Coercive subtyping in type theory[*]

Zhaohui Luo[†]
Department of Computer Science, Durham University
South Road, Durham DH1 3LE, U.K.
Tel: (+44)191-3743657, Fax: (+44)191-3742560, Email: zhaohui.luo@durham.ac.uk

Abstract

We propose and study *coercive subtyping*, a formal extension with subtyping of dependent type theories such as Martin-Löf's type theory [NPS90] and the type theory UTT [Luo94]. In this approach, subtyping with specified implicit coercions is treated as a feature at the level of the logical framework; in particular, subsumption and coercion are combined in such a way that the meaning of an object being in a supertype is given by *coercive definition rules* for the definitional equality. It is shown that this provides a conceptually simple and uniform framework to understand subtyping and coercion relations in type theories with sophisticated type structures such as inductive types and universes. The use of coercive subtyping in formal development and in reasoning about subsets of objects is discussed in the context of computer-assisted formal reasoning.

1 Introduction

A type in type theory is often intuitively thought of as a set. For example, types in Martin-Löf's type theory [ML84, NPS90] can be considered as inductively defined sets. A fundamental difference between type theory and set theory is that in the former we do not have a notion of subtype that corresponds to the notion of subset in the latter. The lack of useful subtyping mechanisms in dependent type theories with inductive types [CPM90, Dyb91, Luo94] and the associated proof development systems [MN94, C^+86, D^+91, LP92] is one of the obstacles in their applications to large-scale formal development.

Although subtyping is conceptually natural and pragmatically important, it has not been clear how useful and suitable subtyping mechanisms can be introduced into dependent type theories. Particularly, in the presence of inductive types which include types of natural numbers, lists, and trees, and types of mathematical structures such as Σ-types, it is not clear how subtyping should be introduced to reason about subsets and represent inheritance, without compromising with good proof-theoretic properties. More recently, in Aczel's project on formalising abstract algebraic theories (Galois theory), Bailey has implemented various forms of coercions in the Lego system [LP92], which are very useful in practical large-scale development of mathematical theories [Bai96].

In this paper, we present an equational formulation of *coercive subtyping*, a novel approach to introducing user-specified implicit coercions into dependent type theories that can be formulated in a logical framework. Examples of such type theories include Martin-Löf's intensional type theory [NPS90] and the type theory UTT [Luo94], which are the underlying type theories of the proof systems ALF and Lego, respectively. Our approach has the following features:

- General subsumption and specified implicit coercions are combined in such a way that the meaning of an object being in a supertype is given by *coercive definition rules* for the definitional equality. Introducing new judgement forms for *principal typing*, this gives a proof-theoretic (and direct meaning-theoretic) treatment of subtyping as implicit coercions, as compared with its possible model-theoretic semantic counterpart (cf., [BCGS91]).

- Subtyping is treated as an extension of the underlying logical framework—the meta-language in which type theories are formulated. Making essential use of a *typed* logical framework LF [Luo94] (a typed version of Martin-Löf's logical framework, see below), it gives a general extension of various intensional

[*]Proc. of CSL'96, the 1996 Annual Conference of the European Association for Computer Science Logic, Utrecht.
[†]This work is partly supported by the UK EPSRC grant on 'Subtyping, Inheritance and Reuse' (GR/K79130).

type theories with subtyping. As a formal system, the extended framework is just a simple extension of LF.

- The extended framework provides a uniform setting to understand various forms of coercions in type systems (e.g., those for structured types and universes and those found in Bailey's implementation in Lego) and some other useful syntactic forms such as type-casting.

Coercive subtyping can be seen to represent a conceptually simple but powerful approach to introducing subtyping into type theory. In the practice of computer-assisted formal reasoning, we believe that coercive subtyping provides easier and more powerful reasoning mechanisms for reasoning about subsets of objects as well as for reusing proven results in developed formal theories (cf., [Acz94]).

In the following section, we briefly introduce the logical framework and explain how to use it as a meta-language to specify type theories. In Section 3, the basic ideas of coercive subtyping are further explained and the extended logical framework is formally presented with discussions of its properties. The use of coercive subtyping is considered in Section 4. Related work and further research topics are discussed in the Conclusion.

2 The logical framework LF and formulation of type theories

The logical framework LF is a *typed* version of Martin-Löf's logical framework (see Chapter 19 of [NPS90] for a presentation of the latter). We should also point out that LF is different from the Edinburgh Logical Framework (ELF) [HHP87].

The presentation of LF and discussions on how it should be used in specifying type theories can be found in Chapter 9 of [Luo94]. The inference rules of LF are given in Figure 1, which include general rules, the rules for the kind of all types (**Type**, which represents the conceptual universe of types), and the rules for dependent product kinds of the form $(x{:}K)K'$ (kinds of functional operations). In the following, we give a brief introduction to LF and its use in specifying type theories, with discussions on several aspects with which we do not assume the familiarity of the reader.

2.1 Functional operations in LF

As in Martin-Löf's meaning explanation for his type theory, a functional operation of kind $(x{:}K)K'$ in LF can be applied to any object k of kind K to yield an object of kind $[k/x]K'$. The meaning of a functional operation is given by explaining its application results. For example, abstractions are special forms of functional operations whose meaning is essentially given by the definitional equality rule (β).

Remark In LF, the functional operations that express abstraction are of the form $[x{:}K]k$, rather than the untyped $[x]k$ as found in Martin-Löf's logical framework. In other words, we regard the meta-level functional operations as having specific domains (and codomains).[1] This feature, as we shall see below, is essential in the formulation of coercive subtyping.

The functional operations, in the form of abstraction or those introduced by declaring constants for a specified type theory (see below), are weakly extensional in the sense that the following rule is derivable by means of the (ξ) and (η) rules:

$$(Ext) \qquad \frac{\Gamma \vdash f : (x{:}K)K' \quad \Gamma \vdash g : (x{:}K)K' \quad \Gamma, x{:}K \vdash f(x) = g(x) : K'}{\Gamma \vdash f = g : (x{:}K)K'}$$

This reflects the idea that LF provides meta-level schematic and definitional mechanisms, and the fact that definitional equality for abbreviations is weakly extensional (in particular, the η-rule holds). For example, as in ordinary mathematical practice, a definition $f(g, x) = g(x)$ has the same effect as $f(g) = g$. This is in contrast with the functions of Π-types in type theory for which η-rule should not hold since it makes little sense to be a computation rule (see below).

Remark In fact, in the presence of the (β)-rule, the above rule (Ext) is equivalent to $(\eta)(\xi)$ as equational

[1]One may want to consider a more philosophical argument of whether an operation should be considered as typed. See, for example, [Bee85] for some relevant discussions.

