Graduality and Parametricity: Together Again for the First Time

MAX S. NEW, Northeastern University, USA
DUSTIN JAMNER, Northeastern University, USA
AMAL AHMED, Northeastern University, USA

Parametric polymorphism and gradual typing have proven to be a difficult combination, with no language yet produced that satisfies the fundamental theorems of each: parametricity and graduality. Notably, Toro, Labrada, and Tanter (POPL 2019) conjecture that for any gradual extension of System F that uses dynamic type generation, graduality and parametricity are "simply incompatible". However, we argue that it is not graduality and parametricity that are incompatible per se, but instead that combining the syntax of System F with dynamic type generation as in previous work necessitates type-directed computation, which we show has been a common source of graduality and parametricity violations in previous work.

We then show that by modifying the syntax of universal and existential types to make the type name generation explicit, we remove the need for type-directed computation, and get a language that satisfies both graduality and parametricity theorems. The language has a simple runtime semantics, which can be explained by translation to a statically typed language where the dynamic type is interpreted as a dynamically extensible sum type. Far from being in conflict, we show that the parametricity theorem follows as a direct corollary of a relational interpretation of the graduality property.

CCS Concepts: • **Theory of computation → Type structures**.

Additional Key Words and Phrases: gradual typing, graduality, polymorphism, parametricity, logical relation

ACM Reference Format:
Max S. New, Dustin Jamner, and Amal Ahmed. 2020. Graduality and Parametricity: Together Again for the First Time. *Proc. ACM Program. Lang.* 4, POPL, Article 46 (January 2020), 32 pages. https://doi.org/10.1145/3371114

1 INTRODUCTION

Gradually typed languages support freely mixing statically typed and dynamically code within a single language and enable a transition from dynamic to static typing [Siek and Taha 2006; Tobin-Hochstadt and Felleisen 2006, 2008]. They allow for stable, typed libraries to be used by ephemeral dynamically typed scripts with no manual programming overhead, streamlining a commonplace pattern in systems software. Furthermore, when some of these dynamically typed scripts inevitably become feature-rich software, static types can be gradually added to help with optimization, refactoring, type-based IDEs and documentation.

Gradually typed languages in the tradition of Siek and Taha [2006] are based on the presence of a dynamic type, written ?, which is the type of dynamically typed code and is treated specially by the type checker. For instance, if f is a statically typed function with type $\mathbb{I} \to \mathbb{B}$—where $\mathbb{I}$ and $\mathbb{B}$ represent integer and boolean types, respectively—and x is a dynamically typed input, then the

Authors' addresses: Max S. New, Khoury College of Computer Sciences, Northeastern University, USA, maxnew@ccs.neu.edu; Dustin Jamner, Khoury College of Computer Sciences, Northeastern University, USA, jamner.d@husky.neu.edu; Amal Ahmed, Khoury College of Computer Sciences, Northeastern University, USA, amal@ccs.neu.edu.

2475-1421/2020/1-ART46
https://doi.org/10.1145/3371114

application fx is allowed by the static type checker because it is "plausible" that x will actually satisfy the type $\mathbb{I}$ at runtime. But note here that since f has type $\mathbb{I} \to \mathbb{B}$, it was written with the expectation that it should only be applied to integers and may, for instance, use arithmetic operations on its argument. In a *sound* gradually typed language, this type information should be reliable: the programmer and compiler should be able to refactor or optimize the function f based on its type, which says it will only be used on values of type $\mathbb{I}$. In order to ensure that this expectation is met at runtime, the application fx is elaborated to a core language called a *cast calculus* where a cast is inserted and the application becomes $f(\langle \mathbb{I} \Leftarrow ? \rangle x)$. If at runtime x is a value that is incompatible with the type $\mathbb{I}$, such as a function value, then the cast will error and signal that the input failed to meet the function's type. While this means that the gradual language admits runtime errors, it ensures the soundness of the type for programmer reasoning and compiler optimization.

When designing the semantics of a gradual language, we must consider not just how programs *run*, but how their behavior *changes* throughout the development process. Specifically, a gradual language should ensure a smooth transition from dynamic to static typing, which is formalized in two properties called the *static* and *dynamic gradual guarantee* [Siek et al. 2015]. The static gradual guarantee states that making types more precise in a program makes it less likely that the program type-checks. Our focus in this paper is on the dynamic gradual guarantee, also called *graduality* [New and Ahmed 2018]. The graduality theorem provides a formalization for the intuition that making types more precise should not impact the *partial* correctness of the program itself. Specifically, it says that if the types in a program are made more precise, then either the more precise program errors, or exhibits the same behavior as before. This means that a programmer can add types to a portion of their program and know that the program as a whole still operates the same way, unless a new dynamic error is raised, in which case there is a flaw either in the code or in the new annotation that was introduced.

Languages can fail to satisfy the graduality theorem for a variety of reasons but a common culprit is *type-directed computation*. Whenever a form in a gradual language has behavior that is defined by inspection of the *type* of an argument, rather than by its *behavior*, there is a potential for a graduality violation, because the computation must be ensured to be *monotone* in the type. For instance, the Grace language supports a construct the designers call "structural type tests". That is, it includes a form M `is` A that checks if M has type A at runtime. Boyland [2014] show that care must be taken in designing the semantics of this construct if A is allowed to be an arbitrary type. For instance, it might seem reasonable to say that $(\lambda x : ?.x)$ `is` $\mathbb{I} \to \mathbb{I}$ should run to false because the function has type $? \to ?$. However, if we increase the precision of the types by changing the annotation, we get $(\lambda x : \mathbb{I}.x)$ `is` $\mathbb{I} \to \mathbb{I}$ which should clearly evaluate to true, violating the graduality principle. In such a system, we can't think of types as just properties whose precision can be tuned up or down: we also need to understand how changing the type might influence our use of type tests at runtime.

Gradual typing researchers have designed languages that support reasoning principles enabled by a variety of advanced static features—such as objects [Siek and Taha 2007; Takikawa et al. 2012], refinement types [Lehmann and Tanter 2017], union and intersection types [Castagna and Lanvin 2017], typestates [Wolff et al. 2011], effect tracking [Bañados Schwerter et al. 2014], subtyping [Garcia et al. 2016], ownership [Sergey and Clarke 2012], session types [Igarashi et al. 2017b], and secure information flow [Disney and Flanagan 2011; Fennell and Thiemann 2013; Toro et al. 2018]. As these typing features become more complicated, the behavior of casts can become sophisticated as well, and the graduality principle is a way of ensuring that these sophisticated mechanisms stay within programmer expectations.

1.1 Polymorphism and Runtime Sealing

Parametric polymorphism, in the form of universal and existential types, allows for abstraction over types within a program. Universal types, written $\forall X.A$, allow for the definition of functions that can be used at many different types. Dually, existential types provide a simple model of a module system. A value of type $\exists X.A$ can be thought of as a module that exports a newly defined type X and then a value A that may include X that gives the interface to the type. Languages with parametric polymorphism provide very strong reasoning principles regarding data abstraction, formalized by the relational parametricity theorem [Reynolds 1983].

The relational parametricity theorem captures the idea that an abstract type is truly opaque to its users: for instance, a consumer of a value of existential type $\exists X.A$ can only interact with X values using the capabilities provided by the interface type A. This allows programmers to use existential types to model abstract data types [Mitchell and Plotkin 1985]. For instance, the existential type $\exists X.X \times (X \to X) \times (X \to \mathbb{I})$ represents the type of an abstract functional counter. The X represents the state, the first component of the tuple is the initial state, the second component is an increment function, and the final component reads out an observable integer value from the state. One obvious example implementation would use $\mathbb{I}$ for X, 0 as the initial state, addition by 1 as the increment, and the identity function as the read-out. In a language with proper data abstraction, we should be able to guarantee that with this implementation, the read-out function should only ever produce positive numbers, because even though the type $\mathbb{I}$ allows for negative numbers, the interface only enables the construction of positive numbers. This pattern of reasoning naturally generalizes to sophisticated data structure invariants such as balanced trees, sorted lists, etc.

Polymorphic languages can fail to satisfy the parametricity theorem for a variety of reasons but one common culprit is *type-directed computation* on abstract types. For instance in Java, values of a generic type T can be cast to an arbitrary object type. If the type T happens to be instantiated with the same type as the cast, then all information about the value will be revealed, and data abstraction is entirely lost. The problem is that the behavior of this runtime type-cast is directed by the type of the input: at runtime the input must carry some information indicating its type so that this cast can be performed. A similar problem arises when naïvely combining gradual typing with polymorphism, as we will see in §2.

While parametric polymorphism ensures data abstraction by means of a static type discipline, *dynamic sealing* provides a means of ensuring data abstraction even in a dynamically typed language. To protect abstract data from exposure, a fresh "key" is generated and implementation code must "seal" any abstract values before sending them to untrusted parties, "unsealing" them when they are passed into the exposed interface. For instance, we can ensure data abstraction for an untyped abstract functional counter by generating a fresh key σ, and producing a tuple where the first component is a 0 sealed with σ, and the increment and read-out function unseal their inputs and the increment function seals its output appropriately. If this is the only way the seal σ is used in the program, then the abstraction is ensured. While the programmer receives less support from the static type checker, this runtime sealing mechanism gives much of the same abstraction benefits.

One ongoing research area has been to satisfactorily combine the static typing discipline of parametric polymorphism with the runtime mechanism of dynamic sealing in a gradually typed language [Ahmed et al. 2011, 2017; Igarashi et al. 2017a; Ina and Igarashi 2011; Toro et al. 2019; Xie et al. 2018]. However, no such language design so far proposed has satisfied both of the desired fundamental theorems: graduality for gradual typing and relational parametricity for parametric polymorphism. Recent work by Toro et al. [2019] claims to prove that graduality and parametricity are inherently incompatible, which backed by analogous difficulties for secure information flow [Toro et al. 2018] has led to the impression that the graduality property is incompatible with

parametric reasoning. This would be the *wrong conclusion to draw*, for the following two reasons. First, the claimed proof has a narrow applicability. It is based on the definition of their logical relation, which we show in §2.3 does not capture a standard notion of parametricity. Second, and more significantly, we should be careful not to conclude that graduality and parametricity are incompatible properties, and that language designs must choose one. In this paper, we reframe the problem: both are desirable, and should be demanded of any gradual or parametric language. The failure of graduality and parametricity in previous work can be interpreted not as an indictment of these properties, but rather points us to reconsider the combination of System F's syntax with runtime semantics based on dynamic sealing. In this paper, we will show that graduality and parametricity are not in conflict *per se*, by showing that by modifying System F's syntax to make the sealing visible, both properties are achieved. Far from being in opposition to each other, both graduality and parametricity can be proven using a *single* logical relation theorem (§6).

1.2 Overview

We summarize the contributions of this work as follows

- We identify type-directed computation as the common cause of graduality and parametricity violations in previous work on gradual polymorphism.
- We show that certain polymorphic programs in Toro et al. [2019]'s language GSF exhibit non-parametric behavior.
- We present a new surface language PolyG$^\nu$ that supports a novel form of universal and existential types where the creation of fresh types is exposed in a controlled way. The semantics of PolyG$^\nu$ is similar to previous gradual parametric languages, but the explicit type creation and sealing eliminates the need for type-directed computation.
- We elaborate PolyG$^\nu$ into an explicit cast calculus PolyC$^\nu$. We then give a translation from PolyC$^\nu$ into a typed target language, CBPV$_{\mathrm{OSum}}$, essentially call-by-push-value with polymorphism and an extensible sum type.
- We develop a novel logical relation that proves both graduality and parametricity for PolyG$^\nu$. Thus, we show that parametricity and graduality are compatible, and we strengthen the connection alluded to by New and Ahmed [2018] that graduality and parametricity are analogous properties.

Complete typing rules, definitions, and proofs are in the technical appendix [New et al. 2020].

2 GRADUALITY AND PARAMETRICITY, FRIENDS OR ENEMIES?

Next, we review the issues in constructing a polymorphic gradual language that satisfies parametricity and graduality that have arisen in previous work. We see in each case that the common obstacle to parametricity and graduality is the presence of type-directed computation. This motivates our own language design, which obviates the need for type-directed computation by making dynamic sealing explicit in code.

2.1 "Naïve" Attempt

Before considering any dynamic sealing mechanisms, let's see why the most obvious combination of polymorphism with gradual typing produces a language that does not maintain data abstraction. Consider a polymorphic function of type $\forall X.X \to \mathbb{B}$. In a language satisfying relational parametricity, we know that the function must treat its input as having abstract type X and so this input cannot have any influence on what value is returned. However, in a gradually typed language, any value can be *cast* using type ascriptions, such as in the function $\Lambda X.\lambda x : X.(x :: ?) :: \mathbb{B}$. Here :: represents a type ascription. In a gradually typed language, a term M of type A can be ascribed a type B if it

is "plausible" that an A is a B. This is typically formalized using a type consistency relation $\sim$ or more generally consistent subtyping relation $\lesssim$, but in either case, it is always plausible that an A is a ? and vice-versa, so in effect a value of any type can be cast to any other by taking a detour through the dynamic type. These ascriptions would then be elaborated to casts producing the term $\Lambda X.\lambda x : X.\langle \mathbb{B} \Leftarrow ? \rangle \langle ? \Leftarrow X \rangle x$ If this function is applied to any value that is not compatible with $\mathbb{B}$, then the function will error, but if passed a boolean, the natural substitution-based semantics would result in the value being completely revealed:

$$(\Lambda X.\lambda x : X.\langle \mathbb{B} \Leftarrow ? \rangle \langle ? \Leftarrow X \rangle x)[\mathbb{B}]\texttt{true} \mapsto^* \langle \mathbb{B} \Leftarrow ? \rangle \langle ? \Leftarrow \mathbb{B} \rangle \texttt{true} \mapsto^* \texttt{true}$$

The root-cause of this parametricity violation is that we allow casts like $\langle ? \Leftarrow X \rangle$ whose behavior depends on how X is instantiated. To construct a gradual language with strong data abstraction we must somehow avoid the dependency of $\langle ? \Leftarrow X \rangle$ on X. One option, is to ban casts like $\langle ? \Leftarrow X \rangle$ altogether. Syntactically, this means changing the notion of plausibility to say that ascribing a term of type X with the dynamic type ? is not allowed. This is possible using the system presented by Igarashi et al. [2017a] if you only allow Λs that use the "static" label. This is compatible with parametricity and graduality, but is somewhat against the spirit of gradual typing, where typically all programs could be written as dynamically typed programs, and dynamically typed functions can be used on values of any type. An alternative is to use dynamic sealing to allow casts like $\langle ? \Leftarrow X \rangle$, but ensure that their behavior does not depend on *how X is instantiated*.

2.2 Type-directed Sealing

In sealing-based gradual parametric languages like λB[Ahmed et al. 2011, 2017], we ensure that casts of abstract type do not depend on their instantiation by adding a layer of indirection. Instead of the usual β rule for polymorphic functions

$$(\Lambda X.M)[A] \mapsto M[A/X],$$

in λB, we dynamically generate a fresh type α and pass that in for X. This first of all means the runtime state must include a store of fresh types, written Σ. When reducing a type application, we generate a fresh type α and instantiate the function with this new type

$$\Sigma; (\Lambda X.M)[A] \mapsto \Sigma, \alpha := A; M[\alpha/X]$$

In this case, we interpret α as being a new tag on the dynamic type that tags values of type A but is different from all previously used tags. The casts involving α are treated like a new base type, incompatible with all existing types. However, if we look at the resulting term, it is not well-typed: if the polymorphic function has type $\forall X.B$, then $M[\alpha/X]$ has type $B[\alpha/X]$, but the context of this term expects it to be of type $B[A/X]$. To paper over this difference, λB wraps the substitution with a *type-directed coercion*, distinct from casts, that mediates between the two types:

$$\Sigma; (\Lambda X.M)[A] \mapsto \Sigma, \alpha := A; M[\alpha/X] : B[\alpha/X] \overset{+\alpha}{\Longrightarrow} B[A/X]$$

This type-directed coercion $M[\alpha/X] : B[\alpha/X] \overset{+\alpha}{\Longrightarrow} B[A/X]$ is the part of the system that performs the actual sealing and unsealing, and is defined by recursion on the type B. The $+\alpha$ indicates that we are *unsealing* values in positive positions and sealing at negative positions. For instance if $B = X \times \mathbb{B}$, and $X = \mathbb{B}$, then on a pair $(\texttt{seal}_\alpha\texttt{true}, \texttt{false})$ the coercion will unseal the sealed boolean on the left and leave the boolean on the right alone. If B is of function type, the definition will involve the dual coercion using $-\alpha$, which *seals* at positive positions. So for instance applying the polymorphic identity function will reduce as follows

$$\begin{array}{ll}
\Sigma; (\Lambda X.\lambda x : X.x)[\mathbb{B}]\texttt{true} & \mapsto \Sigma, \alpha := \mathbb{B}; (\lambda x : \alpha.x : \alpha \to \alpha \overset{+\alpha}{\Longrightarrow} \mathbb{B} \to \mathbb{B})\texttt{true} \\
\mapsto \Sigma, \alpha := \mathbb{B}; (\lambda x : \alpha.x)(\texttt{true} : X \overset{-\alpha}{\Longrightarrow} \alpha) : \alpha \overset{+\alpha}{\Longrightarrow} X & \mapsto \Sigma, \alpha := \mathbb{B}; (\lambda x : \alpha.x)(\texttt{seal}_\alpha \texttt{true}) : \alpha \overset{+\alpha}{\Longrightarrow} X \\
\mapsto \Sigma, \alpha := \mathbb{B}; \texttt{seal}_\alpha \texttt{true} : \alpha \overset{+\alpha}{\Longrightarrow} X & \mapsto \texttt{true}
\end{array}$$

While this achieves the goal of maintaining data abstraction, it unfortunately violates graduality, as first pointed out by Igarashi et al. [2017a]. The reason is that the coercion is a type-directed computation, this time directed by the type $\forall X.B$ of the polymorphic function, whose behavior observably differs at type X from its behavior at type ?. Specifically, a coercion $M : X \overset{-\alpha}{\Longrightarrow} \alpha$ results in sealing the result of M, whereas if X is replaced by dynamic, then $M : ? \overset{-\alpha}{\Longrightarrow} \alpha$ is an identity function. An explicit counter-example is given by modifying the identity function to include an explicit annotation. The term $M_1 = (\Lambda X.\lambda x : X.x :: X)[\mathbb{B}]\texttt{true}$ reduces by generating a seal α, sealing the input `true` with α, then unsealing it, finally producing `true`. On the other hand, if the type of the input were dynamic rather than X, we would get a term $M_2 = (\Lambda X.\lambda x : ?.(x :: X))[\mathbb{B}]\texttt{true}$. In this case, the input is *not* sealed by the implementation, and the ascription of X results in a failed cast since $\mathbb{B}$ is incompatible with α. The only difference between the two terms is a type annotation, meaning that $M_1 \sqsubseteq M_2$ in the term precision ordering (M_1 is more precise than M_2), and so the graduality theorem states that if M_1 does not error, it should behave the same as M_2, but in this case M_2 errors while M_1 does not. The problem here is that the type of the polymorphic function determines whether to seal or unseal the inputs and outputs, but graduality says that the behavior of the dynamic type must align with both abstract types X (indicating sealing/unsealing) and concrete types like $\mathbb{B}$ (indicating no sealing/unsealing). These demands are contradictory since dynamic code would have to simultaneously be opaque until unsealing and available to interact with. So we see that the attempt to remove the type-directed casts which break parametricity by using dynamic sealing led to the need for a type-directed coercion which breaks graduality.

2.3 To Seal, or not to Seal

The language GSF was introduced by Toro et al. [2019] to address several criticisms of the type system and semantics of λB. We agree with the criticisms of the type system and so we will focus on the semantic differences. GSF by design has the same violation of graduality as λB, but has different behavior when using casts.

