

Indexed Kind Checking for Hierarchical Database Schemas

Anonymous Author(s)

Abstract

Types are an appropriate tool for describing the structure of hierarchical databases. But two common database idioms, *foreign keys* and *secondary indexes*, cannot be expressed using standard type systems. We present a language in which a database schema is described using a type layered over a dependently sorted index language. The index language consists of keys and relations on keys. Multiple occurrences of the same index-level relation inside of a schema specify a correlation between two distinct portions of a database instance; both secondary indexes and foreign keys can be specified in this manner. We present an algorithmic system of judgments for deciding an indexed kind relation on this schema language and provide it with an intrinsic denotational semantics.

Keywords: Indexed types, Database schemas, Dependent types

ACM Reference Format:

Anonymous Author(s). 2022. Indexed Kind Checking for Hierarchical Database Schemas. In *Proceedings of ACM Conference (Conference'17)*. ACM, New York, NY, USA, 14 pages. <https://doi.org/10.1145/nnnnnnnn.nnnnnnn>

1 Introduction

Programmers increasingly choose document database systems due to their high performance and intuitive hierarchical structure. Schemas for such databases are often intentionally omitted under the premise that they inhibit rapid iteration. However, we disagree with this premise. We believe that the arguments in favor of database schemas are analogous to those in favor of static types in programming languages. By pairing data with its intended meaning, a schema increases programmer comprehension: a programmer may not know what to do with a record field called *widgetId*, but a good

schema could convey this information, specifying that *widgetId* refers to a structure stored in a specific location of the database.

The type syntax of a standard programming language such as Typescript is for the most part an appropriate tool to describe the structure of modern hierarchical databases. However, two common features of hierarchical databases lie beyond the reach of standard type systems. First, a database may store *foreign keys*: values, such as the aforementioned *widgetId*, used to refer to entities stored elsewhere in the database. Second, a database may replicate a dataset in multiple configurations (a.k.a. *secondary indexes*) to accommodate multiple access patterns in an efficient, spatially local manner.

The prevalence of foreign keys in hierarchical databases constrains sharply against the in-memory data structures one typically finds in general purpose languages. A pointer in a language such as C++ is not a foreign key, because statically our concern is the type of referenced data rather than its location. In contrast, a foreign key is drawn from a specific subset of references of a given type. For example, consider an e-commerce database which records purchases performed by customers. Each purchase holds a *creditCardId* foreign key referring to the credit card used for the payment; this card is not drawn from a global pool of all cards, but instead the pool of cards owned by the customer which made the purchase.

Secondary indexes are also common and useful database structure beyond the reach of traditional type systems. Consider this example demonstrating the utility of secondary indexes. An e-commerce company wishes to issue a recall on a product. To do so, it must obtain a list of all customers that have purchased an item known to be defective; to compute this list efficiently, the set of all customers which purchased each item must be stored in nearby addresses. At another time, a customer may wish to obtain a list of all items he has purchased from the site; to perform this operation efficiently, we must store the set of all of a customer's purchased items in nearby addresses. These two requirements are at odds if our database merely stores a single collection of all purchases processed. However, if we redundantly store customer-to-item and item-to-customers map, we can satisfy both requirements at once. Such redundant data structures are called secondary indexes.

In practice, foreign keys often refer to deleted entities. Such a foreign key is a time-bomb which will raise an exception the next time the data is accessed. Secondary indexes

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Conference'17, July 2017, Washington, DC, USA

© 2022 Association for Computing Machinery.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM. . \$15.00

<https://doi.org/10.1145/nnnnnnnn.nnnnnnn>

can become out-of-sync due to programmer errors. A formal schema language including notions of foreign keys and indexes could be compiled to validation routines to detect these errors.

We make the following contributions.

- We design an expressive schema language for hierarchical databases featuring foreign keys and secondary indexes.
- We provide a category-theoretic characterization of the dependently-sorted analog of simple products, which play a central role in our schema language. This characterization is similar to, but more general than, existing characterizations e.g. by Palmgren [4]; in particular, we identify a *transpose distribution* property necessary for dependent simple products in models whose total categories, unlike Palmgren's, are not thin.

Since we are working with two distinct concepts which are both typically referred to as *index* – indexes in the sense of database indexes and indices in the sense of indexed type checking – we refer to database secondary indices as *reconfigurations* in the sequel.

2 Example

2.1 Foreign Keys

The purpose of a schema is to identify a set of database instances, so before examining our first schema we briefly define some formalisms for describing database instances. We consider *database instances* (also called *instances*) as partial mappings from lists of strings to strings. Given a database instance f and a list of strings σ , we obtain an instance $f|_{\sigma}$ called *the restriction of f to σ* , defined for lists σ' as $f|_{\sigma}(\sigma') \stackrel{\text{def}}{=} f(\sigma ++ \sigma')$, where $\sigma ++ \sigma'$ is the concatenation of σ and σ' . By an abuse of language, we say “ f maps σ to $f|_{\sigma}$ ”. We write $[s_1, \dots, s_n]$ for the list containing the n strings s_1, \dots, s_n .

Figure 1 shows a schema fragment for an e-commerce database. It begins by declaring five predicates: first four unary predicates for item ids, customer ids, purchase ids, and card types, and then a ternary predicate for card ids. The predicate variables and their classifiers have been bolded and colored blue to indicate that they are part of the *index language*. This simple language defines conceptual entities such as predicates and relations, divorced from the physical details of where these entities are stored. The index language is layered beneath our schema language, i.e. schemas may depend on indices but not vice versa.

Lines 7-10 use the record type constructor to define a subschema representing credit cards. It denotes the set of instances which map the list [“billingAddr”] to any string and the list [“cardType”] to any string satisfying the **CardType** predicate. We consider the latter a foreign key; rather than

```

1   $\forall \lambda (\text{ItemId} : \text{str} \rightarrow \text{prop}).$ 
2   $\forall \lambda (\text{CustId} : \text{str} \rightarrow \text{prop}).$ 
3   $\forall \lambda (\text{PurchaseId} : \text{str} \rightarrow \text{prop}).$ 
4   $\forall \lambda (\text{CardType} : \text{str} \rightarrow \text{prop}).$ 
5   $\forall \lambda (\text{CardId} : (x:\text{str}) \rightarrow \text{prf } (\text{CustId } x) \rightarrow \text{str} \rightarrow \text{prop}).$ 
6
7  type Card = {
8    billingAddr : str
9    cardType    : { x : str | CardType x }
10 }
11
12 type Purchase =  $\lambda (\text{cust} : \text{str}). \lambda (\text{prf } (\text{CustId } \text{cust})). \{$ 
13   itemId : { x : str | ItemId x },
14   cardId  : { x : str | CardId cust x }
15 }
16
17 type Customer =  $\lambda (\text{cust} : \text{str}). \lambda (\text{prf } (\text{CustId } \text{cust})). \{$ 
18   purchases : {
19     [p : str] : PurchaseId p > Purchase cust
20   },
21   cards      : {
22     [card : str] : CardId cust card > Card
23   }
24 }
25
26 {
27   cardTypes : { [x : str] : CardType x > "*" }
28   customers : { [x : str] : CustId x > Customer x }
29 }
```

Figure 1. Schema for e-commerce database with foreign keys

taking the naive view that a foreign key is a reference to a specific location in the database, we instead take it to be a string satisfying a predicate in our index-level context (i.e. an element of a conceptual set). We believe our approach is cleaner and more declarative than the naive approach. Because a conceptual set may be replicated at multiple locations in a database, it would not make sense for a foreign key to refer to one of these locations rather than another.

Lines 12-15 define an index-to-type operator called *Purchase*, which maps a string **cust** and a proof that **cust** satisfies the **CustId** predicate to a type representing purchases made by customer **cust**. In the application **CardId cust card**, the second argument to **CardId** is missing. This is because it is a proof; since any two proofs of the same proposition are interchangeable, we apply proofs implicitly to reduce visual clutter.

Lines 17-24 define an index-to-type operator *Customer* mapping a string **cust** and a proof of **CustId cust** to a type representing customers with id **cust**. It denotes the set of instances which

```

221 1  ∨ λ (ItemId : str → prop).
222 2  ∨ λ (CustId : str → prop).
223 3  ∨ λ (Purchased :
224 4      (x : str) → prf (CustId x) →
225 5      (y : str) → prf (ItemId y) → prop).
226 6
227 7  {
228 8      custToItem : {
229 9          [c, i : str] : CustId c, ItemId i, Purchased c i > "*"
230 10     }
231 11     itemToCust : {
232 12         [i, c : str] : CustId c, ItemId i, Purchased c i > "*"
233 13     }
234 14 }

```

Figure 2. Reconfigurations (secondary indices) in an e-commerce database

- Map all lists ["purchases", **p**] such that **purchaseId p** holds to instances of type Purchase **cust**, and
- Map all lists ["cards", **card**] such that **CardId cust card** holds to instances of type Card.

Finally, lines 26-29 define the schema using a record type. It denotes the set of instances which

- Map ["cardTypes"] to a subinstance that stores the set of all card types by mapping each string **x** satisfying the **CardType** predicate to the string "*", and
- Map ["customers"] to a dictionary that maps every **x** satisfying the **CustId** predicate to a subschema of type Customer **x**.