Contexts and assumptions

$$\frac{}{\langle\rangle\ \mathbf{valid}} \qquad \frac{\Gamma \vdash K\ \mathbf{kind} \quad x \notin FV(\Gamma)}{\Gamma, x{:}K\ \mathbf{valid}} \qquad \frac{\Gamma, x{:}K, \Gamma'\ \mathbf{valid}}{\Gamma, x{:}K, \Gamma' \vdash x : K}$$

General equality rules

$$\frac{\Gamma \vdash K\ \mathbf{kind}}{\Gamma \vdash K = K} \qquad \frac{\Gamma \vdash K = K'}{\Gamma \vdash K' = K} \qquad \frac{\Gamma \vdash K = K' \quad \Gamma \vdash K' = K''}{\Gamma \vdash K = K''}$$

$$\frac{\Gamma \vdash k : K}{\Gamma \vdash k = k : K} \qquad \frac{\Gamma \vdash k = k' : K}{\Gamma \vdash k' = k : K} \qquad \frac{\Gamma \vdash k = k' : K \quad \Gamma \vdash k' = k'' : K}{\Gamma \vdash k = k'' : K}$$

Equality typing rules

$$\frac{\Gamma \vdash k : K \quad \Gamma \vdash K = K'}{\Gamma \vdash k : K'} \qquad \frac{\Gamma \vdash k = k' : K \quad \Gamma \vdash K = K'}{\Gamma \vdash k = k' : K'}$$

Substitution rules

$$\frac{\Gamma, x{:}K, \Gamma'\ \mathbf{valid} \quad \Gamma \vdash k : K}{\Gamma, [k/x]\Gamma'\ \mathbf{valid}}$$

$$\frac{\Gamma, x{:}K, \Gamma' \vdash K'\ \mathbf{kind} \quad \Gamma \vdash k : K}{\Gamma, [k/x]\Gamma' \vdash [k/x]K'\ \mathbf{kind}} \qquad \frac{\Gamma, x{:}K, \Gamma' \vdash K'\ \mathbf{kind} \quad \Gamma \vdash k = k' : K}{\Gamma, [k/x]\Gamma' \vdash [k/x]K' = [k'/x]K'}$$

$$\frac{\Gamma, x{:}K, \Gamma' \vdash k' : K' \quad \Gamma \vdash k : K}{\Gamma, [k/x]\Gamma' \vdash [k/x]k' : [k/x]K'} \qquad \frac{\Gamma, x{:}K, \Gamma' \vdash k' : K' \quad \Gamma \vdash k_1 = k_2 : K}{\Gamma, [k_1/x]\Gamma' \vdash [k_1/x]k' = [k_2/x]k' : [k_1/x]K'}$$

$$\frac{\Gamma, x{:}K, \Gamma' \vdash K' = K'' \quad \Gamma \vdash k : K}{\Gamma, [k/x]\Gamma' \vdash [k/x]K' = [k/x]K''} \qquad \frac{\Gamma, x{:}K, \Gamma' \vdash k' = k'' : K' \quad \Gamma \vdash k : K}{\Gamma, [k/x]\Gamma' \vdash [k/x]k' = [k/x]k'' : [k/x]K'}$$

The kind Type

$$\frac{\Gamma\ \mathbf{valid}}{\Gamma \vdash \mathbf{Type}\ \mathbf{kind}} \qquad \frac{\Gamma \vdash A : \mathbf{Type}}{\Gamma \vdash El(A)\ \mathbf{kind}} \qquad \frac{\Gamma \vdash A = B : \mathbf{Type}}{\Gamma \vdash El(A) = El(B)}$$

Dependent product kinds

$$\frac{\Gamma \vdash K\ \mathbf{kind} \quad \Gamma, x{:}K \vdash K'\ \mathbf{kind}}{\Gamma \vdash (x{:}K)K'\ \mathbf{kind}} \qquad \frac{\Gamma \vdash K_1 = K_2 \quad \Gamma, x{:}K_1 \vdash K_1' = K_2'}{\Gamma \vdash (x{:}K_1)K_1' = (x{:}K_2)K_2'}$$

$$\frac{\Gamma, x{:}K \vdash k : K'}{\Gamma \vdash [x{:}K]k : (x{:}K)K'} \qquad (\xi)\ \frac{\Gamma \vdash K_1 = K_2 \quad \Gamma, x{:}K_1 \vdash k_1 = k_2 : K}{\Gamma \vdash [x{:}K_1]k_1 = [x{:}K_2]k_2 : (x{:}K_1)K}$$

$$(app)\ \frac{\Gamma \vdash f : (x{:}K)K' \quad \Gamma \vdash k : K}{\Gamma \vdash f(k) : [k/x]K'} \qquad (appEq)\ \frac{\Gamma \vdash f = f' : (x{:}K)K' \quad \Gamma \vdash k_1 = k_2 : K}{\Gamma \vdash f(k_1) = f'(k_2) : [k_1/x]K'}$$

$$(\beta)\ \frac{\Gamma, x{:}K \vdash k' : K' \quad \Gamma \vdash k : K}{\Gamma \vdash ([x{:}K]k')(k) = [k/x]k' : [k/x]K'} \qquad (\eta)\ \frac{\Gamma \vdash f : (x{:}K)K' \quad x \notin FV(f)}{\Gamma \vdash [x{:}K]f(x) = f : (x{:}K)K'}$$

Figure 1: The inference rules of LF.

rules. However, it is easy to see that (Ext) cannot be used as a reduction rule because of the symmetry between f and g.

2.2 Specifying type theories in LF

In general, a specification of a type theory in the logical framework consists of a collection of declarations of new constants and a collection of computation rules. Formally, declaring a new constant k to be of kind K is to introduce the following inference rule to the specified type theory:

$$\frac{\Gamma\ \mathbf{valid}}{\Gamma \vdash k : K}$$

and, for a kind K which is either **Type** or of the form $El(A)$, one can assert computation rules by writing

$$k = k' : K \quad where \quad k_i : K_i \ (i = 1, ..., n),$$

which introduces the following rule for *computational equality*:

$$(*) \qquad \frac{\Gamma \vdash k_i : K_i \ (i=1,...,n) \quad \Gamma \vdash k : K \quad \Gamma \vdash k' : K}{\Gamma \vdash k = k' : K}$$

For example, one can introduce Π-types by declaring the following constants (in this paper, we shall often omit El to write A for $El(A)$ when no confusion may occur and may write $(K)K'$ for $(x{:}K)K'$ when x does not occur free in K'):

$$\begin{array}{rcl} \Pi & : & (A{:}\mathbf{Type})((A)\mathbf{Type})\mathbf{Type} \\ \lambda & : & (A{:}\mathbf{Type})(B{:}(A)\mathbf{Type})((x{:}A)B(x))\Pi(A,B) \\ \mathbf{E}_\Pi & : & (A{:}\mathbf{Type})(B{:}(A)\mathbf{Type})(C{:}(\Pi(A,B))\mathbf{Type}) \\ & & ((g{:}(x{:}A)B(x))C(\lambda(A,B,g)))\ (z{:}\Pi(A,B))C(z) \end{array}$$

and asserting the following computation rule:

$$\mathbf{E}_\Pi(A,B,C,f,\lambda(A,B,g)) = f(g) : C(\lambda(A,B,g)),$$

where $A : \mathbf{Type}$, $B : (A)\mathbf{Type}$, $C : (\Pi(A,B))\mathbf{Type}$, $f : (g{:}(x{:}A)B(x))C(\lambda(A,B,g))$, and $g : (x{:}A)B(x)$. Then, the usual application operator can be defined as

$$\begin{array}{rcl} \mathbf{app} & =_{\mathrm{df}} & [A{:}\mathbf{Type}][B{:}(A)\mathbf{Type}][F{:}\Pi(A,B)][a{:}A] \\ & & \mathbf{E}_\Pi(A,B,[G{:}\Pi(A,B)]B(a),[g{:}(x{:}A)B(x)]g(a),F). \end{array}$$

and we have $\mathbf{app}(A,B,\lambda(A,B,f)) = f$. However, the η-rule does not hold: $\lambda(A,B,\mathbf{app}(A,B,F)) \neq F$ when $F{:}\Pi(A,B)$ is a variable. In other words, functions of a Π-type are not weakly extensional. Note that a Π-type $\Pi(A,B)$ is different from the kind $(x{:}A)B(x)$ (or more formally, $(x{:}El(A))El(B(x))$), whose objects are weakly extensional functional operations (see above).

Notation For types A and B, when x does not occur free in B, we shall write $A \to B$ for $\Pi(A,[x{:}A]B)$, the type of functions from A to B. Similarly, $A \to B$ is different from the kind $(A)B$.

