{aligned}
$$

$$\boxed{\psi_\tau : U_{\mathsf{T}}[\![\tau]\!]^{\mathrm{n}} \to T[\![\tau]\!]^{\mathrm{v}}}$$

$$
\begin{aligned}
\psi_{\textbf{unit}} &= id_{T1} \ :\ T1 \to T1 \\
\psi_{\tau_1 \times \tau_2} &= seq_{[\![\tau_1]\!]^{\mathrm{v}}, [\![\tau_2]\!]^{\mathrm{v}}} \circ \langle \psi_{\tau_1} \circ \pi_1, \psi_{\tau_2} \circ \pi_2 \rangle \ :\ U_{\mathsf{T}}[\![\tau_1]\!]^{\mathrm{n}} \times U_{\mathsf{T}}[\![\tau_2]\!]^{\mathrm{n}} \to T([\![\tau_1]\!]^{\mathrm{v}} \times [\![\tau_2]\!]^{\mathrm{v}}) \\
\psi_{\textbf{bool}} &= id_{T2} \ :\ T2 \to T2 \\
\psi_{\tau \to \tau'} &= \eta_{[\![\tau \to \tau']\!]^{\mathrm{v}}} \circ (\hat{\phi}_\tau \Rightarrow \psi_{\tau'}) \ :\ U_{\mathsf{T}}[\![\tau]\!]^{\mathrm{n}} \Rightarrow U_{\mathsf{T}}[\![\tau']\!]^{\mathrm{n}} \to T([\![\tau]\!]^{\mathrm{v}} \Rightarrow T[\![\tau']\!]^{\mathrm{v}})
\end{aligned}
$$

Figure 7 Semantic morphisms ϕ from call-by-value to call-by-name and ψ from call-by-name to call-by-value

For the rest of this section, we show that in every model of CBPV that satisfies certain conditions (the assumptions of Theorem 22 below), the functions

$$\phi_\tau \circ - : \mathbf{C}(W, T[\![\tau]\!]^{\mathrm{v}}) \to \mathbf{C}(W, U_{\mathsf{T}}[\![\tau]\!]^{\mathrm{n}}) \quad \psi_\tau \circ - : \mathbf{C}(W, U_{\mathsf{T}}[\![\tau]\!]^{\mathrm{n}}) \to \mathbf{C}(W, T[\![\tau]\!]^{\mathrm{v}})$$

form a Galois connection for every τ and $W \in |\mathbf{C}|$. This is the key step in the proof of our reasoning principle (Theorem 25).

First we note that we cannot expect these maps to form Galois connections in general. Consider what happens when we convert a lazy pair $N = \lambda\{1.\, N_1, 2.\, N_2\}$ of type $(\!|\mathbf{unit} \times \mathbf{unit}|\!)^{\mathrm{n}} = \mathbf{F}\,\mathbf{unit} \mathbin{\underline{\times}} \mathbf{F}\,\mathbf{unit}$ into call-by-value, and then back into call-by-name:

$$\begin{aligned}\Phi_{\mathbf{unit}\times\mathbf{unit}}(\Psi_{\mathbf{unit}\times\mathbf{unit}}N) \;\equiv\; & \lambda\{1.\, N_1 \;\mathbf{to}\; z_1.\, N_2 \;\mathbf{to}\; z_2.\, \mathbf{return}\, z_1,\\ & \quad 2.\, N_1 \;\mathbf{to}\; z_1.\, N_2 \;\mathbf{to}\; z_2.\, \mathbf{return}\, z_2\}\end{aligned}$$

The ith projection of N evaluates only N_i, but the ith projection of $\Phi_{\mathbf{unit}\times\mathbf{unit}}(\Psi_{\mathbf{unit}\times\mathbf{unit}}N)$ evaluates both N_1 and N_2. We therefore cannot expect $\Phi_{\mathbf{unit}\times\mathbf{unit}}(\Psi_{\mathbf{unit}\times\mathbf{unit}}N) \preccurlyeq^{\Gamma}_{\mathrm{ctx}} N$ to hold in general (and semantically, cannot expect $\phi_{\mathbf{unit}\times\mathbf{unit}} \circ \psi_{\mathbf{unit}\times\mathbf{unit}} \circ [\![N]\!] \sqsubseteq [\![N]\!]$ to hold), because moving from the left-hand side to the right-hand side discards side-effects. There is also a problem in the other direction. Converting a strict pair M of type $\mathbf{F}\,(\!|\mathbf{unit} \times \mathbf{unit}|\!)^{\mathrm{v}} = \mathbf{F}\,(\mathbf{unit} \times \mathbf{unit})$ to call-by-name and back duplicates the side-effects of M:

$$\begin{aligned}\Psi_{\mathbf{unit}\times\mathbf{unit}}(\Phi_{\mathbf{unit}\times\mathbf{unit}}M) \;\equiv\; & M \;\mathbf{to}\; z\,.\,\mathbf{match}\; z \;\;\mathbf{with}\; (z_1, z_2).\\ & M \;\mathbf{to}\; z'.\,\mathbf{match}\; z' \;\mathbf{with}\; (z'_1, z'_2).\\ & \mathbf{return}\,(z_1, z'_2)\end{aligned}$$

Hence converting between the two interpretations of pairs can both discard and duplicate computations (just as call-by-name does).

