

Type-Level Computation One Step at a Time

Foo Bar Baz

The University of Foo

{foo,bar,baz}@foo.edu

Abstract

Many type systems support a conversion rule that allows type-level computation. In such type systems ensuring the *decidability* of type checking requires type-level computation to terminate. For calculi where the syntax of types and terms is the same, the decidability of type-checking is usually dependent on the strong normalization of the calculus, which ensures termination. An unfortunate consequence of this coupling between decidability and strong normalization is that adding (unrestricted) general recursion to such calculi is not possible.

This paper proposes an alternative to the conversion rule that allows the same syntax for types and terms, type-level computation, and preserves decidability of type-checking under the presence of general recursion. The key idea, which is inspired by the traditional treatment of *iso-recursive types*, is to make each type-level computation step explicit. Each beta reduction or expansion at the type-level is introduced by a language construct. This allows control over the type-level computation and ensures decidability of type-checking even in the presence of non-terminating programs at the type-level. We realize this idea by presenting a variant of the calculus of constructions with general recursion and recursive types. Furthermore we show how many advanced programming language features of state-of-the-art functional languages (such as Haskell) can be encoded in our minimalistic core calculus.

Categories and Subject Descriptors D.3.1 [Programming Languages]: Formal Definitions and Theory

General Terms Languages, Design

Keywords Dependent types, Intermediate language

1. Introduction

Modern statically typed functional languages (such as ML, Haskell, Scala or OCaml) have increasingly expressive type systems. Often these large source languages are translated into a much smaller typed core language. The choice of the core language is essential to ensure that all the features of the source language can be encoded. For a simple polymorphic functional language it is possible, for example, to pick a variant of System F as a core language. However, the desire for more expressive type system features puts pressure on

the core languages, often requiring them to be extended to support new features. For example, if the source language supports *higher-kinded types* or *type-level functions* then System F is not expressive enough and can no longer be used as the core language. Instead another core language that does provide support for higher-kinded types, such as System F_ω , needs to be used. However System F_ω is significantly more complex than System F and thus harder to maintain. If later a new feature, such as *kind polymorphism*, is desired the core language may need to be changed again to account for the new feature, introducing at the same time new sources of complexity. Indeed the core language for modern versions of functional languages are quite complex, having multiple syntactic sorts (such as terms, types and kinds), as well as dozens of language constructs [BRUNO: F_C].

The more expressive type systems become, the more types become similar to the terms. Therefore a natural idea is to unify terms and types. There are obvious benefits in this approach: only one syntactic level (terms) is needed; and there are much less language constructs, making the core language easier to implement and maintain. At the same time the core language becomes more expressive, giving us for free many useful language features. *Pure type systems* [] build on this observation and they show how a whole family of type systems (including System F and System F_ω) can be implemented using just a single syntactic form. With the added expressiveness it is even possible to have type-level programs expressed using the same syntax as terms as well as dependently typed programs [].

However having the same syntax for types and terms can also be problematic. If arbitrary type-level computation is allowed then type-level programs can use the same language constructs as terms. Usually type systems have a conversion rule to support type-level computation. In such type systems ensuring the *decidability* of type checking requires type-level computation to terminate. For calculi where the syntax of types and terms is the same, the decidability of type-checking is usually dependent on the strong normalization of the calculus, which ensures termination. An unfortunate consequence of this coupling between decidability and strong normalization is that adding (unrestricted) general recursion to such calculi is not possible. There is a clear tension between decidability of type-checking and allowing general recursion at the type-level.

This paper proposes λ_\star^μ : a variant of the calculus of constructions allows the same syntax for types and terms, supports type-level computation, and preserves decidability of type-checking under the presence of general recursion. In λ_\star^μ , each type-level computation step is explicit. **BRUNO: emphasis on the advantages: a minimal core language?** The key idea, which is inspired by the traditional treatment of *iso-recursive types*, is to introduce each beta reduction or expansion at the type-level by a *type-safe cast*. The casts allow control over the type-level computation. For example, if a type-level program requires two beta-reductions to reach normal form, then two casts are needed in the program. If a non-terminating program is used at the type-level, it is not pos-

sible to cause non-termination in the type-checker, because that would require a program with an infinite number of casts. Therefore, since single beta-steps are trivially terminating, decidability of type-checking is possible even in the presence of non-terminating programs at the type-level.

Our motivation to develop λ_\star^μ is to use it as a simpler alternative to existing core languages for languages such as Haskell. The paper shows how many of programming language features of Haskell, including some of the latest extensions, can be encoded in λ_\star^μ via a source language. In particular the source language supports *algebraic datatypes*, *higher-kinded types*, *nested datatypes*, *kind polymorphism* [] and *datatype promotion* []. This result is interesting because λ_\star^μ is a minimal calculus with only 8 language constructs and a single syntactic sort. In contrast the latest versions of System F_C (Haskell’s core language) have multiple syntactic sorts and dozens of language constructs. Even if support for equality and coercions, which constitutes a significant part of System F_C , would be removed the resulting language would still be significantly larger and more complex than λ_\star^μ .

BRUNO: λ_\star^μ sacrifices the convenience of use of type-level computation to gain the ability of doing arbitrary general recursion at the term level. We believe λ_\star^μ is particularly well-suited as a core for Haskell-like languages. In particular the treatment of type-level computation shares similar ideas with Haskell. Although Haskell’s surface language provides a rich set of mechanisms to do type-level computation [], the core language lacks fundamental mechanisms too type-level computation. In particular, like in λ_\star^μ , type equality is purely syntactic (modulo alpha-conversion).

In summary, the contributions of this work are:

- Decidable type checking and managed type-level computation by replacing implicit conversion rule of CoC with generalized fold/unfold semantics.
- A core language based on Calculus of Constructions (CoC) that collapses terms, types and kinds into the same hierarchy, supports general recursion...
- General recursion by introducing recursive types for both terms and types by the same μ primitive.
- Surface language that supports datatypes, pattern matching and other language extensions for Haskell, and can be encoded into the core language.

2. Overview

This section informally introduces the main features of λ_\star^μ . In particular, this section shows how the explicit casts in λ_\star^μ can be used instead of the typical conversion rule present in calculi such as the calculus of constructions. The formal details of λ_\star^μ are presented in Section 4. **JEREMY:** to distinguish code from **Fun** and λ_\star^μ , we may want to use different fonts, e.g., typewriter font for **Fun**

2.1 The Calculus of Constructions and the Conversion Rule

The calculus of constructions (λC) [10] is a powerful higher-order typed lambda calculus supporting dependent types (among various other features). A crucial feature of λC is the so-called *conversion* rule:

$$\frac{\Gamma \vdash e : \tau_1 \quad \Gamma \vdash \tau_2 : s \quad \tau_1 =_\beta \tau_2}{\Gamma \vdash e : \tau_2} \text{ TCC.CONV}$$

The conversion rule allows one to derive $e : \tau_2$ from the derivation of $e : \tau_1$ and the β -equality of τ_1 and τ_2 . This rule is important to *automatically* allow terms with equivalent types to be considered type-compatible. The following example illustrates the

use of the conversion rule:

$$f \equiv \lambda x : (\lambda y : \star.y) \text{Int}.x$$

Here f is a simple identity function. Notice that the type of x (i.e., $(\lambda y : \star.y) \text{Int}$), which is the argument of f , is interesting: it is an identity function on types, applied to an integer. Without the conversion rule, f cannot be applied to, say 3 in λC . However, given that f is actually β -convertible to $\lambda x : \text{Int}.x$, the conversion rule allows the application of f to 3 by implicitly converting $\lambda x : (\lambda y : \star.y) \text{Int}.x$ to $\lambda x : \text{Int}.x$.

Decidability of Type-Checking and Strong Normalization While the conversion rule in λC brings a lot of convenience, an unfortunate consequence is that it couples decidability of type-checking with strong normalization of the calculus [8]. However strong normalization does not hold with general recursion. This is because due to the conversion rule, any non-terminating term would force the type checker to go into an infinitely loop (by constantly applying the conversion rule without termination), thus rendering the type system undecidable.

To illustrate the problem of the conversion rule with general recursion, let us consider a simple example. Suppose that d is a “dependent type” that has type $\text{Int} \rightarrow \star$. With general recursion at hand, we can image a term z that has type $d \text{ loop}$, where loop stands for any diverging computation of type Int . What would happen if we try to type check the following application:

$$(\lambda x : d \ 3.x) z$$

Under the normal typing rules of λC , the type checker would get stuck as it tries to do β -equality on two terms: $d \ 3$ and $d \text{ loop}$, where the latter is non-terminating.

2.2 An Alternative to the Conversion Rule: Explicit Casts

In contrast to the implicit reduction rules of λC , λ_\star^μ makes it explicit as to when and where to convert one type to another. Type conversions are explicit by introducing two language constructs: cast_\downarrow (beta reduction) and cast^\uparrow (beta expansion). The benefit of this approach is that decidability of type-checking no longer is coupled with strong normalization of the calculus.

Beta Reduction The first of the two type conversions (cast_\downarrow), allows a type conversion provided that the resulting type is a *beta reduction* of the original type of the term. The use of cast_\downarrow is better explained by the following simple example. Assume a definition g :

$$g \equiv \lambda x : \text{Int}.x$$

and term z with type

$$(\lambda y : \star.y) \text{Int}$$

In λ_\star^μ , and in contrast to the calculus of constructions, $g \ z$ is an ill-typed application. To type-check the application of g to z an explicit type conversion must be used:

$$g (\text{cast}_\downarrow z)$$

In this case cast_\downarrow is used because the program requires a (type-level) beta-reduction: $(\lambda y : \star.y) \text{Int} \rightarrow_\beta \text{Int}$.

Beta Expansion The dual operation of cast_\downarrow is cast^\uparrow , which allows a type conversion provided that the resulting type is a *beta expansion* of the original type of the term. Let us revisit the example from Section 2.1. In λ_\star^μ , $f \ 3$ is an ill-typed application. Instead we must write the application as

$$f (\text{cast}^\uparrow [(\lambda y : \star.y) \text{Int}] 3)$$

Intuitively, cast^\uparrow is doing a type conversion, as the type of 3 is Int , and $(\lambda y : \star.y) \text{Int}$ is the beta expansion of Int (witnessed by $(\lambda y : \star.y) \text{Int} \rightarrow_\beta \text{Int}$). Notice that for cast^\uparrow to work, we need

to provide the resulting type as argument. This is because for the same term, there are more than one choices for beta expansions (e.g., $1 + 2$ and $2 + 1$ are both the beta expansions of 3).

One-Step A final point to make is that the cast rules specify *one-step* reductions/expansions. **BRUNO: show an example with two beta-reductions and show that two casts would be needed then.** This enables us to have more control over type-level computation. The full technical details about cast rules are presented in Section 4.

2.3 Decidability without Strong Normalization

With explicit type conversion rules the decidability of type-checking no longer depends on the normalization property. A nice consequence of this is that the type system remains decidable even in the presence of non-terminating programs at type level.

To illustrate, let us consider the same example discussed in Section 2.1. Now the type checker will not get stuck when type-checking the following application:

$$(\lambda x : d\ 3.x) z$$

where the type of z is $d\ loop$. This is because in λ_\star^μ , the type checker only does syntactic comparison between $d\ 3$ and $d\ loop$, instead of β -equality. Therefore it rejects the above application as ill-typed. Indeed it is impossible to type-check the application even with the use of cast^\uparrow and/or cast_\downarrow : one would need to write infinite number of cast_\downarrow 's to make the type checker loop forever (e.g., $(\lambda x : d\ 3.x)(\text{cast}_\downarrow(\text{cast}_\downarrow \dots z))$). But it is impossible to write such program in reality.

In summary, λ_\star^μ achieves the decidability of type checking by explicitly controlling type-level computation, which is independent of the normalization property, while supporting general recursion at the same time.

2.4 Recursion and Recursive Types

λ_\star^μ supports general recursion, and allows writing standard recursive programs at the term level. At the same time, the recursive construct can also be used to model recursive types at the type-level. Therefore, λ_\star^μ differs from other programming languages in that it *unifies* both recursion and recursive types by the same μ primitive. With a single language construct we get two powerful features!

Recursion The μ primitive can be used to define recursive functions. For example, the factorial function is written in λ_\star^μ as:

$$\text{fact} = \mu f : \text{Int} \rightarrow \text{Int}. \text{if } x == 0 \text{ then } 1 \text{ else } x \times f(x - 1)$$

BRUNO: x is not bound!**BRUNO: show how to use fact** The above recursive definition works because of the dynamic semantics of the μ primitive:

$$\frac{}{\mu x : \tau. e \longrightarrow e[x \mapsto \mu x : \tau. e]} \text{ S_MU}$$

which is exactly doing recursive unfolding of itself.

It is worth noting that the type τ in S_MU is not restricted to function types. This extra freedom allows us to define a record of mutually recursive functions as the fixed point of a function on records.

Recursive types In the literature on type systems, there are two approaches to recursive types, namely *equi-recursive* and *iso-recursive***BRUNO: reference.** The *iso-recursive* approach treats a recursive type and its unfolding as different, but isomorphic. The isomorphism between a recursive type and its one step unfolding is witnessed by fold and unfold operations. In λ_\star^μ , the isomorphism is witnessed by first cast^\uparrow , then cast_\downarrow . In fact, cast^\uparrow and cast_\downarrow actually generalize fold and unfold: they can convert any types, not just recursive types.

To demonstrate the use of the cast rules with recursive types, let us consider a classic example of a recursive type, the so-called “hungry” type [20]:

$$H = \mu \sigma : \star. \text{Int} \rightarrow \sigma$$

A term z of type H can accept any number of integers and return a new function that is hungry for more, as illustrated below:

$$\begin{aligned} \text{cast}_\downarrow z : \text{Int} &\rightarrow H \\ \text{cast}_\downarrow(\text{cast}_\downarrow z) : \text{Int} &\rightarrow \text{Int} \rightarrow H \\ \text{cast}_\downarrow(\text{cast}_\downarrow \dots z) : \text{Int} &\rightarrow \text{Int} \rightarrow \dots \rightarrow H \end{aligned}$$

Call-by-Name Due to the unification, the *call-by-value* evaluation strategy does not fit in our setting. In call-by-value evaluation, recursion can be expressed by the recursive binder μ as $\mu f : T \rightarrow T. E$ (note that the type of f is restricted to function types). Since we don’t want to pose restrictions on the types, the *call-by-name* evaluation is a sensible choice. **BRUNO: Probably needs to be improved. I’ll come back to this later!**

2.5 Logical Inconsistency

Although the decidability of type-checking is preserved, a consequence having general recursion in λ_\star^μ is that the logical consistency is lost. In λ_\star^μ every type is inhabited, which is problematic if we want to view λ_\star^μ as a logic. Indeed, unlike many other dependently calculi which have been proposed in the past, λ_\star^μ cannot be used directly to do theorem proving.

Nevertheless the loss of logical consistency is a deliberate design decision. Although there are dependently typed languages that support general recursion and still preserve logical consistency, this is done at the cost of additional complexity in the system [**BRUNO: references.** In λ_\star^μ we trade the loss of logical consistency by a significantly simpler system.

Since our goal is use λ_\star^μ as a foundational calculus for languages like Haskell, logical consistency is not an important property: traditional functional languages like Haskell are logically inconsistent as well. For example, in Haskell, we can write a “false” type:

$$\text{type False} = \forall a. a$$

With general recursion, a value with type *False* is given:

$$\begin{aligned} \text{false} &:: \text{False} \\ \text{false} &= \text{false} \end{aligned}$$

whose denotational semantics is \perp , which corresponds to inconsistency in logic.

Type in Type Since logical consistency is already lost due to general recursion, λ_\star^μ also uses the $\star : \star$ axiom**BRUNO: reference?**. As a consequence, having this rule adds expressiveness and simplifies our system (e.g., it will be easy to explicitly quantify over kinds). We return to this issue in Section 7.

2.6 Encoding Datatypes

The explicit type conversion rules and the μ primitive facilitates the encoding of recursive datatypes and recursive functions over datatypes. While inductive datatypes can be encoded using either the Church [20] or the Scott encoding [18], we adopt the Scott encoding as it encodes case analysis, making it more convenient to encode pattern matching. We demonstrate the encoding method using a simple datatype as a running example: Peano numbers.