One can similarly specify various types or type constructors, including inductive types, predicative or impredicative type universes, etc. [CPM90, Dyb91, Luo94]. The following gives a brief description of how inductive schemata can be introduced for introducing a large class of inductive data types (for formal details, see Chapter 9 of [Luo94]), which we shall use later to generalise subtyping for parameterised inductive types.

First, we say that a kind is *small* if it is either of the form $El(A)$ or of the form $(x{:}K_1)K_2$ with K_1 and K_2 small kinds. An *inductive schema* Θ with respect to a type variable X is of one of the following forms:

1. $\Theta \equiv X$, or
2. $\Theta \equiv (x{:}K)\Theta_0$, where K is a small kind and Θ_0 is an inductive schema, or

3. $\Theta \equiv (\Phi)\Theta_0$, where Θ_0 is an inductive schema and Φ is of the form $(x_1{:}K_1)...(x_m{:}K_m)X$, with $m \geq 0$ and K_i being small kinds in which X does not occur free.

An inductive type $\mathcal{M}[\bar{\Theta}]$ is generated by a sequence of inductive schemata $\bar{\Theta} \equiv \Theta_1, ..., \Theta_n$ (with respect to the same type variable X which becomes bound in $\mathcal{M}[\bar{\Theta}]$). Associated with the inductive type are its introduction operators $\iota_i[\bar{\Theta}]$ $(i = 1, ..., n)$ and an elimination operator $\mathbf{E}[\bar{\Theta}]$ with appropriate computation rules.

For example, a type Nat of natural numbers can be defined as $Nat =_{\mathrm{df}} \mathcal{M}[\bar{\Theta}_0]$, where $\bar{\Theta}_0 \equiv X, (X)X$. The associated introduction operators are $0 =_{\mathrm{df}} \iota_1[\bar{\Theta}_0] : Nat$ and $succ =_{\mathrm{df}} \iota_2[\bar{\Theta}_0] : (Nat)Nat$, and the elimination operator is $\mathbf{E}_{Nat} =_{\mathrm{df}} \mathbf{E}[\bar{\Theta}_0] : (C{:}(Nat)\mathbf{Type})(c{:}C(0))(f{:}(x{:}Nat)(C(x))C(succ(x)))(n{:}Nat)C(n)$, with the computation rules $\mathbf{E}_{Nat}(C, c, f, 0) = c$ and $\mathbf{E}_{Nat}(C, c, f, succ(x)) = f(x, \mathbf{E}_{Nat}(C, c, f, x))$.

Note that, as discussed in [Luo94], we can introduce *different* inductive types with isomorphic structure. For instance, a type $Even$ isomorphic to Nat can be introduced as $Even =_{\mathrm{df}} \mathcal{M}'[\bar{\Theta}_0]$, which is just another copy of Nat but with a different name and with different names for its introduction and elimination operators (e.g., $e_0 =_{\mathrm{df}} \iota'_1[\bar{\Theta}_0]$, $e_1 =_{\mathrm{df}} \iota'_2[\bar{\Theta}_0]$, and $\mathbf{E}_{Even} =_{\mathrm{df}} \mathbf{E}'[\bar{\Theta}_0]$). As we shall see below, with coercive subtyping, $Even$ can be regarded as a subtype of Nat—the type of even numbers.

The type theory UTT [Luo94] consists of an impredicative type universe of propositions, inductive data types (and inductive families, not covered above), and predicative type universes. UTT has nice meta-theoretic properties such as Church-Rosser, Subject Reduction, and Strong Normalisation [Gog94]. Implemented in the Lego proof development system, UTT has been applied to program specification and verification (eg, [Bur93, BM92]), data refinement [Luo93], and formalisation of mathematics [Pol94].

2.3 Definitional equality and computational equality

We use LF seriously as a meta-level language (see Section 9.1.2 of [Luo94] for a discussion). Along the same line, we make a distinction between the notion of definitional equality (abbreviational equality, reflected as $\beta\eta$-equality for functional operations in LF) and that of computational equality introduced by asserting computation rules when specifying a type theory.

This distinction is also reflected in our restriction above that new computation rules $(*)$ can only be asserted between two types or two objects of a type, but not between two functional operations with a dependent product kind. For instance, under this restriction, one cannot declare a constant app : $(A{:}\mathbf{Type})(B{:}(A)\mathbf{Type})(\Pi(A, B))(x{:}A)B(x)$ and directly assert $app(A, B, \lambda(A, B, f)) = f$ as a computational equality, though asserting $app(A, B, \lambda(A, B, f), x) = f(x)$ is legitimate. However, taking this latter as a reduction rule (from the left to the right) is problematic since weak extensionality does not hold for the reduction relation and the Church-Rosser property would fail to hold. Another possibility, pointed out to me by Healfdene Goguen, is to declare app : $(A{:}\mathbf{Type})(B{:}(A)\mathbf{Type})(x{:}A)(\Pi(A, B))B(x)$ and assert $app(A, B, x, \lambda(A, B, f)) = f(x)$.

As we shall see below, in the coercive subtyping approach, coercions between types introduce new definitional equalities since implicit coercions are essentially an apparatus for abbreviation. In other words, coercive subtyping is regarded as abbreviational mechanisms similar to definitional mechanisms. The choice of considering coercive subtyping in the meta-level logical framework, rather than in some object-level type systems (say Pure Type Systems such as the Calculus of Constructions) is important and beneficial.

3 Coercive subtyping

In this section, we first introduce the basic ideas and give informal meaning explanations of the judgements in the extended framework with coercive subtyping. Then, a formal presentation is given, followed by discussion of its properties.

3.1 Basic ideas and informal explanation

Introducing subtyping into dependent type theories with inductive types raises new issues that have not been considered before in research on subtyping for simpler type systems. We first consider the basic problems and introduce the idea of coercive subtyping.

3.1.1 Subtyping between inductive types: the problem

An inductive type can be understood as consisting of its canonical objects (values of the type). For instance, the type *Nat* of natural numbers consists of 0 and the successors, and any natural number is regarded as a representation of a canonical natural number, to which it can be computed. If A is a subtype of B, then every object of type A is (regarded as) an object of type B. One of the basic considerations in studying subtyping for inductive types is to look for a suitable approach with which the understanding of types based on the notions of canonical object and computation still applies.

The traditional approaches based on direct overloading (eg, overloading λ-terms to stand for objects of different function types) do not generalise to inductive types. A natural consideration might be to form a subtype A of type B by selecting some (canonical) objects from B, which are regarded as the (canonical) objects of A. For example, we may introduce a subtype *Even* of *Nat* by declaring its canonical objects to be those natural numbers which are either 0 or of the form $succ(succ(e))$, where e is of type *Even*. However, in such a setting, type-checking is difficult (and in general undecidable). It is not clear how one may introduce suitable restrictions on subtype formation to ensure decidable type-checking. One suggestion that has been made in the literature is to specify a subtype by declaring its constructors to be a subset of the constructors of an existing supertype [Coq92], but this would exclude even the example of *Even* and other interesting applications of subtyping such as inheritance between mathematical theories [Acz94]. A related problem is that, in the presence of subtyping, the usual elimination rules for inductive types become inadequate since they do not take into the account (the forms of) the canonical objects in the subtypes. For instance, subtyping between two Σ-types as found in the Extended Calculus of Constructions (ECC) [Luo89, Luo90] is not quite compatible with the general elimination rules as found in Martin-Löf's type theory and UTT. A simple combination would lead to a system for which the subject reduction property fails to hold (see Section 4.3).

3.1.2 Coercive subtyping: informal explanation

There are two basic ideas on which coercive subtyping is based: implicit coercion and coercive rules for definitional equality.