Reconfigurations

Figure 2 shows a schema for a simple database with reconfigurations. An instance satisfies the schema if

- For each customer/item pair **c, i** satisfying **Purchased c i**, it maps ["custToItem", **c, i**] to the string "*" and ["itemToCust", **i, c**] to "*".
- It is not defined anywhere else.

Such an instance allows us to efficiently query both the set of all items purchased by a customer (via **custToItem**) and the set of all customers which purchased an item (via **itemToCust**). Importantly, the two fields both refer to the same set of purchase entities via the index-level relation **Purchased**.

3 Syntax

Index-level terms **i, j, k** consist of string literals, variables, applications, and the proposition constant **true**. An application is annotated with the domain and codomain of the applied function; this eases the definition of our denotational semantics, but the types can be inferred and are thus omitted in our

$TypeVars$	$\stackrel{def}{=}$	the set of all type variables	276
$IndexVars$	$\stackrel{def}{=}$	the set of all index variables	277
x, y, z	\in	$TypeVars$	279
a, b, P	\in	$IndexVars$	280
s, t	\in	$Strings$	281
i, j, k (index)	$::=$	s (string literal)	282
		App _{[a;q],p} (j, k) (index application)	283
		a (index variable)	284
		true (true proposition)	285
p, q, r (sort)	$::=$	str (string sort)	286
		prop (proposition sort)	287
		prf j (proof sort)	288
		(a : q) → r (function sort)	289
τ, σ (pre-type)	$::=$	{ [a : str] : τ } (dictionary)	290
		{ s _i : τ_i ^{$i \in 1..n$} } (record)	291
		{ a : str i } (string refinement)	292
		$\lambda a : q. \tau$ (index-to-type abstr.)	293
		$\lambda x : \kappa. \tau$ (type-to-type abstr.)	294
		$\bigvee \tau$ (union)	295
		$\tau \sigma$ (type app.)	296
		$\tau [j]$ (index-to-type app.)	297
κ, ρ (pre-kind)	$::=$	*	298
		$\kappa \rightarrow \rho$ (proper type pre-kind)	299
		$\forall a : q. \kappa$ (type-to-type op)	300
Ω, Ψ (pre-sort-context)	$::=$	$\Omega, a : q$ (extension)	301
		\diamond (empty)	302
Γ (pre-kind-context)	$::=$	$\Gamma, x : \kappa$ (extension)	303
		\diamond (empty)	304

Figure 3. Syntax

examples, which instead use the notation **j k** as shorthand for **App**_{[a;q],p}(**j, k**).

Following [5], index level classifiers are called *sorts*. Sorts are written using the metavariables **p, q, r**. We include the sorts **str** for strings, **prop** for propositions, **prf j** proofs of the proposition **j**, and the sort constructor **(a : q) → r** for constructing dependent function sorts, which are essentially index-level Π -types.

Our types τ, σ are also called *schemas*. The type { [**a : str**] : τ } represents string-keyed dictionaries of value type τ . The type τ depends on the index variable **a**, and { [**a : str**] : τ } describes the set of database instances which contains exactly those root-level keys **s** such that $\tau[s/a]$ represents a non-empty set of database instances. Record types, type-type applications, and type-type abstractions in our type syntax

$$\begin{array}{c}
\frac{}{\Omega \vdash s : \text{str}} \quad \frac{}{\Omega, a : q, \Omega' \vdash a : q} \\
\frac{\Omega \vdash j : (a : q) \rightarrow r \quad \Omega \vdash k : q}{\Omega \vdash \text{App}_{[a:q],r}(j, k) : r[k/a]} \quad \frac{}{\Omega \vdash \text{true} : \text{prop}} \\
\frac{}{\Omega \vdash \text{str}} \quad \frac{\Omega \vdash q \quad \Omega, a : q \vdash r}{\Omega \vdash (a : q) \Rightarrow r} \quad \frac{\Omega \vdash j : \text{prop}}{\Omega \vdash \text{prf } j} \\
\frac{}{\diamond \vdash} \quad \frac{\Omega \vdash \quad \Omega \vdash p}{\Omega, a : p \vdash}
\end{array}$$

Figure 4. Sorting, sort formation, and sort context formation

are standard. A string refinement type of the form $\{a : \text{str} \mid i\}$ represents the set of all strings s such that $i[s/a]$ holds.

Finally, our type syntax includes index-type abstractions $\lambda(a : q). \tau$, index-type application $\tau [j]$, and union $\bigvee \tau$. In a well-kinded union type $\bigvee \tau$, the type τ is an index-type function whose codomain is $*$ (the kind of proper types). τ then denotes a function which maps a semantic index to a set of database instances; $\bigvee \tau$ denotes the union over the applications of τ to all semantic indices in its domain.

In figures 1 and 2, a dictionary type of the form $\{[a : \text{str}] : i_1, \dots, i_n \vdash \tau\}$ is sugar for

$$\{[a : \text{str}] : \bigvee \lambda(a_1 : \text{prf } i_1) \dots \bigvee \lambda(a_n : \text{prf } i_n). \tau\}$$

4 Static semantics

4.1 Index-level static semantics

Index-level judgment rules are shown in Figure 4. A judgment of the form $\Omega \vdash q$ means that the pre-sort q is a well-formed sort under the context Ω . A judgment of the form $\Omega \vdash$ means that the sort pre-context Ω is a well-formed sort context, i.e. for each binding $a : q$ such that $\Omega = \Omega_1, a : q, \Omega_2$ we have $\Omega_1 \vdash q$. Finally, a judgment $\Omega \vdash j : q$ means that the index j has sort q under context Ω .

4.2 Type-level static semantics

Type-level judgment rules are shown in Figure 5. A judgment of the form $\Omega \vdash \kappa$ means that the pre-kind κ is a well-formed kind under context Ω . A judgment of the form $\Omega \vdash \Gamma$ means that Γ is a well-formed kind context under sort context Ω . Finally, a judgment of the form $\Omega \mid \Gamma \vdash \tau : \kappa$ means that the type τ has kind κ under sort context Ω and kind context Γ .

5 Denotational Semantics

5.1 Index Language

We interpret our index language using Dybjer's *Categories-with-Families* (CwF) framework for semantically modeling

$$\begin{array}{c}
\frac{\Omega, a : \text{str} \mid \Gamma \vdash \tau : *}{\Omega \mid \Gamma \vdash \{[a : \text{str}] : \tau\} : *} \quad \frac{\Omega \mid \Gamma \vdash \tau_i : *^{i \in 1..n}}{\Omega \mid \Gamma \vdash \{s_i : \tau_i^{i \in 1..n}\} : *} \\
\frac{\Omega, a : \text{str} \vdash i : \text{prop}}{\Omega \mid \Gamma \vdash \{a : \text{str} \mid i\} : *} \\
\text{K-INDABS} \quad \frac{\Omega \vdash q \quad \Omega, a : q \mid \Gamma \vdash \tau : \kappa \quad a \notin FV(\Gamma)}{\Omega \mid \Gamma \vdash \lambda a : q. \tau : \forall a : q. \kappa} \\
\frac{\Omega \vdash \kappa \quad \Omega \mid \Gamma, x : \kappa \vdash \tau : \rho}{\Omega \mid \Gamma \vdash \lambda x : \kappa. \tau : \kappa \rightarrow \rho} \quad \frac{\Omega \mid \Gamma \vdash \tau : \forall q. *}{\Omega \mid \Gamma \vdash \bigvee \tau : *} \\
\frac{\Omega \mid \Gamma \vdash \tau : \forall q. \rho \quad \Omega \vdash j : q}{\Omega \mid \Gamma \vdash \tau [j] : \rho} \\
\frac{\Omega \mid \Gamma \vdash \tau : \kappa \rightarrow \rho \quad \Omega \mid \Gamma \vdash \sigma : \kappa}{\Omega \mid \Gamma \vdash \tau \sigma : \rho} \quad \frac{}{\Omega \vdash *} \\
\frac{\Omega \vdash q \quad \Omega, a : q \vdash \kappa}{\Omega \vdash \forall a : q. \kappa} \quad \frac{\Omega \vdash \kappa \quad \Omega \vdash \rho}{\Omega \vdash \kappa \rightarrow \rho} \quad \frac{}{\Omega \vdash \diamond} \\
\frac{\Omega \vdash \Gamma \quad \Omega \vdash \kappa}{\Omega \vdash \Gamma, x : \kappa}
\end{array}$$

Figure 5. Kinding and kind formation

independently typed languages [1]. We first review the definition of CwF, then review the standard set-theoretic CwF, and finally provide an interpretation of the index language with respect to a CwF. The entirety of this section is standard. However, we rename some of the standard terminology to fit into the present setting, e.g. instead of dependent types we have dependent sorts.

5.1.1 Categories with Families. A Category-with-Families (CwF) consists of

- A category \mathbb{C} with a terminal object, called the *category of contexts*. Objects of \mathbb{C} are called semantic contexts and written with the symbols Ω, Ψ, Υ .
- For each semantic context Ω a collection $St(\Omega)$ called the *semantic sorts* of Ω . Semantic sorts are written with the symbols X, Y , and Z .
- For each semantic context Ω and $X \in St(\Omega)$ a collection $In(\Omega, X)$ called the *semantic indices* of sort X .
- For each arrow $f : \Omega \rightarrow \Psi$ a mapping $- \{f\} : St(\Psi) \rightarrow St(\Omega)$. These mappings preserve composition, in the sense that we have $- \{id_\Omega\} = id_{St(\Omega)}$ and for $g : \Psi \rightarrow \Upsilon$ we have $- \{g \circ f\} = - \{g\} \{f\}$.