One motivating example for GSF is what happens when casting the polymorphic identity function to have a dynamically typed output: $(((\Lambda X.\lambda x : X.x) :: \forall X.X \to ?)[\mathbb{I}]1) + 2$. In λB, the input 1 is sealed as dictated by the type, but the dynamically typed output is not unsealed when it is returned from the function, resulting in an error when we try to add it. Ahmed et al. [2011] argue that it should be a free theorem that the behavior of a function of type $\forall X.X \to ?$ should be independent of its argument: it always errors, diverges or it always returns the same dynamic value, based on the intuition that the dynamic type ? does not syntactically contain the free variable X, and that this free theorem holds in System F. This reasoning is suspect since at runtime, the dynamic type does include a case for the freshly allocated type X, so intuitively we should consider ? to include X (and any other abstract types in scope).

Toro et al. [2019] argue on the other hand that intuitively the identity function was written with the intention of having a sealed input that is returned and then unsealed, and so casting the program to be more dynamic should result in the same behavior and so the program should succeed. The function application runs to the equivalent of $\langle ? \Leftarrow \mathbb{I} \rangle 1$ which is then cast to $\mathbb{I}$ and added to 2, resulting in the number 3. The mechanism for achieving this semantics is a system of runtime *evidence*, based on the *Abstracting Gradual Typing* (AGT) framework [Garcia et al. 2016].

An intuition for the behavior is that the sealing is still type-directed, but rather than being directed by the *static* type of the function being instantiated, it is based on the most precise type that the function has had. So here because the function was originally of type $\forall X.X \to X$, the sealing behavior is given by that type.

However, while we agree that the analysis in Ahmed et al. [2011] is incomplete, the behavior in GSF is inherently non-parametric, because the polymorphic program produces values with *different* dynamic type tags based on what the input *type* is. As a user of this function, we should be able to replace the instantiating type $\mathbb{I}$ with $\mathbb{B}$ and give any boolean input and get related behavior at the type ?, but in the program $(((\Lambda X.\lambda x : X.x) :: \forall X.X \to ?)[\mathbb{B}]\texttt{true}) + 2$ the function application reduces to $\langle ? \Leftarrow \mathbb{B}\rangle\texttt{true}$ which errors when cast to $\mathbb{I}$. Intuitively, this behavior is not parametric because the first program places an $\mathbb{I}$ tag on its input, and the second places a $\mathbb{B}$ tag on its input.

The non-parametricity is clearer if we look at a program of type $\forall X.? \to \mathbb{B}$ and consider the following function, a constant function with abstract input type cast to have dynamic input:

$$\text{const} = (\Lambda X.\lambda x : X.\texttt{true}) :: \forall X.? \to \mathbb{B}$$

X now has no effect on static typing, so both const$[\mathbb{I}]3$ and const$[\mathbb{B}]$ are well-typed. However, since the sealing behavior is actually determined by the type $\forall X.X \to \mathbb{B}$, the program will try to seal its input after downcasting it to whatever type X is instantiated at. So the first program casts $\langle \mathbb{I} \Leftarrow ?\rangle\langle ? \Leftarrow \mathbb{I}\rangle 3$, which succeeds and returns `true`, while the second program performs the cast $\langle \mathbb{B} \Leftarrow ?\rangle\langle ? \Leftarrow \mathbb{I}\rangle 3$ which fails. In effect, we have implemented a polymorphic function that for any type X, is a recognizer of dynamically typed values for that type, returning `true` if the input matches X and erroring otherwise. Any implementation of this behavior would clearly require passing of some syntactic representation of types at runtime.

Formally, the GSF language does not satisfy the following *defining principle of relational parametricity*, as found in standard axiomatizations of parametricity such as Dunphy [2002]; Ma and Reynolds [1991]; Plotkin and Abadi [1993]. In a parametric language, the user of a term M of a polymorphic function type $\forall X.A \to B$ should be guaranteed that M will behave uniformly when instantiated multiple times. Specifically, a programmer should be able to instantiate M with two different types B_1, B_2 and choose any relation $R \in Rel[B_1, B_2]$ (where the notion of relation depends on the type of effects present), and be ensured that if they supply related inputs to the functions, they will get related outputs. Formally, for a Kripke-style relation, the following principle should hold:

$$\frac{M : \forall X.A \to B \qquad R \in \text{Rel}[B_1, B_2] \qquad (w, V_1, V_2) \in \mathcal{V}[\![A]\!]\rho[X \mapsto R]}{(w, M[B_1]V_1, M[B_2]V_2) \in \mathcal{E}[\![B]\!]\rho[X \mapsto R]}$$

Here w is a "world" that gives the invariants in the store and ρ is the relational interpretation of free variables. $\mathcal{V}[\![\cdot]\!]$ and $\mathcal{E}[\![\cdot]\!]$ are value and expression relations formalizing an approximation ordering on values and expressions respectively, and $X \mapsto R$ means that the relational interpretation of X is given by R.

Toro et al. [2019] use an unusual logical relation for their language based on a similar relation in Ahmed et al. [2017], so there is no direct analogue of the relational mapping $X \mapsto R$. Instead, the application extends the world with the association of α to R and the interpretation sends X to α. However, we can show that this parametricity principle is violated by *any* ρ we pick for the term const above, using the definition of $\mathcal{E}[\![\cdot]\!]$ given in [Toro et al. 2019][1]. Instantiating the lemma would give us that $(w, \text{const}[\mathbb{I}]3, \text{const}[\mathbb{B}]3) \in \mathcal{E}[\![\mathbb{B}]\!]\rho$ since $(w, 3, 3) \in \mathcal{V}[\![?]\!]\rho$ for any ρ. The definition of $\mathcal{E}[\![\mathbb{B}]\!]\rho$ then says (again for any ρ) that it should be the case that since const$[\mathbb{I}]3$ runs to a value, it should also be the case that const$[\mathbb{B}]3$ runs to a value as well, but in actuality it errors, and so this parametricity principle must be false.

[1] they use slightly different notation, but we use notation that matches the logical relation we present later

How can the above parametricity principle be false when Toro et al. [2019] prove a parametricity theorem for GSF? We have not found a flaw in their *proof*, but rather a mismatch between their *theorem* statement and the expected meaning of parametricity. The definition of $\mathcal{V}[\![\forall X.A]\!]$ in Toro et al. [2019] is not the usual interpretation, but rather is an adaptation of a non-standard definition used in Ahmed et al. [2017]. Neither of their definitions imply the above principle, so we argue that neither paper provides a satisfying proof of parametricity. With GSF, we see that the above behavior violates some expected parametric reasoning, using the definition of $\mathcal{V}[\![?]\!]$ given in Toro et al. [2019]. With λB, we know of no counterexample to the above principle, and we conjecture that it would satisfy a more standard formulation of parametricity.

It is worth noting that the presence of effects—such as nontermination, mutable state, control effects—requires different formulations of the logical relation that defines parametricity. However, those logical relations capture parametricity in that they always formalize *uniformity of behavior* across different type instantiations. For instance, for a language that supports nontermination, the logical relation for parametricity ensures that two different instantiations have the same termination behavior: either both diverge, or they both terminate with related values. Because of this, the presence of effects usually leads to weaker free theorems—in pure System F all inhabitants of $\forall X.X \to X$ are equivalent to the identity function, but in System F with non-termination, every inhabitant is either the identity or *always* errors. Though the free theorems are weaker, parametricity still ensures *uniformity* of behavior. As our counterexample above ($\mathrm{const}[\mathbb{I}]3$ vs. $\mathrm{const}[\mathbb{B}]3$) illustrates, GSF is non-parametric since it does not ensure uniform behavior. However, since the difference in behavior was between error and termination, it is possible that GSF satisfies a property that could be called "partial parametricity" (or parametricity modulo errors) that weakens the notion of uniformity of behavior: either one side errors or we get related behaviors. However, it is not clear to us how to formulate the logical relation for the dynamic type to prove this. We show how this weakened reasoning in the presence of ? compares to reasoning in our language PolyG$^\nu$ in §6.4.

Our counter-example crucially uses the dynamic type, and we conjecture that when the dynamic type does not appear under a quantifier, that the usual parametric reasoning should hold in GSF. This would mean that in GSF once polymorphic functions become "fully static", they support parametric reasoning, but we argue that it should be the goal of gradual typing to support type-based reasoning even in the presence of dynamic typing, since migration from dynamic to static is a gradual process, possibly taking a long time or never being fully completed.

2.4 Resolution: Explicit Sealing

Summarizing the above examples, we see that

(1) The naïve semantics leads to type-directed casts at abstract types, violating parametricity.
(2) λB's type-directed sealing violates graduality because of the ambiguity of whether or not the dynamic type indicates sealing/unsealing or not.
(3) GSF's variant of type-directed sealing based on the most precise type violates graduality and parametricity because the polymorphic function gets to determine which dynamically typed values are sealed (i.e. abstract) and which are not.

We see that in each case, the use of a type-directed computational step leads to a violation of graduality or parametricity. The GSF semantics makes the type-directed sealing of λB more flexible by using the runtime evidence attached to the polymorphic function rather than the type at the instantiation point, but unfortunately this makes it impossible for the continuation to reason about which dynamically typed values it passes will be treated as abstract or concrete. This analysis motivates our own language design PolyG$^\nu$, where

(1) We depart from the syntax of System F.
(2) Sealing/unsealing of values is explicit and programmable, rather than implicit and type-directed.
(3) The party that *instantiates* an abstract type is the party that determines which values are sealed and unsealed. For existential types, this is the package (i.e., the module) and dually for universal types it is the continuation of the instantiation.

The dynamic semantics of PolyG$^\nu$ are similar to λB without the type-directed coercions, removing the obstacle to proving the graduality theorem. By allowing user-programmable sealing and unsealing, more complicated forms of sealing and unsealing are possible: for instance, we can seal every prime number element of a list, which would require a very rich type system to express using type-directed sealing! We conjecture that the language is strictly more expressive than λB in the sense of Felleisen [1990]: λB should be translatable into PolyG$^\nu$ in a way that simulates its operational semantics. Because the sealing is performed by the instantiating party rather than the abstracting party, the expressivity of PolyG$^\nu$ is incomparable to GSF.

3 PolyG$^\nu$: A GRADUAL LANGUAGE WITH POLYMORPHISM AND SEALING

Next, we present a our gradual language, PolyG$^\nu$, that supports a variant of existential and universal quantification while satisfying parametricity and graduality. The language has some unusual features, so we start with an extended example to illustrate what programs look like, and then in § 3.2 introduce the formal syntax and typing rules.

3.1 PolyG$^\nu$ Informally

Let's consider an example of existential types, since they are simpler than universal types in PolyG$^\nu$. In a typed, non-gradual language, we can define an abstract "flipper" type, $\texttt{FLIP} = \exists X.X \times (X \to X) \times (X \to \mathbb{B})$. The first element is the initial state, the second is a "toggle" function and the last element reads out the value as a concrete boolean.

Then we could create an instance of this abstract flipper using booleans as the carrier type X and negation as the toggle function $\texttt{pack}(\mathbb{B}, (\texttt{true}, (\texttt{NOT}, \texttt{ID})))\ \texttt{as}\ \texttt{FLIP}$. Note that we must explicitly mark the existential package with a type annotation, because otherwise we wouldn't be able to tell which occurrences of $\mathbb{B}$ should be hidden and which should be exposed. With different type annotations, the same package could be given types $\exists X.\mathbb{B} \times (\mathbb{B} \to \mathbb{B}) \times (\mathbb{B} \to \mathbb{B})$ or $\exists X.X \times (X \to X) \times (X \to X)$.

The PolyG$^\nu$ language existential type works differently in a few ways. We write $\exists^\nu$ rather than $\exists$ to emphasize that we are only quantifying over fresh types, and not arbitrary types. The equivalent of the above existential package would be written as

$$\texttt{pack}^\nu(X \cong \mathbb{B}, (\texttt{seal}_X\texttt{true}, ((\lambda x : X.\texttt{seal}_X(\texttt{NOT}(\texttt{unseal}_X x))), (\lambda x : X.\texttt{unseal}_X x))) : \texttt{FLIP}$$

The first thing to notice is that rather than just providing a type $\mathbb{B}$ to instantiate the existential, we write a declaration $X \cong \mathbb{B}$. The X here is a *binding position* and the body of the package is typed under the assumption that $X \cong \mathbb{B}$. Then, rather than *substituting* $\mathbb{B}$ for X when typing the body of the package, the type checker checks that the body has type $X \times ((X \to X) \times (X \to \mathbb{B}))$ under the assumption that $X \cong \mathbb{B}$:

$$X \cong \mathbb{B} \vdash (\texttt{seal}_X\texttt{true}, ((\lambda x : X.\texttt{seal}_X(\texttt{NOT}(\texttt{unseal}_X x))), (\lambda x : X.\texttt{unseal}_X x))) : X \times ((X \to X) \times (X \to \mathbb{B}))$$

Crucially, $X \cong \mathbb{B}$ is a *weaker* assumption than $X = \mathbb{B}$. In particular, there are no *implicit* casts from X to $\mathbb{B}$ or vice-versa, but the programmer can *explicitly* "seal" $\mathbb{B}$ values to be X using the form $\texttt{seal}_X M$, which is only well-typed under the assumption that $X \cong A$ for some A consistent with $\mathbb{B}$. We also get a corresponding unseal form $\texttt{unseal}_X M$, and the runtime semantics in § 4.4 defines these to be a bijection. At runtime, X will be a freshly generated type with its own tag on

the dynamic type. An interesting side-effect of making the difference between X and $\mathbb{B}$ explicit in the term is that existential packages do not require type annotations to resolve any ambiguities. For instance, unlike in the typed case, the gradual package above could *not* be ascribed the type $\exists^\nu X.\mathbb{B} \times ((\mathbb{B} \to \mathbb{B}) \times (\mathbb{B} \to \mathbb{B}))$ because the functions explicitly take X values, and not $\mathbb{B}$ values.

The corresponding elimination form for $\exists^\nu$ is a standard unpack: $\texttt{unpack}\ (X, x) = M; N$, where the continuation for the unpack is typed with just X and x added to the context, it doesn't know that $X \cong A$ for any particular A. We call this ordinary type variable assumption an *abstract type variable*, whereas the new assumption $X \cong A$ is a *known type variable* which acts more like a *type definition* than an abstract type. At runtime, when an existential is unpacked, a fresh type X is created that is isomorphic to A but whose behavior with respect to casts is different.

While explicit sealing and unsealing might seem burdensome to the programmer, note that this is directly analogous to a common pattern in Haskell, where modules are used in combination with `newtype` to create a datatype that at runtime is represented in the same way as an existing type, but for type-checking purposes is considered distinct. We give an analogous Haskell module as follows:

```
module Flipper(State, start, toggle, toBool) where
newtype State = Seal { unseal :: Bool }
start :: State
start = Seal True
inc :: State -> State
inc s = Seal (not (unseal s)
toBool :: State -> Bool
toBool = unseal
```

Then a different module that imports `Flipper` is analogous to an unpack, as its only interface to the `State` type is through the functions provided.

We also add universal quantification to the language, using the duality between universals and existentials as a guide. Again we write the type differently, as $\forall^\nu X.A$. In an ordinary polymorphic language, we would write the type of the identity function as $\forall X.X \to X$ and implement it using a Λ form: $\Lambda X.\lambda x : X.x$. The elimination form passes in a type for X. For instance applying the identity function to a boolean would be written as $\texttt{ID}\ [\mathbb{B}]\ \texttt{true}$. And a free theorem tells us that this term must either diverge, error, or return `true`.

The introduction form Λ is dual to the unpack form, and correspondingly looks the same as the ordinary Λ, for example in the identity function $\texttt{ID}^\nu = \Lambda^\nu X.\lambda x : X.x : \forall^\nu X.X \to X$. The body of the Λ^ν is typed with an abstract type variable X in scope. The elimination form of type application is dual to the *pack* form, and so similarly introduces a *known* type variable assumption. Instantiating the identity function as above would be written as $\texttt{unseal}_X(\texttt{ID}^\nu\{X \cong \mathbb{B}\}(\texttt{seal}_X\texttt{true})) : \mathbb{B}$. which introduces a known type variable $X \cong \mathbb{B}$ into the context. Rather than the resulting type being $\mathbb{B} \to \mathbb{B}$, it is $X \to X$ with the assumption $X \cong \mathbb{B}$. Then the argument to the function must be explicitly sealed as an X to be passed to the function. The output of the function is also of type X and so must be explicitly *unsealed* to get a boolean out. However, there is something quite unusual about this term: the $X \cong \mathbb{B}$ binding site is not binding X in a *subterm* of the application, but rather into the *context*: the argument is sealed, and the continuation is performing an unseal! These bindings in $\forall^\nu$ instantiations follow this "inside-out" structure and complicate the typing rules: every term in the language "exports" known type variable bindings that go outwards in addition to the other typing assumptions coming inwards from the context. While unusual, they are intuitively justified by the duality with existentials: we can think of the *continuation* for an instantiation of a $\forall^\nu$ as being analogous to the *body* of the existential package.

To get an understanding of how PolyG$^\nu$ compares to λB and GSF and why it avoids their violation of graduality, let's consider how we might write the examples from the previous section. In PolyG$^\nu$,

$$
\begin{array}{llll}
\text{types} & A & ::= & ? \mid X \mid \mathbb{B} \mid A \times A \mid A \to A \mid \exists^{\nu} X.A \mid \forall^{\nu} X.A \\
\text{Ground types} & G & ::= & X \mid \mathbb{B} \mid ? \times ? \mid ? \to ? \mid \exists^{\nu} X.? \mid \forall^{\nu} X.? \\
\text{terms} & M & ::= & x \mid M :: A \mid \mathtt{seal}_X M \mid \mathtt{unseal}_X M \mid \mathtt{is}(G)?\ M \mid \mathtt{true} \mid \mathtt{false} \\
 & & \mid & \mathtt{if}\ M\ \mathtt{then}\ M\ \mathtt{else}\ M \mid (M, M) \mid \mathtt{let}\ (x, x) = M; M \\
 & & \mid & M\ M \mid \lambda x : A.M \mid \mathsf{pack}^{\nu}(X \cong A, M) \mid \mathsf{unpack}\ (X, x) = M; N \\
 & & \mid & \Lambda^{\nu} X.M \mid M\{X \cong A\} \mid \mathtt{let}\ x = M; M \\
\text{environment} & \Gamma & ::= & \cdot \mid \Gamma, x : A \mid \Gamma, X \mid \Gamma, X \cong A
\end{array}
$$

Fig. 1. PolyG$^{\nu}$ Syntax

if we apply a function of type $\forall X.X \to X$, we have to mark explicitly that the input is sealed, and furthermore if we want to use the output as a boolean, we must *unseal* the output:

$$\mathtt{unseal}_X((\Lambda X.\lambda x : X.x :: X)\{X \cong \mathbb{B}\}(\mathtt{seal}_X \mathtt{true})) \mapsto^* \mathtt{true}$$

Then if we change the type of the input from X to ? the explicit sealing and unsealing remain, so even though the input is dynamically typed it will still be a sealed boolean, and the program exhibits the same behavior:

$$\mathtt{unseal}_X((\Lambda X.\lambda x : ?.x :: X)\{X \cong \mathbb{B}\}(\mathtt{seal}_X \mathtt{true})) \mapsto^* \mathtt{true}$$

If we remove the seal of the input, then the cast to X in the function will fail, giving us the behavior of λB/GSF:

$$\mathtt{unseal}_X((\Lambda X.\lambda x : ?.x :: X)\{X \cong \mathbb{B}\}\mathtt{true}) \mapsto^* \mathfrak{V}$$

but crucially this involved changing the term, not just the type, so the graduality theorem does not tell us that the programs should have related behavior.