Implicit coercions

With coercive subtyping, for any $A < B$, there is a unique coercion κ from A to B and every object a of type A can be regarded as the object $\kappa(a)$ of type B, that is its image under the coercion. Coercions are implicit in the sense that the subsumption rule applies in general (cf., the last two rules in Figure 2). Subtyping relations between basic inductive types are specified by defining coercion functions, which should be definable in the type theory. For instance, the inductive type *Even* with constructors e_0 and e_1, as specified in Section 2.2, can be introduced as an (inductive) subtype of *Nat* (i.e., $Even < Nat$), by giving the coercion defined by means of structural recursion over *Even* as follows: $\kappa(e_0) = 0$ and $\kappa(e_1(e)) = succ(succ(\kappa(e)))$. These basic coercions will then be generalised to other (structured) types.

Judgements and their meaning explanation

When subtyping is introduced, a type theory does not have the property of unique typing (type uniqueness) anymore. Instead, a notion of *principal typing* is the best that one can expect (see, eg, [Luo90] for a definition of the notion of principal type for ECC). Intuitively, K is a *principal kind* of object k if and only if k is of kind K and, for any kind K', k is of kind K' if and only if K is a subkind of K'. Note that being a principal kind is more than just being a minimal or minimum kind, and in general, a principal kind of an object is unique upto the computational equality.

In a type theory specified in LF with coercive subtyping, we introduce explicit judgement forms to represent principal kinding/typing. Hence, besides judgements for context validity, we have the following judgement forms:

- $k : K$ asserts that K is the *principal kind* of k. When $K \equiv El(A)$, it asserts that A is the *principal type* of k, which means that k computes to a canonical object of type A.
- $k :: K$ asserts that k is of kind K, which means that either $k : K$ or $k : K'$ for some $K' < K$.

- $K = K'$ asserts that K and K' are equal kinds, which means that $k : K$ if and only if $k : K'$.
- $K < K'$ asserts that K is a *proper subkind* of K', which means that for a unique $\kappa : (K)K'$, $\kappa(k) : K'$ for every $k : K$.
- $k = k' : K$ asserts that k and k' are equal objects with principal kind K. When $K \equiv El(A)$, it means that k and k' are computationally equal and compute to the same canonical object of type A.
- $k = k' :: K$ asserts that k and k' are equal objects of kind K, which means that either $k = k' : K$ or $k = k' : K'$ for some $K' < K$.

K is a *subkind* of K', notation $K \leq K'$, if and only if either $K = K'$ or $K < K'$.

The meaning explanations of judgements as sketched above coherently extends the meaning theory of Martin-Löf for his intensional type theory (cf., [NPS90]) to subtyping. Our formulation of coercive subtyping is strongly guided by the meaning explanation.

Coercive definition rules

To give meanings for the objects of a subkind being in a superkind, we shall introduce *coercive definition rules*. One of the key points is to design suitable coercive definition rules so that the resulting type theory reflects the intended meanings of judgements and has nice proof-theoretic properties. The following is the basic coercive definition rule, where κ is the (unique) coercion from K_0 to K:

$$\frac{f : (x{:}K)K' \quad k_0 : K_0 \quad K_0 < K}{f(k_0) = f(\kappa(k_0)) : [k_0/x]K'}$$

Intuitively, when a functional operation with domain K is applied to an object k_0 whose principal kind is a proper subkind of K, $f(k_0)$ is definitionally equal to $f(\kappa(k_0))$, that is, the application result is the same as that of applying f to the image of k_0 under the intended coercion. Note that, as to be made precise below, we do not have any coercions (including the indentity function) from a type to itself (K_0 is a *proper* subkind of K in the above rule), for otherwise regarding the above rule (with a premise $K_0 \leq K$) as a reduction rule could lead to infinite reduction sequences.

Note that the above rule conforms to Martin-Löf's meaning explanation for functional operations f, since in the presence of subkinding, the domain of a functional operation also has as objects those of any of its subkinds and the coercive rules explain the effect of applying the functional operation to an object of a subkind of its domain. Note that we do *not* have the following stronger rule, from which our coercive subkinding rule can be derived:

$$(**) \qquad \frac{k :: K \quad K < K'}{k = \kappa(k) :: K'}$$

This rule for dependent functional operations is not appropriate as meaning-giving, since the meaning of a functional operation is not given by its canonical form, but rather by its behaviour when applied to its arguments (this is exactly captured by our coercive equality rules). Furthermore, we note that our coercive definition rules preserve principal kinds when regarded as reduction rules from the left to the right (i.e., the subject reduction property), while the above stronger rule (**) does not. Furthermore, taking the above rule (**) as a reduction rule could also lead to infinite reduction sequences.

3.2 Coercive subtyping: a formal presentation

In this section, we give a formal presentation of type theories with coercive subtyping. We consider how to extend any type theory T specified in LF with coercive subtyping. Examples of such type theories include Martin-Löf's intensional type theory, UTT, and many others.

Let T be any type theory specified in LF (without subkinding). Let $\mathcal{C}$ be a set of pairs of types (ie, objects of kind **Type**) such that

1. for any $(A, B) \in \mathcal{C}$, A and B are distinct, ie, A is not computationally equal to B;
2. the set $\mathcal{C}$ as a relation represents a directed graph which is a forest, that is there is a unique path (up to computational equality) between any two types related by $\mathcal{C}$ (and hence the graph is acyclic);

New rules for application

$$\frac{\Gamma \vdash f \ : \ (x{:}K)K' \quad \Gamma \vdash k \ :: \ K}{\Gamma \vdash f(k) \ : \ [k/x]K'} \quad \frac{\Gamma \vdash f = f' \ : \ (x{:}K)K' \quad \Gamma \vdash k_1 = k_2 \ :: \ K}{\Gamma \vdash f(k_1) = f'(k_2) \ : \ [k_1/x]K'}$$

Coercive definition rule

$$\frac{\Gamma \vdash f \ : \ (x{:}K)K' \quad \Gamma \vdash k_0 \ : \ K_0 \quad \Gamma \vdash K_0 < K}{\Gamma \vdash f(k_0) = f(\kappa[K_0, K](k_0)) \ : \ [k_0/x]K'}$$

Subkinding

$$\frac{\Gamma \ \mathbf{valid} \quad (A, B) \in \mathcal{C}}{\Gamma \vdash El(A) < El(B)} \quad \frac{\Gamma \vdash K < K' \quad \Gamma \vdash K' \leq K''}{\Gamma \vdash K < K''} \quad \frac{\Gamma \vdash K \leq K' \quad \Gamma \vdash K' < K''}{\Gamma \vdash K < K''}$$

$$\frac{\Gamma, x{:}K, \Gamma' \vdash K_1 < K_2 \quad \Gamma \vdash k \ : \ K}{\Gamma, [k/x]\Gamma' \vdash [k/x]K_1 < [k/x]K_2} \quad \frac{\Gamma, x{:}K, \Gamma' \vdash K_1 < K_2 \quad \Gamma \vdash k = k' \ : \ K}{\Gamma, [k/x]\Gamma' \vdash [k/x]K_1 < [k'/x]K_2}$$

$$\frac{\Gamma \vdash K_1' < K_1 \quad \Gamma, x{:}K_1' \vdash K_2 \leq K_2' \quad \Gamma, x{:}K_1 \vdash K_2 \ \mathbf{kind}}{\Gamma \vdash (x{:}K_1)K_2 < (x{:}K_1')K_2'}$$

$$\frac{\Gamma \vdash K_1' \leq K_1 \quad \Gamma, x{:}K_1' \vdash K_2 < K_2' \quad \Gamma, x{:}K_1 \vdash K_2 \ \mathbf{kind}}{\Gamma \vdash (x{:}K_1)K_2 < (x{:}K_1')K_2'}$$

Lifting and subsumption

$$\frac{\Gamma \vdash k \ : \ K}{\Gamma \vdash k \ :: \ K} \quad \frac{\Gamma \vdash k = k' \ : \ K}{\Gamma \vdash k = k' \ :: \ K}$$

$$\frac{\Gamma \vdash k \ :: \ K \quad \Gamma \vdash K < K'}{\Gamma \vdash k \ :: \ K'} \quad \frac{\Gamma \vdash k = k' \ :: \ K \quad \Gamma \vdash K < K'}{\Gamma \vdash k = k' \ :: \ K'}$$

Figure 2: New inference rules in T[$\mathcal{C}$].