The datatype declaration for Peano numbers in Haskell is:

$$\text{data Nat} = Z \mid S \text{ Nat}$$

In the Scott encoding, the encoding of the *Nat* datatype reflects how its two constructors are going to be used. Since *Nat* is a recursive

datatype, we have to use recursive types at some point to reflect its recursive nature. As it turns out, the typed Scott encoding of *Nat* is:

$$\mu X : \star. \Pi B : \star. B \rightarrow (X \rightarrow B) \rightarrow B$$

The function type $B \rightarrow (X \rightarrow B) \rightarrow B$ demystifies the recursive nature of *Nat*: B corresponds to the type of the constructor Z , and $X \rightarrow B$ corresponds to the type of the constructor S . The intuition is that any recursive use of the datatype in the data constructors is replaced with the variable (X in the case) bound by μ , and we make the resulting variable (B in this case) universally quantified so that elements of the datatype with different result types can be used in the same program [15].

The two constructors can be encoded correspondingly via the cast rules:

$$\begin{aligned} Z &= \text{cast}^\uparrow [\text{Nat}] (\lambda B : \star. \lambda z : B. \lambda f : \text{Nat} \rightarrow B. z) \\ S &= \lambda n : \text{Nat}. \text{cast}^\uparrow [\text{Nat}] (\lambda B : \star. \lambda z : B. \lambda f : \text{Nat} \rightarrow B. f n) \end{aligned}$$

BRUNO: be careful with code overflowing margins! Maybe use 2 lines? Intuitively, each constructor selects a different function from the function parameters (z and f in the above example). This provides branching in the process flow, based on the constructors. Note that we use the cast^\uparrow operation to do type conversion between the recursive type and its unfolding.

Finally a recursive function that adds two natural numbers is defined as follows:

$$\begin{aligned} \mu f : \text{Nat} \rightarrow \text{Nat} \rightarrow \text{Nat}. \lambda n : \text{Nat}. \lambda m : \text{Nat}. \\ (\text{cast}_\downarrow n) \text{Nat } m (\lambda n' : \text{Nat}. S (f n' m)) \end{aligned}$$

The above definition resembles case analysis commonly seen in modern functional programming languages. **BRUNO: Explain the use of cast** Indeed, we will formalize an encoding of case analysis in Section 6.

3. Fun: A Surface Language on Top of λ_\star^μ

The main goal of λ_\star^μ is to serve as an expressive core language for functional languages like Haskell or ML. This section shows a number of programs written in the surface language **Fun**, which is built on top of λ_\star^μ . We illustrate the expressiveness of λ_\star^μ by encoding functional programs that require some of the latest features of Haskell, or are non-trivial to encode in dependently typed language like Coq or Agda. All examples shown in this section are runnable in our prototype interpreter. The formalization of the surface language is presented in Section 6.

Datatypes Conventional datatypes like natural numbers or polymorphic lists can be easily defined in **Fun**, as in Haskell. For example, below is the definition of polymorphic lists:

$$\text{data List } (a : \star) = \text{Nil} \mid \text{Cons } (x : a) (xs : \text{List } a);$$

Because **Fun** is explicitly typed, each type parameter needs to be accompanied with a corresponding kind expression. The use of the above datatype is best illustrated by the *length* function:

$$\begin{aligned} \text{letrec length} : (a : \star) \rightarrow \text{List } a \rightarrow \text{nat} = \\ \lambda a : \star. \lambda l : \text{List } a. \text{case } l \text{ of} \\ \quad \text{Nil} \Rightarrow 0 \\ \quad \mid \text{Cons } (x : a) (xs : \text{List } a) \Rightarrow 1 + \text{length } a \text{ } xs \\ \text{in} \\ \text{let test} : \text{List } \text{nat} = \text{Cons } \text{nat } 1 (\text{Cons } \text{nat } 2 (\text{Nil } \text{nat})) \\ \text{in length nat test} \quad \text{-- returns 2} \end{aligned}$$

The *length* function is recursive. **Fun** supports a standard **letrec** construct that facilitates defining recursive functions. Note that due to explicit typing, the program requires quite a few type annotations and type parameters. However, apart from the extra typing, the

program is similar to the code that would be written in a language like Haskell or ML.

HOAS Higher-order abstract syntax **BRUNO: reference** is a representation of abstract syntax where the function space of the meta-language is used to encode the binders of the object language. We show an example of encoding a simple lambda calculus:

$$\begin{aligned} \text{data Exp} &= \text{Num } (n : \text{nat}) \\ &\mid \text{Lam } (f : \text{Exp} \rightarrow \text{Exp}) \\ &\mid \text{App } (a : \text{Exp}) (b : \text{Exp}); \end{aligned}$$

Note that, in the lambda constructor (*Lam*), the recursive occurrence of the datatype appears in a negative position (i.e., in the left side of a function arrow). Systems like Coq and Agda would reject such programs HOAS since it is well-known that such datatypes can lead to logical inconsistencies. Moreover, such logical inconsistencies can be exploited to write non-terminating computations, and make type-checking undecidable. However **Fun** is able to express HOAS in a straightforward way, while preserving decidable type-checking.

Using *Exp* we can write an evaluator for the lambda calculus. As noted by [11], the evaluation function needs an extra function *reify* to invert the result of evaluation. The full code for the evaluator is shown next. **BRUNO: consider removing the error branches to make the code shorter.**

$$\begin{aligned} \text{data Value} &= \text{VI } (n : \text{nat}) \mid \text{VF } (f : \text{Value} \rightarrow \text{Value}); \\ \text{rcrd Eval} &= \text{Ev } \{ \text{eval}' : \text{Exp} \rightarrow \text{Value}, \text{reify}' : \text{Value} \rightarrow \text{Exp} \}; \\ \text{letrec ev} : \text{Eval} = \\ &\quad \text{Ev } (\lambda e : \text{Exp}. \text{case } e \text{ of} \\ &\quad \quad \text{Num } (n : \text{nat}) \Rightarrow \text{VI } n \\ &\quad \mid \text{Lam } (fun : \text{Exp} \rightarrow \text{Exp}) \Rightarrow \\ &\quad \quad \text{VF } (\lambda e' : \text{Value}. \text{eval}' \text{ ev } (fun (\text{reify}' \text{ ev } e'))) \\ &\quad \mid \text{App } (a : \text{Exp}) (b : \text{Exp}) \Rightarrow \\ &\quad \quad \text{case eval}' \text{ ev } a \text{ of} \\ &\quad \quad \quad \text{VI } (n : \text{nat}) \Rightarrow \text{error} \\ &\quad \mid \text{VF } (fun : \text{Value} \rightarrow \text{Value}) \Rightarrow fun (\text{eval}' \text{ ev } b)) \\ &\quad (\lambda v : \text{Value}. \text{case } v \text{ of} \\ &\quad \quad \text{VI } (n : \text{nat}) \Rightarrow \text{Num } n \\ &\quad \mid \text{VF } (fun : \text{Value} \rightarrow \text{Value}) \Rightarrow \\ &\quad \quad \text{Lam } (\lambda e' : \text{Exp}. (\text{reify}' \text{ ev } (fun (\text{eval}' \text{ ev } e'))))) \\ &\quad \text{in let eval} : \text{Exp} \rightarrow \text{Value} = \text{eval}' \text{ ev} \end{aligned}$$

The definition of the evaluator is quite straightforward, although it is worth noting that the evaluator is a partial function that can cause run-time errors. Thanks to the flexibility of the μ primitive, mutual recursion can be encoded by using records **BRUNO: explain better! what are the two functions being defined mutually recursively!**

Evaluation of a lambda expression proceeds as follows:

$$\begin{aligned} \text{let test} : \text{Exp} &= \text{App } (\text{Lam } (\lambda f : \text{Exp}. \text{App } f (\text{Num } 42))) \\ &\quad (\text{Lam } (\lambda g : \text{Exp}. g)) \\ \text{in show } (\text{eval test}) &\quad \text{-- return 42} \end{aligned}$$

Higher-kinded Types Higher-kinded types are types that take other types and produce a new type. To support higher-kinded types, languages like Haskell use core languages to account for kind expressions. (The existing core language of Haskell, System FC, is an extension of System F_ω [12], which naively supports higher-kinded types.) Given that λ_\star^μ subsumes System F_ω , we can easily construct higher-kinded types. We show this by an example of encoding a *functor*:

$$\begin{aligned} \text{rcrd Functor } (f : \star \rightarrow \star) &= \\ \text{Func } \{ \text{fmap} : (a : \star) \rightarrow (b : \star) \rightarrow (a \rightarrow b) \rightarrow f a \rightarrow f b \}; \end{aligned}$$

Here we use a record (by using `rcrd` keyword) to represent a functor, whose only field is a single method, called `fmap`. The functor “instance” of the *Maybe* datatype is:

```
let maybeInst : Functor Maybe =
  Func Maybe (λa : *. λb : *. λf : a → b. λx : Maybe a.
    case x of
      Nothing ⇒ Nothing b
      | Just (z : a) ⇒ Just b (f z))
```

BRUNO: Code needs to be explained! In particular explain the *Func Maybe*.

Fixpoints of Functors Various functional programming techniques employ type-level fixpoints to achieve additional modularity. **BRUNO:** references: *Datatypes a la carte*. Thus, type-level fixpoints are a good example to demonstrate the expressiveness of **Fun**. The definition is:

```
rcrd Fix (f : * → *) = In { out : f (Fix f) };
```

Note that the record notation also introduces the selector function:

```
out : (f : * → *) → Fix f → f (Fix f)
```

The *Fix* datatype is interesting in that now we can define recursive datatypes in a non-recursive way. For instance, a non-recursive definition for natural numbers is:

```
data NatF (self : *) = Zero | Succ self;
```

And the recursive version is just a synonym:

```
let Nat : * = Fix NatF
```

Given *fmap*, many recursive schemes can be defined, for example we can have *catamorphism* [16] or generic function fold:

```
letrec cata : (f : * → *) → (a : *) →
  Functor f → (f a → a) → Fix f → a =
  λf : * → *. λa : *. λm : Functor f. λg : f a → a. λt : Fix f.
    g (fmap f m (Fix f) a (cata f a m g) (out f t))
```

BRUNO: fix formatting of *cata* Unfortunately, in systems like Coq, definitions like *Fix* must be rejected. The problem is related to the possibility of recursive definitions in negative positions. Although the datatype *Fix* appears not to use negative recursive definitions of itself, the type constructor *f* could use the parameter on a negative position. **BRUNO:** Show example that defines a bad functor. Similarly to the HOAS example, this would violate the strictly positive restrictions of Coq. Nevertheless, in **Fun** such definition is also allowed without hindering decidability of type-checking.

Nested Datatypes A nested datatype [5], also known as a *non-regular* datatype, is a parametrised datatype whose definition contains different instances of the type parameters. Functions over nested datatypes usually involve polymorphic recursion. We show that **Fun** is capable of defining nested datatypes and functions over a nested datatype. A simple example would be the type *Pow* of power trees, whose size is exactly a power of two, declared as follows:

```
data PairT (a : *) = P (x : a) (x : a);
data Pow (a : *) = Zero (n : a)
  | Succ (t : Pow (PairT a));
```

Notice that the recursive occurrence of *Pow* does not hold an *a*, but *PairT a*. **BRUNO:** code should be formatted with the proper font! don't use *emphasized for code*. Quite a few changes throughout!. This means every time we use a *Succ* constructor, the size of the pairs doubles. It is instructive to look at the encoding of *Pow* in λ^u_x :

```
let Pow : * → * = mu X : * → *.
  λa : *. (B : *) → (a → B) → (X (PairT a) → B) → B
```

Notice how the higher-kinded type variable $X : * \rightarrow *$ helps encoding nested datatypes. Below is a simple function *toList* that transforms a power tree into a list:

```
letrec toList : (a : *) → Pow a → List a =
  λa : *. λt : Pow a. case t of
    Zero (x : a) ⇒ Cons a x (Nil a)
    | Succ (c : Pow (PairT a)) ⇒
      concatMap (PairT a) a
        (λx : PairT a. case x of
          P (m : a) (n : a) ⇒
            Cons a m (Cons a n (Nil a)))
      (toList (PairT a) c)
```

Kind Polymorphism In Haskell, System FC [30] was proposed to support, among other things **BRUNO:** was this System FC or a later extension of System FC? Please double-check!, kind polymorphism. **BRUNO:** story is not precise here: Previous versions of Haskell, based on System F_ω already had terms, types and kinds. What you want to say is that this was still insufficient for kind polymorphism and yet some more extensions to the core were needed. expressions into terms, types and kinds, which complicates both the implementation and future extensions. **Fun** natively allows definitions to have polymorphic kinds. Here is an example, taken from [30], of a datatype that benefits from kind polymorphism: a higher-kinded fixpoint combinator:

```
data Mu (k : *) (f : (k → *) → k → *) (a : k) =
  Roll (g : f (Mu k f) a);
```

Mu can be used to construct polymorphic recursive types of any kind, for instance:

```
data Listf (f : * → *) (a : *) = Nil
  | Cons (x : a) (xs : (f a));
let List : * → * = λa : *. Mu * Listf a
```

Datatype Promotion Recent versions of Haskell introduced Datatype Promotion [], in order to allow ordinary datatypes as kinds, and data constructors as types. With the full power of dependent types, data promotion is made trivial in **Fun**.

As a last example, we show a representation of a labeled binary tree, where each node is labeled with its depth in the tree. Below is the definition:

```
data PTree (n : Nat) = Empty
  | Fork (z : nat) (x : PTree (S n)) (y : PTree (S n));
```

Notice how the datatype *Nat* is “promoted” to be used in the kind level. Next we can construct such a binary tree that keeps track of its depth statically:

```
Fork Z 1 (Empty (S Z)) (Empty (S Z))
```

If we accidentally write the wrong depth, for example:

```
Fork Z 1 (Empty (S Z)) (Empty Z)
```

The above will fail to pass type checking.

BRUNO: More examples? closed type families; dependent types? **JEREMY:** had hard time thinking of a simple, non-recursive example for type families

4. A Dependently Typed Calculus with Casts

In this section, we present the λ_* calculus. This calculus is very close to the calculus of constructions, except for two key differences: 1) the absence of the \Box constant (due to use of the ‘type-in-type’ axiom); 2) the existence of two cast operators. Like the calculus

e, τ, σ	$::=$	Expressions
x		Variable
\star		Type of types
$e_1 e_2$		Application
$\lambda x : \tau. e$		Abstraction
$\prod x : \tau. \sigma$		Product
$\text{cast}^\uparrow[\tau] e$		Cast up by expansion
$\text{cast}_\downarrow e$		Cast down by reduction
Γ	$::=$	Contexts
\emptyset		Empty
$\Gamma, x : \tau$		Variable binding
v	$::=$	Values
\star		Type of types
$\lambda x : \tau. e$		Abstraction
$\prod x : \tau_1. \tau_2$		Product
$\text{cast}^\uparrow[\tau] e$		Cast up by expansion

Figure 1. Syntax of λ_\star

of constructions, λ_\star has decidable type-checking. However, unlike λC the proof of decidability of type-checking does not require the strong normalization of the calculus. In rest of this section, we demonstrate the syntax, operational semantics, typing rules and meta-theory of λ_\star .

4.1 Syntax

Figure 1 shows the syntax of λ_\star , including expressions, contexts and values. λ_\star uses a unified syntactic representation for different levels of expressions by following the *pure type system* (PTS) representation of λC . Therefore, there is no syntactic distinction between terms, types or kinds. This design brings economy for type checking, since one set of rules can cover all syntactic levels. By convention, we use metavariables τ and σ for an expression on the type-level position and e for one on the term level.

Type of Types Traditionally in λC , there are two distinct sorts \star and \square representing the type of *types* and *sorts* respectively, and an axiom $\star : \square$ specifying the relation between the two sorts. In λ_\star , we further merge types and kinds together by including only a single sort \star and an impredicative axiom $\star : \star$.

Explicit type conversion We introduce two new primitives cast^\uparrow and cast_\downarrow (pronounced as ‘cast up’ and ‘cast down’) to replace the implicit conversion rule of λC with a *one-step* explicit type conversion. They represent two directions of type conversion: cast_\downarrow stands for the β -reduction of types, while cast^\uparrow is the inverse.

Though cast primitives make the syntax verbose when type conversion is heavily used, the implementation of type checking is simplified: the typing rules of λ_\star become syntax-directed without λC ’s implicit conversion rule. **BRUNO: I think the following point should be made in the source language section, not here.** Considering the core language is compiler-oriented, end-users will not directly use them. Some type conversions can be generated through the translation of the source language (§6).

BRUNO: Product should be dependent product.