There is a similar issue for functions. Consider what happens when we convert a CBPV computation $M : \mathbf{F}\,(\!|\mathbf{bool} \to \mathbf{bool}|\!)^{\mathrm{v}} = \mathbf{F}\,(\mathbf{U}\,(\mathbf{bool} \to \mathbf{F}\,\mathbf{bool}))$ to call-by-name and then back to call-by-value. By doing this we obtain a computation that immediately returns:

$$\Psi_{\mathbf{bool}\to\mathbf{bool}}(\Phi_{\mathbf{bool}\to\mathbf{bool}}M) \;\equiv\; \mathbf{return}\,\mathbf{thunk}\,\lambda x\!:\!\mathbf{bool}.\, M \;\mathbf{to}\; z.\, x\,\text{‘}\,\mathbf{force}\, z$$

The computation M may cause side-effects before producing a (thunk of a) function; but $\Psi_{\mathbf{bool}\to\mathbf{bool}}(\Phi_{\mathbf{bool}\to\mathbf{bool}}M)$ does not. Thus in general (e.g. if side-effects include mutable state), we cannot expect to have $[\![M]\!] \sqsubseteq \psi_{\mathbf{bool}\to\mathbf{bool}} \circ \phi_{\mathbf{bool}\to\mathbf{bool}} \circ [\![M]\!]$, and hence cannot expect to have a Galois connection. The round-trip from call-by-value to call-by-name and back thunks the side-effects of M.

Based on these observations, we add further constraints on models to ensure that we get Galois connections. This is where the restrictions on side-effects appear: the constraints are only satisfied in certain cases (they are for each of our three examples, but not e.g. for any adequate model of mutable state). They are that computations are *lax discardable*, *lax copyable* and *lax thunkable*. The non-lax versions of these properties were first defined by Führmann [6]. Table 1 defines the three properties for computations and for morphisms. Given any adequate model, each property on computations $\Gamma \vdash_c M : \mathbf{F}\,A$ holds whenever the corresponding property for the morphism $[\![M]\!]$ holds. The three constraints on side-effects have a similar role to centrality in Section 4.

To obtain Galois connections at function types below, we will assume that *all* morphisms $f : X \to TY$ are thunkable inside the model; this is equivalent to the underlying monad T being *lax idempotent*.

	Computations $\Gamma \vdash_c M : \mathbf{F}\,A$	Morphisms $f : X \to TY$
Lax discardable	M **to** $x.$ **return** $() \preccurlyeq^{\Gamma}_{\text{ctx}}$ **return** $()$	$\begin{tikzcd} X \arrow[r, "\langle\rangle_X"] \arrow[d, "f"'] \arrow[dr, phantom, "\sqsubseteq"] & 1 \arrow[d, "\eta_1"] \\ TY \arrow[r, "T\langle\rangle_Y"'] & T1 \end{tikzcd}$
Lax copyable	M **to** $x.$ **return** (x, x) $\preccurlyeq^{\Gamma}_{\text{ctx}}$ M **to** $x_1.$ M **to** $x_2.$ **return** (x_1, x_2)	$\begin{tikzcd} X \arrow[r, "\langle f,f\rangle"] \arrow[d, "f"'] \arrow[dr, phantom, "\sqsubseteq"] & TY \times TY \arrow[d, "seq_{Y,Y}"] \\ TY \arrow[r, "T\langle id,id\rangle"'] & T(Y \times Y) \end{tikzcd}$
Lax thunkable	M **to** $x.$ **return thunk return** x $\preccurlyeq^{\Gamma}_{\text{ctx}}$ **return thunk** M	$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "f"'] \arrow[dr, phantom, "\sqsubseteq"] & TY \arrow[d, "\eta_{TY}"] \\ TY \arrow[r, "T\eta_Y"'] & T(TY) \end{tikzcd}$

Table 1 The three side-effect axioms, defined for computations, morphisms

▶ **Definition 19.** *A strong* **Poset**-*monad* T *is* lax idempotent[3] *when* $T\eta_Y \sqsubseteq \eta_{TY}$ *for every object* $Y \in |\mathbf{C}|$.

This is a strong property. In particular, it implies all of the others:

▶ **Lemma 20.** *In any CBPV model for which* T *is lax idempotent, every morphism* $f : X \to TY$ *is lax discardable, lax copyable, and lax thunkable.*

To obtain Galois connections on product types, it is enough for morphisms to be lax discardable and lax copyable by the following lemma.