3. each $(A, B) \in \mathcal{C}$ is associated with a unique coercion function $\kappa[A, B]$ (up to computational equality) which is of kind $(A)B$ in T.

The type theory T[$\mathcal{C}$], the extension of T with coercive subtyping with respect to the basic subtyping relation $\mathcal{C}$, is the type system obtained by extending T with new judgement forms $\Gamma \vdash K < K'$, $\Gamma \vdash k \ :: \ K$ and $\Gamma \vdash k = k' \ :: \ K$, and the new inference rules given in Figure 2.

Here, we give informal explanations and some remarks. We first emphasise that the judgement $\Gamma \vdash k \ : \ K$ means that k is an object with *principal* kind K, while $\Gamma \vdash k \ :: \ K$ means that k is of kind K (see informal meaning explanations in the above section). This reflects the fact that, when no subtyping is present, the notions of typing and principal typing coincide.

The basic subtyping relation $\mathcal{C}$ is between *types*, not arbitrary kinds. However, it is not restricted to constant types (such as $Even$ and Nat) but can be between structured types such as Σ-types representing abstract mathematical theories (such as those of rings and groups) possibly with the intended coercions specified by the user of a proof system. The basic coercions are extended to dependent product kinds, giving (unique) coercion operations $\kappa[K, K']$ between every K and K' such that $K < K'$, as follows.

Definition 3.1 (coercion reconstruction) *The following defines a (unique) coercion operation $\kappa_\Gamma[K, K']$ of kind $(K)K'$ (in context Γ) for every K and K' such that $\Gamma \vdash K < K'$.*

1. *For any valid context Γ and $(A, B) \in \mathcal{C}$, $\kappa_\Gamma[A, B]$ is the given coercion function.*

2. *If $\Gamma \vdash K < K'$ and $\Gamma \vdash K' < K''$, $\kappa_\Gamma[K, K''] =_{\mathrm{df}} [x{:}K]\kappa_\Gamma[K', K''](\kappa_\Gamma[K, K'](x))$.*

3. *If* $\Gamma \vdash K_1' < K_1$ *and* $\Gamma, x{:}K_1' \vdash K_2 = K_2'$, *then* $\kappa_\Gamma[(x{:}K_1)K_2, (x{:}K_1')K_2'] =_{\mathrm{df}} [f{:}(x{:}K_1)K_2][x{:}K_1'] f(\kappa_\Gamma[K_1', K_1](x))$.
4. *If* $\Gamma \vdash K_1' = K_1$ *and* $\Gamma, x{:}K_1' \vdash K_2 < K_2'$, *then* $\kappa_\Gamma[(x{:}K_1)K_2, (x{:}K_1')K_2'] =_{\mathrm{df}} [f{:}(x{:}K_1)K_2][x{:}K_1'] \kappa_{\Gamma,x:K_1'}[K_2, K_2'](f(x))$.
5. *If* $\Gamma \vdash K_1' < K_1$ *and* $\Gamma, x{:}K_1' \vdash K_2 < K_2'$, *then* $\kappa_\Gamma[(x{:}K_1)K_2, (x{:}K_1')K_2'] =_{\mathrm{df}} [f{:}(x{:}K_1)K_2][x{:}K_1'] \kappa_{\Gamma,x:K_1'}[K_2, K_2'](f(\kappa_\Gamma[K_1', K_1](x)))$.

For $\kappa_\Gamma[K, K']$, *we shall often omit* Γ, K *and* K' *if they are clear from the context.*

For example, for $Even < Nat$ with coercion κ, we shall have $(Nat)\mathbf{Type} < (Even)\mathbf{Type}$ with coercion $\kappa(f) =_{\mathrm{df}} [x{:}Even]f(\kappa(x))$.

The new rules for application in Figure 2, together with the subsumption rules, make the existing application rules (*app*) and (*appEq*) in Figure 1 redundant. They allow a functional operation $f : (x{:}K)K'$ to be applied to any object k of kind K (ie, $k :: K$), whose principal kind is either K or a proper subkind of K. And, as explained above, the coercive definition rule gives the meaning to the object formed by such an application.

It is obvious that T is a subsystem of T[$\mathcal{C}$]; for instance, if $\Gamma \vdash^{T} k : K$, then $\Gamma \vdash^{T[\mathcal{C}]} k : K$. We note that when the basic subtyping relation $\mathcal{C}$ is empty, we have: $\Gamma \vdash^{T[\emptyset]} k : K$ iff $\Gamma \vdash^{T[\emptyset]} k :: K$ iff $\Gamma \vdash^{T} k : K$; in this sense, T[$\emptyset$] is just the original type theory T. When $\mathcal{C}$ is not empty, there are more typable terms in T[$\mathcal{C}$] than in T.

Typed reduction

The above gives an equational presentation of the extension of type theory with coercive subtyping. The intended notion of computation for the extended type theory is the notion of *typed reduction*. The typed reduction relation $\rhd$ is the reflexive and transitive closure generated from the following rules and the rules for computational equality (($*$) rules in Section 2) taken as reduction rules from the left to the right.

$$\frac{\Gamma \vdash K_1 \rhd K_2 \quad \Gamma, x{:}K_1 \vdash K_1' \rhd K_2'}{\Gamma \vdash (x{:}K_1)K_1' \rhd (x{:}K_2)K_2'} \qquad (\rhd_\xi)\ \frac{\Gamma \vdash K_1 \rhd K_2 \quad \Gamma, x{:}K_1 \vdash k_1 \rhd k_2 : K}{\Gamma \vdash [x{:}K_1]k_1 \rhd [x{:}K_2]k_2 : (x{:}K_1)K}$$

$$(\rhd_\beta)\ \frac{\Gamma, x{:}K \vdash k' : K' \quad \Gamma \vdash k : K}{\Gamma \vdash ([x{:}K]k')(k) \rhd [k/x]k' : [k/x]K'} \qquad (\rhd_\eta)\ \frac{\Gamma \vdash f : (x{:}K)K' \quad x \notin FV(f)}{\Gamma \vdash [x{:}K]f(x) \rhd f : (x{:}K)K'}$$

$$\frac{\Gamma \vdash f \rhd f' : (x{:}K)K' \quad \Gamma \vdash k_1 \rhd k_2 :: K}{\Gamma \vdash f(k_1) \rhd f'(k_2) : [k_1/x]K'} \qquad (\rhd_c)\ \frac{\Gamma \vdash f : (x{:}K)K' \quad \Gamma \vdash k_0 : K_0 \quad \Gamma \vdash K_0 < K}{\Gamma \vdash f(k_0) \rhd f(\kappa[K_0, K](k_0)) : [k_0/x]K'}$$

When the type theory T has nice proof-theoretic properties (e.g., when T is Martin-Löf's intensional type theory or UTT), the typed reduction for T[$\mathcal{C}$] is expected to have good properties as well. There is a coercion completion mapping δ from T[$\mathcal{C}$] to T that inserts all of the appropriate coercions and we have $M \rhd_c \delta(M)$. It is then clear that the extended type system is weakly normalising if type theory T is. Typed reduction also preserves principal kinding (and principal typing).