4.2 Operational Semantics

Figure 2 shows the *call-by-name* operational semantics, defined by one-step reduction. Two base cases include S_BETA for β -reduction and S_CASTDOWNUP for cast canceling. Two inductive cases, S_APP and S_CASTDOWN, define reduction in the head position of an application, and in the cast_\downarrow inner expression respectively. The reduction rules are *weak* in the sense that they are not

$e \longrightarrow e'$	One-step reduction
$(\lambda x : \tau. e_1) e_2 \longrightarrow e_1[x \mapsto e_2]$	S_BETA
$\text{cast}_\downarrow(\text{cast}^\uparrow[\tau] e) \longrightarrow e$	S_CASTDOWNUP
$\frac{e_1 \longrightarrow e'_1}{e_1 e_2 \longrightarrow e'_1 e_2}$	S_APP
$\frac{e \longrightarrow e'}{\text{cast}_\downarrow e \longrightarrow \text{cast}_\downarrow e'}$	S_CASTDOWN

Figure 2. Operational semantics of λ_\star

allowed to reduce inside a λ -term or cast^\uparrow -term which is viewed as a value (see Figure 1).

To evaluate the value of a term-level expression, we apply the one-step reduction multiple times. The number of evaluation steps is not restricted, which is possible to be infinite. The multi-step reduction can be defined as follows:

Definition 4.1 (Multi-step reduction). *The relation \longrightarrow is the transitive and reflexive closure of the one-step reduction \longrightarrow .*

For a consecutive sequence of reductions with n steps, we use the notation \longrightarrow_n to denote the relation between the initial and final expressions:

Definition 4.2 (n -step reduction). *The n -step reduction is denoted by $e_0 \longrightarrow_n e_n$, if there exists a sequence of one-step reductions $e_0 \longrightarrow e_1 \longrightarrow e_2 \longrightarrow \dots \longrightarrow e_n$, where n is a positive integer and e_i ($i = 0, 1, \dots, n$) are valid expressions.*

4.3 Typing

Figure 3 gives the *syntax-directed* typing rules of λ_\star , including rules of context well-formedness $\vdash \Gamma$ and expression typing $\Gamma \vdash e : \tau$. Note that there is only a single set of rules for expression typing, because there is no distinction of different syntactic levels.

Most typing rules are quite standard. We write $\vdash \Gamma$ if a context Γ is well-formed. Note that there is only a single sort \star , we use $\Gamma \vdash \tau : \star$ to check if τ is a well-formed type. Rule T_AX is the ‘type-in-type’ axiom. Rule T_VAR checks the type of variable x from the valid context. Rules T_APP and T_LAM check the validity of application and abstraction. Rules T_PI check the type well-formedness of the dependent function.

The Cast Rules We focus on rules T_CASTUP and T_CASTDOWN that define the semantics of cast operators and replace the conversion rule of λC (see Section 2.1). The relation between the original and converted type is defined by one-step reduction (see Figure 2). Given a judgement $\Gamma \vdash e : \tau_2$ and relation $\tau_1 \longrightarrow \tau_2 \longrightarrow \tau_3$, then $\text{cast}^\uparrow[\tau_1] e$ expands the type of e from τ_2 to τ_1 , while $\text{cast}_\downarrow e$ reduces the type of e from τ_2 to τ_3 . For example, assume $\Gamma \vdash e_1 : \text{Int}$ and $\Gamma \vdash e_2 : (\lambda x : \star.x) \text{Int}$. Note that the following reduction holds:

$$(\lambda x : \star.x) \text{Int} \longrightarrow \text{Int}$$

Thus, we can obtain the following derivations:

$$\frac{\Gamma \vdash e_1 : \text{Int} \quad \Gamma \vdash (\lambda x : \star.x) \text{Int} : \star \quad (\lambda x : \star.x) \text{Int} \longrightarrow \text{Int}}{\Gamma \vdash (\text{cast}^\uparrow[(\lambda x : \star.x) \text{Int}] e_1) : (\lambda x : \star.x) \text{Int}}$$

$$\frac{\Gamma \vdash e_2 : (\lambda x : \star. x) \text{Int} \quad \Gamma \vdash \text{Int} : \star \quad (\lambda x : \star. x) \text{Int} \longrightarrow \text{Int}}{\Gamma \vdash (\text{cast}_\downarrow e_2) : \text{Int}}$$

Importantly, in λ_\star term-level and type-level computation are treated differently. Term-level computation is dealt in the usual way, by using multi-step reduction until a value is finally obtained. Type-level computation, on the other hand, is controlled by the program: each step of the computation is induced by a cast. If a type-level program requires n steps of computation to reach normal form, then it will require n casts to compute a (type-level) value.

Cast Operators in n steps For the brevity purpose, we use syntactic sugar to denote n consecutive cast operators. Suppose we have $e : \tau$ and sequences of reduction $\tau_1 \longrightarrow \tau_2 \longrightarrow \dots \longrightarrow \tau_n \longrightarrow \tau$ and $\tau \longrightarrow \sigma_1 \longrightarrow \sigma_2 \longrightarrow \dots \longrightarrow \sigma_n$. The n -step cast operators can be defined:

$$\begin{aligned} \text{cast}_\uparrow^n[\tau_1, \dots, \tau_n]e &\triangleq \text{cast}_\uparrow^\uparrow[\tau_1](\text{cast}_\uparrow^\uparrow[\tau_2](\dots(\text{cast}_\uparrow^\uparrow[\tau_n]e)\dots)) \\ \text{cast}_\downarrow^n e &\triangleq \underbrace{\text{cast}_\downarrow(\text{cast}_\downarrow(\dots(\text{cast}_\downarrow e)\dots))}_n \end{aligned}$$

By rules T_CASTUP and T_CASTDOWN, we have the following typing results:

$$\begin{array}{ll} \text{cast}_\uparrow^n[\tau_1, \dots, \tau_n]e & : \tau_1 \\ \text{cast}_\downarrow^n e & : \sigma_n \end{array}$$

Benefits of Type in Type In the context of λC , if a term x has the type τ_1 , and τ_2 is a type, i.e. $x : \tau_1 : \star$ and $\tau_2 : \square$, we call the type $\Pi x : \tau_1. \tau_2$ a *dependent product*. λ_\star follows λC to use the same Π -notation to represent dependent function types.

However, a higher-kinded polymorphic function type such as $\Pi x : \square. x \rightarrow x$ is not allowed in λC , because \square is the highest sort, and it cannot be typed. In contrast Π -notation in λ_\star is more expressive and does not have such limitation because of the axiom $\star : \star$. In the surface language, we interchangeably use the arrow form $(x : \tau_1) \rightarrow \tau_2$ of the product for clarity. By convention, we also use the syntactic sugar $\tau_1 \longrightarrow \tau_2$ to represent the product if x does not occur free in τ_2 .

Syntactic Equality Finally, the definition of type equality in λ_\star differs from λC . Without λC 's conversion rule, the type of a term cannot be converted freely against β -equality, unless using cast operators. Thus, types of expressions are equal only if they are syntactically equal, i.e. satisfy the α -equality.

4.4 Meta-theory

We now discuss the meta-theory of λ_\star . We focus on two properties: the decidability of type checking and the type-safety of the language. First, we show that type checking λ_\star is decidable without requiring strong normalization. Second, the language is type safe, proven by standard subject reduction and progress lemmas.

Decidability of Type Checking For the decidability, we need to show there exists a type checking algorithm, which never loops forever and returns a unique type for a well-formed expression e . This is done by induction on the length of e and ranging over typing rules. Most expression typing rules, which have only typing judgements in premises, are already decidable by induction hypothesis. Thus, it is straightforward to follow the syntax-directed judgement to derive a unique type checking result.

The critical case is for rules T_CASTUP and T_CASTDOWN. Both rules contain a premise that needs to judge if two types τ_1 and τ_2 follows the one-step reduction, i.e. if $\tau_1 \longrightarrow \tau_2$ holds. We need to show such τ_2 is *unique* with respect to the one-step reduction, or equivalently, reducing τ_1 by one step will get only a sole result τ_2 . Otherwise, assume $e : \tau_1$ and there exists τ'_2 such that $\tau_1 \longrightarrow \tau'_2$

$\boxed{\vdash \Gamma}$ Well-formed context

$$\frac{}{\vdash \emptyset} \text{ENV_EMPTY} \quad \frac{\vdash \Gamma \quad \Gamma \vdash \tau : \star}{\vdash \Gamma, x : \tau} \text{ENV_VAR}$$

$\boxed{\Gamma \vdash e : \tau}$ Expression typing

$$\begin{aligned} &\frac{\vdash \Gamma}{\Gamma \vdash \star : \star} \text{T_AX} \\ &\frac{\vdash \Gamma \quad x : \tau \in \Gamma}{\Gamma \vdash x : \tau} \text{T_VAR} \\ &\frac{\Gamma \vdash e_1 : (\Pi x : \tau_2. \tau_1) \quad \Gamma \vdash e_2 : \tau_2}{\Gamma \vdash e_1 e_2 : \tau_1[x \mapsto e_2]} \text{T_APP} \\ &\frac{\Gamma, x : \tau_1 \vdash e : \tau_2 \quad \Gamma \vdash (\Pi x : \tau_1. \tau_2) : \star}{\Gamma \vdash (\lambda x : \tau_1. e) : (\Pi x : \tau_1. \tau_2)} \text{T_LAM} \\ &\frac{\Gamma \vdash \tau_1 : \star \quad \Gamma, x : \tau_1 \vdash \tau_2 : \star}{\Gamma \vdash (\Pi x : \tau_1. \tau_2) : \star} \text{T_PI} \\ &\frac{\Gamma \vdash e : \tau_2 \quad \Gamma \vdash \tau_1 : \star \quad \tau_1 \longrightarrow \tau_2}{\Gamma \vdash (\text{cast}_\uparrow^\uparrow[\tau_1] e) : \tau_1} \text{T_CASTUP} \\ &\frac{\Gamma \vdash e : \tau_1 \quad \Gamma \vdash \tau_2 : \star \quad \tau_1 \longrightarrow \tau_2}{\Gamma \vdash (\text{cast}_\downarrow e) : \tau_2} \text{T_CASTDOWN} \end{aligned}$$

Figure 3. Typing rules of λ_\star

and $\tau_1 \longrightarrow \tau'_2$. Then the type of $\text{cast}_\downarrow e$ can be either τ_2 or τ'_2 by rule T_CASTDOWN, which is not decidable. The property is proven by the following lemma:

Lemma 4.3 (Decidability of One-step Reduction). *The one-step reduction \longrightarrow is called decidable if given e there is a unique e' such that $e \longrightarrow e'$ or no such e' .*

Proof. By induction on the structure of e . \square

With this result, we show a decidable algorithm to check whether one-step relation $\tau_1 \longrightarrow \tau_2$ holds. An intuitive algorithm is to reduce the type τ_1 by one step to obtain τ'_1 (which is unique by Lemma 4.3), and compare if τ'_1 and τ_2 are syntactically equal. Thus, checking if $\tau_1 \longrightarrow \tau_2$ is decidable and rules T_CASTUP and T_CASTDOWN are therefore decidable. We can conclude the decidability of type checking:

Theorem 4.4 (Decidability of Type Checking λ_\star). *There is an algorithm which given Γ, e computes the unique τ such that $\Gamma \vdash e : \tau$ or reports there is no such τ .*

Proof. By induction on the structure of e . \square

Note that when proving the decidability of type checking, we do not rely on the on strong normalization. Intuitively, because explicit type conversion rules use one-step reduction, which already has a decidable checking algorithm according to Lemma 4.3. We do not need to further require the normalization of terms. This is different from the proof for λC which requires the language to be strongly normalizing [14]. Because λC 's conversion rule needs to examine

the β -equivalence of terms, which is decidable only if every term has a normal form.

Type-safety Proof of the type-safety (or type soundness) of λ_* is fairly standard by subject reduction (or preservation) and progress lemmas. The subject reduction proof relies on the substitution lemma. We give the proof sketch of related lemmas as follows:

Lemma 4.5 (Substitution). *If $\Gamma_1, x : \sigma, \Gamma_2 \vdash e_1 : \tau$ and $\Gamma_1 \vdash e_2 : \sigma$, then $\Gamma_1, \Gamma_2[x \mapsto e_2] \vdash e_1[x \mapsto e_2] : \tau[x \mapsto e_2]$.*

Proof. By induction on the derivation of $\Gamma_1, x : \sigma, \Gamma_2 \vdash e_1 : \tau$. \square

Theorem 4.6 (Subject Reduction). *If $\Gamma \vdash e : \sigma$ and $e \rightarrow e'$ then $\Gamma \vdash e' : \sigma$.*

Proof. (Sketch) We prove the case for one-step reduction, i.e. $e \rightarrow e'$. The lemma can follow by induction on the number of one-step reductions of $e \rightarrow e'$. The proof is by induction with respect to the definition of one-step reduction \rightarrow . \square

Theorem 4.7 (Progress). *If $\emptyset \vdash e : \sigma$ then either e is a value v or there exists e' such that $e \rightarrow e'$.*

Proof. By induction on the derivation of $\emptyset \vdash e : \sigma$. \square

Simplified form of cast_\uparrow^n Because of the decidability of one-step reduction, we can simplify the form of n -step cast_\uparrow operator. From Lemma 4.3, we immediately have the following corollary for n -step reduction:

Lemma 4.8 (Decidability of n -step Reduction). *The n -step reduction \rightarrow_n is called decidable if given e there is a unique e' such that $e \rightarrow_n e'$ or no such e' .*

Proof. Immediate from Lemma 4.3, by induction on the number of reduction steps. \square

Suppose $e : \tau$ and there exists a reduction sequence $\tau_1 \rightarrow \tau_2 \rightarrow \dots \rightarrow \tau_n \rightarrow \tau$, the following typing of cast_\uparrow^n holds: $\text{cast}_\uparrow^n[\tau_1, \dots, \tau_n]e : \tau_1$.

Note that $\tau_1 \rightarrow_n \tau$ are decidable by Lemma 4.8. Thus, the intermediate types in $\tau_1 \rightarrow_n \tau$, i.e. τ_2, \dots, τ_n , can be uniquely determined. Thus, we can leave them out in the cast_\uparrow^n operator with the following form: $\text{cast}_\uparrow^n[\tau_1]e : \tau_1$.

5. Dependent Types with General Recursion

In this section we present λ_*^μ : an extension of λ_* with a general recursion construct. The general recursion is polymorphic and has a uniform representation on both term level and type level. The same construct works both as a term-level fixpoint and recursive type. The addition of general recursion does not break decidability of type-checking nor type-safety.

5.1 Syntax and Semantics

Figure 4 shows the changes of extending λ_* to λ_*^μ with general recursion. The changes are subtle, since we add only one more primitive, reduction rule and typing rule for general recursion. Nevertheless general recursion allows a large number of programs that can be expressed in programming languages such as Haskell to be expressed in λ_*^μ as well.

For syntax, we add the polymorphic recursion operator μ to represent general recursion on both term and type level in the same form $\mu x : \tau. e$. For operational semantics, we add the rule S_MU to define the unrolling operation of a recursion, which results in $e[x \mapsto \mu x : \tau. e]$. For typing, we add the rule T_MU for checking

e, τ, σ	$::=$	Expressions
	\dots	
	$\mu x : \tau. e$	Polymorphic recursion
$e \rightarrow e'$		One-step reduction
$\mu x : \tau. e \rightarrow e[x \mapsto \mu x : \tau. e]$		S_MU
$\Gamma \vdash e : \tau$		Expression typing
$\Gamma, x : \tau \vdash e : \tau$	$\Gamma \vdash \tau : \star$	T_MU
$\Gamma \vdash (\mu x : \tau. e) : \tau$		

Figure 4. Syntax and semantics changes for general recursion

the validity of a polymorphic recursive term. The rule ensures that the polymorphic recursion $\mu x : \tau. e$ should have the same type τ as the binder x and also the inner expression e .

5.2 Recursion as Term and Type

Term-level Recursion In λ_*^μ , μ -operator works as a *fixpoint* on the term level. By rule S_MU, evaluating a term $\mu x : \tau. e$ will substitute all x in e with the whole μ -term itself, resulting in the unrolling $e[x \mapsto \mu x : \tau. e]$. The μ -term is equivalent to a recursive function that should be allowed to unroll without restriction. Therefore, the definition of values is not changed in λ_*^μ and μ -term is not treated as a value. This is different from conventional term-level fixpoint that is usually treated as values [7].