- For each arrow $f : \Omega \rightarrow \Psi$ and sort $X \in St(\Psi)$ a mapping $- \{f\} : In(\Psi, X) \rightarrow In(\Omega, X \{f\})$. These mappings preserve composition in the same sense as above.
- For each context Ω and sort $X \in St(\Omega)$, we have an object $\Omega.X$ of \mathbb{C} , an arrow $p(X) : \Omega.X \rightarrow \Omega$, and an index $v_X \in In(\Omega.X, X \{p(X)\})$ such that for all $f : \Psi \rightarrow \Omega$ and $M \in In(\Psi, X \{f\})$ there exists a unique morphism $\langle f, M \rangle_X : \Psi \rightarrow \Omega.X$ such that $p(X) \circ \langle f, M \rangle_X = f$ and $v_X \{ \langle f, M \rangle_X \} = M$.

For $M \in In(\Omega, X)$ we write \overline{M} for the arrow $\langle id_\Omega, M \rangle_X : \Omega \rightarrow \Omega.X$.

For $f : \Psi \rightarrow \Omega$ and $X \in St(\Omega)$ we define $q(f, X) : \Psi.X \{f\} \rightarrow \Omega.X$, called the *weakening* f by X as

$$q(f, X) = \langle f \circ p(X \{f\}), v_X \{f\} \rangle_X$$

A *weakening map* is a morphism of the form $p(X) : \Omega.X \rightarrow \Omega$ or of the form $q(w, Y)$ where w is a weakening map. We write f^+ and X^+ for $f \{w\}$ and $X \{w\}$ when w is a weakening map that is clear from context.

A CwF is said to *support* Π -sorts if for any two sorts $X \in St(\Omega)$ and $Y \in St(\Omega.X)$ there is a sort $\Pi(X, Y) \in St(\Omega)$ and a morphism

$$App_{X,Y} : \Omega.X.\Pi(X, Y)^+ \rightarrow \Omega.X.Y$$

such that

$$p(X) \circ App_{X,Y} = p(\Pi(X, Y)^+) \quad App-T$$

and for every $M \in In(\Omega.X, Y)$

$$App_{X,Y} \circ \overline{M \{p(X)\}} = \overline{M} \quad \Pi-C'$$

and for every morphism $f : B \rightarrow \Gamma$

$$\begin{aligned} App_{X,Y} \circ q(q(f, X), \Pi(X, Y) \{q(f, X)\}) \\ = q(q(f, X), Y) \circ App_{X \{f\}, Y \{q(f, \sigma)\}} \end{aligned}$$

5.1.2 Set-theoretic CwF. Our index language will be interpreted using a standard set-theoretic CwF. This CwF's category of contexts is **Sets**, the category of sets and functions. For a set Ω , $St(\Omega)$ is the collection of all Ω -indexed families of sets. For all Ω and $X \in St(\Omega)$, $In(\Omega, X)$ is the collection of families $(x_\omega \in X_\omega)_{\omega \in \Omega}$ selecting an element $x_\omega \in X_\omega$ for each $\omega \in \Omega$. For all $h : \Omega \rightarrow \Psi$, $X \in St(\Psi)$, and $M \in In(\Psi, X)$ we define $X \{h\}_\omega \stackrel{def}{=} X_{f(\omega)}$ and $M \{h\}_\omega \stackrel{def}{=} M_{f(\omega)}$. For each Ω and $X \in St(\Omega)$ we define

$$\Omega.X \stackrel{def}{=} \{(\omega, x) \mid \omega \in \Omega \text{ and } x \in X_\omega\}$$

$p(X)(\omega, x) \stackrel{def}{=} \omega$, and $(v_X)_{(\omega, x)} \stackrel{def}{=} x$. Finally, letting $f : \Psi \rightarrow \Omega$ and $M \in In(\Psi, X \{f\})$, we have $\langle f, M \rangle_X(\psi) = (f(\psi), M_\psi)$.

$$\begin{aligned} \llbracket \diamond \rrbracket &\stackrel{def}{=} \top \\ \llbracket \Omega, a : q \rrbracket &\stackrel{def}{=} \llbracket \Omega \rrbracket, \llbracket \Omega; q \rrbracket \text{ if } x \text{ not in } \Omega, \text{ undefined otherwise.} \\ \llbracket \Omega; (j : q) \Rightarrow r \rrbracket &\stackrel{def}{=} \Pi(\llbracket \Omega; q \rrbracket, \llbracket \Omega, a : q; r \rrbracket) \\ \llbracket \Omega; \text{prf } j \rrbracket &\stackrel{def}{=} (\{*\} \text{ if } \llbracket \Omega; j \rrbracket_\omega = \text{true}, \emptyset \text{ otherwise})_{\omega \in \llbracket \Omega \rrbracket} \\ \llbracket \Omega; \text{prop } j \rrbracket &\stackrel{def}{=} (\{true, false\})_{\omega \in \llbracket \Omega \rrbracket} \\ \llbracket \Omega; \text{str} \rrbracket &\stackrel{def}{=} (\text{the set of strings})_{\omega \in \llbracket \Omega \rrbracket} \\ \llbracket \Omega; \text{App}[a;q].r(i,j) \rrbracket &\stackrel{def}{=} App_{\llbracket \Omega; q \rrbracket, \llbracket \Omega, a;q;r \rrbracket} \circ \langle \llbracket \Omega; i \rrbracket, \llbracket \Omega; j \rrbracket^+ \rangle_{\llbracket \Omega; (a;q) \rightarrow r \rrbracket^+} \\ \llbracket \Omega, a : q; a \rrbracket &\stackrel{def}{=} v_{\llbracket \Omega, a;q \rrbracket} \\ \llbracket \Omega; s \rrbracket &\stackrel{def}{=} (s)_{\omega \in \llbracket \Omega \rrbracket} \end{aligned}$$

Figure 6. Interpretation of index language

$$\begin{aligned} \llbracket \Omega \vdash * \rrbracket_\omega &\stackrel{def}{=} \mathcal{PP}(Inst) \\ \llbracket \Omega \vdash \forall q.k \rrbracket &\stackrel{def}{=} \Pi_{\Omega, q} \llbracket \Omega, a : q \vdash k \rrbracket \\ \llbracket \Omega \vdash \kappa \rightarrow \rho \rrbracket &\stackrel{def}{=} \llbracket \Omega \vdash \kappa \rrbracket \Rightarrow \llbracket \Omega \vdash \rho \rrbracket \end{aligned}$$

Figure 7. Interpretation of kinding judgments

5.1.3 Interpretation. Our index language is a fragment of the calculus of constructions, with the unnotable addition of strings. Its interpretation appears, for example, in Hoffman's tutorial[2], and is displayed in Figure 6 for convenience.

We also state the fragment of the standard soundness theorem which will be relevant in the sequel:

Theorem 5.1. *Our interpretation satisfies the following properties.*

- If $\Omega \vdash$ then $\llbracket \Omega \rrbracket$ is an object of the context category **Sets**.
- If $\Omega \vdash q$ then $\llbracket \Omega; q \rrbracket$ is an element of $St(\llbracket \Omega \rrbracket)$
- If $\Omega \vdash j : q$ then $\llbracket \Omega; j \rrbracket$ is an element of $In(\llbracket \Omega \rrbracket, \llbracket \Omega; q \rrbracket)$

5.2 Type language

Our kinding and kind formation judgments are interpreted in terms of the fibration $Fam(\mathbf{Sets})$. We assume basic knowledge of fibrations, which can be learned from Jacobs [3]. A kind-in-sorting-context $\Omega \vdash \kappa$ is interpreted as an object of $Fam(\mathbf{Sets})_{\llbracket \Omega \rrbracket}$, concretely a $\llbracket \Omega \rrbracket$ -indexed family of sets. For pre-kind contexts $\Gamma = x_1 : \kappa_1, \dots, x_n : \kappa_n$, the kinding context formation judgment $\Omega \vdash \Gamma$ is interpreted as the product $\llbracket \Omega \vdash \kappa_1 \rrbracket \times \dots \times \llbracket \Omega \vdash \kappa_n \rrbracket$ in $Fam(\mathbf{Sets})_{\llbracket \Omega \rrbracket}$. A kinding judgment $\Omega \mid \Gamma \vdash \tau : \kappa$ is interpreted as an arrow of $Fam(\mathbf{Sets})_{\llbracket \Omega \rrbracket}$ from $\llbracket \Omega \vdash \Gamma \rrbracket$ to $\llbracket \Omega \vdash \kappa \rrbracket$.

5.2.1 Dependent Simple Products. We define a CwF-fibration as a pair (C, p) where C is a CwF and $p : \mathbb{E} \rightarrow \mathbb{C}$ is a fibration whose base category \mathbb{C} is C 's context category.