It is worth remarking that, besides the coercive definition rule, the (β) and (η) rules for definitional equality and those for computational equality are all governed by principal kinding requirements, which prevent unintended equalities or reductions. For instance, it is well known that, in the presence of subtyping and untyped $\beta\eta$-reductions, Church-Rosser would fail even for well-typed terms. Typical examples include the terms such as $t \equiv [x{:}Even]([y{:}Nat]y)(x)$ with $Even < Nat$ (and $Even \neq Nat$). As illustrated in Figure 3, with untyped reduction, t β-reduces to $t_1 \equiv [x{:}Even]x$ and η-reduces to $t_2 \equiv [y{:}Nat]y$. However, in our system, we only have

$$t = [x{:}Even]([y{:}Nat]y)(\kappa(x)) = [x{:}Even]\kappa(x) = \kappa,$$

where $\kappa : (Even)Nat$ is the coercion from $Even$ to Nat and, in particular, t is not equal to $[x{:}Even]x$ or $[y{:}Nat]y$, which have principal kinds $(Even)Even$ and $(Nat)Nat$, respectively; they are different from $(Even)Nat$, the principal kind of t. This example also shows that untyped reduction does not preserve

$$\begin{tikzcd}
 & {[y{:}Nat]y} \arrow[rd, "\kappa_1"] & & \\
{[x{:}Even]([y{:}Nat]y)(x)} \arrow[ru, "\text{untyped } \beta"] \arrow[rd, "\text{untyped } \eta"'] \arrow[r, "\rhd_c"] & {[x{:}Even]([y{:}Nat]y)(\kappa(x))} \arrow[r, "\rhd_\beta"] & {[x{:}Even]\kappa(x)} \arrow[r, "\rhd_\eta"] & \kappa \\
 & {[x{:}Even]x} \arrow[ru, "\kappa_2"'] & &
\end{tikzcd}$$

Figure 3: Reduction behaviour for $\beta\eta$: an example.

principal kinding. However, if we apply the coercions κ_1 from $(Even)Even$ to $(Even)Nat$ and κ_2 from $(Nat)Nat$ to $(Even)Nat$, to the incompatible terms t_1 and t_2, respectively, both result in terms that reduce to the term $[x{:}Even]\kappa(x)$ under typed reduction; in other words, the intended reduction result is recovered.

It is interesting to note that certain subtyping relations between types can introduce forms of self-application. An example of this is to consider some types A and B and the basic subtyping relations such that $A < (A \to A) < (B \to B) < B$, where $B \not\leq A$. (Note that the arrow here is the constructor for function *types*, not functional kinds. For any kind K, it is impossible to have in our system, for example, $K < (K)K$ or $(K)K < K$.) Then, for any $x{:}A$, $\mathbf{app}(A, [y{:}A]A, x, x)$ is of type A, because the first occurrence of x is of type $A \to A$ and the second of type A. Now, consider the following term M, which stands for $\lambda x{:}A.xx$ in the usual λ-notation:

$$M \equiv \lambda(A, [y{:}A]A, [x{:}A]\mathbf{app}(A, [y{:}A]A, x, x)),$$

we have that M is of type $A \to A$ and $Y \equiv \mathbf{app}(B, [y{:}B]B, M, M)$ is of type B. It is obvious that the term Y computes to itself under untyped reduction and hence has an infinite reduction sequence. Under typed reduction, the term Y cannot compute to itself, because the principal kinding requirement is not satisfied for the reduction to happen; instead, Y must first compute to $\mathbf{app}(B, [y{:}B]B, \kappa[A \to A, B \to B](M), \kappa[A \to A, B](M))$, with the appropriate coercions inserted.

We conjecture that $\mathrm{T}[\mathcal{C}]$ satisfies the strong normalisation property with respect to the typed reduction, if the original type theory T does.

4 Reasoning with coercive subtyping and applications

In this section, it is shown that the simple framework of coercive subtyping provides a powerful setting in reasoning about subsets of objects, representing inheritance between formal theories in proof development, understanding various forms of coercions as implemented in proof systems, and understanding some issues in implementing type theories such as universe inclusion and type-casting.

4.1 Subtyping between basic inductive types

Besides our earlier example of even and natural numbers, one meets many other applications where subtyping between basic inductive types is useful in practice. For instance, in a formalisation of the syntax of an imperative programming language, the primitive statements form an inductive type which is a subtype of the type of programs.

Coercive subtyping represents such subtyping relations in a natural way and provides new power for expressing and reasoning about subsets in a more concise way. The following gives a simple illustrative example how the induction principles of subtypes can be used to reason about corresponding objects in a supertype. For instance, the induction principle for $Even$ can be used to show that every even natural number of type Nat has certain properties. Let us consider the proof of showing the simple property that 'none of the even natural number is equal to one'. Let D be the predicate defined over natural numbers as $D(n) \equiv (n \neq_{Nat} succ(0))$, where $=_{Nat}$ is a propositional equality over natural numbers. First, without subtyping, we would have to define a predicate $PEven$ over Nat and the above statement can be expressed as the proposition $\forall n{:}Nat.\ PEven(n) \supset D(n)$. With coercive subtyping, it can be expressed more concisely as $\forall e{:}Even.D(e)$, which can then be proved by means of the induction principle expressed by $\mathbf{E}_{Even}$. In such a proof, we only have to consider the base case $D(e_0) = D(0)$ and the induction step $(e{:}Even)(D(e))D(e_1(e)) = (e{:}Even)(D(\kappa(e)))D(succ(succ(\kappa(e))))$, which only concern with the even natural numbers. Note that, without subtyping, to use the induction principle for Nat to prove the statement involving $PEven$, we would have to consider the cases for the other natural numbers (the odd ones) as well.

Such a power in reasoning about subsets of objects through coercive subtyping comes from an effective use of computational functions (coercions) to represent subsets. This represents a typical advantage of type theory being also a computational language, as compared with traditional logical systems.

We note that not every predicate P over Nat can be represented as a function from Nat to Nat in the sense that the images of the function are exactly those natural numbers that satisfy P. In general, not every subset of a type can be represented as a subtype as above to assist inductive reasoning. For every predicate P over a type A, the Σ-type $\Sigma(A,P)$ can be regarded as a subtype of A with the first projection as coercion. However, for example, the subtype $\Sigma(Nat, PEven)$ does not help us reason about even numbers inductively as we have shown above.

4.2 Subtyping between parameterised inductive types

Another class of subtyping relations are between parameterised inductive types. For instance, if $A \leq B$, $List(A) \leq List(B)$. Similarly, since Σ-types are a special kind of inductive types, we naturally expect that, e.g., if $A \leq A'$ and $B \leq B'$, then $A \times B \leq A' \times B'$. In a proof development system, it is natural to assume such extensions of the basic subtyping relation, unless requested otherwise by the user (e.g., by giving a specific coercion between two Σ-types). In [Luo92] we have suggested how subtyping can in general be extended to inductive types on the basis of the inductive schemata. Here, we extend the idea to coercive subtyping, which gives a systematic extension of the user-specified (basic) subtyping to structured inductive types.