Recall the factorial function example (§??), which can be defined as a μ -term as follows:

$\text{fact} \triangleq \mu f : \text{Int} \rightarrow \text{Int}.$

$\lambda x : \text{Int}. \text{if } x \equiv 0 \text{ then } 1 \text{ else } x \times (f(x - 1))$

By rule T_MU, the type of fact is $\text{Int} \rightarrow \text{Int}$. Thus, we can apply fact to an integer, say 1, and will get an integer as the result. By rules S_MU and S_APP, we can evaluate the term $\text{fact } 1$ as follows:

$\text{fact } 1$

$\rightarrow (\lambda x : \text{Int}. \text{if } x \equiv 0 \text{ then } 1 \text{ else } x \times (\text{fact}(x - 1))) 1$

$\rightarrow \text{if } 1 \equiv 0 \text{ then } 1 \text{ else } 1 \times (\text{fact}(1 - 1))$

$\rightarrow 1 \times (\text{fact}(1 - 1))$

$\rightarrow 1 \times (\text{fact } 0) \rightarrow \dots \rightarrow 1 \times 1 \rightarrow 1.$

Note that we never check if a μ -term can terminate or not, which is an undecidable halting problem for general recursive terms. The factorial function example above can stop, while there exist some terms that will loop forever. However, term-level non-termination is only a runtime concern and does not block the type checker. Later we will see type checking λ_*^μ is still decidable. **BRUNO: Show an example of the execution, maybe for $\text{fact}(2)$ LINUS: Done.**

Type-level Recursion On the type level, $\mu x : \tau. e$ works as a *iso-recursive* type, a kind of recursive type that is not equal but only isomorphic to its unrolling. Normally, we need to add two more primitives fold and unfold for the iso-recursive type to map back and forth between the original and unrolled form:

$$\mu x : \tau. \sigma \xrightleftharpoons[\text{fold } [\mu x : \tau. \sigma]]{\text{unfold}} \sigma[x \mapsto \mu x : \tau. \sigma]$$

where the operators satisfy the following typing rules:

$$\frac{\Gamma \vdash e_2 : \sigma[x \mapsto \mu x : \tau. \sigma] \quad \Gamma \vdash (\mu x : \tau. \sigma) : \star}{\Gamma \vdash \text{fold } [\mu x : \tau. \sigma] e_2 : (\mu x : \tau. \sigma)}$$

$$\frac{\Gamma \vdash e_1 : (\mu x : \tau.\sigma) \quad \Gamma \vdash \sigma[x \mapsto \mu x : \tau.\sigma] : \star}{\Gamma \vdash \text{unfold } e_1 : (\sigma[x \mapsto \mu x : \tau.\sigma])}$$

BRUNO: Show the rules for fold and unfold here to help you make the argument. **LINUS:** Figure added to show the relationship.

However, in λ_\star^μ we do not need to introduce fold and unfold operators, because with the rule S_MU, cast^\uparrow and cast_\downarrow generalize fold and unfold and have the same functionality. Assume there exist expressions e_1 and e_2 such that

$$\begin{aligned} e_1 & : \mu x : \tau.e \\ e_2 & : e[x \mapsto \mu x : \tau.e] \end{aligned}$$

Note that e_1 and e_2 have distinct types but the type of e_2 is the unrolling of e_1 's type, which follows the one-step reduction relation by rule S_MU:

$$\mu x : \tau.e \longrightarrow e[x \mapsto \mu x : \tau.e]$$

By applying rules T_CASTUP and T_CASTDOWN, we can obtain the following typing results:

$$\begin{aligned} \text{cast}_\downarrow e_1 & : e[x \mapsto \mu x : \tau.e] \\ \text{cast}^\uparrow [\mu x : \tau.e] e_2 & : \mu x : \tau.e \end{aligned}$$

Thus, cast^\uparrow and cast_\downarrow witness the isomorphism between the original recursive type and its unrolling, which behave the same as fold and unfold operations in iso-recursive types:

$$\mu x : \tau.e \xrightleftharpoons[\text{cast}^\uparrow [\mu x : \tau.e]]{\text{cast}_\downarrow} e[x \mapsto \mu x : \tau.e]$$

5.3 Decidability of Type Checking and Type Safety

BRUNO: I think it is worth restating the lemmas here to aid the discussion. No need to show all lemmas, just the main ones: decidability of type-checking and type-safety. **LINUS:** Revised. We give the proof of the decidability of type checking λ_\star in §4.4 without requiring normalization. The cast rules are critical for decidability, which rely on checking if one type can be reduced to another in one step. When we introduce general recursion into the language, if we can make sure the newly added and original typing rules are still decidable, the decidability of type checking will still follow in λ_\star^μ .

The rule T_MU for checking the well-formedness of polymorphic recursion is decidable because its premises only include decidable typing judgements. However, the rule S_MU changes one-step reduction, which may affect the decidability of cast rules. If the uniqueness of changed reduction rules still holds, by following the same proof tactic of λ_\star , we can show that cast rules are still decidable in λ_\star^μ . Note that given a recursive term $\mu x : \tau.e$, by rule S_MU, there always exists a unique term $e' = e[x \mapsto \mu x : \tau.e]$ such that $\mu x : \tau.e \longrightarrow e'$. Thus, the uniqueness of one-step reduction still holds, i.e. Lemma 4.3 holds in λ_\star^μ . So the decidability of type checking, namely Lemma 4.4 holds in λ_\star^μ :

Theorem 5.1 (Decidability of Type Checking λ_\star^μ). *There is an algorithm which given Γ, e computes the unique τ such that $\Gamma \vdash e : \tau$ or reports there is no such τ .*

Moreover, it is straightforward to show the type-safety of λ_\star^μ by considering rules T_MU and T_MU during the induction of proof. Thus, Lemma 4.6 and 4.7 also hold in λ_\star^μ :

Theorem 5.2 (Subject Reduction). *If $\Gamma \vdash e : \sigma$ and $e \longrightarrow e'$ then $\Gamma \vdash e' : \sigma$.*

Theorem 5.3 (Progress). *If $\emptyset \vdash e : \sigma$ then either e is a value v or there exists e' such that $e \longrightarrow e'$.*

6. Formalization of the Surface language

In this section, we formally present the surface language **Fun**, built on top of λ_\star^μ with features that are convenient for functional pro-

gramming: user-defined datatypes, and pattern matching. Thanks to the expressiveness of λ_\star^μ , all these features can be elaborated into the core language without extending the built-in language constructs of λ_\star^μ . In what follows, we first give the syntax of the surface language, followed by the extended typing rules, then we show the formal translation rules that translates a surface language expression to an expression in λ_\star^μ . Finally we prove the type-safety of the translation.

6.1 Extended Syntax

The full syntax of **Fun** is defined in Figure 5. Compared with λ_\star^μ , **Fun** has a new syntax category: a program, consisting of a list of datatype declarations, followed by an expression. For the purpose of presentation, we sometimes adopt the following syntactic convention:

$$\overline{\tau}^n \rightarrow \tau_r \equiv \tau_1 \rightarrow \cdots \rightarrow \tau_n \rightarrow \tau_r$$

An *algebraic datatype* D is introduced as a top-level **data** declaration with its *data constructors*. The type of a data constructor K takes the form:

$$K : (\overline{u} : \overline{\kappa}^n) \rightarrow (\overline{x} : \overline{T}) \rightarrow D \overline{u}^n$$

The first n quantified type variables \overline{u} appear in the same order in the result type $D \overline{u}$. Note that the use of the dependent product in the data constructor arguments (i.e., $(\overline{x} : \overline{T})$) makes it possible to let the types of some data constructor arguments depend on other data constructor arguments, whereas in Haskell, this is not possible, because the function type (\rightarrow) can be seen as an independent product type. The **case** expression is conventional, used to break up values built with data constructors. The patterns of a case expression are flat (i.e., no nested patterns), and bind variables.

A noticeable difference from λ_\star^μ is that, in **Fun**, we do not allow cast operations to appear in the programs. Indeed, the cast operations are intended to be generated by the compiler for λ_\star^μ , not by the programmers. An unfortunate consequence is that, this, for now, makes the surface language less expressive (e.g., no explicit type-level computation) than λ_\star^μ . However, as we will point in Section 8, we will add new language constructs in **Fun** that help programmers perform type-level computation in a way like type classes or type families do for Haskell programmers.

For the sake of programming, **Fun** employs some syntactic sugar. A non-dependent function type can be written as $T_1 \rightarrow T_2$. A dependent function type $\Pi x : T_1. T_2$ is abbreviated as $(x : T_1) \rightarrow T_2$ for easy reading. We also introduce a Haskell-like record syntax, which is desugared to datatypes with accompanying selector functions.

6.2 Extended Typing Rules

BRUNO: For typing and translation show only one figure (Figure 8), since the typing figure is just a subset. We can use gray to highlight the parts which belong to the translation. **JEREMY:** adjusted!

Figure 6 defines the type system of the surface language (ignore the gray parts for the moment). Several new typing judgments are introduced in the type system. The use of different subscripts of the judgments is to be distinct from the one used in λ_\star^μ . Most rules are standard for systems based on λC , including the rules for the well-formedness of contexts (TENV_EMPTY, TENV_VAR), inferring the types of variables (TR_VAR), and dependent application (TR_APP). Two judgments $\Sigma \vdash_{pg} pgm : T$ and $\Sigma \vdash_d decl : \Sigma'$ are of the essence to the type checking of the surface language. The former type checks a whole program, and the latter type checks datatype declarations.

Rule TR_PGM_PGM type checks a whole program. It first type-checks the declarations, which in return gives a new typing environment. Combined with the original environment, it then contin-

pgm	$::=$ $\overline{decl}; E$	Program Declarations
$decl$	$::=$ $\mathbf{data} D \overline{u} : \overline{\kappa} = \overline{K x : \overline{T}}$	Declarations Datatype declaration
u, v	$::=$ x K	Atoms Variable Data constructor
p	$::=$ $K \overline{x : \overline{T}}$	Patterns Pattern
E, T, S, κ	$::=$ u \star $E_1 E_2$ $\lambda x : T. E$ $\Pi x : T. S$ $\mu x : T. E$ $\mathbf{case} E_1 \mathbf{of} p \Rightarrow \overline{E_2}$	Expressions Atom Type of types Application Abstraction Product General recursion Case analysis
Σ	$::=$ \emptyset $\Sigma, u : T$	Contexts Empty Atom binding
$T_1 \rightarrow T_2$	$\triangleq \Pi x : T_1. T_2$	Syntactic Sugar Function type Dependent function type Record
$(x : T_1) \rightarrow T_2$	$\triangleq \Pi x : T_1. T_2$	
$\mathbf{data} R \overline{u} : \overline{\kappa}^n = K \{ \overline{N : \overline{T}} \}$	$\triangleq \mathbf{data} R \overline{u} : \overline{\kappa}^n = K \overline{x : \overline{T}};$ $\mathbf{let} N_i : (\overline{u} : \overline{\kappa}^n) \rightarrow R \overline{u}^n \rightarrow T_i =$ $\lambda \overline{u} : \overline{\kappa}^n. \lambda y : R \overline{u}^n. \mathbf{case} y \mathbf{of} K \overline{x : \overline{T}} \Rightarrow x_i \mathbf{in}$	

Figure 5. Syntax of the surface language

ues to type check the expression and return the result type. Rule TRPGM.DATA is used to type check datatype declarations. It first ensures the well-formedness of the type of the type constructor application (of kind \star). Note that since our system adopts $\star : \star$, this means we can express kind polymorphism for datatypes. Finally it make sure the types of data constructors are valid.

Rules TR_CASE and TRPAT_ALT handle the type checking of case expressions. The conclusion of TS_CASE binds the right types to the scrutinee E_1 and alternatives $p \Rightarrow \overline{E_2}$. The first premise of TPAT_ALT binds the actual type constructor arguments to \overline{u} . The second premise checks whether the types of the right-hand sides of each alternative, instantiated to the actual type constructor arguments \overline{u} , are equal. Finally the third premise checks the well-formedness of the types of data constructor arguments. **BRUNO: Mention that we do not support refinement, as in GADTs?** **JEREMY: done!**

As can be seen, currently **Fun** does not support refinement on the final result of each data constructor, as in GADTs. However, our encoding method does support some form of GADTs, as is discussed in Section ??.

6.3 Translation Overview

We use a type-directed translation. The basic translation rules have the form:

$$\Sigma \vdash_{\mathbb{S}} E : T \rightsquigarrow e$$

It states that λ_{\star}^{μ} expression e is the translation of the surface expression E of type T . The gray parts in Figure 6 defines the translation rules. **BRUNO: Any particular reasons for this?** **JEREMY: deleted**

Among others, Rules TRDECL_DATA, TRPAT_ALT and TR_CASE are of the essence to the translation. Rule TR_CASE translates case expressions into applications by first translating the scrutinee expression, casting it down to the right type. It is then applied to the result type of the body expression and a list of translated λ_{\star}^{μ} expressions of its alternatives. Rule TRPAT_ALT tells how to translate each alternative. Basically it translates an alternative into a lambda abstraction, where each bound variable in the pattern corresponds to a bound variable in the lambda abstraction in the same order. The body in the alternative is recursively translated and treated as the body of the lambda abstraction. Note that due to the rigidity of the translation, pattern matching must be exhaustive, and the order of patterns matters (the same order as in the datatype declaration).

Rule TRDECL_DATA does the most heavy work and deserves further explanation. First of all, it results in an incomplete expression (as can be seen by the incomplete let expressions), The result expression is supposed to be prepended to the translation of the last expression to form a complete λ_{\star}^{μ} expression, as specified by Rule TRPGM.PGM. Furthermore, each type constructor is translated to a recursive type, of which the body is a type-level lambda abstraction. What is interesting is that each recursive mention of the datatype in the data constructor parameters is replaced with the recursive variable X . Note that for the moment, the result type variable b is restricted to have kind \star . This could pose difficulties when translating GADTs as we will discussion in the future work. Each data constructor is translated to a lambda abstraction. Notice how we use cast^{\dagger} in the lambda body to get the right type.

The rest of the translation rules hold few surprises.

$\boxed{\vdash_{\text{wf}} \Sigma \rightsquigarrow \Gamma}$ Context well-formedness

$$\frac{\overline{\vdash_{\text{wf}} \emptyset \rightsquigarrow \emptyset} \quad \text{TRENV_EMPTY}}{\frac{\vdash_{\text{wf}} \Sigma \rightsquigarrow \Gamma \quad \Sigma \vdash_s T : \star \rightsquigarrow \tau}{\vdash_{\text{wf}} \Sigma, x : T \rightsquigarrow \Gamma, x : \tau}} \text{TRENV_VAR}$$

$\boxed{\Sigma \vdash_{\text{pg}} \text{pgm} : T \rightsquigarrow e}$ Program translation

$$\frac{\overline{\Sigma_0 \vdash_d \text{decl} : \Sigma' \rightsquigarrow e_1} \quad \Sigma = \Sigma_0, \overline{\Sigma'} \quad \Sigma \vdash_s E : T \rightsquigarrow e}{\Sigma_0 \vdash_{\text{pg}} (\overline{\text{decl}}; E) : T \rightsquigarrow \overline{e_1} \uplus e} \text{TRPGM_PGM}$$

$\boxed{\Sigma \vdash_d \text{decl} : \Sigma' \rightsquigarrow e}$ Datatype translation

$$\frac{\Sigma \vdash_s (\overline{u} : \kappa^n) \rightarrow \star : \star \rightsquigarrow (\overline{u} : \rho^n) \rightarrow \star \quad \overline{\Sigma, D : (\overline{u} : \kappa^n) \rightarrow \star, \overline{u} : \kappa^n \vdash_s (\overline{x} : \overline{T}) \rightarrow D \overline{u}^n : \star \rightsquigarrow (\overline{x} : \overline{\tau}) \rightarrow D \overline{u}^n}}{\Sigma \vdash_d (\text{data } D \overline{u} : \kappa^n = \overline{K \overline{x} : \overline{T}}) : (D : (\overline{u} : \kappa^n) \rightarrow \star, \overline{K : (\overline{u} : \kappa^n) \rightarrow (\overline{x} : \overline{T}) \rightarrow D \overline{u}^n}) \rightsquigarrow e} \text{TRDECL_DATA}$$

$$e \triangleq \text{let } D : (\overline{u} : \rho^n) \rightarrow \star = \mu X : (\overline{u} : \rho^n) \rightarrow \star. \lambda \overline{u} : \rho^n. (b : \star) \rightarrow ((\overline{x} : \tau [D \mapsto X]) \rightarrow b) \rightarrow b \text{ in} \\ \text{let } K_i : (\overline{u} : \rho^n) \rightarrow (\overline{x} : \tau) \rightarrow D \overline{u}^n = \lambda \overline{u} : \rho^n. \lambda \overline{x} : \tau. \text{cast}_{\uparrow}^{n+1} [D \overline{u}^n] (\lambda b : \star. \lambda c : (\overline{x} : \tau) \rightarrow b. c_i \overline{x}) \text{ in}$$

$\boxed{\Sigma \vdash_p p \Rightarrow E : T \rightarrow S \rightsquigarrow e}$ Pattern translation

$$\frac{K : (\overline{u} : \kappa^n) \rightarrow (\overline{x} : \overline{T}) \rightarrow D \overline{u}^n \in \Sigma \quad \Sigma, \overline{x : T[\overline{u} \mapsto \overline{v}]} \vdash_s E : S \rightsquigarrow e \quad \Sigma \vdash_s T[\overline{u} \mapsto \overline{v}] : \star \rightsquigarrow \tau}{\Sigma \vdash_p K x : T[\overline{u} \mapsto \overline{v}] \Rightarrow E : D \overline{u}^n \rightarrow S \rightsquigarrow \lambda \overline{x} : \overline{\tau}. e} \text{TRPAT_ALT}$$