We say that a CwF-fibration has *dependent simple products* if

$$\begin{aligned}
\llbracket \{[a : \text{str}] : \tau\} \rrbracket_{\omega Y} &\stackrel{\text{def}}{=} \{f \mid \forall s. f|_s \in \llbracket \tau \rrbracket_{(\omega, s)Y^+} \text{ if } \llbracket \tau \rrbracket_{(\omega, s)Y^+} \text{ is non-empty and } f|_s = \emptyset \text{ otherwise}\} \\
\llbracket \{s_i : \tau_i^{i \in 1..n}\} \rrbracket_{\omega Y} &\stackrel{\text{def}}{=} \{f \mid \forall i \in 1..n. f|_{s_i} \in \llbracket \tau_i \rrbracket_{\omega Y}\} \\
\llbracket [a : \text{str}] \mid i \rrbracket_{\omega Y} &\stackrel{\text{def}}{=} \{f \mid f(\epsilon) \downarrow \wedge \llbracket i \rrbracket_{(\omega, f(\epsilon))} = \text{true}\} \\
\llbracket \vee \tau \rrbracket_{\omega Y} &\stackrel{\text{def}}{=} \bigcup_{M \in \llbracket \mathbf{q} \rrbracket_{\omega}} (\llbracket \tau \rrbracket_{\omega Y})_M
\end{aligned}$$

Figure 8. Semantics for “non-operational” kinding rules

$$\begin{aligned}
f &\stackrel{\text{def}}{=} \llbracket \Omega \mid \Gamma \vdash \tau : \forall \mathbf{q}. \rho \rrbracket : \llbracket \Omega \vdash \Gamma \rrbracket \rightarrow \Pi_{\Omega, \mathbf{q}} \llbracket \Omega, a : \mathbf{q} \vdash \rho \rrbracket \\
M &\stackrel{\text{def}}{=} \llbracket \Omega \vdash j : \mathbf{q} \rrbracket \in \text{In}(\llbracket \Omega \rrbracket, \llbracket \Omega \vdash \mathbf{q} \rrbracket) \\
\hline
\llbracket \Omega \mid \Gamma \vdash \tau [j] : \rho \rrbracket &\stackrel{\text{def}}{=} \overline{M}^* (f^b) \\
X &\stackrel{\text{def}}{=} \llbracket \Omega \vdash \mathbf{q} \rrbracket \in \text{St}(\Omega) \\
f &\stackrel{\text{def}}{=} \llbracket \Omega, a : \mathbf{q} \mid \Gamma \vdash \tau : \kappa \rrbracket : \mathbf{p}(X)^* \llbracket \Omega \vdash \Gamma \rrbracket \rightarrow \llbracket \Omega, a : \mathbf{q} \vdash \kappa \rrbracket \\
\hline
\llbracket \Omega \mid \Gamma \vdash \lambda a : \mathbf{q}. \tau : \forall \mathbf{q}. \kappa \rrbracket &\stackrel{\text{def}}{=} f^\#
\end{aligned}$$

Figure 9. Semantics for abstraction and application

$$\begin{array}{ccc}
\mathbb{E}_\Omega & \xrightarrow{u^*} & \mathbb{E}_\Psi \\
\mathbf{p}(X)^* \left(\begin{array}{c} \uparrow \\ \downarrow \end{array} \right) \Pi_X & & \mathbf{p}(X\{u\})^* \left(\begin{array}{c} \uparrow \\ \downarrow \end{array} \right) \Pi_{X\{u\}} \\
\mathbb{E}_{\Omega.X} & \xrightarrow{q(u, X)^*} & \mathbb{E}_{\Psi.X\{u\}}
\end{array}$$

Figure 10. Some components of Beck-Chevalley for dependent simple products

- For all contexts Ω and all $X \in \text{St}(\Omega)$ the functor $\mathbf{p}(X)^*$ has a right adjoint Π_X . (We write $(-)^{\#}$ for the transposition from $\mathbf{p}(X)^*(A) \rightarrow B$ to $A \rightarrow \Pi_X(B)$, and $(-)^b$ for transposition in the opposite direction.)
- The above adjunction satisfies a *transpose distribution* property: for contexts Ω, Ψ , context arrows $u : \Psi \rightarrow \Omega$, sorts $X \in \text{St}(\Omega)$, total objects Γ over Ω , total objects Δ over $\Omega.X$, and total arrows $f : \mathbf{p}(X)^*(\Gamma) \rightarrow \Delta$. We have $u^* f^{\#} \cong (q(u, X)^* f)^{\#}$.
- For every $u : \Psi \rightarrow \Omega$ and $X \in \text{St}(\Omega)$ the canonical natural transformation $u^* \Pi_X \Rightarrow \Pi_{X\{u\}} q(u, X)^*$ is an isomorphism. This is typically called a *Beck-Chevalley condition*.

It is a standard fact of fibrations that for arrows $u : \Omega \rightarrow \Psi$ and $v : \Psi \rightarrow \Upsilon$ of the base category, we have $u^* v^* \cong (v \circ u)^*$. From this we derive a natural isomorphism $q(u, X)^* \mathbf{p}(X)^* \cong$

$\mathbf{p}(X\{u\})^* u^*$, which will clarify the second and third points:

$$\begin{aligned}
&q(u, X)^* \mathbf{p}(X)^* \\
&\cong \langle u \circ \mathbf{p}^*(X\{u\}), \mathbf{v}_{X\{u\}} \rangle_X^* \mathbf{p}(X)^* \\
&= (\mathbf{p}(X) \circ \langle u \circ \mathbf{p}^*(X\{u\}), \mathbf{v}_{X\{u\}} \rangle_X)^* \\
&= (u \circ \mathbf{p}(X\{u\}))^* \\
&\cong \mathbf{p}(X\{u\})^* u^*
\end{aligned}$$

In the second point, it is not immediately clear that the right-hand side of the isomorphism is “well-typed”. However, applying our isomorphism to the domain of $q(z, X)^* f$ gives $q(u, X)^* \mathbf{p}(X)^*(\Gamma) \cong \mathbf{p}(X\{u\})^* u^*(\Gamma)$. Agreement of codomains follows from the Beck-Chevalley condition.

In the third point, the *canonical transformation* is obtained as the transpose of

$$\mathbf{p}(X\{u\})^* u^* \Pi_X \xrightarrow{\cong} q(u, X)^* \mathbf{p}(X)^* \Pi_X \xrightarrow{q(u, X)^{\epsilon}} q(u, X)^*$$

5.2.2 Set-theoretic Dependent Simple Products. In our set-theoretic model, for semantic contexts Ω and semantic sorts $X \in \text{St}(\Omega)$, we define $\Pi_X : \text{Fam}(\text{Sets})_{\Omega.X} \rightarrow \text{Fam}(\text{Sets})_{\Omega}$ as:

$$\begin{aligned}
\Pi_X \left(\left(B_{(\omega, x)} \right)_{(\omega, x) \in \Omega.X} \right) &\stackrel{\text{def}}{=} \left(\Pi_{X \in X_{\omega}} B_{(\omega, x)} \right)_{\omega \in \Omega} \\
\Pi_X \left(f_{(\omega, x)} : B_{(\omega, x)} \rightarrow C_{(\omega, x)} \right)_{(\omega, x) \in \Omega.X} &\stackrel{\text{def}}{=} \left(\Pi_{X \in X_{\omega}} f_{(\omega, x)} \right)_{\omega \in \Omega}
\end{aligned}$$

We show that Π_X is the right adjoint of $\mathbf{p}(X)^*$ in Appendix A.2, with an underlying correspondence

$$\begin{aligned}
(A_{\omega})_{\omega \in \Omega} &\longrightarrow (\Pi_{X \in X_{\omega}} B_{(\omega, x)})_{\omega \in \Omega} = \Pi_X (B_{(\omega, x)})_{(\omega, x) \in \Omega.X} \\
\hline
\mathbf{p}(X)^* (A_{\omega})_{\omega \in \Omega} &= (A_{\omega})_{(\omega, x) \in \Omega.X} \longrightarrow (B_{(\omega, x)})_{(\omega, x) \in \Omega.X}
\end{aligned}$$

From top to bottom, $(-)^b$ takes an arrow

$$\langle f_{(\omega, x)} : A_{\omega} \rightarrow B_{(\omega, x)} \rangle_{X \in X_{\omega}}_{\omega \in \Omega}$$

to

$$(f_{(\omega, x)} : A_{\omega} \rightarrow B_{(\omega, x)})_{(\omega, x) \in \Omega.X}$$

From bottom to top, $(-)^{\#}$ takes an arrow

$$(f_{(\omega, x)} : A_{\omega} \rightarrow B_{(\omega, x)})_{(\omega, x) \in \Omega.X}$$

to

$$(\Pi_{X \in X_{\omega}} f_{(\omega, x)})_{\omega \in \Omega}$$

In this set-theoretic model, we have that $\Pi_{X\{u\}}q(u, X)^* = u^*\Pi_X$, and furthermore that the canonical transformation is the identity at this functor. Clearly, then, this model satisfies the Beck-Chevalley condition for dependent simple products.

5.3 Interpretation

Blah blah blah.

Our interpretation satisfies the following soundness theorems.