Consider two inductive types with the same number of constructors (introduction operators): $A \equiv \mathcal{M}[\bar{\Theta}]$ and $A' \equiv \mathcal{M}[\bar{\Theta}']$. Intuitively, A can be regarded as a subtype of A', notation $A \leq_p A'$, if the corresponding constructors have the same number of parameters and any parameter kind of a constructor of A is a subkind of that of the corresponding parameter for A'. To give a precise definition of this, we first define a binary relation $\leq_s$ between inductive schemata as the smallest relation generated by the following rules, where Θ, Θ', Θ_0, and Θ_0' stand for arbitrary inductive schemata with respect to type variable X:

$$\frac{}{\Theta \leq_s \Theta} \qquad \frac{K \leq K' \quad \Theta_0 \leq_s \Theta_0'}{(x{:}K)\Theta_0 \leq_s (x{:}K')\Theta_0'} \qquad \frac{\Phi \leq \Phi' \quad \Theta_0 \leq_s \Theta_0'}{(\Phi)\Theta_0 \leq_s (\Phi')\Theta_0'}$$

Let $A \equiv \mathcal{M}[\bar{\Theta}]$ and $A' \equiv \mathcal{M}[\bar{\Theta}']$, where $\bar{\Theta} \equiv \Theta_1, ..., \Theta_n$ and $\bar{\Theta} \equiv \Theta_1', ..., \Theta_n'$. Then, we define: $A \leq_p A'$ if and only if $\Theta_i \leq_s \Theta_i'$ for $i = 1, ..., n$, and $A <_p A'$ if and only if $A \leq_p A'$ and $A \neq A'$. The implicit coercions for the subtyping relation $<_p$ can be defined straightforwardly and we omit the details here. The subscript in the relations $\leq_p$ and $<_p$ is to indicate that this is not necessarily a subtyping relation used in practice because, for example, user-defined implicit coercions may (and should) have higher priority and therefore override such a 'default' relation between inductive types. The following are some examples.

- Lists: $List =_{\mathrm{df}} [A{:}\mathbf{Type}]\ \mathcal{M}[X, (A)(X)X]$. Then, we have $List(A) \leq_p List(B)$ if $A \leq B$.
- Π-types: $\Pi =_{\mathrm{df}} [A{:}\mathbf{Type}][B{:}(A)\mathbf{Type}]\ \mathcal{M}[((x{:}A)B(x))X]$. We have $\Pi(A,B) \leq_p \Pi(A',B')$ if $(x{:}A)B(x) \leq (x{:}A')B'(x)$, that is, if $A' \leq A$ and $B(x) \leq B'(x)$.
- Σ-types: $\Sigma =_{\mathrm{df}} [A{:}\mathbf{Type}][B{:}(A)\mathbf{Type}]\ \mathcal{M}[(x{:}A)(B(x))X]$. We have $\Sigma(A,B) \leq_p \Sigma(A',B')$ if $A \leq A'$ and $B(x) \leq B'(x)$.
- Disjoint union types: $+ =_{\mathrm{df}} [A{:}\mathbf{Type}][B{:}\mathbf{Type}]\ \mathcal{M}[(A)X, (B)X]$. We have $A+B \leq_p A'+B'$ if $A \leq A'$ and $B \leq B'$.
- B-branching trees: $Tree =_{\mathrm{df}} [B{:}\mathbf{Type}]\mathcal{M}[X, ((B)X)X]$. We have $Tree(B) \leq_p Tree(B')$ if $B' \leq B$.

For instance, we have $List(Even) <_p List(Nat)$ and $Tree(Nat) <_p Tree(Even)$.

In practice, the relation $<_p$ can be regarded as a default generalisation of the user-specified subtyping relation ($\mathcal{C}$ in our formulation), with appropriate overriding by the latter (eg, a user-defined coercion between $\Sigma(Even, B)$ and $\Sigma(Nat, B)$ will override the default coercion generated from the coercion between $Even$ and Nat). We note that, when $\mathcal{C}$ is a forest, the above generalisation also guarantees that there is only one unique path between a subtype and a supertype.

4.3 Subtyping between type universes

The inclusion relation between type universes may either be introduced by means of explicit lifting operators (e.g., adopted in the presentation of UTT in [Luo94]) or direct subtyping (e.g., adopted in the presentation of the Extended Calculus of Constructions [Luo90] and used in systems such as Lego), called by Martin-Löf as universes à la Tarski and universes à la Russell, respectively [ML84]. The latter approach is based on overloading common term operators and is not quite compatible with the elimination rules when general inductive types are introduced; in particular, the subject reduction property would fail to hold. For instance, for the product types with introduction operator $pair_\times$ and elimination operator $\mathbf{E}_\times$, we would have, in the following context,

$$x, y{:}Type_0,\ \ C{:}(Type_1 \times Type_1)\mathbf{Type},\ \ f{:}(x, y{:}Type_1)C(pair_\times(Type_1, Type_1, x, y)),$$

that the following well-typed term (since $Type_0 \times Type_0 < Type_1 \times Type_1$)

$$p \equiv \mathbf{E}_\times(Type_1, Type_1, C, f, pair_\times(Type_0, Type_0, x, y)),$$

computes to $f(x, y)$, which is of type $C(pair_\times(Type_1, Type_1, x, y))$ but not of type $C(pair_\times(Type_0, Type_0, x, y))$, a type of p. This has caused a problem in extending subtyping to parameterised inductive types in systems such as Lego. The reason is essentially that the formulation of the elimination rule has not taken into the account that, with subtyping, a supertype also contains canonical objects of its subtypes.[2]

With coercive subtyping, a natural bridge between the (more semantics-oriented) formulation à la Tarski and the more practically useful formulation à la Russell can be established because we can declare subtyping relations between universes and take the explicit lifting operators as the intended implicit coercions. For example, inclusions between predicative universes in UTT are introduced by means of explicit lifting operators as follows (here we omit the introduction of names of inductive types in universes):

$$Type_i : \mathbf{Type},\quad type_i : Type_{i+1},$$

$$\mathbf{T}_i : (Type_i)\mathbf{Type},\quad \mathbf{T}_{i+1}(type_i) = Type_i : \mathbf{Type},$$

$$\mathbf{t}_{i+1} : (Type_i)Type_{i+1},\quad \mathbf{T}_{i+1}(\mathbf{t}_{i+1}(a)) = \mathbf{T}_i(a) : \mathbf{Type}.$$

We can introduce the subtyping relations $Type_i < Type_{i+1}$ with coercions defined as $\kappa[Type_i, Type_{i+1}] =_{\mathrm{df}} \mathbf{t}_{i+1}$. This fits well with the implementation of subtyping in proof development systems such as Lego and solves the above problem.

4.4 Interpreting practical forms of coercion

Anthony Bailey has recently implemented various interesting and useful coercion mechanisms in Lego [Bai96]. (Also see [Sai96] for a related development in the Coq system.) Bailey has given interesting examples to illustrate the use of implicit coercion to make proof development easier (more readable etc.). We briefly discuss how these basic coercion mechanisms may be understood in our setting of coercive subtyping.

On the basis of the Lego system, the implementation of which is not based on a logical framework description of type theory, Bailey has introduced several kinds of implicit coercions, called *argument coercion*, *kind coercion*, and *Π-coercion*. Argument-coercion deals with application of a function of a Π-type $\Pi x{:}A.B(x)$ to an object a whose principal type is a subtype of A. This is similar to our treatment of the application for dependent product kinds. In our setting, argument coercion for Π-types can be understood as a special case when applying $\mathbf{app}(A, B, f)$ to a.

[2]Furthermore, it is not clear how it can be modified to do so, since it may require a substantial extension with new forms of judgements. We do not discuss this issue here.