$\boxed{\Sigma \vdash_s E : T \rightsquigarrow e}$ Expression translation

$$\frac{\vdash_{\text{wf}} \Sigma}{\Sigma \vdash_s \star : \star \rightsquigarrow \star} \text{TR_AX}$$

$$\frac{\vdash_{\text{wf}} \Sigma \rightsquigarrow \Gamma \quad x : T \in \Sigma}{\Sigma \vdash_s x : T \rightsquigarrow x} \text{TR_VAR}$$

$$\frac{\Sigma \vdash_s E_1 : (\Pi x : T_2. T_1) \rightsquigarrow e_1 \quad \Sigma \vdash_s E_2 : T_2 \rightsquigarrow e_2}{\Sigma \vdash_s E_1 E_2 : T_1[x \mapsto E_2] \rightsquigarrow e_1 e_2} \text{TR_APP}$$

$$\frac{\Sigma, x : T_1 \vdash_s E : T_2 \rightsquigarrow e \quad \Sigma \vdash_s (\Pi x : T_1. T_2) : \star \rightsquigarrow \Pi x : \tau_1. \tau_2}{\Sigma \vdash_s (\lambda x : T_1. E) : (\Pi x : T_1. T_2) \rightsquigarrow \lambda x : \tau_1. e} \text{TR_LAM}$$

$$\frac{\Sigma \vdash_s T_1 : \star \rightsquigarrow \tau_1 \quad \Sigma, x : T_1 \vdash_s T_2 : \star \rightsquigarrow \tau_2}{\Sigma \vdash_s (\Pi x : T_1. T_2) : \star \rightsquigarrow \Pi x : \tau_1. \tau_2} \text{TR_PI}$$

$$\frac{\Sigma, x : T \vdash_s E : T \rightsquigarrow e \quad \Sigma \vdash_s T : \star \rightsquigarrow \tau}{\Sigma \vdash_s (\mu x : T. E) : T \rightsquigarrow \mu x : \tau. e} \text{TR_MU}$$

$$\frac{\Sigma \vdash_s E_1 : T \rightsquigarrow e_1 \quad \overline{\Sigma \vdash_p p \Rightarrow E_2 : T \rightarrow S \rightsquigarrow e_2} \quad \Sigma \vdash_s S : \star \rightsquigarrow \sigma}{\Sigma \vdash_s \text{case } E_1 \text{ of } \overline{p \Rightarrow E_2} : S \rightsquigarrow (\text{cast}_{\downarrow}^{n+1} e_1) \sigma \overline{e_2}} \text{TR_CASE}$$

Figure 6. Type directed translation rules of the surface language

6.4 Type-safety of Translation

JEREMY: put Linus’s theorem here

7. Related Work

There is a lot of work on bringing full-spectrum dependent types to the practical programming world.

Unification of Terms, Types, and Kinds BRUNO: This subsection should start by talking about pure type systems and then move on to the work of Henk and others. JEREMY: done!

Pure Type Systems [4] show how a whole family of type systems can be implemented using just a single syntactic form. This line of our work is largely inspired by Henk [13], where they are the first to use the so-called *lambda cube* as a typed intermediate language, unifying all three levels. Since the implicit conversion of the lambda cube is not syntax-directed, they come up with an approach to strategically distribute the conversion rule over the other typing rules. In retrospect, Henk is quite conservative in terms of type-level computation. Actually it is not even a dependently typed language, as they clearly state that they don’t allow types to depend on terms. As for recursion, even though it has a full lambda calculus at the type level, recursion is disallowed.

Another recent work on dependently typed language based on the same syntactic category is *Zombie* [7, 24], where terms, types and the single kind \star all reside in the same level. The language is based on a call-by-value variant of lambda calculus. One beautiful thing about *Zombie* is that it is composed of two fragments: a logical fragment and a programmatic fragment, so that it supports both partial and total programming. Even though *Zombie* has one syntactic category, it is still fairly complicated, as it tries to be both consistent as a logic and pragmatic as a programming language.

$\Pi\Sigma$ [2] is another recently proposed dependently typed core language that resembles λ_\star^μ , as there is no syntactic difference between terms and types. BRUNO:

General Recursion and Managed Type-level Computation As discussed in §5 BRUNO: where? words like *before and after* are a bad smell. In academic writing we use references. Just add the reference to the section. JEREMY: fixed, bringing general recursion blindly into the dependently typed world causes more trouble than convenience. There are many dependently typed languages that allow general recursion. *Zombie* approaches general recursion by separating a consistent sub-language, in which all expressions are known to terminate, from a programmatic language that supports general recursion. What is interesting about *Zombie* is that those two seemingly conflicting worlds can interact with each other nicely, without compromising the consistency property. The key idea of this is to distinguish between these two fragments by using a *consistency classifier* θ . When θ is L , it means the logical part, and P the program part. Like λ_\star^μ , *Zombie* uses *roll* and *unroll* for iso-recursive types. To ensure normalization (in order for decidable type checking), it forbids the use of *unroll* in P , where the potential non-termination could arise.

F^\star [27] also supports writing general-purpose programs with effects (e.g., state, exceptions, non-terminating, etc.) while maintaining a consistent core language. Unlike λ_\star^μ , it has several sub-languages – for terms, proofs and so on. The interesting part of F^\star lies in its kind system, which tracks the sub-languages and controls the interactions between them. The idea is to restrict the use of recursion in specifications and proofs while allowing arbitrary recursion in the program. They use \star to denote program terms that may be effectful and divergent, and P for proofs that identify pure and total functions. In this way, they are able to ensure that fragments in a program used for building proof terms are never mixed with those that are potentially divergent. One difference from λ_\star^μ is that,

types in F^\star can only contain values but no non-value expressions, leading to its less expressiveness than λ_\star^μ .

$\Pi\Sigma$ has a general mechanism for recursion. Like λ_\star^μ , it uses one recursion mechanism for the definition of both types and programs. The key idea relies on lifted types and boxes: definitions are not unfolded inside boxes. The way they achieve decidable type checking is to use boxing to stop the infinite unfolding of the recursive call, at the cost of additional annotations stating where to lift, box and force. One concern of $\Pi\Sigma$ is that its metatheory is not yet formally developed.

Type in Type We are not the first to embrace $\star : \star$ in the system. It has been long known that systems with $\star : \star$ (usually called system λ_\star) is inconsistent as a logic, in the sense that all types are inhabited. In this system, we can encode a variant of Russell’s paradox, known as Girard’s paradox [9].

The core language of the Glasgow Haskell Compiler, System FC [26] has been extended with type promotion [30] and kind equality [29]. The latter one introduces a limited form of dependent types into the system¹, which mixes up types and kinds. This causes no trouble for FC, since all kinds are already inhabited without the above extensions. $\Pi\Sigma$ has an impredicative universe of types with $\text{Type} : \text{Type}$ due to the support of general recursion. The surface language of *Zombie* also has the rule $\Gamma \vdash \text{Type} : \text{Type}$ [25].

The $\star : \star$ axiom makes it convenient to support kind polymorphism, among other language features. One concern is that it often causes type checking to be undecidable, if not dealt with carefully, as it allows to express divergent terms. However, as we explained in §4, this is not the case for λ_\star^μ . Type checking in λ_\star^μ is decidable – all type-level computation is driven by finite cast operations, thus no potentially infinite reductions can happen in reality.

Encoding of Datatypes One thing λ_\star^μ differs from other functional programming languages is that all the high-level features in the surface language like datatypes, pattern matching and so on can be easily encoded into the core language. There is much work on encoding datatypes into various high-level languages. The classic Church encoding of datatypes into System F is detailed in the work of Böhm and Berarducci [6]. The Church encoding excels in implementing iterative or fold-like functions over algebraic datatypes, but is awkward in expressing general recursion, usually in a complex and insufficient way. An alternative encoding of datatypes is the so-called *Scott* encoding. However Scott encoding is not typable in System F, as it needs recursive types to represent recursive datatypes. λ_\star^μ has all it needs to represent polymorphic and recursive datatypes. The explicit cast rules also make it possible to encode GADTs, as can be seen in the last examples in Section 3. Currently we are investigating how the encoding of GADTs interact with the other language constructs. We leave this as future work.

Another line of related work is the *inductive defined types* in the Calculus of Inductive Constructions (CIC) [19], which is the underlying formal language of Coq. In CIC, inductive defined types can be represented by closed types in λC , so are the primitive recursive functionals over elements of the type. The limitation of their work is that functions over inductive defined types are definable only by primitive recursion, not general recursion. Conor McBride’s work on *Observational Type Theory* (OTT) [1] shows the encoding of datatypes via \mathcal{W} -types.

BRUNO: work by Conor McBride on encoding datatypes? That needs to be discussed here. Also the calculus of inductive constructions should deserve some mention. JEREMY: added CIC, which

¹ Richard A. Eisenberg is going to implement kind equality [29] into GHC. The implementation is proposed at <https://phabricator.haskell.org/D808> and related paper is at <http://www.cis.upenn.edu/~eir/papers/2015/equalities/equalities-extended.pdf>.

work of Conor McBride? Towards Observational Type Theory is the paper i searched

8. Discussion

More Type-level Computation

Eliminating Cast Operators Explicit type cast operators cast^\uparrow and cast_\downarrow are generalized from fold and unfold of iso-recursive types. By convention, cast^\uparrow follows fold as a value, so it cannot be further reduced during the evaluation. This may cause semantics and performance issues.

To show the semantics issue, suppose there is a non-terminating term loop . It should loop forever when evaluated. But if it is wrapped by cast^\uparrow like $\text{cast}^\uparrow [\tau] \text{loop}$, the evaluation stops immediately. So the dynamic semantics of a term-level expression wrapped by cast^\uparrow might differ from the original. The performance issue may occur during code generation. Since cast^\uparrow will remain after evaluation, too many cast^\uparrow constructs can increase the size of the program and cause runtime overhead.

In fact, cast operators can be safely eliminated after type checking, because they do not actually transform a term other than changing its type. There are two choices of when to remove cast operators, during evaluation or code generation. The following alternative reduction rules can eliminate cast during evaluation:

$$\frac{}{\text{cast}^\uparrow [\tau] e \longrightarrow e} \text{S_CASTUPE}$$

$$\frac{}{\text{cast}_\downarrow e \longrightarrow e} \text{S_CASTDOWNE}$$

Moreover, cast^\uparrow is no more treated as a value. With such modification, we can obtain $\text{cast}^\uparrow [\tau] \text{loop} \longrightarrow \text{loop}$. So cast^\uparrow will not interfere with the original dynamic semantics of terms. But noting that $\text{cast}^\uparrow [\tau] e$ or $\text{cast}_\downarrow e$ has different types from e , types of terms are not preserved during evaluation. This breaks the subject reduction theorem, and consequently type-safety.

Thus, we stick to the semantics of iso-recursive types for cast operators which has type preservation. And we prefer to eliminate all cast operators during type erasure to address the potential performance issue of code generation.

Encoding of GADTs Our translation rules also open opportunity for encoding GADTs. In our experiment, we have several running examples of encoding GADTs. Below we show a GADT-encoded representation of well-scoped lambda terms using de Bruijn notation.

In this notation, a variable is represented as a number – its de Bruijn index, where the number k stands for the variable bound by the k 's enclosing λ . Using the GADT syntax, below is the definition of lambda terms:

```
data Fin : Nat → * =
  fzero : (n : Nat) → Fin (S n)
  | fsucc : (n : Nat) → Fin n → Fin (S n);
data Term : Nat → * =
  Var : (n : Nat) → Fin n → Term n
  | Lam : (n : Nat) → Term (S n) → Term n
  | App : (n : Nat) → Term n → Term n → Term n;
```

The datatype $\text{Fin } n$ is used to restrict the the de Bruijn index, so that it lies between 0 to $n - 1$. The type of a closed term is simply $\text{Term } Z$, for instance, a lambda term $\lambda x. \lambda y. x$ is represented as (for the presentation, we use decimal numbers instead of Peano numbers):

```
Lam 0 (Lam 1 (Var 2 (fsucc 1 (fzero 0))))
```

If we accidentally write the wrong index, the program would fail to pass type checking.

We do not have space to present a complete encoding, but instead show the encoding of Fin :

```
let Fin : Nat → * =
  μ X : Nat → *. λ a : Nat.
    (B : Nat → *) →
    ((n : Nat) → B (S n)) →
    ((n : Nat) → X n → B (S n)) →
    B a
```

The key issue in encoding GADTs lies in type of variable B . In ordinary datatype encoding, B is fixed to have type $*$, while in GADTs, its type is the same as the variable X (possibly higher-kinded). Currently, we have to manually interpret the type according to the particular use of some GADTs. We are investigating if there exists a general way to do that.

9. Conclusion

Conclusion and related work.

Acknowledgments

Thanks to Blah. This work is supported by Blah.

References

- [1] T. Altenkirch and C. McBride. Towards Observational Type Theory. *Manuscript available online*, 2006. URL <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.117.4451&rep=rep1&type=pdf>.
- [2] T. Altenkirch, N. A. Danielsson, A. Löb, and N. Oury. $\Pi\Sigma$: Dependent types without the sugar. In *Functional and Logic Programming*, pages 40–55. Springer, 2010.
- [3] H. Barendregt. Lambda calculi with types. In *Handbook of Logic in Computer Science*, volume 2, pages 117–309. Oxford University Press, 1992.
- [4] H. Barendregt and H. Barendregt. Introduction to generalized type systems. *Journal of functional programming*, 1:124, 1991. . URL <http://dare.uhn.kun.nl/handle/2066/17240>.
- [5] R. Bird and L. Meertens. Nested datatypes. *Mathematics of program construction*, 1422:52–67, 1998. . URL <http://www.springerlink.com/index/126u6246254u5263.pdf>.
- [6] C. Böhm and A. Berarducci. Automatic synthesis of typed Λ -programs on term algebras, 1985. ISSN 03043975.
- [7] C. Casinghino, V. Sjöberg, and S. Weirich. Combining proofs and programs in a dependently typed language. *ACM SIGPLAN Notices*, 49(1):33–45, 2014.
- [8] T. Coquand. *Une théorie des constructions*. PhD thesis, 1985.
- [9] T. Coquand. An analysis of girard's paradox. 1986.
- [10] T. Coquand and G. Huet. The calculus of constructions. *Inf. Comput.*, 76(2-3):95–120, Feb. 1988. ISSN 0890-5401. . URL [http://dx.doi.org/10.1016/0890-5401\(88\)90005-3](http://dx.doi.org/10.1016/0890-5401(88)90005-3).
- [11] L. Fegaras and T. Sheard. Revisiting catamorphisms over datatypes with embedded functions (or, programs from outer space). In *Proceedings of the 23rd ACM SIGPLAN-SIGACT symposium on Principles of programming languages - POPL '96*, pages 284–294, 1996. ISBN 0897917693. . URL <http://dl.acm.org/citation.cfm?id=237721.237792>.
- [12] J.-Y. Girard. Interprétation fonctionnelle et élimination des coupures de l'arithmétique d'ordre supérieur. 1972.
- [13] S. P. Jones and E. Meijer. Henk: a typed intermediate language. 1997.
- [14] L. Jutting. Typing in pure type systems. *Information and Computation*, 105(1):30–41, 1993.