▶ **Lemma 21.** *Suppose that every morphism* $f : X \to TY$ *is lax discardable and lax copyable. If* $\phi_{\tau_i} \circ \psi_{\tau_i} \sqsubseteq id_{U_{\mathsf{T}}[\![\tau_i]\!]^{\mathrm{n}}}$ *and* $id_{T[\![\tau_i]\!]^{\mathrm{v}}} \sqsubseteq \psi_{\tau_i} \circ \phi_{\tau_i}$ *for each* $i \in \{1, 2\}$*, then*

$$U_{\mathsf{T}}\phi_{\tau_1 \times \tau_2} \circ \psi_{\tau_1 \times \tau_2} \sqsubseteq id_{U_{\mathsf{T}}[\![\tau_1]\!]^{\mathrm{n}} \times U_{\mathsf{T}}[\![\tau_2]\!]^{\mathrm{n}}} \qquad id_{T([\![\tau_1]\!]^{\mathrm{v}} \times [\![\tau_2]\!]^{\mathrm{v}})} \sqsubseteq \psi_{\tau_1 \times \tau_2} \circ \phi_{\tau_1 \times \tau_2}$$

To get Galois connections on function types, we require morphisms to be lax thunkable, i.e. that T is lax idempotent. Since this implies that morphisms are lax discardable and lax copyable, this is also enough for product types. We also trivially have Galois connections for **unit** and **bool**. Hence we arrive at the main theorem of this section:

▶ **Theorem 22.** *Given a CBPV model in which* T *is lax idempotent, the functions*

$$\phi_\tau \circ - : \mathbf{C}(W, T[\![\tau]\!]^{\mathrm{v}}) \to \mathbf{C}(W, U_{\mathsf{T}}[\![\tau]\!]^{\mathrm{n}}) \quad \psi_\tau \circ - : \mathbf{C}(W, U_{\mathsf{T}}[\![\tau]\!]^{\mathrm{n}}) \to \mathbf{C}(W, T[\![\tau]\!]^{\mathrm{v}})$$

form a Galois connection for every source-language type τ *and object* $X \in |\mathbf{C}|$.

Proof sketch. By induction on the type τ. This is trivial for **unit** and **bool**. For product types we use Lemma 21. For function types both of the required inequalities use the fact that T is lax idempotent. ◀

[3] Lax idempotent **Poset**-monads are a special case of lax idempotent 2-monads, which are well-known, and are often called *Kock-Zöberlein* monads [10].

If the model is adequate, it follows from this theorem that the maps Φ_τ and Ψ_τ on terms form a Galois connection (with respect to $\preccurlyeq_{\mathrm{ctx}}$), by Lemma 18.

▶ **Example 23.** Each of our three example models is adequate, and has a lax idempotent T. For no side-effects, we use the identity monad, which is trivially lax idempotent because $T\eta_Y = id_Y = \eta_{TY}$. For divergence, the monad 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, we have $T\eta_Y S = \downarrow\{\downarrow\{y\} \mid y \in S\} \subseteq \downarrow\{S\} = \eta_{TY} S$ because $\downarrow\{y\} \subseteq S$ for every $y \in S$ (intuitively, we can postpone nondeterministic choices). Thus we obtain Galois connections in each of these three cases.

6 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 we compose the call-by-name translation of each e with the maps Φ and Ψ defined above, to form a CBPV computation of the same type as the call-by-value translation:

$$(\!|\Gamma|\!)^{\mathrm{v}} \longrightarrow (\!|\Gamma|\!)^{\mathrm{n}} \xrightarrow{(\!|e|\!)^{\mathrm{n}}} (\!|\tau|\!)^{\mathrm{n}} \longrightarrow \mathbf{F}\,(\!|\tau|\!)^{\mathrm{v}}$$

We first define this composition precisely. The arrow on the right is just given by applying Ψ_τ. The arrow on the left is a substitution $\hat{\Phi}_\Gamma$ from terms in call-by-name contexts $(\!|\Gamma|\!)^{\mathrm{n}}$ to terms in call-by-value contexts $(\!|\Gamma|\!)^{\mathrm{v}}$. Given any source-language typing context $\Gamma = x_1 : \tau_1, \dots x_n : \tau_n$, we define $\hat{\Phi}_\Gamma = x_1 \mapsto \mathbf{thunk}\,(\hat{\Phi}_{\tau_1} x_1), \dots, x_n \mapsto \mathbf{thunk}\,(\hat{\Phi}_{\tau_n} x_n)$. If $\Gamma \vdash e : \tau$ then $(\!|\Gamma|\!)^{\mathrm{v}} \vdash_c \Psi_\tau\big((\!|e|\!)^{\mathrm{n}}[\hat{\Phi}_\Gamma]\big) : \mathbf{F}\,(\!|\tau|\!)^{\mathrm{v}}$, which has the same typing as $(\!|e|\!)^{\mathrm{v}}$. The statement we wish to prove is $(\!|e|\!)^{\mathrm{v}} \preccurlyeq_{\mathrm{ctx}} \Psi_\tau\big((\!|e|\!)^{\mathrm{n}}[\hat{\Phi}_\Gamma]\big)$. Again we reason using the denotational semantics. Given a CBPV model, the interpretation of the substitution $\hat{\Phi}_\Gamma$ is a morphism $\hat{\phi}_\Gamma : [\![\Gamma]\!]^{\mathrm{v}} \to [\![\Gamma]\!]^{\mathrm{n}}$, given by $\hat{\phi}_\diamond = id_1$ and $\hat{\phi}_{\Gamma,x:\tau} = \hat{\phi}_\Gamma \times \hat{\phi}_\tau$. If the model is adequate, to show our goal $(\!|e|\!)^{\mathrm{v}} \preccurlyeq_{\mathrm{ctx}} \Psi_\tau\big((\!|e|\!)^{\mathrm{n}}[\hat{\Phi}_\Gamma]\big)$, it suffices to show that $[\![e]\!]^{\mathrm{v}} \sqsubseteq \psi_\tau \circ [\![e]\!]^{\mathrm{n}} \circ \hat{\phi}_\Gamma$.