Next, let's consider the parametricity violation from GSF. When we instantiate the constant function, we need to decide if the argument is sealed or not. We get the behavior of GSF when we instantiate with $\mathbb{I}$ and seal the input 3:

$$\mathrm{const}\{X \cong \mathbb{I}\}\mathtt{seal}_X 3 \mapsto^* \mathtt{true}$$

However, if we try to write the analogous program with $\mathbb{B}$: instead of $\mathbb{I}$ $\mathrm{const}\{X \cong \mathbb{B}\}\mathtt{seal}_X 3$ then the program is not well typed because $X \cong \mathbb{B}$ and 3 has type $\mathbb{I}$ which is not compatible. We can replicate the outcome of the GSF program by *not* sealing the 3:

$$\mathrm{const}\{X \cong \mathbb{B}\}3 \mapsto^* \mathfrak{V}$$

But this is not a parametricity violation because the 3 here will be embedded at the dynamic type with the $\mathbb{I}$ tag, whereas above the 3 was tagged with the X tag, which is not related.

3.2 PolyG$^{\nu}$ Formal Syntax and Semantics

Figure 1 presents the syntax of the surface language types, terms and environments. Most of the language is a typical gradual functional language, using ? as the dynamic type, and including type ascription $M :: A$. The unusual aspects of the language are the $\mathtt{seal}_X M$ and $\mathtt{unseal}_X M$ forms and the "fresh" existential $\exists^{\nu} X.A$ and universal $\forall^{\nu} X.A$. Note also the non-standard environments Γ, which include ordinary typing assumptions $x : A$, abstract type variable assumptions X and known type variable assumptions $X \cong A$. For simplicity, we assume freshness of all type variable bindings, i.e. when we write Γ, X or $\Gamma, X \cong A$ that X does not occur in Γ.

The typing rules are presented in Figure 2. On a first pass, we suggest ignoring all shaded parts of the rules, which only concern the inside-out scoping needed for the $\forall^{\nu} X.A$ forms and would not be necessary if this type was removed. We follow the usual formulation of gradual surface languages in the style of [Siek and Taha 2006]: type checking is strict when checking compatibility of different connectives, but lax when the dynamic type is involved. The first $M :: B$ form is type-ascription, which is well formed when the types are *consistent* with each other, written $A \sim B$. We define this

$$\frac{\Gamma \vdash M : A; \Gamma_o \qquad A \sim B}{\Gamma \vdash (M :: B) : B; \Gamma_o} \qquad \frac{x : A \in \Gamma}{\Gamma \vdash x : A; \cdot} \qquad \frac{\Gamma \vdash M : A; \Gamma_M \qquad \Gamma, \Gamma_M, x : A \vdash N : B; \Gamma_N}{\Gamma \vdash \mathtt{let}\ x = M; N : B; \Gamma_M, \Gamma_N}$$

$$\frac{\Gamma \vdash M : B; \Gamma_o \qquad X \cong A \in \Gamma, \Gamma_o \qquad B \sim A}{\Gamma \vdash \mathtt{seal}_X M : X; \Gamma_o} \qquad \frac{\Gamma \vdash M : B; \Gamma_o \qquad X \cong A \in \Gamma, \Gamma_o \qquad B \sim X}{\Gamma \vdash \mathtt{unseal}_X M : A; \Gamma_o}$$

$$\frac{\Gamma \vdash M : A; \Gamma_o \qquad \Gamma, \Gamma_o \vdash G}{\Gamma \vdash \mathtt{is}(G)?\ M : \mathbb{B}; \Gamma_o} \qquad \Gamma \vdash \mathtt{true} : \mathbb{B}; \cdot \qquad \Gamma \vdash \mathtt{false} : \mathbb{B}; \cdot$$

$$\frac{\Gamma \vdash M : A; \Gamma_M \qquad A \sim \mathbb{B} \qquad \Gamma, \Gamma_M \vdash N_t : B_t; \Gamma_t \qquad \Gamma, \Gamma_M \vdash N_f : B_f; \Gamma_f}{\Gamma \vdash \mathtt{if}\ M\ \mathtt{then}\ N_t\ \mathtt{else}\ N_f : B_t \sqcap B_f; \Gamma_M, \Gamma_t \cap \Gamma_f}$$

$$\frac{\Gamma \vdash M_1 : A_1; \Gamma_1 \qquad \Gamma, \Gamma_1 \vdash M_2 : A_2; \Gamma_2}{\Gamma \vdash (M_1, M_2) : A_1 \times A_2; \Gamma_1, \Gamma_2} \qquad \frac{\Gamma \vdash M : A; \Gamma_M \qquad \Gamma, \Gamma_M, x : \pi_1(A), y : \pi_2(A) \vdash N : B; \Gamma_N}{\Gamma \vdash \mathtt{let}\ (x, y) = M; N : B; \Gamma_M, \Gamma_N}$$

$$\frac{\Gamma \vdash A \qquad \Gamma, x : A \vdash M : B; \Gamma_o}{\Gamma \vdash \lambda x : A.M : A \to B; \cdot} \qquad \frac{\Gamma \vdash M : A; \Gamma_M \qquad \Gamma, \Gamma_M \vdash N : B; \Gamma_N \qquad \mathrm{dom}(A) \sim B}{\Gamma \vdash M\ N : \mathrm{cod}(A); \Gamma_M, \Gamma_N}$$

$$\frac{\Gamma, X \cong A \vdash M : B; \Gamma_o}{\Gamma \vdash \mathsf{pack}^\nu(X \cong A, M) : \exists^\nu X.B; \cdot} \qquad \frac{\Gamma \vdash M : A; \Gamma_M \qquad \Gamma, \Gamma_M, X, x : \mathrm{un}\exists^\nu(A) \vdash N : B; \Gamma_N \qquad \Gamma, \Gamma_M, \Gamma_N|_X \vdash B}{\Gamma \vdash \mathsf{unpack}\ (X, x) = M; N : B; \Gamma_M, \Gamma_N|_X}$$

$$\frac{\Gamma, X \vdash M : A; \Gamma_o}{\Gamma \vdash \Lambda^\nu X.M : \forall^\nu X.A; \cdot} \qquad \frac{\Gamma \vdash M : A; \Gamma_M \qquad \Gamma, \Gamma_M \vdash B}{\Gamma \vdash M\{X \cong B\} : \mathrm{un}\forall^\nu(A); \Gamma_M, X \cong B}$$

$$? \sim A \qquad A \sim ? \qquad \mathbb{B} \sim \mathbb{B} \qquad X \sim X$$

$$\frac{A_i \sim B_i \qquad A_o \sim B_o}{A_i \to A_o \sim B_i \to B_o} \qquad \frac{A_1 \sim B_1 \qquad A_2 \sim B_2}{A_1 \times A_2 \sim B_1 \times B_2} \qquad \frac{A \sim B}{\exists^\nu X.A \sim \exists^\nu X.B} \qquad \frac{A \sim B}{\forall^\nu X.A \sim \forall^\nu X.B}$$

$$\begin{aligned}
\mathrm{dom}(A \to B) &= A & \mathrm{un}\exists^\nu(\exists^\nu X.A) &= A \\
\mathrm{dom}(?) &= ? & \mathrm{un}\exists^\nu(?) &= ? \\
\mathrm{cod}(A \to B) &= B & & \\
\mathrm{cod}(?) &= ? & & \\
\pi_i(A_1 \times A_2) &= A_i & & \\
\pi_i(?) &= ? & & \\
\mathrm{un}\forall^\nu(\forall^\nu X.A) &= A & & \\
\mathrm{un}\forall^\nu(?) &= ? & &
\end{aligned}$$

$$\cdot|_{\Gamma'} = \cdot$$
$$(X \cong A, \Gamma)|_{\Gamma'} = X \cong A, (\Gamma|_{\Gamma'}) \quad (\mathrm{FV}(A) \cap \Gamma' = \emptyset)$$
$$(X \cong A, \Gamma)|_{\Gamma'} = \Gamma|_{\Gamma', X} \quad (\mathrm{FV}(A) \cap \Gamma' \neq \emptyset)$$

Fig. 2. PolyG$^\nu$ Type System

in the standard way in as being the least congruence relation including equality and rules making ? consistent with every type.

We include variable and let-binding rules, which are standard other than the shaded parts. Next, we include sealing $\mathtt{seal}_X M$ and unsealing $\mathtt{unseal}_X M$ forms. The sealing and unsealing forms are

valid when the assumption $X \cong A$ is in the environment and give the programmer access to an explicit bijection between the types X and A. It is crucial for graduality to hold that this bijection is explicit and not implicit, because the behavior of casts involving X and A are very different. To show that it has no adverse effect on the calculus, we also include a form $\texttt{is}(G)?\ M$ that checks at runtime whether M returns a value that is compatible with the ground type G. M can have any type in this case because it is always a safe operation, but the result is either trivially `true` or `false` unless M has type ?.

Next, we have booleans, whose values are `true` and `false`, and whose elimination form is an if-statement. The if-statement checks that the scrutinee has a type compatible with $\mathbb{B}$, and as in previous work uses gradual meet $B_t \sqcap B_f$ for the output type [Garcia and Cimini 2015]. Gradual meet is only partially defined, since this ensures that if the two sides have different (non-?) head connectives then type checking errors, in keeping with the philosophy of strict checking when precise types are used.

Next, we have pairs and functions, which are fairly standard. We use pattern-matching as the elimination form for pairs. To reduce the number of rules, we present the elimination forms in the style of Garcia and Cimini [2015], using partial functions π_i, dom, cod and later un$\forall^\nu$, un$\exists^\nu$ to extract the subformula from a type "up to ?". For the correct type this extracts the actual subformula, but for ? is defined to be ? and for other connectives is undefined. We define these at the bottom of the figure, where uncovered cases are undefined. Next, we have existentials, which are as described in §3.1.

Finally, we consider the shaded components of the judgment. The full form of the judgment is $\Gamma \vdash M : A; \Gamma_o$ where Γ_o is the list of bindings that are *generated by* M and exported outward. Note that the type A of M can use variables in Γ_o as well as variables in Γ. Also, while we write these as Γ_o, Γ_M, etc., they only contain sequences of known type variables, and never any abstract type variables or typing assumptions $x : A$. These bindings are generated in the $\forall^\nu$ elimination rule, where the instantiation $M\{X \cong B\}$ adds $X \cong B$ to the output context. Rules that produce delayed thunks—the function, existential and universal introduction rules—have bodies that generate bindings, but these are not exported because these bindings will only be generated at the point where the thunk is forced to evaluate. The rest of the rules work similarly to an effect system: for instance in the function application rule $M\ N$ the bindings generated in M are bound in N, and the application produces all of the bindings they generate, and similarly for product introduction. In the unpack form, care must be taken to make sure that the X from the unpack is not leaked in the output Γ_N, in addition to making sure the output type B does not mention X. Any known type variables that mention X are removed from the output context, using the *restriction* form $\Gamma_{\Gamma'}$ defined at the bottom of Figure 2. Finally, in the if form, each branch might export different known type variables, so the if statement as a whole only exports the intersection of the two branches, since these are the only ones guaranteed to be generated.

4 PolyC$^\nu$: CAST CALCULUS

As is standard in gradual languages, rather than giving the surface language an operational semantics directly, we define a *cast calculus* that makes explicit the casts that perform the dynamic type checking in gradual programs. We present the cast calculus syntax in Figure 3. The cast calculus syntax is almost the same as the surface syntax, though the typing is quite different. First, the type ascription form is removed, and several forms are added to replace it. Based on the analysis in [New and Ahmed 2018], we add two cast forms: an upcast $\langle A^{\sqsubseteq} \rangle\!\uparrow M$ and a downcast $\langle A^{\sqsubseteq} \rangle\!\downarrow M$, whereas most prior work includes a single cast form $\langle A \Leftarrow B \rangle$. The $A^{\sqsubseteq}$ used in the upcast and downcast forms here is a proof that $A_l \sqsubseteq A_r$ for some types A_l, A_r, i.e., that A_l is a more precise (less dynamic) type than A_r. This type precision definition is key to formalizing the graduality property, but previous work has shown that it is useful for formalizing the semantics of casts as

$$
\begin{array}{rrcl}
\text{type names} & \alpha & ::= & \sigma \mid X \\
\text{types} & A, B & +::= & \sigma \\
\text{ground types} & G & ::= & \alpha \mid \mathbb{B} \mid ? \times ? \mid ? \to ? \mid \exists^\nu X.? \mid \forall^\nu X.? \\
\text{precision derivations} & A^\sqsubseteq, B^\sqsubseteq & ::= & ? \mid \mathrm{tag}_G(A^\sqsubseteq) \mid \alpha \mid \mathbb{B} \mid A^\sqsubseteq \times A^\sqsubseteq \mid A^\sqsubseteq \to B^\sqsubseteq \\
 & & \mid & \exists^\nu X.A^\sqsubseteq \mid \forall^\nu X.A^\sqsubseteq \\
\text{values} & V & ::= & \mathtt{seal}_\alpha V \mid \mathtt{true} \mid \mathtt{false} \mid x \mid (V, V) \mid \lambda(x : A).M \\
 & & \mid & \Lambda^\nu X.M \mid \mathtt{inj}_G\ V \mid \langle A_1^\sqsubseteq \to A_2^\sqsubseteq \rangle\!\uparrow M \mid \langle A_1^\sqsubseteq \to A_2^\sqsubseteq \rangle\!\downarrow M \\
 & & \mid & \langle \forall^\nu X.A^\sqsubseteq \rangle\!\uparrow M \mid \langle \forall^\nu X.A^\sqsubseteq \rangle\!\downarrow M \mid \mathtt{pack}^\nu(X \cong A', [\overline{A^\sqsubseteq \updownarrow}], M) \\
\text{expressions} & M, N & -::= & (M :: A) \\
 & & +::= & \mho \mid \langle A^\sqsubseteq \rangle\!\uparrow M \mid \langle A^\sqsubseteq \rangle\!\downarrow M \mid \mathtt{hide}\ X \cong A; M \mid \mathtt{inj}_G\ M \\
 & & \mid & \mathtt{pack}^\nu(X \cong A', [A^\sqsubseteq \updownarrow, \ldots], M) \mid \mathtt{seal}_\sigma M \mid \mathtt{unseal}_\sigma M \\
\text{Evaluation Context} & E & ::= & [\,] \mid (E, M) \mid (V, E) \mid E[A] \mid E\ M \mid V\ E \mid \mathtt{inj}_G\ E \\
 & & \mid & \mathtt{if}\ E\ \mathtt{then}\ M\ \mathtt{else}\ M \mid \mathtt{let}\ (x, y) = E; M \mid \langle A^\sqsubseteq \rangle\!\uparrow E \\
 & & \mid & \mathtt{unpack}\ (X, x) = E; M \mid \mathtt{seal}_\alpha E \mid \mathtt{unseal}_\alpha E \mid \langle A^\sqsubseteq \rangle\!\downarrow E
\end{array}
$$

Fig. 3. PolyC$^\nu$ Syntax

$$
\frac{\Gamma \vdash A^\sqsubseteq : A \sqsubseteq G}{\Gamma \vdash \mathrm{tag}_G(A^\sqsubseteq) : A \sqsubseteq ?} \qquad \Gamma \vdash ? : ? \sqsubseteq ? \qquad \Gamma \vdash \mathbb{B} : \mathbb{B} \sqsubseteq \mathbb{B} \qquad \frac{X \in \Gamma}{\Gamma \vdash X : X \sqsubseteq X}
$$

$$
\frac{\Gamma \vdash A_1^\sqsubseteq : A_{l1} \sqsubseteq A_{r1} \qquad \Gamma \vdash A_2^\sqsubseteq : A_{l2} \sqsubseteq A_{r2}}{\Gamma \vdash A_1^\sqsubseteq \times A_2^\sqsubseteq : A_{l1} \times A_{l2} \sqsubseteq A_{r1} \times A_{r2}} \qquad \frac{\Gamma \vdash A^\sqsubseteq : A_l \sqsubseteq A_r \qquad \Gamma \vdash B^\sqsubseteq : B_l \sqsubseteq B_r}{\Gamma \vdash A^\sqsubseteq \to B^\sqsubseteq : A_l \to B_l \sqsubseteq A_r \to B_r}
$$

$$
\frac{\Gamma, X \vdash A^\sqsubseteq : A_l \sqsubseteq A_r}{\Gamma \vdash \exists^\nu X.A^\sqsubseteq : \exists^\nu X.A_l \sqsubseteq \exists^\nu X.A_r} \qquad \frac{\Gamma, X \vdash A^\sqsubseteq : A_l \sqsubseteq A_r}{\Gamma \vdash \forall^\nu X.A^\sqsubseteq : \forall^\nu X.A_l \sqsubseteq \forall^\nu X.A_r}
$$

Fig. 4. PolyC$^\nu$ Type Precision

well. We emphasize the structure of these proofs because the central semantic constructions of this work: the operational semantics of casts, the translation of casts into functions and finally our graduality logical relation are all naturally defined by recursion on these derivations.

4.1 PolyC$^\nu$ Type Precision

We present the definition of type precision in Figure 4. The judgment $\Gamma \vdash A^\sqsubseteq : A_l \sqsubseteq A_r$ is read as "using the variables in Γ", $A^\sqsubseteq$ proves that A_l is more precise/less dynamic than A_r. If you ignore the precision derivations, our definition of type precision is a simple extension of the usual notion: type variables are only related to the dynamic type and themselves, and similarly for $\forall$ and $\exists$. Since we have quantifiers and type variables, we include a context Γ of known and abstract type variables. Crucially, even under the assumption that $X \cong A$, X and A are *unrelated* precision-wise unless A is ?. As before, $X \in \Gamma$ ranges over both known and abstract type variables. It is easy to see that precision reflexive and transitive, and that ? is the greatest element. Finally, ? is the least precise type, meaning for any type A there is a derivation that $A \sqsubseteq ?$. The precision notation is a natural extension of the syntax of types: with base types $?, \mathbb{B}$ serving as the proof of reflexivity at the type and constructors $\times, \to$, etc. serving as syntax for congruence proofs. It is important to note that while we give a syntax for derivations, there is at most *one* derivation $A^\sqsubseteq$ that proves any given $A_l \sqsubseteq A_r$. We prove these and several more lemmas about type precision in the appendix [New et al. 2020].