- [15] Y. Mandelbaum and a. Stump. GADTs for the OCaml masses. *Workshop on ML*, 2009.
- [16] E. Meijer, M. Fokkinga, and R. Paterson. Functional programming with bananas, lenses, envelopes and barbed wire. *Functional Programming Languages and Computer Architecture*, pages 124–144, 1991. . URL http://link.springer.com/chapter/10.1007/3540543961_7.
- [17] A. Middelkoop, A. Dijkstra, and S. D. Swierstra. A lean specification for gads: system f with first-class equality proofs. *Higher-Order and Symbolic Computation*, 23(2):145–166, 2010.
- [18] T. A. Mogensen. Efficient self-interpretation in lambda calculus, 1992. ISSN 0956-7968.
- [19] F. Pfenning and C. Paulin-Mohring. Inductively defined types in the Calculus of Constructions. *Mathematical Foundations of Programming Semantics*, pages 209–228, 1990. . URL papers://cfff96cb1-96b7-4b11-a3e3-f4947c1d45b9/Paper/p6382.
- [20] B. C. Pierce. *Types and programming languages*. MIT press, 2002.
- [21] A. Rastogi, A. Delignat-lavaud, C. Keller, P.-y. Strub, and K. Bhargavan. Semantic Purity and Effects Reunited in F. pages 1–19, 2015.
- [22] J.-W. Roorda and J. Jeuring. Pure type systems for functional programming. 2007.
- [23] P. G. Severi and F.-J. J. de Vries. Pure type systems with corecursion on streams: from finite to infinitary normalisation. In *ACM SIGPLAN Notices*, volume 47, pages 141–152. ACM, 2012.
- [24] V. Sjöberg. *A Dependently Typed Language with Nontermination*. PhD thesis, University of Pennsylvania, 2015.
- [25] V. Sjöberg and S. Weirich. Programming up to congruence. In *Proceedings of the 42Nd Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, POPL ’15*, pages 369–382, New York, NY, USA, 2015. ACM. .
- [26] M. Sulzmann, M. M. Chakravarty, S. P. Jones, and K. Donnelly. System f with type equality coercions. In *Proceedings of the 2007 ACM SIGPLAN international workshop on Types in languages design and implementation*, pages 53–66. ACM, 2007.
- [27] N. Swamy, J. Chen, C. Fournet, P.-Y. Strub, K. Bhargavan, and J. Yang. Secure distributed programming with value-dependent types, 2011. ISSN 03621340.
- [28] J. C. Vanderwaart, D. Dreyer, L. Petersen, K. Crary, R. Harper, and P. Cheng. *Typed compilation of recursive datatypes*, volume 38. ACM, 2003.
- [29] S. Weirich, J. Hsu, and R. A. Eisenberg. Towards dependently typed haskell: System fc with kind equality. In *Proceedings of the 18th ACM SIGPLAN International Conference on Functional Programming, ICFP*, volume 13. Citeseer, 2013.
- [30] B. A. Yorgey, S. Weirich, J. Cretin, S. Peyton Jones, D. Vytiniotis, and J. P. Magalhães. Giving haskell a promotion. In *Proceedings of the 8th ACM SIGPLAN workshop on Types in language design and implementation*, pages 53–66. ACM, 2012.

A. Full Specification of Core Language

A.1 Syntax

e, τ, σ, ρ	$::=$	x \star $e_1 e_2$ $\lambda x : \tau. e$ $\Pi x : \tau. \sigma$ $\text{cast}^\uparrow[\tau] e$ $\text{cast}_\downarrow e$ $\mu x : \tau. e$	Expressions Variable Type of types Application Abstraction Product Cast up by expansion Cast down by reduction General recursion
Γ	$::=$	\emptyset $\Gamma, x : \tau$	Contexts Empty Variable binding
v	$::=$	\star $\lambda x : \tau. e$ $\Pi x : \tau_1. \tau_2$ $\text{cast}^\uparrow[\tau] e$	Values Type of types Abstraction Product Cast up by expansion
Syntactic Sugar			
$T_1 \rightarrow T_2$	\triangleq	$\Pi x : T_1. T_2$	$x \notin \text{FV}(T_2)$
$(x : T_1) \rightarrow T_2$	\triangleq	$\Pi x : T_1. T_2$	
$\text{let } x : \tau = e_2 \text{ in } e_1$	\triangleq	$e_1[x \mapsto e_2]$	$\Gamma \vdash e_2 : \tau$
$\text{cast}_\uparrow^n[\tau_1] e$	\triangleq	$\text{cast}_\uparrow[\tau_1](\text{cast}_\uparrow[\tau_2](\dots(\text{cast}_\uparrow[\tau_n] e) \dots))$	
$\text{cast}_\downarrow^n e$	\triangleq	$\underbrace{\text{cast}_\downarrow(\text{cast}_\downarrow(\dots(\text{cast}_\downarrow e) \dots))}_n$	

A.2 Operational Semantics

$e \longrightarrow e'$	One-step reduction
$\overline{(\lambda x : \tau. e_1) e_2 \longrightarrow e_1[x \mapsto e_2]}$	S_BETA
$\overline{\text{cast}_\downarrow(\text{cast}_\uparrow[\tau] e) \longrightarrow e}$	S_CASTDOWNUP
$\frac{e_1 \longrightarrow e'_1}{e_1 e_2 \longrightarrow e'_1 e_2}$	S_APP
$\frac{e \longrightarrow e'}{\text{cast}_\downarrow e \longrightarrow \text{cast}_\downarrow e'}$	S_CASTDOWN
$\overline{\mu x : \tau. e \longrightarrow e[x \mapsto \mu x : \tau. e]}$	S_MU

A.3 Typing

$\boxed{\vdash \Gamma}$	Well-formed context
$\overline{\vdash \emptyset}$	ENV_EMPTY
$\frac{\vdash \Gamma \quad \Gamma \vdash \tau : \star}{\vdash \Gamma, x : \tau}$	ENV_VAR
$\boxed{\Gamma \vdash e : \tau}$	Expression typing
$\frac{\vdash \Gamma}{\Gamma \vdash \star : \star}$	T_AX
$\frac{\vdash \Gamma \quad x : \tau \in \Gamma}{\Gamma \vdash x : \tau}$	T_VAR

$$\begin{array}{c}
\frac{\Gamma \vdash e_1 : (\Pi x : \tau_2. \tau_1) \quad \Gamma \vdash e_2 : \tau_2}{\Gamma \vdash e_1 e_2 : \tau_1[x \mapsto e_2]} \text{ T_APP} \\
\\
\frac{\Gamma, x : \tau_1 \vdash e : \tau_2 \quad \Gamma \vdash (\Pi x : \tau_1. \tau_2) : \star}{\Gamma \vdash (\lambda x : \tau_1. e) : (\Pi x : \tau_1. \tau_2)} \text{ T_LAM} \\
\\
\frac{\Gamma \vdash \tau_1 : \star \quad \Gamma, x : \tau_1 \vdash \tau_2 : \star}{\Gamma \vdash (\Pi x : \tau_1. \tau_2) : \star} \text{ T_PI} \\
\\
\frac{\Gamma \vdash e : \tau_2 \quad \Gamma \vdash \tau_1 : \star \quad \tau_1 \longrightarrow \tau_2}{\Gamma \vdash (\text{cast}^\uparrow[\tau_1] e) : \tau_1} \text{ T_CASTUP} \\
\\
\frac{\Gamma \vdash e : \tau_1 \quad \Gamma \vdash \tau_2 : \star \quad \tau_1 \longrightarrow \tau_2}{\Gamma \vdash (\text{cast}_\downarrow e) : \tau_2} \text{ T_CASTDOWN} \\
\\
\frac{\Gamma, x : \tau \vdash e : \tau \quad \Gamma \vdash \tau : \star}{\Gamma \vdash (\mu x : \tau. e) : \tau} \text{ T_MU}
\end{array}$$

B. Proofs about Core Language

B.1 Properties

Lemma B.1 (Free Variable). *If $\Gamma \vdash e : \tau$, then $\text{FV}(e) \subseteq \text{dom}(\Gamma)$ and $\text{FV}(\tau) \subseteq \text{dom}(\Gamma)$.*

Proof. By induction on the derivation of $\Gamma \vdash e : \tau$. We only treat cases T_MU, T_CASTUP and T_CASTDOWN (since proofs of other cases are the same as λC [3]):

Case T_MU: From premises of $\Gamma \vdash (\mu x : \tau. e_1) : \tau$, by induction hypothesis, we have $\text{FV}(e_1) \subseteq \text{dom}(\Gamma) \cup \{x\}$ and $\text{FV}(\tau) \subseteq \text{dom}(\Gamma)$. Thus the result follows by $\text{FV}(\mu x : \tau. e_1) = \text{FV}(e_1) \setminus \{x\} \subseteq \text{dom}(\Gamma)$ and $\text{FV}(\tau) \subseteq \text{dom}(\Gamma)$.

Case T_CASTUP: Since $\text{FV}(\text{cast}^\uparrow[\tau] e_1) = \text{FV}(e_1)$, the result follows directly by the induction hypothesis.

Case T_CASTDOWN: Since $\text{FV}(\text{cast}_\downarrow e_1) = \text{FV}(e_1)$, the result follows directly by the induction hypothesis. \square

Lemma B.2 (Thinning). *Let Γ and Γ' be legal contexts such that $\Gamma \subseteq \Gamma'$. If $\Gamma \vdash e : \tau$ then $\Gamma' \vdash e : \tau$.*

Proof. By trivial induction on the derivation of $\Gamma \vdash e : \tau$. \square

Lemma B.3 (Substitution). *If $\Gamma_1, x : \sigma, \Gamma_2 \vdash e_1 : \tau$ and $\Gamma_1 \vdash e_2 : \sigma$, then $\Gamma_1, \Gamma_2[x \mapsto e_2] \vdash e_1[x \mapsto e_2] : \tau[x \mapsto e_2]$.*

Proof. By induction on the derivation of $\Gamma_1, x : \sigma, \Gamma_2 \vdash e_1 : \tau$. Let $e^* \equiv e[x \mapsto e_2]$. Then the result can be written as $\Gamma_1, \Gamma_2^* \vdash e_1^* : \tau^*$. We only treat cases T_MU, T_CASTUP and T_CASTDOWN. Consider the last step of derivation of the following cases:

Case T_MU:
$$\frac{\Gamma_1, x : \sigma, \Gamma_2 \vdash e_1 : \tau \quad \Gamma_1, x : \sigma, \Gamma_2 \vdash \tau : \star}{\Gamma_1, x : \sigma, \Gamma_2 \vdash (\mu y : \tau. e_1) : \tau}$$
 By induction hypothesis, we have $\Gamma_1, \Gamma_2^* \vdash e_1^* : \tau^*$ and $\Gamma_1, \Gamma_2^* \vdash \tau^* : \star$. Then by the deviation rule, $\Gamma_1, \Gamma_2^* \vdash (\mu y : \tau^*. e_1^*) : \tau^*$. Thus we have $\Gamma_1, \Gamma_2^* \vdash (\mu y : \tau. e_1)^* : \tau^*$ which is just the result.

Case T_CASTUP:
$$\frac{\Gamma_1, x : \sigma, \Gamma_2 \vdash e_1 : \tau_2 \quad \Gamma_1, x : \sigma, \Gamma_2 \vdash \tau_1 : \star \quad \tau_1 \longrightarrow \tau_2}{\Gamma_1, x : \sigma, \Gamma_2 \vdash (\text{cast}^\uparrow[\tau_1] e_1) : \tau_1}$$
 By induction hypothesis, we have $\Gamma_1, \Gamma_2^* \vdash e_1^* : \tau_2^*$, $\Gamma_1, \Gamma_2^* \vdash \tau_1^* : \star$ and $\tau_1 \longrightarrow \tau_2$. By the definition of substitution, we can obtain $\tau_1^* \longrightarrow \tau_2^*$ by $\tau_1 \longrightarrow \tau_2$. Then by the deviation rule, $\Gamma_1, \Gamma_2^* \vdash (\text{cast}^\uparrow[\tau_1^*] e_1^*) : \tau_1^*$. Thus we have $\Gamma_1, \Gamma_2^* \vdash (\text{cast}^\uparrow[\tau_1] e_1)^* : \tau_1^*$ which is just the result.

Case T_CASTDOWN:
$$\frac{\Gamma_1, x : \sigma, \Gamma_2 \vdash e_1 : \tau_1 \quad \Gamma_1, x : \sigma, \Gamma_2 \vdash \tau_2 : \star \quad \tau_1 \longrightarrow \tau_2}{\Gamma_1, x : \sigma, \Gamma_2 \vdash (\text{cast}_\downarrow e_1) : \tau_2}$$

By induction hypothesis, we have $\Gamma_1, \Gamma_2^* \vdash e_1^* : \tau_1^*$, $\Gamma_1, \Gamma_2^* \vdash \tau_2^* : \star$ and $\tau_1 \longrightarrow \tau_2$ thus $\tau_1^* \longrightarrow \tau_2^*$. Then by the deviation rule, $\Gamma_1, \Gamma_2^* \vdash (\text{cast}_\downarrow e_1^*) : \tau_2^*$. Thus we have $\Gamma_1, \Gamma_2^* \vdash (\text{cast}_\downarrow e_1)^* : \tau_2^*$ which is just the result. \square

Lemma B.4 (Generation).

- (1) If $\Gamma \vdash x : \sigma$, then there exist an expression τ such that $\tau \equiv \sigma$, $\Gamma \vdash \tau : \star$ and $x : \tau \in \Gamma$.
- (2) If $\Gamma \vdash e_1 e_2 : \sigma$, then there exist expressions τ_1 and τ_2 such that $\Gamma \vdash e_1 : (\Pi x : \tau_1. \tau_2)$, $\Gamma \vdash e_2 : \tau_1$ and $\sigma \equiv \tau_2[x \mapsto e_2]$.
- (3) If $\Gamma \vdash (\lambda x : \tau_1. e) : \sigma$, then there exist an expression τ_2 such that $\sigma \equiv \Pi x : \tau_1. \tau_2$ where $\Gamma \vdash (\Pi x : \tau_1. \tau_2) : \star$ and $\Gamma, x : \tau_1 \vdash e : \tau_2$.
- (4) If $\Gamma \vdash (\Pi x : \tau_1. \tau_2) : \sigma$, then $\sigma \equiv \star$, $\Gamma \vdash \tau_1 : \star$ and $\Gamma, x : \tau_1 \vdash \tau_2 : \star$.
- (5) If $\Gamma \vdash (\mu x : \tau. e) : \sigma$, then $\Gamma \vdash \tau : \star$, $\sigma \equiv \tau$ and $\Gamma, x : \tau \vdash e : \tau$.
- (6) If $\Gamma \vdash (\text{cast}^\uparrow[\tau_1] e) : \sigma$, then there exist an expression τ_2 such that $\Gamma \vdash e : \tau_2$, $\Gamma \vdash \tau_1 : \star$, $\tau_1 \longrightarrow \tau_2$ and $\sigma \equiv \tau_1$.
- (7) If $\Gamma \vdash (\text{cast}_\downarrow e) : \sigma$, then there exist expressions τ_1, τ_2 such that $\Gamma \vdash e : \tau_1$, $\Gamma \vdash \tau_2 : \star$, $\tau_1 \longrightarrow \tau_2$ and $\sigma \equiv \tau_2$.

Proof. Consider a derivation of $\Gamma \vdash e : \sigma$ for one of cases in the lemma. We can follow the process of derivation until expression e is introduced the first time. The last step of derivation can be done by

- rule T_VAR for case 1;
- rule T_APP for case 2;
- rule T_LAM for case 3;
- rule T_PI for case 4;
- rule T_MU for case 5;
- rule T_CASTUP for case 6;
- rule T_CASTDOWN for case 7.

In each case, assume the conclusion of the rule is $\Gamma' \vdash e : \tau'$ where $\Gamma' \subseteq \Gamma$ and $\tau' \equiv \sigma$. Then by inspection of used derivation rules and Lemma B.2, it can be shown that the statement of the lemma holds and is the only possible case. \square

Lemma B.5 (Correctness of Types). *If $\Gamma \vdash e : \tau$ then $\tau \equiv \star$ or $\Gamma \vdash \tau : \star$.*

Proof. Trivial induction on the derivation of $\Gamma \vdash e : \tau$ using Lemma B.4. \square

B.2 Decidability of Type Checking

Lemma B.6 (Decidability of One-step Reduction). *The one-step reduction \longrightarrow is called decidable if given e there is a unique e' such that $e \longrightarrow e'$ or no such e' .*

Proof. By induction on the structure of e :

Case $e = v$: e has one of the following forms: (1) \star , (2) $\lambda x : \tau. e$, (3) $\Pi x : \tau_1. \tau_2$, (4) $\text{cast}^\uparrow[\tau] e$, which cannot match any rules of \longrightarrow . Thus there is no e' such that $e \longrightarrow e'$.

Case $e = (\lambda x : \tau. e_1) e_2$: There is a unique $e' = e_1[x \mapsto e_2]$ by rule S_BETA.

Case $e = \text{cast}_\downarrow(\text{cast}^\uparrow[\tau] e)$: There is a unique $e' = e$ by rule S_CASTDOWNUP.

Case $e = \mu x : \tau. e$: There is a unique $e' = e[x \mapsto \mu x : \tau. e]$ by rule S.MU.

Case $e = e_1 e_2$ and e_1 is not a λ -term: If $e_1 = v$, there is no e'_1 such that $e_1 \rightarrow e'_1$. Since e_1 is not a λ -term, there is no rule to reduce e . Thus there is no e' such that $e \rightarrow e'$. Otherwise, there exists some e'_1 such that $e_1 \rightarrow e'_1$. By the induction hypothesis, e'_1 is unique reduction of e_1 . Thus by rule S.APP, $e' = e'_1 e_2$ is the unique reduction for e .