We show that this is the case directly using the properties of Galois connections, which allow us to push composition with ψ_τ into the structure of terms.

▶ **Lemma 24.** *In every CBPV model for which the functions*

$$\phi_\tau \circ - : \mathbf{C}(W, T[\![\tau]\!]^{\mathrm{v}}) \to \mathbf{C}(W, U_{\mathsf{T}}[\![\tau]\!]^{\mathrm{n}}) \quad \psi_\tau \circ - : \mathbf{C}(W, U_{\mathsf{T}}[\![\tau]\!]^{\mathrm{n}}) \to \mathbf{C}(W, T[\![\tau]\!]^{\mathrm{v}})$$

form Galois connections for all τ, W, we have $[\![e]\!]^{\mathrm{v}} \sqsubseteq \psi_\tau \circ [\![e]\!]^{\mathrm{n}} \circ \hat{\phi}_\Gamma$ for all $\Gamma \vdash e : \tau$.

Proof sketch. By induction on the derivation of $\Gamma \vdash e : \tau$. We give just the case for function applications $e\,e'$; which shows where having Galois connections is useful. The two inequalities below both involve use properties of Galois connections. The equalities follow from properties of products, exponentials, and T-algebras.

$$\begin{aligned}
[\![e\,e']\!]^{\mathrm{v}} &= (ev^\dagger \circ \langle \pi_2, [\![e']\!]^{\mathrm{v}} \circ \pi_1 \rangle)^\dagger \circ \langle id, [\![e]\!]^{\mathrm{v}} \rangle \\
&\sqsubseteq \psi_{\tau'} \circ \phi_{\tau'} \circ (ev^\dagger \circ \langle \pi_2, \psi_\tau \circ [\![e']\!]^{\mathrm{n}} \circ \hat{\phi}_\Gamma \circ \pi_1 \rangle)^\dagger \circ \langle id, \psi_{\tau\to\tau'} \circ [\![e]\!]^{\mathrm{n}} \circ \hat{\phi}_\Gamma \rangle \\
&= \psi_{\tau'} \circ \Lambda^{-1}(\phi_{\tau\to\tau'} \circ \psi_{\tau\to\tau'} \circ [\![e]\!]^{\mathrm{n}}) \circ \langle id, [\![e']\!]^{\mathrm{n}} \rangle \circ \hat{\phi}_\Gamma \\
&\sqsubseteq \psi_{\tau'} \circ \Lambda^{-1}[\![e]\!]^{\mathrm{n}} \circ \langle id, [\![e']\!]^{\mathrm{n}} \rangle \circ \hat{\phi}_\Gamma \\
&= \psi_{\tau'} \circ [\![e\,e']\!]^{\mathrm{n}} \circ \hat{\phi}_\Gamma
\end{aligned}$$

◀

Theorem 22 provides a sufficient condition for the maps between the two evaluation orders to form Galois connections. By combining this sufficient condition with the above lemma, we arrive at our reasoning principle, which we state formally as Theorem 25. Recall that a program relation $\preccurlyeq$ is a family of relations on CBPV programs, and that each program relation induces a contextual preorder $\preccurlyeq_{\mathrm{ctx}}$. Given any program relation $\preccurlyeq$, to show that the call-by-value and call-by-name translations of source-language expressions are related by $\preccurlyeq_{\mathrm{ctx}}$ it is enough to find an adequate model involving a lax idempotent T.

▶ **Theorem 25** (Relationship between call-by-value and call-by-name)**.** *Suppose we are given a program relation $\preccurlyeq$, and a model of CBPV that is adequate with respect to $\preccurlyeq$, and has a lax idempotent T. If $\Gamma \vdash e : \tau$ then*

$$(\!|e|\!)^{\mathrm{v}} \;\; \preccurlyeq_{\mathrm{ctx}} \;\; \Psi_\tau\big((\!|e|\!)^{\mathrm{n}}[\hat{\Phi}_\Gamma]\big)$$

Proof. Theorem 22 implies that ϕ and ψ form Galois connections, and then Lemma 24 implies $[\![e]\!]^{\mathrm{v}} \sqsubseteq \psi_\tau \circ [\![e]\!]^{\mathrm{n}} \circ \hat{\phi}_\Gamma$. The result follows from adequacy of the model and Lemma 18. ◀

The generality of this theorem comes from two sources. First, we consider arbitrary program relations $\preccurlyeq$. 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. We obtain a corollary about source-language *programs* (closed expressions of type **unit** or **bool**). This corollary is closer to the standard results that are proved for specific side-effects.