References

- [1] Peter Dybjer. 1995. Internal type theory. In *International Workshop on Types for Proofs and Programs*. Springer, 120–134.
- [2] Martin Hofmann. 1997. Syntax and Semantics of Dependent Types. In *Semantics and Logics of Computation*. Cambridge University Press, 79–130.
- [3] Bart Jacobs. 1999. *Categorical logic and type theory*. Elsevier.
- [4] Erik Palmgren. 2019. Categories with families and first-order logic with dependent sorts. *Annals of Pure and Applied Logic* 170, 12 (2019), 102715.
- [5] Hongwei Xi. 2007. Dependent ML: An approach to practical programming with dependent types. *Journal of Functional Programming* 17, 2 (2007), 215–286.

A Set-theoretic Dependent Simple Products

A.1 Transpose Distribution

Here we show that dependent simple products in the standard set-theoretic model satisfy the transpose distribution property.

For contexts Ω, Ψ , arrows $u : \Psi \rightarrow \Omega$, semantic sorts $X \in St(\Omega)$, total objects A over Ω , and total objects B over $\Omega.X$, an arrow $f : p(X)^*(A) \rightarrow B$ has the form

$$(f_{(\omega, x)} : A_\omega \rightarrow B_{(\omega, x)})_{(\omega, x) \in \Omega.X}$$

We then have:

$$\begin{aligned} & u^* f^\# \\ & u^* (f_{(\omega, x)})_{(\omega, x) \in \Omega.X}^\# \\ & = u^* (\langle f_{(\omega, x)} \rangle_{x \in X_\omega})_{\omega \in \Omega} \\ & = \left(\langle f(u(\psi), x) \rangle_{x \in X_{u(\psi)}} \right)_{\psi \in \Psi} \\ & = (f_{(u(\psi), x)})_{(\psi, x) \in \Psi.X}^\# \\ & = (\langle u \circ p(X\{u\}), v_{X\{u\}} \rangle_X^* (f_{(\omega, x)}))_{(\omega, x) \in \Omega.X}^\# \\ & = (\langle u \circ p(X\{u\}), v_{X\{u\}} \rangle_X^* f)^\# \\ & = (q(u, X)^* f)^\# \end{aligned}$$

A.2 The Beck-Chevalley Condition

In our set-theoretic model, the canonical natural transformation of dependent simple products is an identity and therefore satisfies the Beck-Chevalley condition. We proceed to demonstrate this.

For semantic contexts Ω and sorts $X \in St(\Omega)$, we define $\Pi_X : Fam(\mathbf{Sets})_{\Omega.X} \rightarrow Fam(\mathbf{Sets})_\Omega$ as

$$\begin{aligned} \Pi_X (B_{(\omega, x)})_{(\omega, x) \in \Omega.X} &= (\Pi_{x \in X_\omega} B_{(\omega, x)})_{\omega \in \Omega} \\ \Pi_X (f_{(\omega, x)} : B_{(\omega, x)} \rightarrow C_{(\omega, x)})_{(\omega, x) \in \Omega.X} &= (\Pi_{x \in X_\omega} f_{(\omega, x)})_{\omega \in \Omega} \end{aligned}$$

We have a bijection of the following form.

$$\frac{(A_\omega)_{\omega \in \Omega} \longrightarrow (\Pi_{x \in X_\omega} B_{(\omega, x)})_{\omega \in \Omega} = \Pi_X (B_{(\omega, x)})_{(\omega, x) \in \Omega.X}}{p(X)^*(A_\omega)_{\omega \in \Omega} = (A_\omega)_{(\omega, x) \in \Omega.X} \longrightarrow (B_{(\omega, x)})_{(\omega, x) \in \Omega.X}}$$

From top to bottom, this bijection, which we'll call $(-)^b$, takes an arrow

$$(\langle f_{(\omega, x)} : A_\omega \rightarrow B_{(\omega, x)} \rangle_{x \in X_\omega})_{\omega \in \Omega}$$

to

$$(f_{(\omega, x)} : A_\omega \rightarrow B_{(\omega, x)})_{(\omega, x) \in \Omega.X}$$

From bottom to top, $(-)^b$'s inverse $(-)^\#$ takes an arrow

$$(f_{(\omega, x)} : A_\omega \rightarrow B_{(\omega, x)})_{(\omega, x) \in \Omega.X}$$

to

$$(\Pi_{x \in X_\omega} f_{(\omega, x)})_{\omega \in \Omega}$$

$(-)^b$ underlies an adjunction $p(X)^* \dashv \Pi_{\Omega.X}$. To prove this, first consider an arrow g , where

$$g = (g_\omega)_{\omega \in \Omega} : (C_\omega)_{\omega \in \Omega} \rightarrow (A_\omega)_{\omega \in \Omega}$$

and a arrow f where

$$f = (f_{(\omega, x)})_{(\omega, x) \in \Omega.X} : (A_\omega)_{(\omega, x) \in \Omega.X} \rightarrow (B_{(\omega, x)})_{(\omega, x) \in \Omega.X}$$

We then have

$$(f \circ p^*(g))^\#$$

$$\begin{aligned}
&= ((f_{(\omega,x)} \circ g_{\omega})_{(\omega,x) \in \Omega.X})^{\sharp} \\
&= (\Pi_{x \in X_{\omega}} f_{(\omega,x)} \circ g_{\omega})_{\omega \in \Omega} \\
&= ((\Pi_{x \in X_{\omega}} f_{(\omega,x)}) \circ g_{\omega})_{\omega \in \Omega} \\
&= f^{\sharp} \circ g
\end{aligned}$$

Next, consider a arrow f , where

$$f = (\langle f_{(\omega,x)} \rangle_{x \in X_{\omega}})_{\omega \in \Omega} : (A_{\omega})_{\omega \in \Omega} \rightarrow (\Pi_{x \in X_{\omega}} B_{(\omega,x)})_{\omega \in \Omega}$$

and a arrow g , where

$$g = (g_{(\omega,x)})_{(\omega,x) \in \Omega.X} : (B_{(\omega,x)})_{(\omega,x) \in \Omega.X} \rightarrow (D_{(\omega,x)})_{(\omega,x) \in \Omega.X}$$

We then have

$$\begin{aligned}
&(\Pi_X(g) \circ f)^b \\
&= ((\langle g_{(\omega,x)} \circ f_{(\omega,x)} \rangle_{x \in X_{\omega}})_{\omega \in \Omega})^b \\
&= (g_{(\omega,x)} \circ f_{(\omega,x)})_{(\omega,x) \in \Omega.X} \\
&= g \circ f^b
\end{aligned}$$

To obtain the counit of this bijection at component $(B_{(\omega,x)})_{(\omega,x) \in \Omega.X}$, writing π_x for the projection $\Pi_{x \in X_{\omega}} B_{(x,\omega)} \rightarrow B_{(x,\omega)}$, we map the identity arrow

$$(\langle \pi_x \rangle_{x \in X_{\omega}})_{\omega \in \Omega} : \Pi_X (B_{(\omega,x)})_{(\omega,x) \in \Omega} \rightarrow \Pi_X (B_{(\omega,x)})_{(\omega,x) \in \Omega}$$

through $(-)^b$, obtaining the counit ϵ as

$$(\pi_{(\omega,x')} : \Pi_{x \in X_{\omega}} B_{(\omega,x)} \rightarrow B_{(\omega,x')})_{(\omega,x') \in \Omega.X}$$

It is a arrow of type

$$\mathbf{p}(X)^* \Pi_X (B_{(\omega,x)})_{(\omega,x) \in \Omega.X} \longrightarrow (B_{(\omega,x)})_{(\omega,x) \in \Omega.X}$$

To prove the Beck-Chevalley condition, we must first concretely describe the canonical natural transformation $u^* \Pi_X \implies \Pi_{X\{u\}} \mathbf{q}(u, X)^*$. As a first step, we concretely describe the natural transformation

$$\mathbf{p}(X\{u\})^* u^* \Pi_X \xrightarrow{\cong} \mathbf{q}(u, X)^* \mathbf{p}(X)^* \Pi_X \xrightarrow{\mathbf{q}(u, X)^{\epsilon}} \mathbf{q}(u, X)^*$$

at component $(B_{(\omega,x)})_{(\omega,x) \in \Omega.X}$

$$\begin{aligned}
&\mathbf{p}(X\{u\})^* u^* \Pi_X (B_{(\omega,x)})_{(\omega,x) \in \Omega.X} \\
&= \mathbf{p}(X\{u\})^* u^* (\Pi_{x \in X_{\omega}} B_{(\omega,x)})_{\omega \in \Omega} \\
&= \mathbf{p}(X\{u\})^* \left(\Pi_{x \in X_{u(\psi)}} B_{(u(\psi), x)} \right)_{\psi \in \Psi} \\
&= \left(\Pi_{x \in X_{u(\psi)}} B_{(u(\psi), x)} \right)_{(\psi, x') \in \Psi.X\{u\}} \\
&= \left(\Pi_{x \in X_{\pi(u \circ \mathbf{p}(X\{u\}), \mathbf{v}_{X\{u\}})(\psi, x')}} B_{(\pi(u \circ \mathbf{p}(X\{u\}), \mathbf{v}_{X\{u\}})(\psi, x'), x)} \right)_{(\psi, x') \in \Psi.X\{u\}} \\
&= \langle u \circ \mathbf{p}(X\{u\}), \mathbf{v}_{X\{u\}} \rangle^* \left(\Pi_{x \in X_{\pi(\omega, x')}} B_{(\pi(\omega, x'), x)} \right)_{(\omega, x') \in \Omega.X} \\
&= \langle u \circ \mathbf{p}(X\{u\}), \mathbf{v}_{X\{u\}} \rangle^* (\Pi_{x \in X_{\omega}} B_{(\omega, x)})_{(\omega, x') \in \Omega.X}
\end{aligned}$$

$$\begin{array}{c}
\downarrow \\
\langle u \circ \mathbf{p}(X\{u\}), \mathbf{v}_{X\{u\}} \rangle^* (\pi_{(\omega, x')})_{(\omega, x') \in \Omega.X}
\end{array}$$

$$\langle u \circ \mathbf{p}(X\{u\}), \mathbf{v}_{X\{u\}} \rangle^* (B_{(\omega, x')})_{(\omega, x') \in \Omega.X}$$