$$\frac{x : A \in \Gamma}{\Gamma \vdash x : A; \cdot} \qquad \frac{\Gamma \vdash M : A_l; \Gamma_M \qquad \Gamma \vdash A^{\sqsubseteq} : A_l \sqsubseteq A_r}{\Gamma \vdash \langle A^{\sqsubseteq} \rangle\!\uparrow M : A_r; \Gamma_M} \qquad \frac{\Gamma \vdash M : A_r; \Gamma_M \qquad \Gamma \vdash A^{\sqsubseteq} : A_l \sqsubseteq A_r}{\Gamma \vdash \langle A^{\sqsubseteq} \rangle\!\downarrow M : A_l; \Gamma_M}$$

$$\frac{\Gamma \vdash M : A; \Gamma_M \qquad \Gamma, \Gamma_M, x : A \vdash N : B; \Gamma_N}{\Gamma \vdash \mathtt{let}\ x = M; N : B; \Gamma_M, \Gamma_N} \qquad \frac{\Gamma \vdash M : A; \Gamma_o \qquad X \cong A \in \Gamma, \Gamma_o}{\Gamma \vdash \mathtt{seal}_X M : X; \Gamma_o}$$

$$\frac{\Gamma \vdash M : X; \Gamma_o \qquad X \cong A \in \Gamma, \Gamma_o}{\Gamma \vdash \mathtt{unseal}_X M : A; \Gamma_o} \qquad \frac{\Gamma \vdash M : ?; \Gamma_o \qquad \Gamma \vdash G}{\Gamma \vdash \mathtt{is}(G)?\ M : \mathbb{B}; \Gamma_o} \qquad \Gamma \vdash \mathtt{true} : \mathbb{B}; \cdot \qquad \Gamma \vdash \mathtt{false} : \mathbb{B}; \cdot$$

$$\frac{\Gamma \vdash M : \mathbb{B}; \Gamma_M \qquad \Gamma, \Gamma_M \vdash N_t : B; \Gamma_N \qquad \Gamma, \Gamma_M \vdash N_f : B; \Gamma_N}{\Gamma \vdash \mathtt{if}\ M\ \mathtt{then}\ N_t\ \mathtt{else}\ N_f : B; \Gamma_M, \Gamma_N} \qquad \frac{\Gamma \vdash M_1 : A_1; \Gamma_1 \qquad \Gamma, \Gamma_1 \vdash M_2 : A_2; \Gamma_2}{\Gamma \vdash (M_1, M_2) : A_1 \times A_2; \Gamma_1, \Gamma_2}$$

$$\frac{\Gamma \vdash M : A_1 \times A_2; \Gamma_M \qquad \Gamma, \Gamma_M, x : A_1, y : A_2 \vdash N : B; \Gamma_N}{\Gamma \vdash \mathtt{let}\ (x, y)\ =\ M; N : B; \Gamma_M, \Gamma_N} \qquad \frac{\Gamma \vdash A \qquad \Gamma, x : A \vdash M : B; \cdot}{\Gamma \vdash \lambda x : A.M : A \to B; \cdot}$$

$$\frac{\Gamma \vdash M : A \to B; \Gamma_M \qquad \Gamma, \Gamma_M \vdash N : A; \Gamma_N}{\Gamma \vdash M\ N : B; \Gamma_M, \Gamma_N} \qquad \frac{\Gamma, X \cong A \vdash M : B; \cdot}{\Gamma \vdash \mathsf{pack}^{\nu}(X \cong A, M) : \exists^{\nu} X.B; \cdot}$$

$$\frac{\Gamma \vdash M : \exists^{\nu} X.A; \Gamma_M \qquad \Gamma, \Gamma_M, X, x : A \vdash N : B; \Gamma_N \qquad \Gamma, \Gamma_M \vdash \Gamma_N \qquad \Gamma, \Gamma_M, \Gamma_N \vdash B}{\Gamma \vdash \mathtt{unpack}\ (X, x) = M; N : B; \Gamma_M, \Gamma_N}$$

$$\frac{\Gamma, X \vdash M : A; \cdot}{\Gamma \vdash \Lambda^{\nu} X.M : \forall^{\nu} X.A; \cdot} \qquad \frac{\Gamma \vdash M : \forall^{\nu} X.A; \Gamma_M \qquad \Gamma, \Gamma_M \vdash B}{\Gamma \vdash M\{X \cong B\} : A; \Gamma_M, X \cong B} \qquad \frac{\Gamma \vdash M : \Gamma_M, X \cong A, \Gamma'_M \qquad \Gamma, \Gamma_M \vdash \Gamma'_M}{\Gamma \vdash \mathtt{hide}\ X \cong A; M; \Gamma_M, \Gamma'_M}$$

Fig. 5. PolyC$^{\nu}$ Typing

4.2 PolyC$^{\nu}$ Type System

The static type system for the cast calculus is given Figure 5. The cast calculus type system differs from the surface language in that all type checking is *strict* and precise. This manifests in two ways. First, the dynamic type is not considered implicitly compatible with other types. Instead, in the translation from PolyG$^{\nu}$ to PolyC$^{\nu}$, we insert casts wherever consistency is used in the judgment. Second, in the if rule, the branches must have the *same* type, and an upcast is inserted in the translation to make the two align. Finally, the outward scoping of known type variables is handled more explicitly. We add a new form $\mathtt{hide}\ X \cong A; M$ that delimits the scope of $X \cong A$ from going further outward, enforced by the side condition that $\Gamma, \Gamma_M \vdash \Gamma'_M$. Then in rules that include delayed computations, i.e., values of function, existential and universal type, whereas in the surface language the delayed term could produce any names, now in PolyC$^{\nu}$, they must all be manually hidden. Similarly in the branches of an if statement, the two sides must have the same generated names, and hides must be used in the elaboration to make them align.

4.3 Elaboration from PolyG$^{\nu}$ to PolyC$^{\nu}$

We define the elaboration of PolyG$^{\nu}$ into the cast calculus PolyC$^{\nu}$ in Figure 6. Following [New and Ahmed 2018], an ascription is interpreted as a cast up to ? followed by a cast down to the ascribed type. Most of the elaboration is standard, with elimination forms being directly translated to the corresponding PolyC$^{\nu}$ form if the head connective is correct, and inserting a downcast if the elimination position has type ?. We formalize this using the metafunction $G \wr M$ defined towards the bottom of the figure. For the if case, in PolyC$^{\nu}$ the two branches of the if have to have the same

$$
\begin{aligned}
(M :: B)^+ &= \langle B?^{\sqsubseteq}\rangle\!\downarrow \langle A?^{\sqsubseteq}\rangle\!\uparrow M^+ && \text{(where } M : A, A?^{\sqsubseteq} : A \sqsubseteq ?, B?^{\sqsubseteq} : B \sqsubseteq ?)\\
x^+ &= x\\
(\texttt{let } x = M; N)^+ &= \texttt{let } x = M^+; N^+\\
(\texttt{seal}_X M)^+ &= \texttt{seal}_X (M :: A)^+ && \text{(where } X \cong A)\\
(\texttt{unseal}_X M)^+ &= \texttt{unseal}_X (X \wr M)\\
(\texttt{is}(G)?\ M)^+ &= \texttt{is}(G)?\ (\langle A?^{\sqsubseteq}\rangle\!\uparrow M) && \text{(where } M : A, A?^{\sqsubseteq} : A \sqsubseteq ?)\\
b^+ &= b && (b \in \{\texttt{true}, \texttt{false}\})\\
(\texttt{if } M \texttt{ then } N_t \texttt{ else } N_f)^+ &= \texttt{if } \mathbb{B} \wr M \texttt{ then } (\langle B_t^{\sqsubseteq}\rangle\!\downarrow \texttt{hide } \Gamma_t \subseteq \Gamma_t \cap \Gamma_f; N_t^+)\\
&\qquad \texttt{else } (\langle B_f^{\sqsubseteq}\rangle\!\downarrow \texttt{hide } \Gamma_f \subseteq \Gamma_t \cap \Gamma_f; N_f^+)\\
&\qquad \text{(where } \Gamma \vdash \texttt{if } M \texttt{ then } N_t \texttt{ else } N_f : B_t \sqcap B_f; \Gamma_M, \Gamma_t \cap \Gamma_f)\\
&\qquad \text{(and } B_t^{\sqsubseteq} : B_t \sqcap B_f \sqsubseteq B_t, B_f^{\sqsubseteq} : B_t \sqcap B_f \sqsubseteq B_f)\\
(M_1, M_2)^+ &= (M_1^+, M_2^+)\\
(\texttt{let } (x, y) = M; N)^+ &= \texttt{let } (x, y) = ? \times ? \wr M; N^+\\
(\lambda x : A.M)^+ &= \lambda x : A.\texttt{hide } \Gamma_o \subseteq \Gamma_o; M^+ && \text{(where } \Gamma, x : A \vdash M : B; \Gamma_o)\\
(M\ N)^+ &= (? \to ? \wr M)\ (N :: \mathrm{dom}(A))^+ && \text{(where } M : A)\\
(\texttt{pack}^\nu (X \cong A, M))^+ &= \texttt{pack}^\nu (X \cong A, \texttt{hide } \Gamma_o \subseteq \Gamma_o; M^+) && \text{(where } M : B; \Gamma_o)\\
(\texttt{unpack } (X, x) = M; N)^+ &= \texttt{unpack } (X, x) = \exists^\nu X.? \wr M; \texttt{hide } \Gamma_N|_X \subseteq \Gamma_N; N^+\\
\Lambda^\nu X.M^+ &= \Lambda^\nu X.\texttt{hide } \Gamma_o \subseteq \Gamma_o; M^+ && \text{(where } M : A; \Gamma_o)\\
M\{X \cong B\}^+ &= (\forall^\nu X.? \wr M)\{X \cong B\}\\
G \wr M &= \langle \mathrm{tag}_G(G)\rangle\!\downarrow M^+ && \text{(when } M : ?)\\
G \wr M &= M^+ && \text{(otherwise)}\\
\texttt{hide } \Gamma_s \subseteq (\Gamma_b, X \cong A); M &= \texttt{hide } \Gamma_s \subseteq \Gamma_b; \texttt{hide } X \cong A; M && (X \notin \Gamma_s)\\
\texttt{hide } (\Gamma_s, X \cong A) \subseteq (\Gamma_b, X \cong A); M &= \texttt{hide } \Gamma_s \subseteq \Gamma_b; M\\
\texttt{hide } \cdot \subseteq \cdot; M &= M
\end{aligned}
$$

Fig. 6. Elaborating PolyG$^\nu$ to PolyC$^\nu$

output type and export the same names, so we downcast each branch, but also we hide any names not generated by both sides, using the metafunction $\texttt{hide } \Gamma \subseteq \Gamma'; M$, defined at the bottom of the figure, which hides all names present in Γ' that are not in Γ. Finally, in the values that are thunks (pack, λ and Λ), the bodies of the thunks must not generate names in PolyC$^\nu$, so we hide names there as well.

4.4 PolyC$^\nu$ Dynamic Semantics

The dynamic semantics of PolyC$^\nu$, presented in Figure 7, extends traditional cast semantics with appropriate rules for our name-generating universals and existentials. The runtime state is a pair of a term M and a *case store* Σ. A case store Σ represents the set of cases allocated so far in the program. Formally, a store Σ is just a pair of a number $\Sigma.n$ and a function $\Sigma.f : [n] \to \mathrm{Ty}$ where Ty is the set of all types and $[n] = \{m \in \mathbb{N} \mid m < n\}$ is from some prefix of natural numbers to types. All rules take configurations $\Sigma \triangleright M$ to configurations $\Sigma' \triangleright M'$. When the step does not change the store, we write $M \mapsto M'$ for brevity.

The first rule states that all non-trivial evaluation contexts propagate errors. Next, unsealing a seal gets out the underlying value, and $\texttt{is}(G)?\ V$ literally checks if the tag of V is G. The hide

$$
\begin{array}{lcl}
E[\mathfrak{V}] & \mapsto & \mathfrak{V} \text{ where } E \neq [\,] \\
E[\mathtt{unseal}_\sigma(\mathtt{seal}_\sigma V)] & \mapsto & E[V] \\
E[\mathtt{is}(G)?\ (\mathtt{inj}_G\ V)] & \mapsto & E[\mathtt{true}] \\
E[\mathtt{is}(G)?\ (\mathtt{inj}_H\ V)] & \mapsto & E[\mathtt{false}] \text{ where } G \neq H \\
\Sigma \triangleright E[\mathtt{hide}\ X \cong A; M] & \mapsto & \Sigma, \sigma : A \triangleright E[M[\sigma/X]] \\
\Sigma \triangleright E\left[\begin{array}{l}\mathtt{unpack}\ (X,x) = \mathsf{pack}^\nu(X \cong A', [\overline{A^\sqsubseteq \updownarrow}], M); \\ N\end{array}\right] & \mapsto & \Sigma, \sigma : A' \triangleright E\left[\begin{array}{l}\mathtt{let}\ x = \overline{\langle A^\sqsubseteq[\sigma/X]\rangle\updownarrow}\ M[\sigma/X]; \\ N[\sigma/X]\end{array}\right] \\
E\,[\mathsf{pack}^\nu(X \cong A, M)] & \mapsto & E[\mathsf{pack}^\nu(X \cong A', [\,], M)] \\
E[(\Lambda^\nu X.M)\{\sigma \cong A\}] & \mapsto & E[M[\sigma/X]] \\
E[\langle A^\sqsubseteq\rangle\updownarrow\ V] & \mapsto & E[V] \text{ where } A^\sqsubseteq \in \{\mathbb{B}, \sigma, ?\} \\
E[\langle A_1^\sqsubseteq \times A_2^\sqsubseteq\rangle\updownarrow\ (V_1, V_2)] & \mapsto & E[(\langle A_1^\sqsubseteq\rangle\updownarrow\ V_1, \langle A_2^\sqsubseteq\rangle\updownarrow\ V_2)] \\
E[(\langle A_1^\sqsubseteq \to A_2^\sqsubseteq\rangle\updownarrow\ V_1)\ V_2] & \mapsto & E[\langle A_2^\sqsubseteq\rangle\updownarrow\ (V_1\ \langle A_1^\sqsubseteq\rangle\updownarrow^-\ V_2)] \\
E[\langle \exists^\nu X.A^\sqsubseteq\rangle\updownarrow\ \mathsf{pack}^\nu(X \cong A', [A'^\sqsubseteq\ \updownarrow', \ldots], M)] & \mapsto & E[\mathsf{pack}^\nu(X \cong A', [A^\sqsubseteq\ \updownarrow, A'^\sqsubseteq\ \updownarrow', \ldots], M)] \\
E[(\langle \forall^\nu X.A^\sqsubseteq\rangle\updownarrow\ V)\{\sigma \cong A\}] & \mapsto & E[\langle A^\sqsubseteq[\sigma/X]\rangle\updownarrow\ (V\{\sigma \cong A\})] \\
E[\langle \mathrm{tag}_G(A^\sqsubseteq)\rangle\uparrow\ V] & \mapsto & E[\mathtt{inj}_G\ \langle A^\sqsubseteq\rangle\uparrow\ V] \\
E[\langle \mathrm{tag}_G(A^\sqsubseteq)\rangle\downarrow\ \mathtt{inj}_G\ V] & \mapsto & E[\langle A^\sqsubseteq\rangle\downarrow\ V] \\
E[\langle \mathrm{tag}_G(A^\sqsubseteq)\rangle\downarrow\ \mathtt{inj}_H\ V] & \mapsto & \mathfrak{V} \text{ where } H \neq G
\end{array}
$$

Fig. 7. PolyC$^\nu$ Operational Semantics (fragment)

form generates a fresh seal $\sigma : A$ and substitutes it into the continuation. The pack form steps to an intermediate state used for building up a stack of casts that will be used again in the existential cast rule. The unpack rule generates a fresh seal for the $X \cong A$ and then applies all of the accumulated casts to the body of the pack. Here we use $\updownarrow$ to indicate one of $\uparrow$ and $\downarrow$. For $\forall^\nu$ instantiation, we do not need to generate the seal, because it must have already been generated by a hide form further up the term, so the rule is just a substitution. As is typical for a cast calculus, the remaining types have ordinary call-by-value β reduction and so we elide them.

The remaining rules give the behavior of casts. Other than the use of type precision derivations, the behavior of our casts is mostly standard: identity casts for $\mathbb{B}$, σ and ? are just the identity, and the product cast proceeds structurally. Function casts are values, and when applied to a value, the cast is performed on the output and the oppositely oriented case on the input. We use $\updownarrow^-$ to indicate the opposite arrow, so $\uparrow^-=\downarrow$ and $\downarrow^-=\uparrow$ to cut down the number of rules. Next, the $\forall^\nu$ casts are also values that reduce when the instantiating type is supplied. As with existentials, the freshly generated type σ is substituted for X in the precision derivation guiding the cast. Finally, the upcast case for $\mathrm{tag}_G(A^\sqsubseteq)$ simply injects the result of upcasting with $A^\sqsubseteq$ into the dynamic type using the tag G. For the downcast case, the opposite is done if the input has the right tag, and otherwise a dynamic type error is raised.

In the appendix, we extend the typing to runtime terms which are typed with respect to a case-store and prove a standard type safety theorem for the language that terms either take a step or are values or errors [New et al. 2020].

5 TYPED INTERPRETATION OF THE CAST CALCULUS

In the previous section we developed a cast calculus with an operational semantics defining the behavior of the name generation and gradual type casts. However, this ad hoc design addition of new type connectives and inside-out scoping of $\forall^\nu$-instantiations make the cast calculus less than ideal for proving meta-theoretic properties of the system.

Instead of directly proving metatheoretic properties of the cast calculus, we give a *contract translation* of the cast calculus into a statically typed core language, translating the gradual type casts to ordinary terms in the typed language that raise errors. The key benefit of the typed

$$
\begin{array}{rrcl}
\text{value types} & A & ::= & X \mid \mathsf{Case}\ A \mid \mathsf{OSum} \mid A \times A \mid \mathbb{B} \mid \exists X.A \mid UB \\
\text{computation types} & B & ::= & A \to B \mid \forall X.B \mid FA \\
\text{values} & V & ::= & \sigma \mid \mathtt{inj}_V\ V \mid \mathtt{pack}(A,V)\ \mathtt{as}\ \exists X.A \mid x \mid (V,V) \\
 & & \mid & \mathtt{true} \mid \mathtt{false} \mid \mathtt{thunk}\ M \\
\text{computations} & M & ::= & \mho \mid \mathtt{force}\ V \mid \mathtt{ret}\ V \mid x \leftarrow M; N \mid M\ V \mid \lambda x : A.M \\
 & & \mid & \mathtt{newcase}_A\ x; M \mid \mathtt{match}\ V\ \mathtt{with}\ V\{\mathtt{inj}\ x.M \mid N\} \\
 & & \mid & \mathtt{unpack}\ (X,x) = V; M \mid \mathtt{let}\ (x,x) = V; M \mid \Lambda X.M \mid M[A] \\
 & & \mid & \mathtt{if}\ V\ \mathtt{then}\ M\ \mathtt{else}\ M \\
\text{stacks} & S & ::= & \bullet \mid S\,V \mid S\,[A] \mid x \leftarrow S; M \\
\text{value typing context} & \Gamma & ::= & \cdot \mid \Gamma, x : A \\
\text{type variable context} & \Delta & ::= & \cdot \mid \Delta, X
\end{array}
$$

Fig. 8. $\mathrm{CBPV}_{\mathrm{OSum}}$ Syntax

language is that it does not have built-in notions of fresh existential and universal quantification. Instead, the type translation decomposes these features into the combination of *ordinary* existential and universal quantification combined with a somewhat well-studied programming feature: a dynamically extensible "open" sum type we call OSum. Finally, it gives a static type interpretation of the dynamic type: rather than being a finitary sum of a few statically fixed cases, the dynamic type is implemented as the open sum type which includes those types allocated at runtime.

5.1 Typed Metalanguage

We present the syntax of our typed language $\mathrm{CBPV}_{\mathrm{OSum}}$ in Figure 8, an extension of Levy's Call-by-push-value calculus [Levy 2003], which we use as a convenient metalanguage to extend with features of interest. Call-by-push-value (CBPV) is a typed calculus with highly explicit evaluation order, providing similar benefits to continuation-passing style and A-normal form [Sabry and Felleisen 1992]. The main distinguishing features of CBPV are that values V and effectful computations M are distinct syntactic categories, with distinct types: value types A and computation types B. The two "shift" types U and F mediate between the two worlds. A value of type UB is a first-class "thUnk" of a computation of type B that can be forced, behaving as a B. A computation of type FA is a computation that can perform effects and return a value of type A, and whose elimination form is a monad-like *bind*. Notably while sums and (strict) tuples are value types, function types $A \to B$ are computations since a function interacts with its environment by receiving an argument. We include existentials as value type and universals as computation types, that in each case quantify over *value* types because we are using it as the target of a translation from a call-by-value language.