▶ **Corollary 26.** *If the assumptions of Theorem 25 hold, then for every closed expression e of type τ, where $\tau \in \{\mathbf{unit}, \mathbf{bool}\}$, we have $(\!|e|\!)^{\mathrm{v}} \preccurlyeq (\!|e|\!)^{\mathrm{n}}$.*

Proof. We have $[\![e]\!]^{\mathrm{v}} \sqsubseteq \psi_\tau \circ [\![e]\!]^{\mathrm{n}} \circ \hat{\phi}_\diamond = [\![e]\!]^{\mathrm{n}}$ because both ψ_τ and $\hat{\phi}_\diamond$ are identities. Adequacy implies $(\!|e|\!)^{\mathrm{v}} \preccurlyeq_{\mathrm{ctx}} (\!|e|\!)^{\mathrm{n}}$, and hence $(\!|e|\!)^{\mathrm{v}} \preccurlyeq (\!|e|\!)^{\mathrm{n}}$. ◀

Our reasoning principle also has a partial converse:

▶ **Lemma 27.** *If $[\![e]\!]^{\mathrm{v}} \sqsubseteq \psi_\tau \circ [\![e]\!]^{\mathrm{n}} \circ \hat{\phi}_\Gamma$ for each $\Gamma \vdash e : \tau$, then for $X \in \{1, 2\}$ we have $T\eta_X \sqsubseteq \eta_{TX}$, and every morphism $Y \to TX$ is lax thunkable.*

Proof sketch. The first step is to show that $id \sqsubseteq \psi_\tau \circ \phi_\tau$ for every τ, by applying the assumption to the expression $x : \mathbf{unit} \to \tau \vdash x\,() : \tau$. It follows for each $\tau' \in \{\mathbf{unit}, \mathbf{bool}\}$ that $id_{T(1 \Rightarrow T[\![\tau']\!]^{\mathrm{v}})} \sqsubseteq \psi_{\mathbf{unit} \to \tau'} \circ \phi_{\mathbf{unit} \to \tau'} = \eta_{1 \Rightarrow T[\![\tau']\!]^{\mathrm{v}}} \circ id_{1 \Rightarrow T[\![\tau']\!]^{\mathrm{v}}}{}^{\ddagger}$, and hence

$$T\eta_{[\![\tau']\!]^{\mathrm{v}}} \;=\; id_{T(T[\![\tau']\!]^{\mathrm{v}})} \circ T\eta_{[\![\tau']\!]^{\mathrm{v}}} \;\sqsubseteq\; \eta_{T[\![\tau']\!]^{\mathrm{v}}} \circ id_{T[\![\tau']\!]^{\mathrm{v}}}{}^{\dagger} \circ T\eta_{[\![\tau']\!]^{\mathrm{v}}} \;=\; \eta_{T[\![\tau']\!]^{\mathrm{v}}}$$

as required. ◀

We now return to our three examples. For each example, we take the adequate model defined in Section 3; in all three cases, the strong **Poset**-monad T is lax idempotent. After extending the inductive proof of Lemma 24 with cases for the extra syntax, we can apply our relationship between call-by-value and call-by-name (Theorem 25).

In particular, we can apply Corollary 26 to relate source-language programs. For no side-effects, this shows for each $e : \mathbf{bool}$ that there is some V such that $(\!|e|\!)^{\mathrm{v}} \Downarrow \mathbf{return}\,V$ and $(\!|e|\!)^{\mathrm{n}} \Downarrow \mathbf{return}\,V$. In other words, e evaluates to the same result in call-by-value and in call-by-name (since evaluation is deterministic). For divergence and for nondeterminism, the corollary says that $(\!|e|\!)^{\mathrm{v}} \Downarrow \mathbf{return}\,V$ implies $(\!|e|\!)^{\mathrm{n}} \Downarrow \mathbf{return}\,V$ for all V. Hence for divergence, if the call-by-value execution terminates with some result, the call-by-name execution terminates with the same result. For nondeterminism, all possible results of call-by-value executions are possible results of call-by-name executions.

6.1 Equivalent reasoning principles

Our reasoning principle (Theorem 25) shows that a call-by-value computation can be replaced with a call-by-name computation when the call-by-name computation is composed with our maps between evaluation orders. Here we have made an arbitrary choice: by composing the call-by-value computation with the maps instead, there are three other reasoning principles we can derive.