The above arrow is equal to

$$(\pi_{(u(\psi), x')} : \Pi_{x \in X_{u(\psi)}} B_{(u(\psi), x)} \rightarrow B_{(u(\psi), x')})_{(\psi, x') \in \Psi.X\{u\}}$$

Transposing gives

$$\left(\prod_{x' \in X_{u(\psi)}} (\pi_{(u(\psi), x')} : (\prod_{x \in X_{u(\psi)}} B_{(u(\psi), x)}) \rightarrow B_{(u(\psi), x')}) \right)_{\psi \in \Psi}$$

This is the identity arrow on the object

$$\left(\prod_{x' \in X_{u(\psi)}} B_{(u(\psi), x')} \right)_{\psi \in \Psi}$$

which has the following “type signature”

$$\prod_{\{u\}} q(u, X)^* (B_{(\omega, x)})_{(\omega, x) \in \Omega.X} \longrightarrow u^* \prod_X (B_{(\omega, x)})_{(\omega, x) \in \Omega.X}$$

Because it is an identity, it is clearly invertible.

B Substitution lemmas

We first reproduce some standard definitions and theorems from the semantics of dependent types. These definitions and lemmas will subsequently be used to establish our own soundness proofs.

Lemma B.1. *If $\Omega, a : p, \Psi \vdash q$ and $\Omega \vdash j : p$ then $\Omega, \Psi[j/a] \vdash q[j/a]$.*

Proof. TODO □

For pre-contexts Ω, Ψ and pre-sorts q, r we define the expression $P(\Omega; p; \Psi)$ inductively by

$$\begin{aligned} P(\Omega; q; \diamond) &\stackrel{\text{def}}{=} p(\llbracket \Omega; p \rrbracket) \\ P(\Omega; q; \Psi, a : r) &\stackrel{\text{def}}{=} q(P(\Omega; q; \Psi), \llbracket \Omega, \Psi; r \rrbracket) \end{aligned}$$

The idea is that $P(\Omega; q; \Psi)$ is a morphism from $\llbracket \Omega, a : q, \Psi \rrbracket$ to $\llbracket \Omega, \Psi \rrbracket$ projecting the q part.

Now let Ω, Ψ, p, q be as before and i a pre-index. We define

$$\begin{aligned} T(\Omega; q; \diamond; i) &\stackrel{\text{def}}{=} \overline{\llbracket \Omega; i \rrbracket} \\ T(\Omega; q; \Psi, a : r; i) &\stackrel{\text{def}}{=} q(T(\Omega; q; \Psi; i), \llbracket \Omega, b : q, \Psi; r \rrbracket) \quad b \text{ fresh} \end{aligned}$$

The idea here is that $T(\Omega; q; \Psi; i)$ is a morphism from $\llbracket \Omega, \Psi[i/b] \rrbracket$ to $\llbracket \Omega, b : q, \Psi \rrbracket$ yielding $\llbracket \Omega; i \rrbracket$ at the $b : q$ position and variables otherwise.

The above ideas must be proven simultaneously in the form of weakening and substitution lemmas.

Lemma B.2. (Weakening) *Let Ω, Ψ be pre-contexts, p, q pre-sorts, i a pre-index, and b a fresh variable. Let $A \in \{p, i\}$. The expression $P(\Omega; p; \Psi)$ is defined iff $\llbracket \Omega, p : q, \Psi \rrbracket$ and $\llbracket \Omega, \Psi \rrbracket$ are defined and in this case is a morphism from the former to the latter. If $\llbracket \Omega, \Psi; A \rrbracket$ is defined then*

$$\llbracket \Omega, b : p, \Psi; A \rrbracket \simeq \llbracket \Omega, \Psi; A \rrbracket \{P(\Omega; p; \Psi)\}$$

Lemma B.3. (Substitution) *Let Ω, Ψ be pre-contexts, p, q pre-sorts, i, j pre-indices, and b a fresh variable. Let $A \in \{p, i\}$ and suppose that $\llbracket \Omega; i \rrbracket$ is defined.*

The expression $T(\Omega; p; \Psi; i)$ is defined iff $\llbracket \Omega, \Psi[i/b] \rrbracket$ and $\llbracket \Omega, b : p, \Psi \rrbracket$ are both defined and in this case is a morphism from the former to the latter. If $\llbracket \Omega, b : p, \Psi; A \rrbracket$ is defined then

$$\llbracket \Omega, \Psi[i/b]; A[i/b] \rrbracket \simeq \llbracket \Omega, b : p, \Psi; A \rrbracket \{T(\Omega; p; \Psi; i)\}$$

Proof. TODO □

Now that the above standard lemmas have been established, we state and prove our substitution lemmas.

Lemma B.4. *If $\Omega, a : p, \Psi \vdash \kappa$ and $\Omega \vdash j : p$ then $\Omega, \Psi[j/a] \vdash \kappa[j/a]$ and*

$$T(\Omega; a : p; \Psi; j)^* \llbracket \Omega, a : p, \Psi \vdash \kappa \rrbracket \cong \llbracket \Omega, \Psi[j/a] \vdash \kappa[j/a] \rrbracket$$

Proof. By induction on the proof of $\Omega, a : p, \Psi \vdash \kappa$.

Case WFK-INDABS:

We have $\kappa = \forall \mathbf{q}.\kappa'$. Our premises are $\Omega, \mathbf{a} : \mathbf{p}, \Psi \vdash \mathbf{q}$ and $\Omega, \mathbf{a} : \mathbf{p}, \Psi, \mathbf{b} : \mathbf{q} \mid \Gamma \vdash \kappa'$. Applying lemma B.1 we have $\Omega, \Psi[\mathbf{j}/\mathbf{a}] \vdash \mathbf{q}[\mathbf{j}/\mathbf{a}]$ and hence $\llbracket \Omega, \Psi[\mathbf{j}/\mathbf{a}] \vdash \mathbf{q}[\mathbf{j}/\mathbf{a}] \rrbracket \downarrow$. By lemma B.3 we have

$$\llbracket \Omega, \mathbf{a} : \mathbf{p}, \Psi \vdash \mathbf{q} \rrbracket \{T(\Omega; \mathbf{p}; \Psi; \mathbf{j})\} = \llbracket \Omega, \Psi[\mathbf{j}/\mathbf{a}] \vdash \mathbf{q}[\mathbf{j}/\mathbf{a}] \rrbracket$$

Applying the IH to the second premise we have

$$\Omega, \Psi[\mathbf{j}/\mathbf{a}], \mathbf{b} : \mathbf{q}[\mathbf{j}/\mathbf{a}] \mid \Gamma[\mathbf{j}/\mathbf{a}] \vdash \kappa'[\mathbf{j}/\mathbf{a}]$$

and

$$T(\Omega; \mathbf{p}; \Psi, \mathbf{b} : \mathbf{q}; \mathbf{j})^* \llbracket \Omega, \mathbf{a} : \mathbf{p}, \Psi, \mathbf{b} : \mathbf{q} \mid \Gamma \vdash \kappa' \rrbracket \cong \llbracket \Omega, \Psi[\mathbf{j}/\mathbf{a}], \mathbf{b} : \mathbf{q}[\mathbf{j}/\mathbf{a}] \mid \Gamma[\mathbf{j}/\mathbf{a}] \vdash \kappa'[\mathbf{j}/\mathbf{a}] \rrbracket$$

Applying WFK-INDABS we get

$$\frac{\Omega, \Psi[\mathbf{j}/\mathbf{a}] \vdash \mathbf{q}[\mathbf{j}/\mathbf{a}] \quad \Omega, \Psi[\mathbf{j}/\mathbf{a}], \mathbf{b} : \mathbf{q}[\mathbf{j}/\mathbf{a}] \mid \Gamma[\mathbf{j}/\mathbf{a}] \vdash \kappa'[\mathbf{j}/\mathbf{a}]}{\Omega, \Psi[\mathbf{j}/\mathbf{a}] \mid \Gamma[\mathbf{j}/\mathbf{a}] \vdash \forall \mathbf{q}[\mathbf{j}/\mathbf{a}].\kappa'[\mathbf{j}/\mathbf{a}]}$$