Case $e = \text{cast}_\downarrow e_1$ and e_1 is not a cast^\uparrow -term: If $e_1 = v$, there is no e'_1 such that $e_1 \rightarrow e'_1$. Since e_1 is not a cast^\uparrow -term, there is no rule to reduce e . Thus there is no e' such that $e \rightarrow e'$. Otherwise, there exists some e'_1 such that $e_1 \rightarrow e'_1$. By the induction hypothesis, e'_1 is unique reduction of e_1 . Thus by rule S.CASTDOWN, $e' = \text{cast}_\downarrow e'_1$ is the unique reduction for e .

□

Lemma B.7 (Decidability of n -step Reduction). *The n -step reduction \rightarrow_n is called decidable if given e there is a unique e' such that $e \rightarrow_n e'$ or no such e' .*

Proof. Immediate from Lemma B.6, by induction on the number of reduction steps. □

Theorem B.8 (Decidability of Type Checking). *There is an algorithm which given Γ, e computes the unique τ such that $\Gamma \vdash e : \tau$ or reports there is no such τ .*

Proof. By induction on the structure of e :

Case $e = \star$: Trivial by applying T.AX and $\tau \equiv \star$.

Case $e = x$: Trivial by rule T.VAR and τ is the unique type of x if $x : \tau \in \Gamma$.

Case $e = e_1 e_2$: By rule T.APP and introduction hypothesis, there exist unique expressions τ_1 and τ_2 such that $\Gamma \vdash e_1 : (\Pi x : \tau_1. \tau_2)$, $\Gamma \vdash e_2 : \tau_1$. Thus, from Lemma B.4, $\tau_2[x \mapsto e_2]$ is the unique type of e .

Case $e = \lambda x : \tau_1. e_1$: By rule T.LAM and introduction hypothesis, there exist unique expressions τ_2 such that $\Gamma \vdash (\Pi x : \tau_1. \tau_2) : \star$ and $\Gamma, x : \tau_1 \vdash e_1 : \tau_2$. Thus, from Lemma B.4, $\Pi x : \tau_1. \tau_2$ is the unique type of e .

Case $e = \Pi x : \tau_1. \tau_2$: By rule T.PI and introduction hypothesis, we have $\Gamma \vdash \tau_1 : \star$ and $\Gamma, x : \tau_1 \vdash \tau_2 : \star$. Thus, from Lemma B.4, \star is the unique type of e .

Case $e = \mu x : \tau. e_1$: By rule T.MU and introduction hypothesis, we have $\Gamma \vdash \tau : \star$ and $\Gamma, x : \tau \vdash e_1 : \tau$. Thus, from Lemma B.4, τ is the unique type of e .

Case $e = \text{cast}^\uparrow[\tau_1] e_1$: From the premises of rule T.CASTUP, by induction hypothesis, we can derive the type of e_1 as τ_2 , and check whether τ_1 is legal, i.e. its sorts is \star . If τ_1 is legal, by Lemma B.6, there is at most one τ'_1 such that $\tau_1 \rightarrow \tau'_1$. If such τ'_1 does not exist, then we report type checking fails. Otherwise, we examine if τ'_1 is syntactically equal to τ_2 , i.e. $\tau'_1 \equiv \tau_2$. If the equality holds, we obtain the unique type of e which is τ_1 . Otherwise, we report e fails to type check.

Case $e = \text{cast}_\downarrow e_1$: From the premises of rule T.CASTDOWN, by induction hypothesis, we can derive the type of e_1 as τ_1 . By Lemma B.6, there is at most one τ_2 such that $\tau_1 \rightarrow \tau_2$. If such τ_2 exists and its sorts is \star , we have found the unique type of e is τ_2 . Otherwise, we report e fails to type check.

□

B.3 Soundness

Definition B.9 (Multi-step reduction). *The relation \rightarrow is the transitive and reflexive closure of \rightarrow .*

Theorem B.10 (Subject Reduction). *If $\Gamma \vdash e : \sigma$ and $e \rightarrow e'$ then $\Gamma \vdash e' : \sigma$.*

Proof. We prove the case for one-step reduction, i.e. $e \rightarrow e'$. The lemma can follow by induction on the number of one-step reductions of $e \rightarrow e'$. The proof is by induction with respect to the definition of one-step reduction \rightarrow as follows:

Case $\frac{}{(\lambda x : \tau. e_1) e_2 \rightarrow e_1[x \mapsto e_2]}$ S.BETA:

Suppose $\Gamma \vdash (\lambda x : \tau_1. e_1) e_2 : \sigma$ and $\Gamma \vdash e_1[x \mapsto e_2] : \sigma'$. By Lemma B.4(2), there exist expressions τ'_1 and τ_2 such that

$$\begin{aligned} \Gamma \vdash (\lambda x : \tau_1. e_1) : (\Pi x : \tau'_1. \tau_2) \quad (1) \\ \Gamma \vdash e_2 : \tau'_1 \\ \sigma \equiv \tau_2[x \mapsto e_2] \end{aligned}$$

By Lemma B.4(3), the judgement (1) implies that there exists an expression τ'_2 such that

$$\begin{aligned} \Pi x : \tau'_1. \tau_2 \equiv \Pi x : \tau_1. \tau'_2 \quad (2) \\ \Gamma, x : \tau_1 \vdash e_1 : \tau'_2 \end{aligned}$$

Hence, by (2) we have $\tau_1 \equiv \tau'_1$ and $\tau_2 \equiv \tau'_2$. Then we can obtain $\Gamma, x : \tau_1 \vdash e_1 : \tau_2$ and $\Gamma \vdash e_2 : \tau_1$. By Lemma B.3, we have $\Gamma \vdash e_1[x \mapsto e_2] : \tau_2[x \mapsto e_2]$. Therefore, we conclude with $\sigma' \equiv \tau_2[x \mapsto e_2] \equiv \sigma$.

Case $\frac{e_1 \rightarrow e'_1}{e_1 e_2 \rightarrow e'_1 e_2}$ S.APP:

Suppose $\Gamma \vdash e_1 e_2 : \sigma$ and $\Gamma \vdash e'_1 e_2 : \sigma'$. By Lemma B.4(2), there exist expressions τ_1 and τ_2 such that

$$\begin{aligned} \Gamma \vdash e_1 : (\Pi x : \tau_1. \tau_2) \\ \Gamma \vdash e_2 : \tau_1 \\ \sigma \equiv \tau_2[x \mapsto e_2] \end{aligned}$$

By induction hypothesis, we have $\Gamma \vdash e'_1 : (\Pi x : \tau_1. \tau_2)$. By rule T.APP, we obtain $\Gamma \vdash e'_1 e_2 : \tau_2[x \mapsto e_2]$. Therefore, $\sigma' \equiv \tau_2[x \mapsto e_2] \equiv \sigma$.

Case $\frac{e \rightarrow e'}{\text{cast}_\downarrow e \rightarrow \text{cast}_\downarrow e'}$ S.CASTDOWN:

Suppose $\Gamma \vdash \text{cast}_\downarrow e : \sigma$ and $\Gamma \vdash \text{cast}_\downarrow e' : \sigma'$. By Lemma B.4(7), there exist expressions τ_1, τ_2 such that

$$\begin{aligned} \Gamma \vdash e : \tau_1 \quad \Gamma \vdash \tau_2 : \star \\ \tau_1 \rightarrow \tau_2 \quad \sigma \equiv \tau_2 \end{aligned}$$

By induction hypothesis, we have $\Gamma \vdash e' : \tau_1$. By rule T.CASTDOWN, we obtain $\Gamma \vdash \text{cast}_\downarrow e' : \tau_2$. Therefore, $\sigma' \equiv \tau_2 \equiv \sigma$.

Case $\frac{}{\text{cast}_\downarrow(\text{cast}^\uparrow[\tau] e) \rightarrow e}$ S.CASTDOWNUP:

Suppose $\Gamma \vdash \text{cast}_\downarrow(\text{cast}^\uparrow[\tau] e) : \sigma$ and $\Gamma \vdash e : \sigma'$. By Lemma B.4(7), there exist expressions τ'_1, τ_2 such that

$$\Gamma \vdash (\text{cast}^\uparrow[\tau] e) : \tau'_1 \quad (3)$$

$$\tau'_1 \rightarrow \tau_2 \quad (4)$$

$$\sigma \equiv \tau_2 \quad (5)$$

By Lemma B.4(6), the judgement (3) implies that there exists an expression τ'_2 such that

$$\Gamma \vdash e : \tau'_2 \quad (6)$$

$$\tau_1 \longrightarrow \tau'_2 \quad (7)$$

$$\tau'_1 \equiv \tau_1 \quad (8)$$

By (4, 7, 8) and Lemma B.6 we obtain $\tau_2 \equiv \tau'_2$. From (6) we have $\sigma' \equiv \tau'_2$. Therefore, by (5), $\sigma' \equiv \tau'_2 \equiv \tau_2 \equiv \sigma$.

Case $\frac{\mu x : \tau.e \longrightarrow e[x \mapsto \mu x : \tau.e]}{\text{S_MU:}}$

Suppose $\Gamma \vdash (\mu x : \tau.e) : \sigma$ and $\Gamma \vdash e[x \mapsto \mu x : \tau.e] : \sigma'$. By Lemma B.4(5), we have $\sigma \equiv \tau$ and $\Gamma, x : \tau \vdash e : \tau$. Then we obtain $\Gamma \vdash (\mu x : \tau.e) : \tau$. Thus by Lemma B.3, we have $\Gamma \vdash e[x \mapsto \mu x : \tau.e] : \tau[x \mapsto \mu x : \tau.e]$.

Note that $x : \tau$, i.e. the type of x is τ , then $x \notin \text{FV}(\tau)$ holds implicitly. Hence, by the definition of substitution, we obtain $\tau[x \mapsto \mu x : \tau.e] \equiv \tau$. Therefore, $\sigma' \equiv \tau[x \mapsto \mu x : \tau.e] \equiv \tau \equiv \sigma$.

□

Theorem B.11 (Progress). *If $\emptyset \vdash e : \sigma$ then either e is a value v or there exists e' such that $e \longrightarrow e'$.*

Proof. By induction on the derivation of $\emptyset \vdash e : \sigma$ as follows:

Case $e = x$: Impossible, because the context is empty.

Case $e = v$: Trivial, since e is already a value that has one of the following forms: (1) \star , (2) $\lambda x : \tau.e$, (3) $\Pi x : \tau_1.\tau_2$, (4) $\text{cast}^\uparrow[\tau] e$.

Case $e = e_1 e_2$: By Lemma B.4(2), there exist expressions τ_1 and τ_2 such that $\emptyset \vdash e_1 : (\Pi x : \tau_1.\tau_2)$ and $\emptyset \vdash e_2 : \tau_1$. Consider whether e_1 is a value:

- If $e_1 = v$, by Lemma B.4(3), it must be a λ -term such that $e_1 \equiv \lambda x : \tau_1.e'_1$ for some e'_1 satisfying $\emptyset \vdash e'_1 : \tau_2$. Then by rule S_BETA, we have $(\lambda x : \tau_1.e'_1) e_2 \longrightarrow e'_1[x \mapsto e_2]$. Thus, there exists $e' \equiv e'_1[x \mapsto e_2]$ such that $e \longrightarrow e'$.
- Otherwise, by induction hypothesis, there exists e'_1 such that $e_1 \longrightarrow e'_1$. Then by rule S_APP, we have $e_1 e_2 \longrightarrow e'_1 e_2$. Thus, there exists $e' \equiv e'_1 e_2$ such that $e \longrightarrow e'$.

Case $e = \text{cast}_\downarrow e_1$: By Lemma B.4(7), there exist expressions τ_1 and τ_2 such that $\emptyset \vdash e_1 : \tau_1$ and $\tau_1 \longrightarrow \tau_2$. Consider whether e_1 is a value:

- If $e_1 = v$, by Lemma B.4(6), it must be a cast^\uparrow -term such that $e_1 \equiv \text{cast}^\uparrow[\tau_1] e'_1$ for some e'_1 satisfying $\emptyset \vdash e'_1 : \tau_2$. Then by rule S_CASTDOWNUP, we can obtain $\text{cast}_\downarrow(\text{cast}^\uparrow[\tau_1] e'_1) \longrightarrow e'_1$. Thus, there exists $e' \equiv e'_1$ such that $e \longrightarrow e'$.
- Otherwise, by induction hypothesis, there exists e'_1 such that $e_1 \longrightarrow e'_1$. Then by rule S_CASTDOWN, we have $\text{cast}_\downarrow e_1 \longrightarrow \text{cast}_\downarrow e'_1$. Thus, there exists $e' \equiv \text{cast}_\downarrow e'_1$ such that $e \longrightarrow e'$.

Case $e = \mu x : \tau.e_1$: By rule S_MU, there always exists $e' \equiv e_1[x \mapsto \mu x : \tau.e_1]$.

□

C. Full Specification of Surface Language

C.1 Syntax

See Figure 7.

C.2 Expression Typing

See Figure 8.

C.3 Translation to the Core

See Figure 9.

D. Proofs about Surface Language

D.1 Type-safety of Translation

Theorem D.1 (Type-safety of Expression Translation). *If $\Sigma \vdash_5 E : T \rightsquigarrow e$, $\Sigma \vdash_5 T : \star \rightsquigarrow \tau$ and $\vdash_{\text{wf}} \Sigma \rightsquigarrow \Gamma$, then $\Gamma \vdash e : \tau$.*

Proof. By induction on the derivation of $\Sigma \vdash_5 E : T \rightsquigarrow e$. Suppose there is a core language context Γ such that $\vdash_{\text{wf}} \Sigma \rightsquigarrow \Gamma$.

Case TR_AX: Trivial. $e = \tau = \star$ and $\Sigma \vdash_5 \star : \star$ holds by rule T_AX.

Case TR_VAR: Trivial. By rule T_VAR, we have $\vdash_{\text{wf}} \Sigma \rightsquigarrow \Gamma$, then $x : \tau \in \Gamma$ where $\Sigma \vdash_5 T : \star \rightsquigarrow \tau$.

Case TR_APP: Suppose

$$\begin{aligned} \Sigma \vdash_5 E_1 E_2 : T_1[x \mapsto E_2] &\rightsquigarrow e_1 e_2 \\ \Sigma \vdash_5 T_1[x \mapsto E_2] : \star &\rightsquigarrow \tau_1[x \mapsto e_2]. \end{aligned}$$

By induction hypothesis, we have $\Gamma \vdash e_1 : (\Pi x : \tau_2.T_1)$, $\Gamma \vdash e_2 : \tau_2$, where

$$\begin{aligned} \Sigma \vdash_5 E_1 : (\Pi x : T_2.T_1) &\rightsquigarrow e_1 \\ \Sigma \vdash_5 (\Pi x : T_2.T_1) : \star &\rightsquigarrow (\Pi x : \tau_2.T_1) \\ \Sigma \vdash_5 E_2 : T_2 &\rightsquigarrow e_2 \\ \Sigma \vdash_5 T_2 : \star &\rightsquigarrow \tau_2. \end{aligned}$$

Thus by rule T_APP, we have $\Gamma \vdash e_1 e_2 : \tau_1[x \mapsto e_2]$.

Case TR_LAM: Suppose

$$\begin{aligned} \Sigma \vdash_5 (\lambda x : T_1.E) : (\Pi x : T_1.T_2) &\rightsquigarrow \lambda x : \tau_1.e \\ \Sigma \vdash_5 \Pi x : T_1.T_2 : \star &\rightsquigarrow \Pi x : \tau_1.T_2. \end{aligned}$$

By induction hypothesis, we have $\Gamma, x : \tau_1 \vdash e : \tau_2$, $\Gamma \vdash \Pi x : \tau_1.T_2 : \star$ where

$$\begin{aligned} \Sigma, x : T_1 \vdash_5 E : T_2 &\rightsquigarrow e \\ \Sigma \vdash_5 T_1 : \star &\rightsquigarrow \tau_1 & \Sigma \vdash_5 T_2 : \star &\rightsquigarrow \tau_2 \\ \Sigma \vdash_5 (\Pi x : T_1.T_2) : \star &\rightsquigarrow \Pi x : \tau_1.T_2 \end{aligned}$$

Thus by rule T_LAM, we have $\Gamma \vdash (\lambda x : \tau_1.e) : (\Pi x : \tau_1.T_2)$.

Case TR_PI: Suppose

$$\Sigma \vdash_5 (\Pi x : T_1.T_2) : \star \rightsquigarrow \Pi x : \tau_1.T_2.$$

By induction hypothesis, we have $\Gamma \vdash \tau_1 : \star$, $\Gamma, x : \tau_1 \vdash \tau_2 : \star$ where $\Sigma \vdash_5 T_1 : \star \rightsquigarrow \tau_1$, $\Sigma, x : T_1 \vdash_5 T_2 : \star \rightsquigarrow \tau_2$. Thus by rule T_PI we have $\Gamma \vdash (\Pi x : \tau_1.T_2) : \star$.