We furthermore extend CBPV with two new value types: OSum and Case A, which add an open sum type similar to the extensible exception types in ML, but with an expression-oriented interface more suitable to a core calculus. The open sum type OSum is initially empty, but can have new cases allocated at runtime. A value of Case A is a first class representative of a case of OSum. The introduction form $\mathtt{inj}_{V_c}\ V$ for OSum uses a case V_c : Case A to inject a value $V : A$ into OSum. The elimination form $\mathtt{match}\ V_o\ \mathtt{with}\ V_c\{\mathtt{inj}\ x.M \mid N\}$ for OSum is to use a V_c : Case A to do a pattern match on a value V_o : OSum. Since OSum is an *open* sum type, it is unknown what cases V_o might use, so the pattern-match has two branches: the one $\mathtt{inj}\ x.M$ binds the underlying value to $x : A$ and proceeds as M and the other is a catch-all case N in case V_o was not constructed using V_c. Finally, there is a form $\mathtt{newcase}_A\ x; M$ that allocates a fresh Case A, binds it to x and proceeds as M. In addition to the similarity to ML exception types, they are also similar to the dynamically typed sealing mechanism introduced in Sumii and Pierce [2004].

$$\frac{}{\Delta;\Gamma \vdash \mho : B} \qquad \frac{}{\Delta;\Gamma, x : A, \Gamma' \vdash x : A} \qquad \frac{\Delta;\Gamma \vdash V_c : \text{Case } A \qquad \Delta;\Gamma \vdash V : A}{\Delta;\Gamma \vdash \text{inj}_{V_c}\ V : \text{OSum}}$$

$$\frac{\Delta;\Gamma \vdash V : \text{OSum} \qquad \Delta;\Gamma \vdash V_c : \text{Case } A \qquad \Delta;\Gamma, x : A \vdash M : B \qquad \Delta;\Gamma \vdash N : B}{\Delta;\Gamma \vdash \texttt{match}\ V\ \texttt{with}\ V_c\{\texttt{inj}\ x.M \mid N\} : B}$$

$$\frac{\Delta \vdash A \qquad \Delta;\Gamma, x : \text{Case } A \vdash M : B}{\Delta;\Gamma \vdash \texttt{newcase}_A\ x; M : B} \qquad \frac{\Delta;\Gamma \vdash V : A[A'/X]}{\Delta;\Gamma \vdash \texttt{pack}(A', V)\ \texttt{as}\ \exists X.A : \exists X.A}$$

$$\frac{\Delta;\Gamma \vdash V : \exists X.A \qquad \Delta \vdash B \qquad \Delta, X;\Gamma, x : A \vdash M : B}{\Delta;\Gamma \vdash \texttt{unpack}\ (X, x) = V; M : B} \qquad \frac{\Delta;\Gamma \vdash M : B}{\Delta;\Gamma \vdash \texttt{thunk}\ M : UB} \qquad \frac{\Delta;\Gamma \vdash V : UB}{\Delta;\Gamma \vdash \texttt{force}\ V : B}$$

$$\frac{\Delta, X;\Gamma \vdash M : B}{\Delta;\Gamma \vdash \Lambda X.M : \forall X.B} \qquad \frac{\Delta;\Gamma \vdash M : \forall X.B \qquad \Delta \vdash A}{\Delta;\Gamma \vdash M[A] : B[A/X]} \qquad \frac{\Delta;\Gamma \vdash V : A}{\Delta;\Gamma \vdash \texttt{ret}\ V : FA} \qquad \frac{\Delta;\Gamma \vdash M : FA \qquad \Delta;\Gamma \vdash N : B}{\Delta;\Gamma \vdash x \leftarrow M; N : B}$$

Fig. 9. $\text{CBPV}_{\text{OSum}}$ Type System (fragment)

5.2 Static and Dynamic Semantics

We show a fragment of the typing rules for $\text{CBPV}_{\text{OSum}}$ in Figure 9. There are two judgments corresponding to the two syntactic categories of terms: $\Delta;\Gamma \vdash V : A$ for typing a value and $\Delta;\Gamma \vdash M : A$ for typing a computation. Δ is the environment of type variables and Γ is the environment for term variables. Unlike in PolyG^ν and PolyC^ν, these are completely standard, and there is no concept of a known type variable.

First, an error $\mho$ is a computation and can be given any type. Variables are standard and the OSum/Case forms are as described above. Existentials are a value form and are standard as in CBPV using ordinary substitution in the pack form. In all of the value type elimination rules, the discriminee is restricted to be a *value*. A computation $M : B$ can be thunked to form a value $\texttt{thunk}\ M : UB$, which can be forced to run as a computation. Like the existentials, the universal quantification type is standard, using substitution in the elimination form. Finally, the introduction form for FA returns a value $V : A$, and the elimination form is a bind, similar to a monadic semantics of effects, except that the continuation can have any computation type B, rather than restricted to FA.

A representative fragment of the operational semantics is given in Figure 10, the full semantics are in the appendix [New et al. 2020]. S represents a *stack*, the CBPV analogue of an evaluation context, defined in Figure 8. Here Σ is like the Σ in PolyC^ν, but maps to *value types*. The semantics is standard, other than the fact that we assign a count to each step of either 0 or 1. The only steps that count for 1 are those that introduce non-termination to the language, which is used later as a technical device in our logical relation in §6.

5.3 Translation

Next, we present the "contract translation" of PolyC^ν into $\text{CBPV}_{\text{OSum}}$. This translation can be thought of as an alternate semantics to the operational semantics for PolyC^ν, but with a tight correspondence given in §5.4. Since $\text{CBPV}_{\text{OSum}}$ is a typed language that uses ordinary features like functions, quantification and an open sum type, this gives a simple explanation of the semantics of PolyC^ν in terms of fairly standard language features.

In the left side of Figure 11, we present the type translation from PolyC^ν to $\text{CBPV}_{\text{OSum}}$. Since PolyC^ν is a call-by-value language, types are translated to $\text{CBPV}_{\text{OSum}}$ *value* types. Booleans and pairs are translated directly, and the function type is given the standard CBPV translation for call-by-value functions, $U(\llbracket A \rrbracket \to F\llbracket B \rrbracket)$: a call-by-value function is a thunked computation of

$$
\begin{aligned}
S[\mho] &\mapsto^0 \mho \\
\Sigma \triangleright S[\mathtt{newcase}_A\ x; M] &\mapsto^0 \Sigma, \sigma : A \triangleright S[M[\sigma/x]] \\
S[\mathtt{match}\ \mathtt{inj}_\sigma\ V\ \mathtt{with}\ \sigma\{\mathtt{inj}\ x.M \mid N\}] &\mapsto^1 S[M[V/x]] \\
S[\mathtt{match}\ \mathtt{inj}_{\sigma_1}\ V\ \mathtt{with}\ \sigma_2\{\mathtt{inj}\ x.M \mid N\}] &\mapsto^1 S[N] \qquad (\text{where } \sigma_1 \neq \sigma_2) \\
S[\mathtt{force}\ (\mathtt{thunk}\ M)] &\mapsto^0 S[M] \\
S[\mathtt{unpack}\ (X,x) = \mathtt{pack}(A,V); M] &\mapsto^0 S[M[A/X, V/x]] \\
S[(\Lambda X.M)[A]] &\mapsto^0 S[M[A/X]] \\
S[x \leftarrow \mathtt{ret}\ V; N] &\mapsto^0 S[N[V/x]]
\end{aligned}
$$

Fig. 10. $\mathrm{CBPV}_{\mathrm{OSum}}$ Operational Semantics (fragment)

$$
\begin{aligned}
[\![\Sigma;\Gamma \vdash ?]\!] &= \mathrm{OSum} \\
[\![\Sigma;\Gamma \vdash X]\!] &= X \quad (\text{where } X \in \Gamma) \\
[\![\Sigma;\Gamma \vdash X]\!] &= [\![A]\!] \quad (\text{where } X \cong A \in \Gamma) \\
[\![\Sigma;\Gamma \vdash \sigma]\!] &= [\![A]\!] \quad (\text{where } \sigma : A \in \Sigma) \\
[\![\Sigma;\Gamma \vdash \mathbb{B}]\!] &= \mathbb{B} \\
[\![\Sigma;\Gamma \vdash A \to B]\!] &= U([\![\Sigma;\Gamma \vdash A]\!] \to F[\![\Sigma;\Gamma \vdash B]\!]) \\
[\![\Sigma;\Gamma \vdash A_1 \times A_2]\!] &= [\![\Sigma;\Gamma \vdash A_1]\!] \times [\![\Sigma;\Gamma \vdash A_2]\!] \\
[\![\Sigma;\Gamma \vdash \exists^\nu X.A]\!] &= \exists X.U(\mathrm{Case}\ X \to F[\![\Sigma;\Gamma, X \vdash A]\!]) \\
[\![\Sigma;\Gamma \vdash \forall^\nu X.A]\!] &= U(\forall X.\mathrm{Case}\ X \to F[\![\Sigma;\Gamma, X \vdash A]\!])
\end{aligned}
$$

$$
\begin{aligned}
[\![\Sigma \vdash \cdot]\!] &= \cdot;\cdot \\
[\![\Sigma \vdash \Gamma, x : A]\!] &= \Delta', \Gamma', x : [\![\Sigma;\Gamma \vdash A]\!] \\
&\quad (\text{where } [\![\Sigma \vdash \Gamma]\!] = \Delta';\Gamma') \\
[\![\Sigma \vdash \Gamma, X]\!] &= \Delta', X, \Gamma', c_X : \mathrm{Case}\ X \\
&\quad (\text{where } [\![\Sigma \vdash \Gamma]\!] = \Delta';\Gamma') \\
[\![\Sigma \vdash \Gamma, X \cong A]\!] &= \Delta', \Gamma', c_X : \mathrm{Case}\ [\![\Sigma;\Gamma \vdash A]\!] \\
&\quad (\text{where } [\![\Sigma \vdash \Gamma]\!] = \Delta';\Gamma')
\end{aligned}
$$

Fig. 11. PolyC^ν type and environment translation

a function that takes an $[\![A]\!]$ as input and may return a $[\![B]\!]$ as output. The dynamic type ? is interpreted as the open sum type. The meaning of a type variable depends on the context: if it is an abstract type variable, it is translated to a type variable, but if it is a known type variable $X \cong A$, it is translated to $[\![A]\!]$! That is, at runtime, values of a known type variable X are just values of the type isomorphic to X, and as we will see later, sealing and unsealing are no-ops. Similarly, a runtime type tag σ is translated to the type that the corresponding case maps to. These are inductively well-defined because Σ stays constant in the type translation and Γ only adds abstract type variables.

The final two cases are the most revealing. First the fresh universal quantifier, $\forall^\nu X.A$, translates to *not just* a thunk of a universally quantified computation, but also takes in a Case X as input. The body of a Λ will then use that Case X in order to interpret casts involving X. This is precisely why parametricity is more complex for our source language: if it were translated to just $U(\forall X.F[\![A]\!])$, then parametricity would follow directly by parametricity for $\mathrm{CBPV}_{\mathrm{OSum}}$, but the Case X represents additional information that the function is being passed that potentially provides information about the type X. It is only because code translated from PolyC^ν always generates a fresh case that this extra input is benign. The fresh existential $\exists^\nu X.A$ is translated to a real existential of a thunk that expects a Case X and returns a $[\![A]\!]$. Note that while the quantification is the dual of the $\forall^\nu$ case, both of them receive a Case X from the environment, which is freshly generated.

Next, while PolyG^ν and PolyC^ν have a single environment Γ that includes type variables and term variables, in $\mathrm{CBPV}_{\mathrm{OSum}}$, these are separated into a type variable environment Δ and a term variable environment Γ. For this reason in the right side of Figure11 we define the translation of an

$$
\begin{array}{lcl}
[\![M]\!]_p & = & \mathsf{newcase}_{[\![\mathbb{B}]\!]}\ c_{\mathrm{Bool}};\\
& & \mathsf{newcase}_{[\![?\to ?]\!]}\ c_{\mathrm{Fun}};\\
& & \mathsf{newcase}_{[\![?\times ?]\!]}\ c_{\mathrm{Times}};\\
& & \mathsf{newcase}_{[\![\exists^{\nu}X.?]\!]}\ c_{\mathrm{Ex}};\\
& & \mathsf{newcase}_{[\![\forall^{\nu}X.?]\!]}\ c_{\mathrm{All}};\\
& & [\![M]\!]
\end{array}
\qquad
\begin{array}{lcl}
\mathrm{case}(\mathbb{B}) & = & c_{\mathrm{Bool}}\\
\mathrm{case}(A\to B) & = & c_{\mathrm{Fun}}\\
\mathrm{case}(A\times B) & = & c_{\mathrm{Times}}\\
\mathrm{case}(\exists^{\nu}X.A) & = & c_{\mathrm{Ex}}\\
\mathrm{case}(\forall^{\nu}X.A) & = & c_{\mathrm{All}}\\
\mathrm{case}(X) & = & c_X\\
\mathrm{case}(\sigma) & = & \sigma
\end{array}
$$

$$\Gamma_p \quad = \quad \mathrm{Bool}\cong\mathbb{B}, \mathrm{Fun}\cong ?\to ?, \mathrm{Times}\cong ?\times ?, \mathrm{Ex}\cong\exists^{\nu}X.?, \mathrm{All}\cong\forall^{\nu}X.?$$

Fig. 12. Ground type tag management

environment Γ to be a pair of environments in $\mathrm{CBPV}_{\mathrm{OSum}}$. The term variable $x : A$ is just translated to a variable $x : [\![A]\!]$, but the type variables are more interesting. An abstract type variable X is translated to a pair of a type variable X but also an associated term variable $c_X : \mathrm{Case}\ X$, which represents the case of the dynamic type that will be instantiated with a freshly generated case. On the other hand, since known type variables $X \cong A$ are translated to $[\![A]\!]$, we do not extend Δ with a new variable, but still produce a variable $c_X : [\![A]\!]$ as with an unknown type variable. Finally, the empty context $\cdot$ is translated to a pair of empty contexts.

To translate a whole program, written $[\![\Sigma; \cdot \vdash M]\!]_p$, we insert a preamble that generates the cases of the open sum type for each ground type. In Figure 12, we show our whole-program translation which inserts a preamble to generate a case of the OSum type for each ground type. This allows us to assume the existence of these cases in the rest of the translation. These can be conveniently modeled as a sequence of "global" definitions of some known type variables, which we write as Γ_p. We also define a function $\mathrm{case}(\cdot)$ from types to their corresponding case value, which is a case variable for all types except those generated at runtime σ.

Next, we consider the term translation, which is defined with the below type preservation Theorem 5.1 in mind. First, all PolyC^{ν} terms of type A are translated to $\mathrm{CBPV}_{\mathrm{OSum}}$ *computations*, with type $F[\![A]\!]$, which is standard for translating CBV to CBPV. Also, note that the *output* environment of fresh type names in a term is just translated as an extension to the input environment, the difference is irrelevant in the translated code, because the names themselves are actually generated in the translation of the *hide* form. Finally, we include the preamble context Γ_p to the front of the terms to account for the fact that all terms can use the cases generated in the preamble.

Theorem 5.1. *If* $\Gamma_1 \vdash M : A, \Gamma_2$ *then* $\Delta; \Gamma \vdash [\![M]\!] : F[\![\Sigma; \Gamma_1, \Gamma_2 \vdash A]\!]$ *where* $[\![\Gamma_p, \Gamma_1, \Gamma_2]\!] = \Delta; \Gamma$.

We show most of the term translation in Figure 13. To reduce the context clutter in the translations, we elide the contexts Σ, Γ in the definition of the semantics. While they are technically needed to translate type annotations, they do not affect the operational semantics and so can be safely ignored. We put the bool, pair, and our pack-cast intermediate form cases in the appendix [New et al. 2020].

Variables translate to a return of the variable, let is translated bind, and errors are translated to errors. Since the type translation maps known type variables to their bound types, the target language `seal` and `unseal` disappear in the translation. Injection into the dynamic type translates to injection into the open sum type and ground type checks in PolyC^{ν} are implemented using pattern matching on OSum in $\mathrm{CBPV}_{\mathrm{OSum}}$. Next, the hide form is translated to a newcase form.

Next, we cover the cases involving thunks. As a warmup, the functions follow the usual CBV translation into CBPV: a CBV λ is translated to a thunk of a CBPV λ, and the application translation makes the evaluation order explicit and forces the function with an input. We translate existential packages in the cast calculus to $\mathrm{CBPV}_{\mathrm{OSum}}$ packages containing functions from a case of the open

$$
\begin{array}{lcl}
[\![x]\!] & = & \mathtt{ret}\ x \\
[\![\mathtt{let}\ x = M; N]\!] & = & x \leftarrow [\![M]\!]; [\![N]\!] \\
[\![\mathcal{U}_A]\!] & = & \mathcal{U} \\
[\![\mathtt{seal}_X M]\!] & = & [\![M]\!] \\
[\![\mathtt{unseal}_X M]\!] & = & [\![M]\!] \\
[\![\mathtt{inj}_G\ M]\!] & = & r \leftarrow [\![M]\!]; \mathtt{ret}\ \mathtt{inj}_{\mathrm{case}(G)}\ r \\
[\![\mathtt{is}(G)?\ M]\!] & = & r \leftarrow [\![M]\!]; \mathtt{match}\ r\ \mathtt{with}\ \mathrm{case}(G)\{\mathtt{inj}\ y.\mathtt{ret}\ \mathtt{true} \mid \mathtt{ret}\ \mathtt{false}\} \\
[\![\mathtt{hide}\ X \cong A; M]\!] & = & \mathtt{newcase}_{[\![A]\!]}\ c_X; [\![M]\!] \\
[\![\langle A^{\sqsubseteq}\rangle\!\updownarrow\ M]\!] & = & [\![A^{\sqsubseteq}]\!]\!\updownarrow [\![M]\!] \\
[\![\lambda(x:A).M]\!] & = & \mathtt{ret}\ \mathtt{thunk}\ \lambda(x:[\![A]\!]).[\![M]\!] \\
[\![M\ N]\!] & = & f \leftarrow [\![M]\!]; a \leftarrow [\![N]\!]; (\mathtt{force}\ f)\ a \\
[\![\mathtt{pack}^{\nu}(X \cong A, M)]\!] & = & \mathtt{ret}\ \mathtt{pack}(A, \mathtt{thunk}\ (\lambda c_X : \mathrm{Case}\ A.[\![M]\!])) \\
[\![\mathtt{unpack}\ (X,x) = M; N]\!] & = & r \leftarrow [\![M]\!]; \mathtt{unpack}\ (X,f) = r; \mathtt{newcase}_X\ c_X; x \leftarrow (\mathtt{force}\ f)\ c_X; [\![N]\!] \\
[\![\Lambda^{\nu} X.M]\!] & = & \mathtt{ret}\ (\mathtt{thunk}\ (\Lambda X.\lambda(c_X : \mathrm{Case}\ X).[\![M]\!])) \\
[\![M\{X \cong A\}]\!] & = & f \leftarrow [\![M]\!]; (\mathtt{force}\ f)[A]\ c_X
\end{array}
$$

$$
\begin{array}{lcl}
[\![G]\!]\!\updownarrow & = & \bullet \quad (\text{when } G = ?,\ \alpha,\ \text{or } \mathbb{B}) \\
[\![\mathrm{tag}_G(A^{\sqsubseteq})]\!]\!\updownarrow & = & r \leftarrow [\![A^{\sqsubseteq}]\!]\!\updownarrow [\bullet]; \mathtt{ret}\ \mathtt{inj}_{\mathrm{case}(G)}\ r \\
[\![\mathrm{tag}_G(A^{\sqsubseteq})]\!]\!\downarrow & = & x \leftarrow \bullet; \mathtt{match}\ x\ \mathtt{with}\ \mathrm{case}(G)\{\mathtt{inj}\ y.[\![A^{\sqsubseteq}]\!]\!\downarrow [\mathtt{ret}\ y] \mid \mathcal{U}\} \\
[\![A_1^{\sqsubseteq} \times A_2^{\sqsubseteq}]\!]\!\updownarrow & = & x \leftarrow \bullet; \mathtt{let}\ (x_1,x_2) = x; x_1' \leftarrow [\![A_1^{\sqsubseteq}]\!]\!\updownarrow [\mathtt{ret}\ x_1]; x_2' \leftarrow [\![A_2^{\sqsubseteq}]\!]\!\updownarrow [\mathtt{ret}\ x_2]; \mathtt{ret}\ (x_1', x_2') \\
[\![A_1^{\sqsubseteq} \to A_2^{\sqsubseteq}]\!]\!\updownarrow & = & x \leftarrow \bullet; \mathtt{ret}\ \mathtt{thunk}\ (\lambda y : A'.a \leftarrow [\![A_1^{\sqsubseteq}]\!]\!\updownarrow^{-} [\mathtt{ret}\ y]; [\![A_2^{\sqsubseteq}]\!]\!\updownarrow [\mathtt{force}\ x\ a]) \\
 & & \text{where } A_1^{\sqsubseteq} : A_{1l} \sqsubseteq A_{1r} \text{ and if } \updownarrow = \nwarrow, A' = A_{1r}, \text{ else } A' = A_{1l} \\
[\![\forall^{\nu} X.A^{\sqsubseteq}]\!]\!\updownarrow & = & x \leftarrow \bullet; \mathtt{ret}\ \mathtt{thunk}\ (\Lambda X.\lambda c_X : \mathrm{Case}\ X.[\![A^{\sqsubseteq}]\!]\!\updownarrow [\mathtt{force}\ x\ [X]\ c_X]) \\
[\![\exists^{\nu} X.A^{\sqsubseteq}]\!]\!\updownarrow & = & x \leftarrow \bullet; \mathtt{unpack}\ (Y,f) = x; \mathtt{ret}\ \mathtt{pack}(Y, \mathtt{thunk}\ (\lambda c_X : \mathrm{Case}\ Y.[\![A^{\sqsubseteq}]\!]\!\updownarrow [(\mathtt{force}\ f)\ c_X]))
\end{array}
$$

Fig. 13. PolyC$^{\nu}$ term translation (fragment)

sum type to the body of the package. In PolyC$^{\nu}$ we delay execution of pack bodies, so the translation inserts a thunk to make the order of execution explicit. Since pack bodies translate to functions, the translation of an unpack must provide a case of the open sum type to the package it unwraps. Type abstractions (Λ^{ν}), like packs, wrap their bodies in functions that, on instantiation, expect a case of the open sum type matching the instantiating type. Since hide generates the requisite type name, it translates to a newcase. A type application then simply plugs its given type and the tag associated with its type variable into the supplied type abstraction.