Additionally, we have

$$\begin{aligned} & T(\Omega; \mathbf{a} : \mathbf{p}; \Psi; \mathbf{j})^* \llbracket \Omega, \mathbf{a} : \mathbf{p}, \Psi \vdash \forall \mathbf{q}.\kappa' \rrbracket \\ = & \{ \text{Interpretation of KF-FORALL} \} \\ & T(\Omega; \mathbf{a} : \mathbf{p}; \Psi; \mathbf{j})^* \Pi_{\llbracket \Omega, \mathbf{a} : \mathbf{p}, \Psi \vdash \mathbf{q} \rrbracket} \llbracket \Omega, \mathbf{a} : \mathbf{p}, \Psi, \mathbf{b} : \mathbf{q} \vdash \kappa' \rrbracket \\ \cong & \{ \text{Beck-Chevalley For Dependent Simple Products, taking } T(\Omega; \mathbf{a} : \mathbf{p}; \Psi; \mathbf{j}) \text{ as } u \} \\ & \Pi_{\llbracket \Omega, \mathbf{a} : \mathbf{p}, \Psi \vdash \mathbf{q} \rrbracket \{T(\Omega; \mathbf{a} : \mathbf{p}; \Psi; \mathbf{j})\}} \mathbf{q}(T(\Omega; \mathbf{a} : \mathbf{p}; \Psi; \mathbf{j}), \llbracket \Omega, \mathbf{a} : \mathbf{p}, \Psi \vdash \mathbf{q} \rrbracket)^* \llbracket \Omega, \mathbf{a} : \mathbf{p}, \Psi, \mathbf{b} : \mathbf{q} \vdash \kappa' \rrbracket \\ = & \{ \text{Lemma B.3} \} \\ & \Pi_{\llbracket \Omega, \Psi[\mathbf{j}/\mathbf{a}] \vdash \mathbf{q}[\mathbf{j}/\mathbf{a}] \rrbracket} \mathbf{q}(T(\Omega; \mathbf{a} : \mathbf{p}; \Psi; \mathbf{j}), \llbracket \Omega, \mathbf{a} : \mathbf{p}, \Psi \vdash \mathbf{q} \rrbracket)^* \llbracket \Omega, \mathbf{a} : \mathbf{p}, \Psi, \mathbf{b} : \mathbf{q} \vdash \kappa' \rrbracket \\ = & \{ \text{Definition of } T \} \\ & \Pi_{\llbracket \Omega, \Psi[\mathbf{j}/\mathbf{a}] \vdash \mathbf{q}[\mathbf{j}/\mathbf{a}] \rrbracket} T(\Omega; \mathbf{p}; (\Psi, \mathbf{a} : \mathbf{q}); \mathbf{j})^* \llbracket \Omega, \mathbf{a} : \mathbf{p}, \Psi, \mathbf{b} : \mathbf{q} \vdash \kappa' \rrbracket \\ = & \{ \text{IH} \} \\ & \Pi_{\llbracket \Omega, \Psi[\mathbf{j}/\mathbf{a}] \vdash \mathbf{q}[\mathbf{j}/\mathbf{a}] \rrbracket} \llbracket \Omega, \Psi[\mathbf{j}/\mathbf{a}], \mathbf{b} : \mathbf{q}[\mathbf{j}/\mathbf{a}] \vdash \kappa'[\mathbf{j}/\mathbf{a}] \rrbracket \\ = & \{ \text{Interpretation of KF-FORALL} \} \\ & \llbracket \Omega, \Psi[\mathbf{j}/\mathbf{a}] \vdash \forall \mathbf{q}[\mathbf{j}/\mathbf{a}].\kappa'[\mathbf{j}/\mathbf{a}] \rrbracket \\ = & \{ \text{Definition of Index-to-Type Substitution} \} \\ & \llbracket \Omega, \Psi[\mathbf{j}/\mathbf{a}] \vdash (\forall \mathbf{q}.\kappa')[\mathbf{j}/\mathbf{a}] \rrbracket \end{aligned}$$

Other cases:

TODO

□

Lemma B.5. If $\Omega, \mathbf{a} : \mathbf{p}, \Psi \mid \Gamma \vdash \tau : \kappa$ and $\Omega \vdash \mathbf{j} : \mathbf{p}$ then $\Omega, \Psi[\mathbf{j}/\mathbf{a}] \mid \Gamma[\mathbf{j}/\mathbf{p}] \vdash \tau[\mathbf{j}/\mathbf{p}] : \kappa[\mathbf{j}/\mathbf{p}]$ and

$$T(\Omega; \mathbf{a} : \mathbf{p}; \Psi; \mathbf{j})^* \llbracket \Omega, \mathbf{a} : \mathbf{p}, \Psi \mid \Gamma \vdash \tau : \kappa \rrbracket \cong \llbracket \Omega, \Psi[\mathbf{j}/\mathbf{a}] \mid \Gamma[\mathbf{j}/\mathbf{a}] \vdash \tau[\mathbf{j}/\mathbf{a}] : \kappa[\mathbf{j}/\mathbf{a}] \rrbracket$$

Proof. By induction on the proof of $\Omega, \mathbf{a} : \mathbf{p}, \Psi \mid \Gamma \vdash \tau : \kappa$.

Case K-INDABS:

We have $\tau = \lambda \mathbf{b} : \mathbf{q}.\tau'$ and $\kappa = \forall \mathbf{q}.\kappa'$. Our premises are $\Omega, \mathbf{a} : \mathbf{p}, \Psi \vdash \mathbf{q}$ and $\Omega, \mathbf{a} : \mathbf{p}, \Psi, \mathbf{b} : \mathbf{q} \mid \Gamma \vdash \tau' : \kappa'$. Applying lemma B.1 we have $\Omega, \Psi[\mathbf{j}/\mathbf{a}] \vdash \mathbf{q}[\mathbf{j}/\mathbf{a}]$ and hence $\llbracket \Omega, \Psi[\mathbf{j}/\mathbf{a}] \vdash \mathbf{q}[\mathbf{j}/\mathbf{a}] \rrbracket \downarrow$. By lemma B.3 we have

$$\llbracket \Omega, \mathbf{a} : \mathbf{p}, \Psi \vdash \mathbf{q} \rrbracket \{T(\Omega; \mathbf{p}; \Psi; \mathbf{j})\} = \llbracket \Omega, \Psi[\mathbf{j}/\mathbf{a}] \vdash \mathbf{q}[\mathbf{j}/\mathbf{a}] \rrbracket$$

Applying the IH to the second premise we have

$$\Omega, \Psi[\mathbf{j}/\mathbf{a}], \mathbf{b} : \mathbf{q}[\mathbf{j}/\mathbf{a}] \mid \Gamma[\mathbf{j}/\mathbf{a}] \vdash \tau'[\mathbf{j}/\mathbf{a}] : \kappa'[\mathbf{j}/\mathbf{a}]$$

and

$$T(\Omega; \mathbf{p}; \Psi, \mathbf{b} : \mathbf{q}; \mathbf{j})^* \llbracket \Omega, \mathbf{a} : \mathbf{p}, \Psi, \mathbf{b} : \mathbf{q} \mid \Gamma \vdash \tau' : \kappa' \rrbracket \cong \llbracket \Omega, \Psi[\mathbf{j}/\mathbf{a}], \mathbf{b} : \mathbf{q}[\mathbf{j}/\mathbf{a}] \mid \Gamma[\mathbf{j}/\mathbf{a}] \vdash \tau'[\mathbf{j}/\mathbf{a}] : \kappa'[\mathbf{j}/\mathbf{a}] \rrbracket$$

Now, applying the K-INDABS rule, we get

$$\frac{\Omega, \Psi[j/a] \vdash q[j/a] \quad \Omega, \Psi[j/a], b : q[j/a] \mid \Gamma[j/a] \vdash \tau'[j/a] : \kappa'[j/a]}{\Omega, \Psi[j/a] \mid \Gamma[j/a] \vdash \lambda b : q[j/a].\tau'[j/a] : \forall q[j/a].\kappa'[j/a]}$$

We also have

$$\begin{aligned} & \llbracket \Omega, \Psi'[j/a] \mid \Gamma[j/a] \vdash (\lambda b : q.\tau')[j/a] : (\forall q.\kappa')[j/a] \rrbracket \\ = & \quad \{\text{Definition of substitution}\} \\ & \llbracket \Omega, \Psi'[j/a] \mid \Gamma[j/a] \vdash \lambda b : q[j/a].\tau'[j/a] : \forall q[j/a].\kappa'[j/a] \rrbracket \\ = & \quad \{\text{Interpretation of K-INDABS}\} \\ & \llbracket \Omega, \Psi'[j/a], b : q[j/a] \mid \Gamma[j/a] \vdash \tau'[j/a] : \kappa'[j/a] \rrbracket^\# \\ \cong & \quad \{\text{IH}\} \\ & (T(\Omega; p; (\Psi, b : q); j)^* \llbracket \Omega, a : p, \Psi, b : q \mid \Gamma \vdash \tau' : \kappa' \rrbracket)^\# \\ \cong & \quad \{\text{Transpose distribution}\} \\ & T(\Omega; p; \Psi; j)^* \llbracket \Omega, a : p, \Psi, b : q \mid \Gamma \vdash \tau' : \kappa' \rrbracket^\# \\ = & \quad \{\text{Interpretation of K-INDABS}\} \\ & T(\Omega; p; \Psi; j)^* \llbracket \Omega, a : p, \Psi \mid \Gamma \vdash \lambda b : q.\tau' : \forall q.\kappa' \rrbracket \end{aligned}$$

Other cases:

TODO.