We use the properties of Galois connections to derive the three alternative reasoning principles from the one above. The first alternative, which converts the result of the call-by-value computation to call-by-name, is almost immediate: $\Phi_\tau (\!| e |\!)^{\mathrm{v}} \preccurlyeq_{\mathrm{ctx}} (\!| e |\!)^{\mathrm{n}}[\hat{\Phi}_\Gamma]$ because

$$\phi_\tau \circ [\![e]\!]^{\mathrm{v}} \ \sqsubseteq\ \phi_\tau \circ \psi_\tau \circ [\![e]\!]^{\mathrm{n}} \circ \hat{\phi}_\Gamma \ \sqsubseteq\ [\![e]\!]^{\mathrm{n}} \circ \hat{\phi}_\Gamma \tag{1}$$

where the first inequality is the reasoning principle we have already given, and the second is a property of Galois connections. The other two involve converting call-by-name contexts into call-by-value contexts. These are more difficult because call-by-name contexts contain unevaluated computations, but call-by-value contexts contain results of computations. Hence to convert a call-by-name context to a call-by-value context we must force the evaluation of computations.

To derive these last two reasoning principles, we first define a morphism $\psi_\Gamma : [\![\Gamma]\!]^{\mathrm{n}} \to T[\![\Gamma]\!]^{\mathrm{v}}$ for each source-language context Γ:

$$\psi_\diamond \ =\ \eta_1 \ :\ 1 \to T1 \qquad \psi_{\Gamma,x:\tau} \ =\ \mathit{seq} \circ (\psi_\Gamma \times \psi_\tau) \ :\ [\![\Gamma]\!]^{\mathrm{n}} \times U_{\mathsf{T}}[\![\tau]\!]^{\mathrm{n}} \to T([\![\Gamma]\!]^{\mathrm{v}} \times [\![\tau]\!]^{\mathrm{v}})$$

These morphisms map call-by-name contexts to call-by-value contexts in the denotational semantics.

The four reasoning principles are summarised in Figure 8. We give each of them syntactically, in the form of an instance of $\preccurlyeq_{\mathrm{ctx}}$, and semantically as an instance of $\sqsubseteq$. Each square in the figure is the interpretation of the instance of $\preccurlyeq_{\mathrm{ctx}}$ above it, so adequacy enables us to derive the reasoning principles on the syntax by showing these instances of $\sqsubseteq$. The reasoning principle on the top left is the one we derived as Theorem 25. The top right is Equation (1). The bottom two are the more complex ones that involve converting call-by-name contexts into call-by-value contexts.

The bottom two reasoning principles can be derived from the top left if T is lax idempotent. This can be shown by applying the following lemma to $f = (\!| e |\!)^{\mathrm{v}}$ and $g = (\!| e |\!)^{\mathrm{n}}$.

▶ **Lemma 28.** *If T is lax idempotent, then for each $f : [\![\Gamma]\!]^{\mathrm{v}} \to T[\![\tau]\!]^{\mathrm{v}}$ and $g : [\![\Gamma]\!]^{\mathrm{n}} \to U_{\mathsf{T}}[\![\tau]\!]^{\mathrm{n}}$, the following four inequalities are equivalent:*

$$f \sqsubseteq \psi_\tau \circ g \circ \hat{\phi}_\Gamma \qquad \phi_\tau \circ f \sqsubseteq g \circ \hat{\phi}_\Gamma \qquad \phi_\tau \circ f^\dagger \circ \psi_\Gamma \sqsubseteq g \qquad f^\dagger \circ \psi_\Gamma \sqsubseteq \psi_\tau \circ g$$

7 Related work

Comparing evaluation orders Plotkin [22] and many others (e.g. [8]) relate call-by-value and call-by-name. Crucially, they consider lambda-calculi with no side-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 side-effects. Other evaluation orders (such as call-by-need) have also been compared in similarly restricted settings [16, 17, 7]. We suspect our technique could also be adapted to these. It might also be possible to recast some of our work in terms of the *duality* between call-by-value and call-by-name [4, 2, 27].

$$\llparenthesis e \rrparenthesis^{\mathrm{v}} \precsim_{\mathrm{ctx}} \Psi_\tau\big(\llparenthesis e \rrparenthesis^{\mathrm{n}}[\hat{\Phi}_\Gamma]\big)$$

$$\begin{tikzcd}
\llbracket \Gamma \rrbracket^{\mathrm{v}} \arrow[r, "\hat{\phi}_\Gamma"] \arrow[d, "\llbracket e \rrbracket^{\mathrm{v}}"'] \arrow[dr, phantom, "\sqsubseteq"] & \llbracket \Gamma \rrbracket^{\mathrm{n}} \arrow[d, "\llbracket e \rrbracket^{\mathrm{n}}"] \\
T\llbracket \tau \rrbracket^{\mathrm{v}} & U_{\mathsf{T}}\llbracket \tau \rrbracket^{\mathrm{n}} \arrow[l, "\psi_\tau"]
\end{tikzcd}$$

$$\Phi_\tau \llparenthesis e \rrparenthesis^{\mathrm{v}} \precsim_{\mathrm{ctx}} \llparenthesis e \rrparenthesis^{\mathrm{n}}[\hat{\Phi}_\Gamma]$$