Case TR_MU: Suppose

$$\begin{aligned} \Sigma \vdash_5 (\mu x : T.E) : T &\rightsquigarrow \mu x : \tau.e \\ \Sigma \vdash_5 T : \star &\rightsquigarrow \tau. \end{aligned}$$

By induction hypothesis, we have

$$\Gamma, x : \tau \vdash e : \tau, \text{ where } \Sigma, x : T \vdash_5 E : T \rightsquigarrow e.$$

Thus by rule T_MU, we have $\Gamma \vdash (\mu x : \tau.e) : \tau$.

Case TR_CASE: Suppose

$$\begin{aligned} \Sigma \vdash_5 \text{case } E_1 \text{ of } \overline{p \Rightarrow E_2} : S &\rightsquigarrow (\text{cast}_\downarrow^{n+1} e_1) \sigma \overline{e_2} \\ \Sigma \vdash_5 S : \star &\rightsquigarrow \sigma. \end{aligned}$$

By induction hypothesis, we have

$$\begin{aligned} \Sigma \vdash_5 E_1 : T_1 &\rightsquigarrow e_1 & \Sigma \vdash_5 T_1 : \star &\rightsquigarrow \tau_1 \\ \Gamma \vdash e_1 : \tau_1 & & \Sigma \vdash_5 p \Rightarrow E_2 : T_1 \rightarrow S &\rightsquigarrow e_2 \end{aligned}$$

pgm	$::=$ $ \quad \overline{decl}; E$	Program Declarations
$decl$	$::=$ $ \quad \mathbf{data} \ D \ \overline{u : \kappa} = \overline{K \ x : T}$	Declarations Datatype declaration
u, v	$::=$ $ \quad x$ $ \quad K$	Atoms Variable Data constructor
p	$::=$ $ \quad K \ \overline{x : T}$	Patterns Pattern
E, T, S, κ	$::=$ $ \quad u$ $ \quad \star$ $ \quad E_1 E_2$ $ \quad \lambda x : T. E$ $ \quad \Pi x : T. S$ $ \quad \mu x : T. E$ $ \quad \mathbf{case} \ E_1 \ \mathbf{of} \ p \Rightarrow \overline{E_2}$	Expressions Atom Type of types Application Abstraction Product General recursion Case analysis
Σ	$::=$ $ \quad \emptyset$ $ \quad \Sigma, u : T$	Contexts Empty Atom binding
$T_1 \rightarrow T_2$	$\triangleq \Pi x : T_1. T_2$	$x \notin \text{FV}(T_2)$ Syntactic Sugar
$(x : T_1) \rightarrow T_2$	$\triangleq \Pi x : T_1. T_2$	Function type
$\mathbf{data} \ R \ \overline{u : \kappa^n} = K \{ \overline{N : T} \}$	$\triangleq \mathbf{data} \ R \ \overline{u : \kappa^n} = K \ \overline{x : T};$ $\mathbf{let} \ N_i : (\overline{u : \kappa^n}) \rightarrow R \ \overline{u^n} \rightarrow T_i =$ $\lambda \overline{u : \kappa^n}. \lambda y : R \ \overline{u^n}. \mathbf{case} \ y \ \mathbf{of} \ K \ \overline{x : T} \Rightarrow x_i \ \mathbf{in}$	$i = 1, 2, \dots, n$ Dependent function type Record

Figure 7. Syntax of the surface language

By rule TRPAT_ALT, we have

$$\begin{aligned}
 p &\equiv K \ x : T[\overline{u \mapsto v}] \\
 T_1 &\equiv D \ \overline{v^n} \\
 \overline{e_2} &\equiv \overline{\lambda x : \tau'. e}
 \end{aligned}$$

where

$$\frac{\Sigma \vdash_{\mathbb{S}} E_2 : S \rightsquigarrow e}{\Sigma \vdash_{\mathbb{S}} T[\overline{u \mapsto v}] : \star \rightsquigarrow \tau[u \mapsto u']} \quad \frac{\Gamma \vdash e : \sigma}{\Sigma \vdash_{\mathbb{S}} v : \star \rightsquigarrow u'}$$

$\tau' \equiv \tau[u \mapsto u']$

By rule TRDECL_DATA, we have $D \equiv \mu X : (\overline{u : \rho^n}) \rightarrow$

$\star. \lambda \overline{u : \rho^n}. (b : \star) \rightarrow ((x : \tau[D \mapsto X])[X \mapsto D]) \rightarrow b) \rightarrow b$. Thus,

$$\tau_1 \equiv D \ \overline{u'^n}, \text{ where } \Gamma \vdash u' : \rho.$$

Note that by operational semantics, the following reduction sequence follows for τ_1 :

$$\begin{aligned}
 D \ \overline{u'^n} &\longrightarrow (\lambda \overline{u : \rho^n}. (b : \star) \rightarrow ((x : \tau[D \mapsto X])[X \mapsto D]) \rightarrow b) \rightarrow b) \overline{u'^n} \\
 &\longrightarrow_n (b : \star) \rightarrow (\overline{x : \tau'}) \rightarrow b \rightarrow b
 \end{aligned}$$

Then by rule T_CASTDOWN and the definition of n -step cast operator, the type of $\text{cast}_{\downarrow}^{n+1} e_1$ is

$$(b : \star) \rightarrow (\overline{x : \tau'}) \rightarrow b \rightarrow b.$$

Note that by rule T_LAM, $\Gamma \vdash e_2 : (\overline{x : \tau'}) \rightarrow \sigma$. Therefore, by rule T_APP, we obtain $\Gamma \vdash (\text{cast}_{\downarrow}^{n+1} e_1) \sigma \ \overline{e_2} : \sigma$, which follows the result.

$\boxed{\vdash_{\text{wf}} \Sigma}$ Context well-formedness

$$\frac{}{\vdash_{\text{wf}} \emptyset} \text{TSENV_EMPTY}$$

$$\frac{\vdash_{\text{wf}} \Sigma \quad \Sigma \vdash_{\text{s}} T : \star}{\vdash_{\text{wf}} \Sigma, x : T} \text{TSENV_VAR}$$

$\boxed{\Sigma \vdash_{\text{pg}} \text{pgm} : T}$ Program typing

$$\frac{\Sigma_0 \vdash_{\text{d}} \text{decl} : \Sigma' \quad \Sigma = \Sigma_0, \Sigma' \quad \Sigma \vdash_{\text{s}} E : T}{\Sigma_0 \vdash_{\text{pg}} (\text{decl}; E) : T} \text{TSPGM_PGM}$$

$\boxed{\Sigma \vdash_{\text{d}} \text{decl} : \Sigma'}$ Datatype typing

$$\frac{\Sigma \vdash_{\text{s}} (\overline{u} : \overline{\kappa}^n) \rightarrow \star : \star \quad \overline{\Sigma, D : (\overline{u} : \overline{\kappa}^n) \rightarrow \star, \overline{u} : \overline{\kappa}^n \vdash_{\text{s}} (\overline{x} : \overline{T}) \rightarrow D \overline{u}^n : \star}}{\Sigma \vdash_{\text{d}} (\mathbf{data} D \overline{u} : \overline{\kappa}^n = \overline{K} \overline{x} : \overline{T}) : (D : (\overline{u} : \overline{\kappa}^n) \rightarrow \star, K : (\overline{u} : \overline{\kappa}^n) \rightarrow (\overline{x} : \overline{T}) \rightarrow D \overline{u}^n)} \text{TSDECL_DATA}$$

$\boxed{\Sigma \vdash_{\text{p}} p \Rightarrow E : T \rightarrow S}$ Pattern typing

$$\frac{K : (\overline{u} : \overline{\kappa}^n) \rightarrow (\overline{x} : \overline{T}) \rightarrow D \overline{u}^n \in \Sigma \quad \Sigma, \overline{x} : T[\overline{u} \mapsto \overline{v}] \vdash_{\text{s}} E : S \quad \Sigma \vdash_{\text{s}} T[\overline{u} \mapsto \overline{v}] : \star}{\Sigma \vdash_{\text{p}} K \overline{x} : T[\overline{u} \mapsto \overline{v}] \Rightarrow E : D \overline{v}^n \rightarrow S} \text{TSPAT_ALT}$$

$\boxed{\Sigma \vdash_{\text{s}} E : T}$ Expression typing

$$\frac{\vdash_{\text{wf}} \Sigma}{\Sigma \vdash_{\text{s}} \star : \star} \text{TS_AX}$$

$$\frac{\vdash_{\text{wf}} \Sigma \quad x : T \in \Sigma}{\Sigma \vdash_{\text{s}} x : T} \text{TS_VAR}$$

$$\frac{\Sigma \vdash_{\text{s}} E_1 : (\Pi x : T_2. T_1) \quad \Sigma \vdash_{\text{s}} E_2 : T_2}{\Sigma \vdash_{\text{s}} E_1 E_2 : T_1[x \mapsto E_2]} \text{TS_APP}$$

$$\frac{\Sigma, x : T_1 \vdash_{\text{s}} E : T_2 \quad \Sigma \vdash_{\text{s}} (\Pi x : T_1. T_2) : \star}{\Sigma \vdash_{\text{s}} (\lambda x : T_1. E) : (\Pi x : T_1. T_2)} \text{TS_LAM}$$

$$\frac{\Sigma \vdash_{\text{s}} T_1 : \star \quad \Sigma, x : T_1 \vdash_{\text{s}} T_2 : \star}{\Sigma \vdash_{\text{s}} (\Pi x : T_1. T_2) : \star} \text{TS_PI}$$

$$\frac{\Sigma, x : T \vdash_{\text{s}} E : T \quad \Sigma \vdash_{\text{s}} T : \star}{\Sigma \vdash_{\text{s}} (\mu x : T. E) : T} \text{TS_MU}$$

$$\frac{\Sigma \vdash_{\text{s}} E_1 : T \quad \overline{\Sigma \vdash_{\text{p}} p \Rightarrow E_2 : T \rightarrow \overline{S}} \quad \Sigma \vdash_{\text{s}} S : \star}{\Sigma \vdash_{\text{s}} \mathbf{case} E_1 \mathbf{of} \overline{p \Rightarrow E_2} : S} \text{TS_CASE}$$

Figure 8. Typing rules of the surface language

$\boxed{\vdash_{\text{wf}} \Sigma \rightsquigarrow \Gamma}$ Context well-formedness

$$\frac{\overline{\vdash_{\text{wf}} \emptyset \rightsquigarrow \emptyset} \quad \text{TRENV_EMPTY}}{\frac{\vdash_{\text{wf}} \Sigma \rightsquigarrow \Gamma \quad \Sigma \vdash_s T : \star \rightsquigarrow \tau}{\vdash_{\text{wf}} \Sigma, x : T \rightsquigarrow \Gamma, x : \tau}} \text{TRENV_VAR}$$

$\boxed{\Sigma \vdash_{\text{pg}} \text{pgm} : T \rightsquigarrow e}$ Program translation

$$\frac{\overline{\Sigma_0 \vdash_d \text{decl} : \Sigma' \rightsquigarrow e_1} \quad \Sigma = \Sigma_0, \overline{\Sigma'} \quad \Sigma \vdash_s E : T \rightsquigarrow e}{\Sigma_0 \vdash_{\text{pg}} (\overline{\text{decl}}; E) : T \rightsquigarrow \overline{e_1} \uplus e} \text{TRPGM_PGM}$$

$\boxed{\Sigma \vdash_d \text{decl} : \Sigma' \rightsquigarrow e}$ Datatype translation

$$\frac{\Sigma \vdash_s (\overline{u : \kappa^n}) \rightarrow \star : \star \rightsquigarrow (\overline{u : \rho^n}) \rightarrow \star \quad \overline{\Sigma, D : (\overline{u : \kappa^n}) \rightarrow \star, \overline{u : \kappa^n} \vdash_s (\overline{x : \overline{T}}) \rightarrow D \overline{u^n} : \star \rightsquigarrow (\overline{x : \overline{\tau}}) \rightarrow D \overline{u^n}}}{\Sigma \vdash_d (\text{data } D \overline{u : \kappa^n} = \overline{K \overline{x : \overline{T}}}) : (D : (\overline{u : \kappa^n}) \rightarrow \star, \overline{K : (\overline{u : \kappa^n}) \rightarrow (\overline{x : \overline{T}}) \rightarrow D \overline{u^n}}) \rightsquigarrow e} \text{TRDECL_DATA}$$

$$e \triangleq \text{let } D : (\overline{u : \rho^n}) \rightarrow \star = \mu X : (\overline{u : \rho^n}) \rightarrow \star. \lambda \overline{u : \rho^n}. (b : \star) \rightarrow ((\overline{x : \tau} [D \mapsto X]) \rightarrow b) \rightarrow b \text{ in} \\ \text{let } K_i : (\overline{u : \rho^n}) \rightarrow (\overline{x : \tau}) \rightarrow D \overline{u^n} = \lambda \overline{u : \rho^n}. \lambda \overline{x : \tau}. \text{cast}_{\uparrow}^{n+1} [D \overline{u^n}] (\lambda b : \star. \lambda c : (\overline{x : \tau}) \rightarrow b. c_i \overline{x}) \text{ in}$$

$\boxed{\Sigma \vdash_p p \Rightarrow E : T \rightarrow S \rightsquigarrow e}$ Pattern translation

$$\frac{K : (\overline{u : \kappa^n}) \rightarrow (\overline{x : \overline{T}}) \rightarrow D \overline{u^n} \in \Sigma \quad \Sigma, \overline{x : T[\overline{u \mapsto v}]} \vdash_s E : S \rightsquigarrow e \quad \Sigma \vdash_s T[\overline{u \mapsto v}] : \star \rightsquigarrow \tau}{\Sigma \vdash_p K x : T[\overline{u \mapsto v}] \Rightarrow E : D \overline{u^n} \rightarrow S \rightsquigarrow \lambda \overline{x : \overline{\tau}}. e} \text{TRPAT_ALT}$$

$\boxed{\Sigma \vdash_s E : T \rightsquigarrow e}$ Expression translation

$$\frac{\vdash_{\text{wf}} \Sigma}{\Sigma \vdash_s \star : \star \rightsquigarrow \star} \text{TR_AX}$$

$$\frac{\vdash_{\text{wf}} \Sigma \rightsquigarrow \Gamma \quad x : T \in \Sigma}{\Sigma \vdash_s x : T \rightsquigarrow x} \text{TR_VAR}$$

$$\frac{\Sigma \vdash_s E_1 : (\Pi x : T_2. T_1) \rightsquigarrow e_1 \quad \Sigma \vdash_s E_2 : T_2 \rightsquigarrow e_2}{\Sigma \vdash_s E_1 E_2 : T_1[x \mapsto E_2] \rightsquigarrow e_1 e_2} \text{TR_APP}$$

$$\frac{\Sigma, x : T_1 \vdash_s E : T_2 \rightsquigarrow e \quad \Sigma \vdash_s (\Pi x : T_1. T_2) : \star \rightsquigarrow \Pi x : \tau_1. \tau_2}{\Sigma \vdash_s (\lambda x : T_1. E) : (\Pi x : T_1. T_2) \rightsquigarrow \lambda x : \tau_1. e} \text{TR_LAM}$$

$$\frac{\Sigma \vdash_s T_1 : \star \rightsquigarrow \tau_1 \quad \Sigma, x : T_1 \vdash_s T_2 : \star \rightsquigarrow \tau_2}{\Sigma \vdash_s (\Pi x : T_1. T_2) : \star \rightsquigarrow \Pi x : \tau_1. \tau_2} \text{TR_PI}$$

$$\frac{\Sigma, x : T \vdash_s E : T \rightsquigarrow e \quad \Sigma \vdash_s T : \star \rightsquigarrow \tau}{\Sigma \vdash_s (\mu x : T. E) : T \rightsquigarrow \mu x : \tau. e} \text{TR_MU}$$

$$\frac{\Sigma \vdash_s E_1 : T \rightsquigarrow e_1 \quad \overline{\Sigma \vdash_p p \Rightarrow E_2 : T \rightarrow S \rightsquigarrow e_2} \quad \Sigma \vdash_s S : \star \rightsquigarrow \sigma}{\Sigma \vdash_s \text{case } E_1 \text{ of } \overline{p \Rightarrow E_2} : S \rightsquigarrow (\text{cast}_{\downarrow}^{n+1} e_1) \sigma \overline{e_2}} \text{TR_CASE}$$

Figure 9. Translation rules of the surface language