□

C An Order-Theoretic Model (and Redesign)

To avoid dangling foreign keys, industrial databases often avoid removing data entries, giving their evolution over time an inflationary character. In such a situation we can interpret the sort **prop** as the ordinal 2, containing the two elements *known* and *unknown*, ordered such that *unknown* \leq *known*. A predicate $P : \text{str} \rightarrow \text{prop}$ then represents a set of elements such that membership in the set is either known (definitely a member) or undetermined (may or may not be a member).

C.1 Contexts, Sorts, Indices, and Substitution

We capture the above intuition with a CwF. Its category of contexts is **Posets**. For semantic contexts P , we define $St(\Omega)$ (read the *semantic sorts* of context Ω) as the collection of Ω -indexed families of posets. Such a family $(X_\omega)_{\omega \in \Omega}$ is a poset-indexed family rather than a set-indexed family; i.e., for $\omega_1, \omega_2 \in \Omega$ with $\omega_1 \leq \omega_2$ we have a chosen monotone injection $i_{\omega_1 \leq \omega_2} : X_{\omega_1} \rightarrow X_{\omega_2}$ such that for $\omega \in \Omega$ we have $i_{\omega \leq \omega} = id_{X_\omega}$ and for $\omega_1 \leq \omega_2 \leq \omega_3$ we have $i_{\omega_1 \leq \omega_3} = i_{\omega_2 \leq \omega_3} \circ i_{\omega_1 \leq \omega_2}$. A poset-indexed family of posets $(X_\omega)_{\omega \in \Omega}$ can itself be considered a poset whose elements are families $(M_\omega \in X_\omega)_{\omega \in \Omega}$ and $(M_\omega)_{\omega \in \Omega} \leq (M'_\omega)_{\omega \in \Omega} \Leftrightarrow (M_\omega \leq M'_\omega)$ for all $\omega \in \Omega$.

For semantic contexts Ω and all $X \in St(P)$ we define $In(P, X)$ as the collection of all Ω -indexed families $(M_\omega \in X_\omega)_{\omega \in \Omega}$ such that for $\omega_1 \leq_P \omega_2$ we have $i(M_{\omega_1}) \leq_{X_{\omega_2}} M_{\omega_2}$.

For each monotone function $f : \Omega \rightarrow \Psi$ we define a sort-level semantic substitution operator $- \{f\} : St(\Psi) \rightarrow St(\Omega)$ as

$$X \{f\} \stackrel{\text{def}}{=} (X_{f(\omega)})_{\omega \in \Omega}$$

For $\omega_1 \leq \omega_2$ we have $f(\omega_1) \leq f(\omega_2)$ and so our chosen monotone injection is

$$i_{\omega_1 \leq \omega_2} : X \{f\}_{\omega_1} \rightarrow X \{f\}_{\omega_2} \stackrel{\text{def}}{=} i_{f(\omega_1) \leq f(\omega_2)}$$

For $X \in St(\Psi)$ a index-level semantic substitution operator $- \{f\} : In(\Psi, X) \rightarrow In(\Omega, X \{f\})$.

C.2 Comprehensions

Let Ω be a semantic context and X a semantic sort in context Ω . The comprehension $\Omega.X$ is the poset of pairs (ω, x) with $\omega \in \Omega$ and $x \in X_\omega$ such that

$$(\omega_1, x_1) \leq_{\Omega.X} (\omega_2, x_2) \stackrel{\text{def}}{\Leftrightarrow} (\omega_1 \leq \omega_2) \wedge (i(x_1) \leq x_2)$$

We have a monotone function $\mathbf{p}(X) : \Omega.X \rightarrow \Omega$ defined as

$$\mathbf{p}(X)(\omega, x) \stackrel{\text{def}}{=} \omega$$

and also a semantic index term $\mathbf{v}_X \in \text{In}(\Omega.X, X\{\mathbf{p}(X)\})$ defined as

$$(\mathbf{v}_X)_{(\omega, x)} \stackrel{\text{def}}{=} x$$

Lemma C.1. *Let $f : \Omega \rightarrow \Psi$ be a monotone function, $X \in \text{St}(\Psi)$, and $M \in \text{In}(\Omega, X\{f\})$. Then there exists a unique morphism $\langle f, M \rangle_X : \Omega \rightarrow \Psi.X$ satisfying $\mathbf{p}(X) \circ \langle f, M \rangle_X = f$ and $\mathbf{v}_X\{\langle f, M \rangle_X\} = M$.*

Proof. The morphism is

$$\omega \xrightarrow{\langle f, M \rangle_X} (f(\omega), M_\omega)$$

Clearly we have $\mathbf{p}(X) \circ \langle f, M \rangle_X = f$ since

$$\omega \xrightarrow{\langle f, M \rangle_X} (f(\omega), M_\omega) \xrightarrow{\mathbf{p}(X)} f(\omega)$$

Also, we have $\mathbf{v}_X\{\langle f, M \rangle_X\} = M$ since

$$\mathbf{v}_X\{\langle f, M \rangle_X\}_\omega = (\mathbf{v}_X)_{\langle f, M \rangle_X(\omega)} = (\mathbf{v}_X)_{(f(\omega), M_\omega)} = M_\omega$$

Also, $\langle f, M \rangle_X$ is monotone since if $\omega \leq \omega'$ we have

$$\langle f, M \rangle_X(\omega) = (f(\omega), M_\omega) \leq (f(\omega'), M_{\omega'}) = \langle f, M \rangle_X(\omega')$$

(Reminder: Since $M \in X\{f\}$ we have $M_\omega \in X_{f(\omega)}$.)

TODO: prove uniqueness □

C.3 Π -sorts

This model does not have arbitrary Π -sorts. However, identifying the *pointed sorts* $PSt(\Omega)$ as those families of posets over Ω whose component posets all have \perp (minimum) elements, our model has those Π -sorts whose codomains are pointed.

Theorem C.2. *For all contexts $\Omega, X \in \text{St}(\Omega)$, $Y \in PSt(\Omega.X)$, our model supports the Π -sort $\Pi(X, Y)$.*

Proof. UNFINISHED. TODO. Given $X \in \text{St}(\Omega)$ and $Y \in \text{St}(\Omega.X)$ we define the semantic sort $\Pi(X, Y) \in \text{St}(\Omega)$ as

$$\omega \in \Omega \mapsto (Y_{(\omega, x)})_{x \in X_\omega}$$

where above, the right-hand-side is a poset-indexed family of posets considered as a poset. Indeed, given $x \leq x' \in X_\omega$ we have $(\omega, i(x)) = (\omega, x) \leq (\omega, x') = (\omega, i(x'))$ and hence a monotone injection $i : Y_{(\omega, x)} \rightarrow Y_{(\omega, x')}$. □

D An Order-Theoretic Model (and Redesign) with Multiplicities

I could compose two fibrations where strings are in the bottom layer, formulas in the middle, types on top.

I could restrict Π sorts so that only strings may occur in the domain. We're still giving coefficients to all indices, though.

D.1 Syntax

D.2 Static semantics

$TypeVars$	$\stackrel{def}{=}$	the set of all type variables
$FormulaVars$	$\stackrel{def}{=}$	the set of all index variables
$SubjectVars$	$\stackrel{def}{=}$	the set of all index variables
x, y, z	\in	$TypeVars$
a, b, P	\in	$FormulaVars$
s, t	\in	$Strings$
u, v	\in	$SubjectVars$
i, j, k (formula)	$::=$	$App_{[a:q],p}(j, k)$ (formula application)
		a (formula variable)
		$true$ (true formula)
m, n (multiplicity)	$::=$	$one \mid lone \mid some \mid set$
g, h (subject)	$::=$	s (string)
		u (subject var)
d, e, f (sort)	$::=$	str (string sort)
p, q, r (pre-signature)	$::=$	$prop$ (prop sig)
		$prf\ j$ (proof sig)
		$(u : d) \xrightarrow{m} r$ (subject-to-formula function sort)
		$(a : q) \rightarrow r$ (formula-to-formula function sort)
Ω, Ψ (pre-formula-context)	$::=$	$\Omega, a : q$ (extension)
		\diamond (empty)
Ξ, Θ (pre-sort-context)	$::=$	$\Xi, u :^m d$ (extension)
		\diamond (empty)

Figure 11. Syntax

$\overline{\Omega \vdash s : str}$	$\overline{\Omega, a : q, \Omega' \vdash a : q}$	$\frac{\Omega \vdash j : (a : q) \rightarrow r \quad \Omega \vdash k : q}{\Omega \vdash App_{[a:q],r}(j, k) : r[k/a]}$	$\overline{\Omega \vdash true : prop}$	$\overline{\Omega \vdash str}$
$\frac{\Omega \vdash q \quad \Omega, a : q \vdash r}{\Omega \vdash (a : q) \Rightarrow r}$	$\frac{\Omega \vdash j : prop}{\Omega \vdash prf\ j}$	$\overline{\diamond \vdash}$	$\frac{\Omega \vdash \quad \Omega \vdash p}{\Omega, a : p \vdash}$	

Figure 12. Sorting, sort formation, and sort context formation