$$\begin{tikzcd}
\llbracket \Gamma \rrbracket^{\mathrm{v}} \arrow[r, "\hat{\phi}_\Gamma"] \arrow[d, "\llbracket e \rrbracket^{\mathrm{v}}"'] \arrow[dr, phantom, "\sqsubseteq"] & \llbracket \Gamma \rrbracket^{\mathrm{n}} \arrow[d, "\llbracket e \rrbracket^{\mathrm{n}}"] \\
T\llbracket \tau \rrbracket^{\mathrm{v}} \arrow[r, "\phi_\tau"'] & U_{\mathsf{T}}\llbracket \tau \rrbracket^{\mathrm{n}}
\end{tikzcd}$$

$$\begin{pmatrix} \Psi_{\tau_1}(\mathbf{force}\, x_1)\ \mathbf{to}\ x'_1. \\ \dots \\ \Psi_{\tau_n}(\mathbf{force}\, x_n)\ \mathbf{to}\ x'_n. \\ \Phi_\tau(\llparenthesis e \rrparenthesis^{\mathrm{v}}[\overline{x_i \mapsto x'_i}]) \end{pmatrix} \precsim_{\mathrm{ctx}} \llparenthesis e \rrparenthesis^{\mathrm{n}}$$

$$\begin{tikzcd}
T\llbracket \Gamma \rrbracket^{\mathrm{v}} \arrow[d, "(\llbracket e \rrbracket^{\mathrm{v}})^\dagger"'] \arrow[dr, phantom, "\sqsubseteq"] & \llbracket \Gamma \rrbracket^{\mathrm{n}} \arrow[l, "\psi_\Gamma"'] \arrow[d, "\llbracket e \rrbracket^{\mathrm{n}}"] \\
T\llbracket \tau \rrbracket^{\mathrm{v}} \arrow[r, "\phi_\tau"'] & \llbracket \tau \rrbracket^{\mathrm{n}}
\end{tikzcd}$$

$$\begin{pmatrix} \Psi_{\tau_1}(\mathbf{force}\, x_1)\ \mathbf{to}\ x'_1. \\ \dots \\ \Psi_{\tau_n}(\mathbf{force}\, x_n)\ \mathbf{to}\ x'_n. \\ \llparenthesis e \rrparenthesis^{\mathrm{v}}[\overline{x_i \mapsto x'_i}] \end{pmatrix} \precsim_{\mathrm{ctx}} \Psi_\tau \llparenthesis e \rrparenthesis^{\mathrm{n}}$$

$$\begin{tikzcd}
T\llbracket \Gamma \rrbracket^{\mathrm{v}} \arrow[d, "(\llbracket e \rrbracket^{\mathrm{v}})^\dagger"'] \arrow[dr, phantom, "\sqsubseteq"] & \llbracket \Gamma \rrbracket^{\mathrm{n}} \arrow[l, "\psi_\Gamma"'] \arrow[d, "\llbracket e \rrbracket^{\mathrm{n}}"] \\
T\llbracket \tau \rrbracket^{\mathrm{v}} & U_{\mathsf{T}}\llbracket \tau \rrbracket^{\mathrm{n}} \arrow[l, "\psi_\tau"]
\end{tikzcd}$$

Figure 8 Four equivalent reasoning principles (where $\Gamma \vdash e : \tau$ and $\Gamma = x_1 : \tau_1, \dots, x_n : \tau_n$)

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 [24] to relate direct and continuation semantics of the lambda calculus, and later used by others (e.g. [19, 11, 1, 5]). There are several differences with our approach. Reynolds constructs a logical relation between the two semantics, and uses this to establish a relationship with the two maps. We skip the logical relation step. Reynolds also 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 [25, 12, 26] 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.

8 Conclusions

In this paper, we give a general reasoning principle (Theorem 25) that relates the observable behaviour of terms under call-by-value and call-by-name. The reasoning principle works for various collections of side-effects, in particular, it enables us to obtain theorems about divergence and nondeterminism. It is about open expressions, and allows us to change evaluation order *within* programs. We obtain a result about call-by-value and call-by-name evaluations of *programs* as a corollary (Corollary 26). 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 side-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 side-effects (thunkable, etc.) that give rise to relationships between evaluation orders.

References

1 Robert Cartwright and Matthias Felleisen. Extensible denotational language specifications. In *Proceedings of the International Conference on Theoretical Aspects of Computer Software*, pages 244–272. Springer, 1994. doi:10.1007/3-540-57887-0_99.

2 Pierre-Louis Curien and Hugo Herbelin. The duality of computation. In *Proceedings of the Fifth ACM SIGPLAN International Conference on Functional Programming*, pages 233–243. ACM, 2000. doi:10.1145/351240.351262.

3 Marco Devesas Campos and Paul Blain Levy. A syntactic view of computational adequacy. In Christel Baier and Ugo Dal Lago, editors, *Foundations of Software Science and Computation Structures*, pages 71–87. Springer, 2018. doi:10.1007/978-3-319-89366-2_4.

4 Andrzej Filinski. Declarative continuations and categorical duality. Master's thesis, University of Copenhagen, 1989.

5 Andrzej Filinski. *Controlling Effects*. PhD thesis, Carnegie Mellon University, 1996.

6 Carsten Führmann. Direct models of the computational lambda-calculus. *Electronic Notes in Theoretical Computer Science*, 20:245–292, 1999. doi:10.1016/S1571-0661(04)80078-1.