Next, we define the implementation of casts as "contracts", i.e., ordinary functions in the CBPV$_{\mathrm{OSum}}$. Reflexive casts at atomic types, ?, α, and $\mathbb{B}$, translate away. Structural casts at composite types, pair types, function types, universals, and existentials, push casts for their sub-parts into terms of each type. Function and product casts are entirely standard, noting that we use $r(A^{\sqsubseteq}) = A_r$. Universal casts delay until type application and then cast the output. Existential casts push their subcasts into whatever package they are given.

5.4 Simulation

In §6, we establish graduality and parametricity theorems for PolyG$^{\nu}$/PolyC$^{\nu}$ by analysis of the semantics of terms translated into CBPV$_{\mathrm{OSum}}$. But since we take the operational semantics of PolyC$^{\nu}$ as definitional, we need to bridge the gap between the operational semantics in CBPV$_{\mathrm{OSum}}$ and PolyC$^{\nu}$ by proving the following adequacy theorem that says that the final behavior of terms in PolyC$^{\nu}$ is the same as the behavior of their translations:

Theorem 5.2 (Adequacy). *If $\cdot \vdash M : A; \cdot$, then $M \Uparrow$ if and only if $[\![M]\!]_p \Uparrow$ and $M \Downarrow V$ if and only if $[\![M]\!]_p \Downarrow V'$ and $M \Downarrow \mathcal{U}$ if and only if $[\![M]\!]_p \Downarrow \mathcal{U}$.*

$$\frac{\Gamma^{\sqsubseteq} \vdash M_l \sqsubseteq M_r : A^{\sqsubseteq}; \Gamma^{\sqsubseteq\prime} \qquad \Gamma^{\sqsubseteq}, \Gamma^{\sqsubseteq\prime} \vdash B^{\sqsubseteq} : B_l \sqsubseteq B_r}{\Gamma^{\sqsubseteq} \vdash (M_l :: B_l) \sqsubseteq (M_r :: B_r) : B^{\sqsubseteq}; \Gamma^{\sqsubseteq\prime}} \qquad \frac{\Gamma^{\sqsubseteq}, X \vdash M_l \sqsubseteq M_r : A^{\sqsubseteq}; \Gamma_o^{\sqsubseteq}}{\Gamma^{\sqsubseteq} \vdash \Lambda^{\nu} X.M_l \sqsubseteq \Lambda^{\nu} X.M_r : \forall^{\nu} X.A^{\sqsubseteq}; \cdot}$$

$$\frac{\Gamma^{\sqsubseteq} \vdash M_l \sqsubseteq M_r : A^{\sqsubseteq}; \Gamma_M^{\sqsubseteq} \qquad \Gamma^{\sqsubseteq} \vdash B^{\sqsubseteq} : B_l \sqsubseteq B_r}{\Gamma^{\sqsubseteq} \vdash M_l\{X \cong B_l\} \sqsubseteq M_r\{X \cong B_r\} : \mathrm{un}\forall^{\nu}(A^{\sqsubseteq}); \Gamma_M^{\sqsubseteq}, X \cong B^{\sqsubseteq}}$$

Fig. 14. PolyG$^{\nu}$ Term Precision (fragment)

The proof of the theorem follows by a forward *simulation* argument given in the appendix, adapting a similar CBPV simulation given by Levy [2003], and proves that the V and V' in the adequacy proof are in the simulation relation [New et al. 2020].

6 GRADUALITY AND PARAMETRICITY

In this section we prove the central metatheoretic results of the paper: that our surface language satisfies both graduality and parametricity theorems. Each of these is considered a technical challenge to prove: parametricity is typically proven by a logical relation and graduality is proven either by a simulation argument [Siek et al. 2015] or a logical relation [New and Ahmed 2018; New et al. 2019], so in the worst case this would require two highly technical developments. However, we show that this is not necessary: the logical relations proof for graduality is general enough that the parametricity theorem is a corollary of the *reflexivity* of the logical relation. This substantiates the analogy between parametricity and graduality originally proposed in [New and Ahmed 2018].

The key to sharing this work is that we give a novel *relational* interpretation of type precision derivations. That is, our logical relation is indexed not by types, but by type precision derivations. For any derivation $A^{\sqsubseteq} : A_l \sqsubseteq A_r$, we define a relation $\mathcal{V}[\![A^{\sqsubseteq}]\!]$ between values of A_l and A_r. By taking the reflexivity case $A : A \sqsubseteq A$, we recover the parametricity logical relation. Previous logical relations proofs of graduality defined a logical relation indexed by types, and used casts to define a second relation based on type precision judgments, but the direct relational approach simplifies the proofs and immediately applies to parametricity as well.

6.1 Term Precision

To state the graduality theorem, we begin by formalizing the syntactic *term* precision relation. The intuition behind a precision relation $M \sqsubseteq M'$ is that M' is a (somewhat) dynamically typed term and we have changed some of its type annotations to be more precise, producing M. This is one of the main intended use cases for a gradual language: hardening the types of programs over time. Restated in a less directed way, a term M is (syntactically) more precise than M' when the types and annotations in M are more precise than M' and otherwise the terms have the same structure. We formalize this as a judgment $\Gamma^{\sqsubseteq} \vdash M_l \sqsubseteq M_r : A^{\sqsubseteq}; \Gamma^{\sqsubseteq\prime}$, where $\Gamma^{\sqsubseteq}, \Gamma^{\sqsubseteq\prime} \vdash A^{\sqsubseteq} : A_l \sqsubseteq A_r$ is a type precision derivation and $\Gamma^{\sqsubseteq} : \Gamma_l \sqsubseteq \Gamma_r$ and $\Gamma^{\sqsubseteq\prime} : \Gamma_l' \sqsubseteq \Gamma_r'$ are precision *contexts* and $\Gamma_l \vdash M_l : A_l; \Gamma_l'$ and $\Gamma_r \vdash M_r : A_r; \Gamma_r'$. A precision context $\Gamma^{\sqsubseteq}$ is like a precision derivation between two contexts: everywhere a type would be in an ordinary context, a precision derivation is used instead.

We show term precision rules for annotations and $\forall^{\nu}$ introduction and elimination for the surface language in Figure 14, with full rules in the appendix [New et al. 2020]. The rules are all completely structural: just check that the two terms have the same term constructor and all of the corresponding arguments of the rule are $\sqsubseteq$. As exhibited by the $\forall^{\nu}$ elimination rule, the metafunctions dom, cod, un$\forall^{\nu}$, un$\exists^{\nu}$ are extended in the obvious way to work on precision derivations. We define a similar notion of term precision for PolyC$^{\nu}$. Again we show the rules for casts and $\forall^{\nu}$ in Figure 15, the full definition is in the appendix [New et al. 2020]. The main difference is that, following [New et al. 2019], we include four rules involving casts: two for downcasts and

$$\frac{\Gamma^{\sqsubseteq} \vdash M_l \sqsubseteq M_r : AC^{\sqsubseteq}; \Gamma^{\sqsubseteq\prime} \qquad \Gamma^{\sqsubseteq}, \Gamma^{\sqsubseteq\prime} \vdash AC^{\sqsubseteq} : A \sqsubseteq C \qquad \Gamma, \Gamma' \vdash AB^{\sqsubseteq} : A \sqsubseteq B \qquad \Gamma^{\sqsubseteq}, \Gamma^{\sqsubseteq\prime} \vdash BC^{\sqsubseteq} : B \sqsubseteq C}{\Gamma^{\sqsubseteq} \vdash \langle AB^{\sqsubseteq} \rangle\!\uparrow M_l \sqsubseteq M_r : BC^{\sqsubseteq}; \Gamma^{\sqsubseteq\prime}}$$

$$\frac{\Gamma^{\sqsubseteq} \vdash M_l \sqsubseteq M_r : BC^{\sqsubseteq}; \Gamma^{\sqsubseteq\prime} \qquad \Gamma^{\sqsubseteq}, \Gamma^{\sqsubseteq\prime} \vdash AC^{\sqsubseteq} : A \sqsubseteq C \qquad \Gamma, \Gamma' \vdash AB^{\sqsubseteq} : A \sqsubseteq B \qquad \Gamma^{\sqsubseteq}, \Gamma^{\sqsubseteq\prime} \vdash BC^{\sqsubseteq} : B \sqsubseteq C}{\Gamma^{\sqsubseteq} \vdash \langle AB^{\sqsubseteq} \rangle\!\downarrow M_l \sqsubseteq M_r : AC^{\sqsubseteq}; \Gamma^{\sqsubseteq\prime}}$$

$$\frac{\Gamma^{\sqsubseteq} \vdash M_l \sqsubseteq M_r : AB^{\sqsubseteq}; \Gamma^{\sqsubseteq\prime} \qquad \Gamma^{\sqsubseteq}, \Gamma^{\sqsubseteq\prime} \vdash AC^{\sqsubseteq} : A \sqsubseteq C \qquad \Gamma, \Gamma' \vdash BC^{\sqsubseteq} : B \sqsubseteq C \qquad \Gamma^{\sqsubseteq}, \Gamma^{\sqsubseteq\prime} \vdash AB^{\sqsubseteq} : A \sqsubseteq B}{\Gamma^{\sqsubseteq} \vdash M_l \sqsubseteq \langle BC^{\sqsubseteq} \rangle\!\uparrow M_r : AC^{\sqsubseteq}; \Gamma^{\sqsubseteq\prime}}$$

$$\frac{\Gamma^{\sqsubseteq} \vdash M_l \sqsubseteq M_r : AC^{\sqsubseteq}; \Gamma^{\sqsubseteq\prime} \qquad \Gamma^{\sqsubseteq}, \Gamma^{\sqsubseteq\prime} \vdash AC^{\sqsubseteq} : A \sqsubseteq C \qquad \Gamma, \Gamma' \vdash BC^{\sqsubseteq} : B \sqsubseteq C \qquad \Gamma^{\sqsubseteq}, \Gamma^{\sqsubseteq\prime} \vdash AB^{\sqsubseteq} : A \sqsubseteq B}{\Gamma^{\sqsubseteq} \vdash M_l \sqsubseteq \langle BC^{\sqsubseteq} \rangle\!\downarrow M_r : AB^{\sqsubseteq}; \Gamma^{\sqsubseteq\prime}}$$

$$\frac{\Gamma^{\sqsubseteq} \vdash M_l \sqsubseteq M_r : \forall^{\nu} X.A^{\sqsubseteq}; \Gamma^{\sqsubseteq\prime} \qquad \Gamma^{\sqsubseteq} \vdash B^{\sqsubseteq} : B_l \sqsubseteq B_r}{\Gamma^{\sqsubseteq} \vdash M_l\{X \cong B_l\} \sqsubseteq M_r\{X \cong B_r\} : A^{\sqsubseteq}; \Gamma^{\sqsubseteq\prime}, X \cong B^{\sqsubseteq}} \qquad \frac{\Gamma^{\sqsubseteq}, X \vdash M_l \sqsubseteq M_r : A^{\sqsubseteq}; \cdot}{\Gamma^{\sqsubseteq} \vdash \Lambda^{\nu} X.M_l \sqsubseteq \Lambda^{\nu} X.M_r : \forall^{\nu} X.A^{\sqsubseteq}; \cdot}$$

Fig. 15. PolyC$^{\nu}$ Term Precision (fragment)

two for upcasts. We can summarize all four by saying that if $M_l \sqsubseteq M_r$, then adding a cast to either M_l or M_r still maintains that the left side is more precise than the right, as long as the type on the left is more precise than the right. Semantically, these are the most important term precision rules, as they bridge the worlds of type and term precision.

Then the key lemma is that the elaboration process from PolyG$^{\nu}$ to PolyC$^{\nu}$ preserves term precision. The proof, presented in the appendix, follows by induction on term precision proofs [New et al. 2020].

Lemma 6.1. *If* $\Gamma^{\sqsubseteq} \vdash M_l \sqsubseteq M_r : A^{\sqsubseteq}; \Gamma^{\sqsubseteq\prime}$ *in the surface language, then* $\Gamma^{\sqsubseteq} \vdash M_l^+ \sqsubseteq M_r^+ : A^{\sqsubseteq}; \Gamma^{\sqsubseteq\prime}$

6.2 Graduality Theorem

The graduality theorem states that if a term M is *syntactically* more precise than a term M', then M *semantically* refines the behavior of M': it may error, but otherwise it has the same behavior as M': if it diverges so does M' and if it terminates at V, M' terminates with some V' as well. If we think of M as the result of hardening the types of M', then this shows that hardening types semantically only increases the burden of runtime type checking and doesn't otherwise interfere with program behavior. We call this *operational* graduality, as we will consider some related notions later.

Theorem 6.2 (Operational Graduality). *If* $\cdot \vdash M_l \sqsubseteq M_r : A^{\sqsubseteq}; \cdot$, *then either* $M_l^+ \Downarrow \mathfrak{U}$ *or both terms diverge* $M_l^+, M_r^+ \Uparrow$ *or both terms terminate successfully* $M_l^+ \Downarrow V_l$ *and* $M_r^+ \Downarrow V_r$.

6.3 Logical Relation

The basic idea of the logical relations proof to proving graduality is to interpret a term precision judgment $\Gamma^{\sqsubseteq} \vdash M_l \sqsubseteq M_r : A^{\sqsubseteq}; \Gamma_o^{\sqsubseteq}$ in a *relational* manner. That is, to every type precision derivation $A^{\sqsubseteq} : A_l \sqsubseteq A_r$, we associate a relation $\mathcal{V}[\![A^{\sqsubseteq}]\!]$ between closed values of types A_l and A_r. Then we define a semantic version of the term precision judgment $\Gamma^{\sqsubseteq} \vDash M_l \sqsubseteq M_r \in A^{\sqsubseteq}; \Gamma_o^{\sqsubseteq}$ which says that given inputs satisfying the relations in $\Gamma^{\sqsubseteq}, \Gamma_o^{\sqsubseteq}$, then either M_l will error, both sides diverge, or M_l and M_r will terminate with values in the relation $\mathcal{V}[\![A^{\sqsubseteq}]\!]$. We define this relation over CBPV$_{\text{OSum}}$ translations of PolyC$^{\nu}$ terms, rather than PolyC$^{\nu}$ terms because the operational semantics is simpler.

More precisely, we use the now well established technique of Kripke, step-indexed logical relations [Ahmed et al. 2009]. Because the language includes allocation of fresh type names at runtime, the set of values that should be in the relation grows as the store increases. This is modeled *Kripke* structure, which indexes the relation by a "possible world" that attaches invariants to the allocated cases. Because our language includes diverging programs (due to the open sum type), we need to use a *step-indexed* relation that decrements when pattern matching on OSum, and "times out" when the step index hits 0. Finally, following [New and Ahmed 2018; New et al. 2019], to model graduality we need to associate two relations to each type precision derivation: one which

$$
\begin{array}{rcl}
\mathrm{Atom}_n[A_l, A_r] & = & \{(w, V_l, V_r) \mid w \in \mathrm{World}_n \wedge (w.\Sigma_l \mid \cdot \vdash V_l : A_l) \wedge (w.\Sigma_r \mid \cdot \vdash V_r : A_r)\} \\
\mathrm{CAtom}_n[A_l, A_r] & = & \{(w, V_l, V_r) \mid w \in \mathrm{World}_n \wedge w.\Sigma_l \mid \cdot \vdash M_l : FA_l \wedge w.\Sigma_r \mid \cdot \vdash M_r : FA_r\} \\
\mathrm{Rel}_n[A_l, A_r] & = & \{R \subseteq \mathrm{Atom}_n[A_l, A_r] \mid \forall (w, V_l, V_r) \in R, w' \sqsupseteq w.(w', V_l, V_r) \in R\} \\
\mathrm{World}_n & = & \{(j, \Sigma_l, \Sigma_r, \eta) \mid j < n \wedge \eta \in \mathrm{Interp}_j[\Sigma_l, \Sigma_r]\} \\
\mathrm{Interp}_n[\Sigma_l, \Sigma_r] & = & \{(size, f, \rho) \mid size \in \mathbb{N} \wedge f \in [size] \to ([\Sigma_l.size] \times [\Sigma_2.size]) \\
 & & \quad \wedge \rho : (i < size) \to Rel_n[\Sigma_l(f(i)); \Sigma_r(f(i))] \\
 & & \quad \wedge \forall i < j < size. f(i)_l \neq f(j)_l \wedge f(i)_r \neq f(j)_r\} \\
w' \sqsupset w. & = & (w' \sqsupseteq w) \wedge w'.j > w.j \\
w' \sqsupseteq w & = & w'.j \leq w.j \wedge w'.\Sigma_l \sqsupseteq w.\Sigma_l \wedge w'.\Sigma_r \sqsupseteq w.\Sigma_r \wedge w'.\eta \sqsupseteq \lfloor w.\eta \rfloor_{w'.j} \\
\Sigma' \sqsupseteq \Sigma & = & \Sigma'.size \geq \Sigma.size \wedge \forall i < \Sigma.size.\ \Sigma'(i) = \Sigma(i) \\
\eta' \sqsupseteq \eta & = & \eta'.size \geq \eta \wedge \forall i < \eta.size.\ \eta'.f(i) = \eta.f(i) \wedge \eta'.\rho(i) = \eta.\rho(i) \\
\triangleright R & = & \{(w, V_l, V_r) \mid \forall w' \sqsupset w.(w', V_l, V_r) \in R\} \\
\mathrm{Rel}_\omega[A_l, A_r] & = & \{R \subseteq \bigcup_{n \in \mathbb{N}} \mathrm{Atom}_n[A_l, A_r] \mid \forall n \in \mathbb{N}.\ \lfloor R \rfloor_n \in \mathrm{Rel}_n[A_l, A_r]\} \\
\lfloor \eta \rfloor_n & = & (\eta.size, \eta.f, \lambda i. \lfloor \rho(i) \rfloor_n) \\
\lfloor R \rfloor_n & = & \{(w, V_l, V_r) \mid w.j < n \wedge (w, V_l, V_r) \in R\} \\
\eta \vDash (\sigma_l, \sigma_r, R) & = & \exists i < \eta.size.\ \eta.f(i) = (\sigma_l, \sigma_r) \wedge \eta.\rho(i) = R\}
\end{array}
$$

Fig. 16. Logical Relation Auxiliary Definitions

times out when the left hand hand term runs out of steps, but allows the right hand side to take any number of steps and vice-versa one that times out when the right runs out of steps.