7 Jennifer Hackett and Graham Hutton. Call-by-need is clairvoyant call-by-value. *Proc. ACM Program. Lang.*, 3(ICFP):114:1–114:23, 2019. doi:10.1145/3341718.

8 Jun Inoue and Walid Taha. Reasoning about multi-stage programs. *Journal of Functional Programming*, 26(e22), 2016. doi:10.1017/S0956796816000253.

9 Anders Kock. Monads on symmetric monoidal closed categories. *Arch. Math.*, 21(1):1–10, 1970. doi:10.1007/bf01220868.

10 Anders Kock. Monads for which structures are adjoint to units. *Journal of Pure and Applied Algebra*, 104(1):41–59, 1995. doi:10.1016/0022-4049(94)00111-U.

11 Jakov Kučan. Retraction approach to cps transform. *Higher Order Symbol. Comput.*, 11(2):145–175, 1998. doi:10.1023/A:1010012532463.

12 Julia L. Lawall and Olivier Danvy. Separating stages in the continuation-passing style transformation. In *Proceedings of the 20th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, pages 124–136. ACM, 1993. doi:10.1145/158511.158613.

13 Paul Blain Levy. Call-by-push-value: A subsuming paradigm. In Jean-Yves Girard, editor, *Typed Lambda Calculi and Applications*, pages 228–243. Springer, 1999. doi:10.1007/3-540-48959-2_17.

14 Paul Blain Levy. Adjunction models for call-by-push-value with stacks. *Electronic Notes in Theoretical Computer Science*, 69:248–271, 2003. CTCS'02, Category Theory and Computer Science. doi:https://doi.org/10.1016/S1571-0661(04)80568-1.

15 Paul Blain Levy. Call-by-push-value: Decomposing call-by-value and call-by-name. *Higher-Order and Symbolic Computation*, 19(4):377–414, 2006. doi:10.1007/s10990-006-0480-6.

16 John Maraist, Martin Odersky, David N. Turner, and Philip Wadler. Call-by-name, call-by-value, call-by-need, and the linear lambda calculus. In *Proceedings of the Eleventh Annual Mathematical Foundations of Programming Semantics Conference*, pages 370–392, 1995. doi:10.1016/S1571-0661(04)00022-2.

17 Dylan McDermott and Alan Mycroft. Extended call-by-push-value: Reasoning about effectful programs and evaluation order. In Luís Caires, editor, *Programming Languages and Systems*, pages 235–262. Springer, 2019. doi:10.1007/978-3-030-17184-1_9.

18 Austin Melton, David A Schmidt, and George E Strecker. Galois connections and computer science applications. In *Category Theory and Computer Programming*, pages 299–312. Springer, 1986. `doi:10.1007/3-540-17162-2_130`.

19 Albert R. Meyer and Mitchell Wand. Continuation semantics in typed lambda-calculi. In Rohit Parikh, editor, *Logics of Programs*, pages 219–224. Springer, 1985. `doi:10.1007/3-540-15648-8_17`.

20 Eugenio Moggi. Notions of computation and monads. *Inf. Comput.*, 93(1):55–92, 1991. `doi:10.1016/0890-5401(91)90052-4`.

21 A. M. Pitts. Operationally-based theories of program equivalence. In P. Dybjer and A. M. Pitts, editors, *Semantics and Logics of Computation*, Publications of the Newton Institute, pages 241–298. Cambridge University Press, 1997.

22 G.D. Plotkin. Call-by-name, call-by-value and the λ-calculus. *Theoretical Computer Science*, 1(2):125–159, 1975. `doi:10.1016/0304-3975(75)90017-1`.

23 John Power and Edmund Robinson. Premonoidal categories and notions of computation. *Mathematical Structures in Comp. Sci.*, 7(5):453–468, 1997. `doi:10.1017/S0960129597002375`.

24 John C. Reynolds. On the relation between direct and continuation semantics. In *Proceedings of the 2nd Colloquium on Automata, Languages and Programming*, pages 141–156. Springer, 1974. `doi:10.1007/978-3-662-21545-6_10`.

25 Amr Sabry and Matthias Felleisen. Reasoning about programs in continuation-passing style. In *Proceedings of the 1992 ACM Conference on LISP and Functional Programming*, pages 288–298. ACM, 1992. `doi:10.1145/141471.141563`.

26 Amr Sabry and Philip Wadler. A reflection on call-by-value. In *Proceedings of the First ACM SIGPLAN International Conference on Functional Programming*, pages 13–24. ACM, 1996. `doi:10.1145/232627.232631`.

27 Philip Wadler. Call-by-value is dual to call-by-name. In *Proceedings of the Eighth ACM SIGPLAN International Conference on Functional Programming*, pages 189–201. ACM, 2003. `doi:10.1145/944705.944723`.

28 Glynn Winskel. *The Formal Semantics of Programming Languages: An Introduction*. MIT Press, 1993. `doi:10.7551/mitpress/3054.001.0001`.