Figure 16 includes preliminary definitions we need for the logical relation. First, $\mathrm{Atom}_n[A_l, A_r]$ and $\mathrm{CAtom}_n[A_l, A_r]$ define the world-term-term triples that the relations are defined over. A relation $R \in \mathrm{Rel}_n[A_l, A_r]$ at stage n consists of triples of a world, and a value of type A_l and a value of type A_r (ignore the n for now) such that it is monotone in the world. The world $w \in \mathrm{World}_n$ contains the number of steps remaining $w.j$, the current state of each side $w.\Sigma_l$, $w.\Sigma_r$, and finally an interpretation of how the cases in the two stores are related $w.\eta$. An interpretation $\eta \in \mathrm{Interp}_n[\Sigma_l, \Sigma_r]$ consists of a cardinality $\eta.size$ which says how many cases are related and a function $\eta.f$ which says *which* cases are related, i.e., for each $i \in \eta.size$ it gives a pair of cases, one valid in the left hand store and one in the right. Finally, $\eta.\rho$ gives a relation between the types of these two cases. The final side-condition says this association is a *partial bijection*: a case on one side is associated to at most one case on the other side. Staging the relations and worlds is necessary due to a circularity here: a relation is (contravariantly) dependent on the current world, which contains relations. A relation in Rel_n is indexed by a World_n, but a World_n contains relations in $\mathrm{Rel}_{w.j}$, and $w.j < n$. In particular, $\mathrm{World}_0 = \emptyset$, so the definition is well-founded.

The next portion of the figure contains the definition of *world* extension $w' \sqsupseteq w$[2], representing the idea that w' is a possible "future" of w: the step index j is smaller and the states of the two sides have increased the number of allocated cases, but the old invariants are still there. We define strict extension $w' \sqsupset w$ to mean extension where the step has gotten strictly smaller. This allows us to define the *later* relation $\triangleright R$ which is used to break circular definitions in the logical relation. Next, we define our notion of non-indexed relation Rel_ω, which is what we quantify over in the interpretation of $\forall^v, \exists^v$. Then we define the restriction of interpretations and relations to a stage n. An infinitary relation can be "clamped" to any stage n using $\lfloor R \rfloor_n$. Finally, we define when two cases are related in an interpretation as $\eta \vDash (\sigma_l, \sigma_r, R)$.

The top of Figure 17 contains the definition of the logical relation on values and computations, except for the standard cases for booleans, products and functions, which are included in the appendix [New et al. 2020]. First, we write $\sim$ as a metavariable that ranges over two symbols: $\prec$ which indicates that we are counting steps on the left side, and $\succ$ which indicates we are counting

[2]there is a clash of notation between precision $\sqsubseteq$ and world extension $\sqsupseteq$ but it should be clear which is meant at any time.

steps on the right side. We then define the value relation $\mathcal{V}_n^{\sim}[\![A^{\sqsubseteq}]\!]\gamma\delta \in \mathrm{Rel}_n[\delta_l(A_l), \delta_r(A_r)]$. Here γ maps the free term variables to pairs of values and δ maps free type variables to triples of two types and a relation between them. First, the definition for type variables looks up the relation in the relational substitution δ. Next, two values in ? are related when they are both injections into OSum, and the "payloads" of the injections are *later* related in the relation R which the world associates to the corresponding cases. The $\triangleright$ is used because we count pattern matching on OSum as a step. This also crucially lines up with the fact that pattern matching on the open sum type is the only reduction that consumes a step in our operational semantics. Note that this is a generalization of the logical relation definition for a recursive sum type, where each injection corresponds to a case of the sum. Here since the sum type is open, we must look in the world to see what cases are allocated. Next, the $\mathrm{tag}_G(A^{\sqsubseteq})$ case relates values on the left at some type A_l and values on the right of type ?. The definition states that the dynamically typed value must be an injection using the tag given by G, and that the payload of that injection must be related to V_l with the relation given by $A^{\sqsubseteq}$. This case splits into two because we are pattern matching on a value of the open sum type, and so in the $\succ$ case we must decrement because we are consuming a step on the right, whereas in the $\prec$ case we do not decrement because we are only counting steps on the left. In the $\forall^{\nu}X.B^{\sqsubseteq}$ case, two values are related when in any future world, and any relation $R \in \mathrm{Rel}_{\omega}[A_l, A_r]$, and any pair of cases σ_l, σ_r that have $\lfloor R \rfloor_{w.j}$ as their associated relation, if the values are instantiated with A_l, σ_l and A_r, σ_r respectively, then they behave as related computations. The intuition is that values of type $\forall^{\nu}X.B$ are parameterized by a type A and a tag for that type σ, but the relational interpretation of the two must be the *same*. This is the key to proving the $\mathtt{seal}_X$ and $\mathtt{unseal}_X$ cases of graduality. The fresh existential is dual in the choice of relation, but the same in its treatment of the case σ.

Next, we define the relation on expressions. The two expression relations, $\mathcal{E}^{\prec}[\![A^{\sqsubseteq}]\!]$ and $\mathcal{E}^{\succ}[\![A^{\sqsubseteq}]\!]$ capture the semantic idea behind graduality: either the left expression raises an error, or the two programs have the same behavior: diverge or return related values in $\mathcal{V}^{\sim}[\![A^{\sqsubseteq}]\!]$. However, to account for step-indexing, each is an *approximation* to this notion where $\mathcal{E}^{\prec}[\![A^{\sqsubseteq}]\!]$ times out if the left side consumes all of the available steps $w.j$ (where $(\Sigma, M) \mapsto^j$ is shorthand for saying it steps to something in j steps), and $\mathcal{E}^{\succ}[\![A^{\sqsubseteq}]\!]$ times out if the right side consumes all of the available steps. relation is that when We define the *infinitary* version of the relations $\mathcal{V}^{\sim}[\![A^{\sqsubseteq}]\!]$ and $\mathcal{E}^{\sim}[\![A^{\sqsubseteq}]\!]$ a the union of all of the level n approximations.

Next, we give the relational interpretation of environments. The interpretation of the empty environment are empty substitutions with a valid world w. Extending with a value variable $x : A^{\sqsubseteq}$ means extending γ with a pair of values related by $\mathcal{V}^{\sim}[\![A^{\sqsubseteq}]\!]$. For an abstract type variable X, first δ is extended with a pair of types and a relation between them. Then, γ must also be extended with a pair of *cases* encoding how these types are injected into the dynamic type. Crucially, just as with the $\forall^{\nu}, \exists^{\nu}$ value relations, these cases must be associated by w to the $w.j$ approximation of the *same* relation with which we extend δ. The interpretation of the *known* type variables $X \cong A^{\sqsubseteq}$ has the same basic structure, the key difference is that rather than using an arbitrary, δ is extended with the value relation $\mathcal{V}^{\sim}[\![A^{\sqsubseteq}]\!]$.

With all of that preparation finished, we finally define the semantic interpretation of the graduality judgment $\Gamma^{\sqsubseteq} \vDash M_l \sqsubseteq M_r \in A^{\sqsubseteq}; \Gamma'^{\sqsubseteq}$ in the bottom of Figure 17. First, it says that both $M_l \sqsubseteq_{\prec} M_r$ and $M_l \sqsubseteq_{\succ} M_r$ hold, where we define $\sqsubseteq_{\sim}$ to mean that for any valid instantiation of the environments (including the preamble Γ_p), we get related computations. We can then define the "logical" Graduality theorem, that syntactic term precision implies semantic term precision, briefly, $\vdash$ implies $\vDash$.

Theorem 6.3 (Logical Graduality). *If* $\Gamma^{\sqsubseteq} \vdash M_l \sqsubseteq M_r : A^{\sqsubseteq}; \Gamma'^{\sqsubseteq}$, *then* $\Gamma^{\sqsubseteq} \vDash M_l \sqsubseteq M_r \in A^{\sqsubseteq}; \Gamma'^{\sqsubseteq}$

The proof is by induction on the term precision derivation. Each case is proven as a separate lemma in the appendix [New et al. 2020]. The cases of $\forall^{\nu}, \exists^{\nu}$, sealing and unsealing follow because

$$
\begin{array}{rcl}
\mathcal{V}_n^{\sim}[\![X]\!]\gamma\delta & = & \lfloor \delta(X) \rfloor_n \\
\mathcal{V}_n^{\sim}[\![?]\!]\gamma\delta & = & \{(w, \mathtt{inj}_{\sigma_l}\ V_l, \mathtt{inj}_{\sigma_r}\ V_r) \in \mathrm{Atom}_n[?]\delta \mid w.\eta \vDash (\sigma_l, \sigma_r, R) \wedge (w, V_l, V_r) \in \triangleright R\} \\
\mathcal{V}_n^{\prec}[\![\mathrm{tag}_G(A^{\sqsubseteq})]\!]\gamma\delta & = & \{(w, V_l, \mathtt{inj}_{\gamma_r(\mathrm{case}(G))}\ V_r) \in \mathrm{Atom}_n[\mathrm{tag}_G(A^{\sqsubseteq})]\delta \mid (w, V_l, V_r) \in \mathcal{V}_n^{\prec}[\![A^{\sqsubseteq}]\!]\gamma\delta\} \\
\mathcal{V}_n^{\succ}[\![\mathrm{tag}_G(A^{\sqsubseteq})]\!]\gamma\delta & = & \{(w, V_l, \mathtt{inj}_{\gamma_r(\mathrm{case}(G))}\ V_r) \in \mathrm{Atom}_n[\mathrm{tag}_G(A^{\sqsubseteq})]\delta \mid (w, V_l, V_r) \in (\triangleright \mathcal{V}^{\succ}[\![A^{\sqsubseteq}]\!]\gamma\delta)\} \\
\mathcal{V}_n^{\sim}[\![\forall^{\nu} X.B^{\sqsubseteq}]\!]\gamma\delta & = & \{(w, V_l, V_r) \in \mathrm{Atom}_n[\forall^{\nu} X.A^{\sqsubseteq}]\delta \mid \\
 & & \quad \forall R \in \mathrm{Rel}_{\omega}[A_l, A_r].\forall w' \sqsupseteq w.\forall \sigma_l, \sigma_r.w'.\eta \vDash (\sigma_l, \sigma_r, \lfloor R \rfloor_{w'.j}) \implies \\
 & & \quad (w', \mathtt{force}\ V_l\ [A_l]\ \sigma_l, \mathtt{force}\ V_r\ [A_r]\ \sigma_r) \in \mathcal{E}_n^{\sim}[\![B^{\sqsubseteq}]\!]\gamma'\delta' \\
 & & \quad (\text{where } \gamma' = \gamma, c_X \mapsto (\sigma_l, \sigma_r), \text{ and } \delta' = \delta, X \mapsto (A_l, A_r, R))\} \\
\mathcal{V}_n^{\sim}[\![\exists^{\nu} X.B^{\sqsubseteq}]\!]\gamma\delta & = & \{(w, \mathtt{pack}\ (A_l, V_l), \mathtt{pack}\ (A_r, V_r)) \in \mathrm{Atom}_n[\exists^{\nu} X.B^{\sqsubseteq}]\delta \mid \\
 & & \quad \exists R \in \mathrm{Rel}_{\omega}[A_l, A_r].\forall w' \sqsupseteq w.\forall \sigma_l, \sigma_r.w'.\eta \vDash (\sigma_l, \sigma_r, \lfloor R \rfloor_{w'.j}) \implies \\
 & & \quad (\mathtt{force}\ V_l\ \sigma_l, \mathtt{force}\ V_r\ \sigma_r) \in \mathcal{E}_n^{\sim}[\![B^{\sqsubseteq}]\!]\gamma'\delta' \\
 & & \quad (\text{where } \gamma' = \gamma, c_X \mapsto (\sigma_l, \sigma_r), \text{ and } \delta' = \delta, X \mapsto (A_l, A_r, R))\} \\
\mathcal{E}_n^{\prec}[\![A^{\sqsubseteq}]\!]\gamma\delta & = & \{(w, M_l, M_r) \in \mathrm{CAtom}_n[A^{\sqsubseteq}]\delta \mid (w.\Sigma_l, M_l) \mapsto^{w.j} \vee((w.\Sigma_l, M_l) \mapsto^{<w.j} (\Sigma_l', \mathfrak{V})) \\
 & & \vee(\exists w' \sqsupseteq w.(w', V_l, V_r) \in \mathcal{V}_n^{\prec}[\![A^{\sqsubseteq}]\!]\gamma\delta. \\
 & & \quad (w.\Sigma_l, M_l) \mapsto^{w'.j - w.j} (w'.\Sigma_l, \mathtt{ret}\ V_l) \wedge (w.\Sigma_r, M_r) \mapsto^{*} (w'.\Sigma_r, \mathtt{ret}\ V_r))\} \\
\mathcal{E}_n^{\succ}[\![A^{\sqsubseteq}]\!]\gamma\delta & = & \{(w, M_l, M_r) \in \mathrm{CAtom}_n[A^{\sqsubseteq}]\delta \mid (w.\Sigma_r, M_r) \mapsto^{w.j} \vee((w.\Sigma_l, M_l) \mapsto^{*} (\Sigma_l', \mathfrak{V})) \\
 & & \vee\exists w' \sqsupseteq w.(w', V_l, V_r) \in \mathcal{V}_n^{\prec}[\![A^{\sqsubseteq}]\!]\gamma\delta. \\
 & & \quad (w.\Sigma_l, M_l) \mapsto^{*} (w'.\Sigma_l, \mathtt{ret}\ V_l) \wedge (w.\Sigma_r, M_r) \mapsto^{w'.j - w.j} (w'.\Sigma_r, \mathtt{ret}\ V_r)\}
\end{array}
$$

$$
\mathcal{V}^{\sim}[\![A^{\sqsubseteq}]\!]\gamma\delta = \bigcup_{n \in \mathbb{N}} \mathcal{V}_n^{\sim}[\![A^{\sqsubseteq}]\!]\gamma\delta \qquad\qquad \mathcal{E}^{\sim}[\![A^{\sqsubseteq}]\!]\gamma\delta = \bigcup_{n \in \mathbb{N}} \mathcal{E}_n^{\sim}[\![A^{\sqsubseteq}]\!]\gamma\delta
$$

$$
\begin{array}{rcl}
\mathcal{G}^{\sim}[\![\cdot]\!] & = & \{(w, \emptyset, \emptyset) \mid \exists n.w \in \mathrm{World}_n\} \\
\mathcal{G}^{\sim}[\![\Gamma^{\sqsubseteq}, x : A^{\sqsubseteq}]\!] & = & \{(w, (\gamma, x \mapsto (V_l, V_r)), \delta) \mid (w, \gamma, \delta) \in \mathcal{G}^{\sim}[\![\Gamma^{\sqsubseteq}]\!] \wedge (w, V_l, V_r) \in \mathcal{V}^{\sim}[\![A^{\sqsubseteq}]\!]\gamma\delta\} \\
\mathcal{G}^{\sim}[\![\Gamma^{\sqsubseteq}, X]\!] & = & \{(w, (\gamma, c_X \mapsto (\sigma_l, \sigma_r)), \delta, X \mapsto (A_l, A_r, R)) \mid (w, \gamma, \delta) \in \mathcal{G}^{\sim}[\![\Gamma^{\sqsubseteq}]\!] \\
 & & \quad \wedge R \in \mathrm{Rel}_{\omega}[A_l, A_r] \wedge (\sigma_l, \sigma_r, \lfloor R \rfloor_{w.j}) \in w\} \\
\mathcal{G}^{\sim}[\![\Gamma^{\sqsubseteq}, X \cong A^{\sqsubseteq}]\!] & = & \{(w, (\gamma, c_X \mapsto (\sigma_l, \sigma_r)), \delta, X \mapsto (A_l, A_r, \mathcal{V}^{\sim}[\![A^{\sqsubseteq}]\!]\gamma\delta)) \mid (w, \gamma, \delta) \in \mathcal{G}^{\sim}[\![\Gamma^{\sqsubseteq}]\!] \\
 & & \quad w \vDash (\sigma_l, \sigma_r, \mathcal{V}_{w.j}^{\sim}[\![A^{\sqsubseteq}]\!]\gamma\delta)\}
\end{array}
$$

$$
\begin{array}{rcl}
\Gamma^{\sqsubseteq} \vDash M_l \sqsubseteq M_r \in A^{\sqsubseteq}; \Gamma^{\sqsubseteq\prime} & = & \Gamma^{\sqsubseteq} \vDash M_l \sqsubseteq_{\prec} M_r \in A^{\sqsubseteq}; \Gamma^{\sqsubseteq\prime} \wedge \Gamma^{\sqsubseteq} \vDash M_l \sqsubseteq_{\succ} M_r \in A^{\sqsubseteq}; \Gamma^{\sqsubseteq\prime} \\
\Gamma^{\sqsubseteq} \vDash M_l \sqsubseteq_{\sim} M_r \in A^{\sqsubseteq}; \Gamma^{\sqsubseteq\prime} & = & \forall (w, \gamma, \delta) \in \mathcal{G}^{\sim}[\![\Gamma_p, \Gamma^{\sqsubseteq}, \Gamma^{\sqsubseteq\prime}]\!]. \\
 & & \quad (w, [\![M_l]\!][\gamma_l][\delta_l], [\![M_r]\!][\gamma_l][\delta_l]) \in \mathcal{E}^{\sim}[\![A^{\sqsubseteq}]\!]\gamma\delta
\end{array}
$$

Fig. 17. Graduality/Parametricity Logical Relation (fragment)

the treatment of type variables between the value and environment relations is the same. In the hide case, the world is extended with $\mathcal{V}^{\sim}[\![A^{\sqsubseteq}]\!]$ as the relation between new cases. The cast cases are the most involved, following by two lemmas proven by induction over precision derivations: one for when the cast is on the left, and the other when the cast is on the right.

Finally, we prove the operational graduality theorem as a corollary of the logical graduality Theorem 6.3 and the adequacy Theorem 5.2. By constructing a suitable starting world w_{pre} that allocates the globally known tags, we ensure the operational graduality property holds for the code translated to $\mathrm{CBPV}_{\mathrm{OSum}}$, and then the simulation theorem implies the analogous property holds for the PolyC$^{\nu}$ operational semantics.

6.4 Parametricity and Free Theorems

Our relational approach to proving the graduality theorem is not only elegant, it also makes the theorem more general, and in particular it *subsumes* the parametricity theorem that we want for the language, because we already assign arbitrary relations to abstract type variables. Then the parametricity theorem is just the reflexivity case of the graduality theorem.

Theorem 6.4 (Parametricity). *If* $\Gamma \vdash M : A; \Gamma'$, *then* $\Gamma \vDash M^{+} \sqsubseteq M^{+} : A; \Gamma'$.

$$\frac{M : \forall^\nu X.X \to X \qquad V_A : A \qquad V_B : B}{\lambda\_ : ?.\mathtt{unseal}_X(M\{X \cong A\}(\mathtt{seal}_X V_A)) \approx^{\mathrm{ctx}} \lambda\_ : ?.\mathtt{let}\ y = M\{X \cong B\}V_B; V_A}$$

$$\frac{M : \forall^\nu X.\forall^\nu Y.(X \times Y) \to (Y \times X) \qquad V_A : A \qquad V_B : B}{\begin{array}{c}\lambda\_ : ?.\mathtt{let}\ (y,x) = (M\{X \cong A\}\{Y \cong B\}\,(\mathtt{seal}_X V_A, \mathtt{seal}_Y V_B)); (\mathtt{unseal}_X x, \mathtt{unseal}_Y y) \\ \approx^{\mathrm{ctx}} \lambda\_ : ?.\mathtt{let}\ (y,x) = (M\{X \cong B\}\{Y \cong A\}\,(\mathtt{seal}_X V_B, \mathtt{seal}_Y V_A)); (\mathtt{unseal}_Y y, \mathtt{unseal}_X x)\end{array}}$$

$$\frac{\begin{array}{c}\mathrm{NOT} = \lambda b : \mathbb{B}.\mathtt{if}\ b\ \mathtt{then}\ \mathtt{false}\ \mathtt{else}\ \mathtt{true} \\ \mathrm{WRAPNOT} = \lambda x : X.\mathtt{seal}_X(\mathrm{NOT}(\mathtt{unseal}_X x))\end{array}}{\begin{array}{c}\mathtt{pack}^\nu(X \cong \mathbb{B}, (\mathtt{seal}_X \mathtt{true}, (\mathrm{WRAPNOT}, \lambda x : X.\mathtt{unseal}_X x))) \\ \approx^{\mathrm{ctx}} \mathtt{pack}^\nu(X \cong \mathbb{B}, (\mathtt{seal}_X \mathtt{false}, (\mathrm{WRAPNOT}, \lambda x : X.\mathrm{NOT}\,(\mathtt{unseal}_X x))))\end{array}}$$

Fig. 18. Free Theorems without ?

To demonstrate that this really is a parametricity theorem, we show that from this theorem we can prove "free theorems" that are true in polymorphic languages. These free theorems are naturally stated in terms of *contextual equivalence*, the gold standard for operational equivalence of programs, which we define as both programs diverging, erroring, or terminating successfully when plugged into an arbitrary context. The appendix contains a formal definition [New et al. 2020].

To use our logical relation to prove contextual equivalence, we need the following lemma, which says that semantic term precision both ways is *sound* for PolyG$^\nu$ contextual equivalence.

Lemma 6.5. *If* $\Gamma \vDash M_l \sqsubseteq M_2 \in A; \Gamma_M$ *and* $\Gamma \vDash M_2 \sqsubseteq M_1 \in A; \Gamma_M$, *then* $\Gamma \vDash M_l \approx^{ctx} M_2 \in A; \Gamma_M$.

We now substantiate that this is a parametricity theorem by proving a few contextual equivalence results. First we present adaptations of some standard free theorems from typed languages in Figure 18. The first equivalence shows that the behavior of any term with the "identity function type" $\forall^\nu X.X \to X$ must be independent of the input it is given. We place a λ on each side to delimit the scope of the X outward. Without the X (or a similar thunking feature like $\forall^\nu$ or $\exists^\nu$), the two programs would not have the same (effect) typing. In a more realistic language, this corresponds to wrapping each side in a module boundary. The next result shows that a function $\forall^\nu X.\forall^\nu Y.(X \times Y) \to (Y \times X)$, if it terminates, must flip the values of the pair, and furthermore whether it terminates, diverges or errors does not depend on the input values. Finally, we give an example using existential types. That shows that an abstract "flipper" which uses `true` for on and `false` for off in its internal state is equivalent to one using `false` and `true`, respectively as long as they return the same value in their "readout" function.

Next, to give a flavor of what kind of relational reasoning is possible in the presence of the dynamic type, we consider what free theorems are derivable for functions of type $\forall X.? \to X$. A good intuition for this type is that the only possible outputs of the function are sealed values that are contained within the dynamically typed input. It is difficult to summarize this in a single statement, so instead we give the following three examples:

Theorem 6.6 ($\forall^\nu X.? \to X$ Free Theorems). *Let* $\cdot \vdash M : \forall^\nu X.? \to X$

(1) *For any* $\cdot \vdash V : ?$, *then* $\lambda\_ : ?.\mathtt{unseal}_X(M\{X \cong A\}\,V))\ \mathtt{true}$ *either diverges or errors.*
(2) *For any* $\cdot \vdash V : A$ *and* $\cdot \vdash V' : B$,
$\lambda\_ : ?.\mathtt{unseal}_X(M\{X \cong A\}\,(\mathtt{seal}_X V)) \approx^{ctx} \lambda\_ : ?.\mathtt{let}\ y = (\mathtt{unseal}_X(M\{X \cong B\}\,(\mathtt{seal}_X V')))\mathtt{;}\, V$
(3) *For any* $\cdot \vdash V : A, \cdot \vdash V' : B, \cdot \vdash V_d : ?$,
$\lambda\_ : ?.\mathtt{unseal}_X(M\{X \cong A\}\,(\mathtt{seal}_X V, V_d)) \approx^{ctx} \lambda\_ : ?.\mathtt{let}\ y = (\mathtt{unseal}_X(M\{X \cong B\}\,(\mathtt{seal}_X V', V_d)))\mathtt{;}\, V$

The first example passes in a value V that does not use the seal X, so we know that the function cannot possibly return a value of type X. The second example mimics the identity function's free

theorem. It passes in a sealed value V and the equivalence shows that V''s effects do not depend on what V was sealed and the only value that V' can return is the one that was passed in. The third example illustrates that there are complicated ways in which sealed values might be passed as a part of a dynamically typed value, but the principle remains the same: since there is only one sealed value that's part of the larger dynamically typed value, it is the only possible return value, and the effects cannot depend on its actual value. The proof of the first case uses the relational interpretation that X is empty. The latter two use the interpretation that X includes a single value.

Compare this reasoning to what is available in GSF, where the polymorphic *function* determines which inputs are sealed and which are not, rather than the caller. Because of this, Toro et al. [2019] only prove "cheap" theorems involving ? where the polymorphic function is known to be a literal Λ function and not a casted function. As an example, for arbitrary $M : \forall X.? \to X$, consider the application $M\,[\mathbb{B}]\,(\texttt{true}, \texttt{false})$. The continuation of this call has no way of knowing if neither, one or both booleans are members of the abstract type X. The following examples of possible terms for M illustrate these three cases:

(1) $M_1 = (\Lambda X.\lambda x : \mathbb{B} \times \mathbb{B}.\texttt{if or}\, x\, \texttt{then } \mho\, \texttt{else } \Omega) :: \forall X.? \to X$
(2) $M_2 = (\Lambda X.\lambda x : X \times \mathbb{B}.\texttt{if snd}\, x\, \texttt{then fst}\, x\, \texttt{else } \Omega) :: \forall X.? \to X$
(3) $M_3 = (\Lambda X.\lambda x : X \times X.\texttt{snd}\, x) :: \forall X.? \to X$

If $M = M_1$, both booleans are concrete so X is empty, but from the inputs the function can determine whether to diverge or error. If $M = M_2$, the first boolean is abstract and the second is concrete, so only the first can inhabit X, but the second can be used to determine whether to return a value or not. Finally if $M = M_3$, both booleans are abstract so the function cannot inspect them, but either can be returned. It is unclear what reasoning the continuation has here: it must anticipate every possible way in which the function might decide which values to seal, and so has to consider every dynamically typed value of the instantiating type as possibly abstract and possibly concrete.

7 DISCUSSION AND RELATED WORK

Dynamically typed PolyG$^\nu$ and Design Alternatives. Most gradually typed languages are based on adding types to an existing dynamically typed language, with the static types capturing some feature already existing in the dynamic language that can be migrated to use static typing. PolyG$^\nu$ was designed as a proof-of-concept standalone gradual language, so it might not be clear what dynamic typing features it supports migration of. In particular, since all sealing is explicit, PolyG$^\nu$ does not model migration from programming without seals entirely to programming with them, so its types are relevant to languages that include some kind of nominal data type generation.

The fresh existential types correspond to a particular mode of use of a module system that supports creation of nominal types. The package itself corresponds to a module with a fresh type declaration. Then sealing corresponds to the constructor of the fresh datatype, and unsealing to pattern matching against it. For example, in Racket structs can be used to make fresh nominal types and *units* provide first-class modules. It would be interesting future work to see if our logical relation can usefully be adapted to Typed Racket's typed units [Tobin-Hochstadt et al. [n. d.]].

Our fresh polymorphic types are more exotic than the fresh existentials, and don't clearly correspond to any existing programming features, but they model abstraction over nominal datatypes where the datatype is guaranteed to be freshly generated. One issue with adding this feature to a realistic language is that the outward scoping of type variables may be undesirable, so it is useful to consider alternative designs that achieve the same abstraction principles. One possible design would be to force an ANF-like [Sabry and Felleisen 1992] restriction on instantiations of polymorphic functions, where all instantiations have to be of the form $\texttt{let } M\{X \cong A\} = f; N$, where $X \cong A$ is bound in N. This makes the scope of the X explicit: it is only bound in N. Our translation could

easily be modified to accommodate this feature, we chose instead to consider the outward scoping since it makes it easier to compare to programming in the style of System F.

Our use of abstract and known type variables was directly inspired by Neis et al. [2009], who present a language with a fresh type creation mechanism which they show enables parametric reasoning though the language overall does not satisfy a traditional parametricity theorem. This suggests an alternative language design, where $\forall$ and $\exists$ behave normally and we add a `newtype` facility, analogous to that feature of Haskell, where `newtype` allocates a new case of the open sum type for each type it creates. Such a language would not have known type variables or PolyG$^\nu$'s inside-out scoping of type instantiation, but it would also not be parametric by default. Instead, programmers could manually create fresh types and know that they are abstract to other modules.

Since sealing is explicit in PolyG$^\nu$, it does not provide a drop-in replacement for System F, and so the additional syntactic overhead of sealing and unsealing can be quite heavy, especially when using higher-order combinators. For instance a higher-order function composition combinator has type $\forall^\nu X.\forall^\nu Y.\forall^\nu Z.(Y \to Z) \to (X \to Y) \to X \to Z$ and the System F composition $\mathsf{compose}[\mathbb{I}][\mathbb{B}][\mathbb{B}]\,\mathsf{not}\,(>0)$ would in PolyG$^\nu$ be written

$$\begin{aligned}\lambda n : \mathbb{I}.\, \mathtt{unseal}_Z(\mathsf{compose}\{X \cong \mathbb{I}\}\{Y \cong \mathbb{B}\}\{Z \cong \mathbb{B}\}(\lambda y : Y.\mathtt{seal}_Z(\mathsf{not}\,(\mathtt{unseal}_Y y)))\\ (\lambda x : X.\mathtt{seal}_Y(>0\,(\mathtt{unseal}_X x)))\\ \mathtt{seal}_X n)\end{aligned}$$

This syntactic overhead can be mitigated via generic wrapping functions using dynamic typing:

$$\begin{aligned}\mathsf{wrap}\,\mathtt{seal}_X\,\mathtt{unseal}_Z\,(\mathsf{compose}\{X \cong \mathbb{I}\}\{Y \cong \mathbb{B}\}\{Z \cong \mathbb{B}\}(\mathsf{wrap}\,\mathtt{unseal}_Y\,\mathtt{seal}_Z\,\mathsf{not})\\ (\mathsf{wrap}\,\mathtt{unseal}_X\,\mathtt{seal}_Y\,(>0))\end{aligned}$$

But the syntactic overhead cannot be completely removed or done entirely with static typing.

Tag Checking. Siek et al. [2015] claim that graduality demands that tag-checking functions like our `is(𝔹)?` form must error when applied to sealed values, and used this as a criticism of the design of Typed Racket. However, in our language, `is(𝔹)?` will simply return `false`, which matches Typed Racket's behavior. This is desirable if we are adding types to an existing dynamic language, because typically a runtime tag check should be a *safe* operation in a dynamically typed language. Explicit sealing avoids this graduality issue, an advantage over previous work.

Logical Relations. Our use of explicit sealing eliminates much of the complexity of prior logical relations [Ahmed et al. 2017; Toro et al. 2019]. To accommodate dynamic conversion and evidence insertion, those relations adopted complex value relations for universal types that in turn restricted the ways in which they could treat type variables. Additionally, we are the first to give a logical relation for fresh *existential* types, and it is not clear how to adapt the non-standard relation for universals to existentials [Ahmed et al. 2017; Toro et al. 2019].

Next, while we argue that our logical relation more fully captures parametricity than previous work on gradual polymorphism, this is not a fully formal claim. To formalize it, in future work we could show that PolyG$^\nu$ is a model of an effectful variant of an axiomatic parametricity formulation such as Dunphy [2002]; Ma and Reynolds [1991]; Plotkin and Abadi [1993].

ACKNOWLEDGMENTS

We thank the anonymous reviewers for their in-depth comments. We also thank Matthias Felleisen for explaining the importance that tag checks be total operations. This material is based on research supported by the National Science Foundation under grants CCF-1910522, CCF-1816837, and CCF-1453796. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Amal Ahmed, Derek Dreyer, and Andreas Rossberg. 2009. State-Dependent Representation Independence. In *ACM Symposium on Principles of Programming Languages (POPL), Savannah, Georgia.*

Amal Ahmed, Robert Bruce Findler, Jeremy Siek, and Philip Wadler. 2011. Blame for All. In *ACM Symposium on Principles of Programming Languages (POPL), Austin, Texas.* 201–214.

Amal Ahmed, Dustin Jamner, Jeremy G. Siek, and Philip Wadler. 2017. Theorems for Free for Free: Parametricity, With and Without Types. In *International Conference on Functional Programming (ICFP), Oxford, United Kingdom.*

Felipe Bañados Schwerter, Ronald Garcia, and Éric Tanter. 2014. A Theory of Gradual Effect Systems. In *Proceedings of the 19th ACM SIGPLAN International Conference on Functional Programming (ICFP '14).* 283–295.

John Boyland. 2014. The Problem of Structural Type Tests in a Gradually-Typed Language. In *21st Workshop on Foundations of Object-Oriented Languages.*

Giuseppe Castagna and Victor Lanvin. 2017. Gradual Typing with Union and Intersection Types. *Proc. ACM Program. Lang.* 1, ICFP, Article 41 (Aug. 2017), 28 pages. https://doi.org/10.1145/3110285

Tim Disney and Cormac Flanagan. 2011. Gradual Information Flow Typing. In *Workshop on Script-to-Program Evolution (STOP).*

Brian Patrick Dunphy. 2002. *Parametricity As a Notion of Uniformity in Reflexive Graphs.* Ph.D. Dissertation. Champaign, IL, USA. Advisor(s) Reddy, Uday.

Matthias Felleisen. 1990. On the expressive power of programming languages. *ESOP'90* (1990).

Luminous Fennell and Peter Thiemann. 2013. Gradual Security Typing with References. In *CSF.* IEEE Computer Society, 224–239.

Ronald Garcia and Matteo Cimini. 2015. Principal Type Schemes for Gradual Programs *(POPL '15).*

Ronald Garcia, Alison M. Clark, and Éric Tanter. 2016. Abstracting Gradual Typing. In *ACM Symposium on Principles of Programming Languages (POPL).*

Atsushi Igarashi, Peter Thiemann, Vasco Vasconcelos, and Philip Wadler. 2017b. Gradual Session Types. In *International Conference on Functional Programming (ICFP).*

Yuu Igarashi, Taro Sekiyama, and Atsushi Igarashi. 2017a. On Polymorphic Gradual Typing. In *International Conference on Functional Programming (ICFP), Oxford, United Kingdom.*

Lintaro Ina and Atsushi Igarashi. 2011. Gradual typing for generics. In *Proceedings of the 2011 ACM international conference on Object oriented programming systems languages and applications (OOPSLA '11).*

Nico Lehmann and Éric Tanter. 2017. Gradual Refinement Types. In *ACM Symposium on Principles of Programming Languages (POPL).*

Paul Blain Levy. 2003. *Call-By-Push-Value: A Functional/Imperative Synthesis.* Springer.

QingMing Ma and John C. Reynolds. 1991. Types, Abstractions, and Parametric Polymorphism, Part 2. In *Mathematical Foundations of Programming Semantics, 7th International Conference, Pittsburgh, PA, USA.*

John C. Mitchell and Gordon D. Plotkin. 1985. Abstract types have existential type. In *ACM Symposium on Principles of Programming Languages (POPL), New Orleans, Louisiana.*

Georg Neis, Derek Dreyer, and Andreas Rossberg. 2009. Non-Parametric Parametricity. In *International Conference on Functional Programming (ICFP).* 135–148.

Max S. New and Amal Ahmed. 2018. Graduality from Embedding-Projection Pairs. In *International Conference on Functional Programming (ICFP), St. Louis, Missouri.*

Max S. New, Dustin Jamner, and Amal Ahmed. 2020. Technical Appendix to Graduality and Parametricity: Together Again for the First Time. http://www.ccs.neu.edu/home/amal/papers/gradparam-tr.pdf

Max S. New, Daniel R. Licata, and Amal Ahmed. 2019. Gradual Type Theory *(POPL '19).*

Gordon Plotkin and Martín Abadi. 1993. A logic for parametric polymorphism. *Typed Lambda Calculi and Applications* (1993), 361–375.

John C. Reynolds. 1983. Types, Abstraction and Parametric Polymorphism. In *Information Processing 83, Proceedings of the IFIP 9th World Computer Congress, Paris, France.*

Amr Sabry and Matthias Felleisen. 1992. Reasoning about Programs in Continuation-Passing Style. In *Conf. on LISP and functional programming, LFP '92.*

Ilya Sergey and Dave Clarke. 2012. Gradual Ownership Types. In *ESOP (Lecture Notes in Computer Science)*, Vol. 7211. Springer, 579–599.

Jeremy Siek, Micahel Vitousek, Matteo Cimini, and John Tang Boyland. 2015. Refined Criteria for Gradual Typing. In *1st Summit on Advances in Programming Languages.*

Jeremy G. Siek and Walid Taha. 2006. Gradual Typing for Functional Languages. In *Scheme and Functional Programming Workshop (Scheme).* 81–92.

Jeremy G. Siek and Walid Taha. 2007. Gradual Typing for Objects. In *European Conference on Object-Oriented Programming (ECOOP).*

Eijiro Sumii and Benjamin C. Pierce. 2004. A Bisimulation for Dynamic Sealing. In *ACM Symposium on Principles of Programming Languages (POPL), Venice, Italy*.

Asumu Takikawa, T. Stephen Strickland, Christos Dimoulas, Sam Tobin-Hochstadt, and Matthias Felleisen. 2012. Gradual typing for first-class classes *(ACM Symposium on Object Oriented Programming: Systems, Languages, and Applications (OOPSLA))*.

Sam Tobin-Hochstadt and Matthias Felleisen. 2006. Interlanguage Migration: From Scripts to Programs. In *Dynamic Languages Symposium (DLS)*. 964–974.

Sam Tobin-Hochstadt and Matthias Felleisen. 2008. The Design and Implementation of Typed Scheme. In *ACM Symposium on Principles of Programming Languages (POPL), San Francisco, California*.

Sam Tobin-Hochstadt, Vincent St-Amour, Eric Dobson, and Asumu Takikawa. [n. d.]. Typed Racket Reference. https://docs.racket-lang.org/ts-reference/Typed_Units.html Accessed: 2019-10-30.

Matías Toro, Ronald Garcia, and Éric Tanter. 2018. Type-Driven Gradual Security with References. *ACM Transactions on Programming Languages and Systems* 40, 4 (Dec. 2018). http://doi.acm.org/10.1145/3229061

Matías Toro, Elizabeth Labrada, and Éric Tanter. 2019. Gradual Parametricity, Revisited. *Proc. ACM Program. Lang.* 3, POPL, Article 17 (Jan. 2019), 30 pages. https://doi.org/10.1145/3290330

Roger Wolff, Ronald Garcia, Éric Tanter, and Jonathan Aldrich. 2011. Gradual Typestate. In *Proceedings of the 25th European Conference on Object-oriented Programming (ECOOP'11)*.

Ningning Xie, Xuan Bi, and Bruno C. d. S. Oliveira. 2018. Consistent Subtyping for All. In *Programming Languages and Systems*, Amal Ahmed (Ed.). Springer International Publishing, Cham, 3–30.