Bailey's notion of kind-coercion may better be called type-coercion in our terminology. It is introduced to deal with the following situation. For instance, suppose *Group* is the Σ-type representing the theory of groups and $G : Group$. One may hope to write, for example, $\Pi x{:}G.C(x)$ for 'for all x in (the carrier type of) G, $C(x)$', though this is not allowed since G is not its carrier type. This becomes possible when a kind-coercion *el* from *Group* to types in a universe U is declared to obtain the carrier type of every group structure, and the above Π-type would stand for $\Pi x{:}el(G).C(x)$. Kind-coercion can be understood as a special case of argument coercion in our formulation. For instance, for the above example, $\Pi x{:}G.C(x)$ abbreviates $\Pi(G, C) = \Pi(el(G), C)$ when $Group < U$ with coercion *el*.

The notion of Π-coercion is to deal with the cases where an object F supposed to be of a Π-type is applied to an object but F does not have the right Π-type. This can also be regarded as a special case of coercive definition: with an implicit coercion κ from the type of F to the expected Π-type $\Pi(A, B)$, we have $\mathbf{app}(A, B, F, a) = \mathbf{app}(A, B, \kappa(F), a)$.

4.5 Coercive subtyping and type-casting

In a private communication with the author, Peter Aczel has recently pointed out a close relationship between coercive subtyping and type-casting. Type-casting is a way to form terms in proof systems such as Lego, which was introduced primarily for dealing with type ambiguity introduced by subtyping (eg, universe inclusion) or omission of syntax to provide user-friendliness. For example, in the Lego system, one may write a term of the form `a:A`, which stands for 'the term `a` with principal type `A`'.

We can understand `a:A` as the term $([x{:}A]x)(a)$ in our system. When $a : A_0$ and $A_0 < A$, we have, by the coercive definition rule and (β), $([x{:}A]x)(a) = \kappa(a)$, whose principal type is A. For instance, with universe subtyping introduced above, we will have $(type_0{:}Type_3) = ([x{:}Type_3]x)(type_0) = \mathbf{t}_3(\mathbf{t}_2(type_0))$.

In other words, in a type theory where every object has a principal type, type-casting is definable with coercive subtyping. Alternatively, on the basis of the above understanding, one may directly introduce type-casting as basic terms and introduce the following rules (modified from a suggestion by Aczel), which analyse the coercive definition rules into two parts:

$$\frac{k_0 \ :: \ K}{(k_0{:}K) : K} \qquad \frac{k_0 : K_0}{(k_0{:}K_0) = k_0 : K_0} \qquad \frac{K_0 < K \quad k_0 : K_0}{(k_0{:}K) = \kappa(k_0) : K} \qquad \frac{f : (x{:}K)K' \ \ k_0 : K_0 \ \ K_0 < K}{f(k_0) = f(k_0{:}K) : [k_0/x]K'}$$

These rules reflect the meaning of type-casting directly. If $(k_0{:}K)$ is defined as $([x{:}K]x)(k_0)$, the above rules are all derivable.

Type-casting pairs: an example

The above understanding can be used to understand the use of type-casting in resolving the possible type ambiguity of pairs (objects of Σ-types) as well. For instance, for $a : A$, the (untyped) pair (A, a) may have type $Type_0 \times A$ or $\Sigma X{:}Type_0.X$, which are incompatible. In the Lego system, the former is the default type and if it is the second type that is the intended one, one has to use type-casting to write explicitly `(A,a : <X:Type(0)>X)`. Such a decision seems to be ad hoc, though very useful in practice.

To analyse this problem with coercive subtyping, we can consider a special family of unit types $\mathbf{1}(A, a)$ indexed by types $A : \mathbf{Type}$ and objects $a : A$, each of which has only one constructor $*(A, a) : \mathbf{1}(A, a)$. Then, for any type A, any family of types $B : (A)\mathbf{Type}$, and any $a : A$, we can define two coercions, κ_1 from $\mathbf{1}(A, a) \times B(a)$ to $A \times B(a)$ and κ_2 from $\mathbf{1}(A, a) \times B(a)$ to $\Sigma(A, B)$ (ie, $\Sigma x{:}A.B(x)$), as follows:

$$\begin{aligned} \kappa_1(pair_\times(\mathbf{1}(A, a), B(a), *(A, a), b)) &= pair_\times(A, B(a), a, b), \\ \kappa_2(pair_\times(\mathbf{1}(A, a), B(a), *(A, a), b)) &= pair_\Sigma(A, B, a, b). \end{aligned}$$

where $pair_\times$ and $pair_\Sigma$ are the introduction operators for the product types and the Σ-types, respectively. Then, for $a : A$ and $b : B(a)$, we can regard the untyped pair (a, b) as standing for $pair_\times(\mathbf{1}(A, a), B(a), *(A, a), b)$. An implementation based on this understanding will then figure out the appropriate typing for an untyped pair by choosing one of the coercions to apply. Note that according to this analysis, the treatment in the Lego system is not quite adequate, while the implementation of type-checking pairs in the system PVS is more flexible and seems to be very close to our analysis.[3]

[3]Thanks to Paul Jackson for an interesting conversation on this topic, which has helped to spot an error in an earlier version of this paper.

The above discussion on type-casting is an example of understanding implicit syntax [Pol90] (via coercive subtyping). It would be interesting to study whether other forms of implicit syntax can be understood in such a manner.

5 Conclusions, related work, and further research

The central idea of coercive subtyping is to introduce coercive definition rules so that general subsumption and implicit coercion can be combined together and obtain a uniform proof-theoretic treatment in type theory. This is different from using model-theoretic (denotational) semantics in understanding subtyping. Our approach is more syntactic and gives direct meaning-theoretic treatment of subtyping and coercions (and extends Martin-Löf's meaning theory of type theory in a coherent way). It offers the opportunity for subtyping to be introduced into the current proof development systems such as Coq, ALF, and Lego, and to make the task of formal development easier. Although the formal treatment deals with type theories formulated in LF directly, it is not difficult to be modified so that it can be applied to other type theories such as those implemented in the Coq system or the NuPRL system. We believe that direct inheritance supported by coercive subtyping is a very useful mechanism that provides a powerful tool in applications such as specification and data refinement (with refinement maps between specifications [Luo93] as coercions), development of mathematical theories in proof development (with theory morphisms [Luo91] as coercions [Bai96]), and library structuring for proof reuse [Luo95].

Subtyping is in general a subtle issue partly because, in the presence of (arbitrary) subtyping, a judgement of the form $k :: K$ is a synthetic judgement in the sense of Martin-Löf [ML94]; that is, the judgement form is essentially existential and hence in general undecidable, unless the formulation of the system has certain restrictions. Coercive subtyping offers one such restriction. It remains to see how further development can be made in this direction. For instance, with the new form of judgement $k :: K$, it may be reasonable to consider introduction of assumptions of the form $x{::}K$ in contexts and hence quantifications of the form $(x{::}K)K'$ or $\Pi x{::}A.B$. Also, we have not considered 'dependent' coercions between a type and a family of types, whose kind can be of the form $(x{:}K)K'$ where x does occur free in K'. These coercions may be useful in understanding other forms of implicit syntax such as those offered by argument synthesis (e.g., in the Lego system); but they can be used to introduce infinitely many coercions, which may be difficult to implement properly to guarantee uniqueness of coercions. Similar implementation problems exist for parameterised coercions and are among the interesting further research topics.

In this paper, we have restricted the basic subtyping relation to be a forest to guarantee the uniqueness (up to computational equality) of a coercion between any two types or kinds. Note that we allow multiple inheritance in the sense that a type can have more than one subtype or supertype, though it is not allowed to have two different coercions (which are not computationally equal) between two types. For instance, two different mappings from a type of rings to a type of monoids, which map a ring to its different monoids, are not allowed to be coercions at the same time. The only possibility is to regard them as coercions from the ring type to two different types of monoids.

Among the closely related work, Pollack and Pierce's suggestion (private communication) to consider coercions as a basic mechanism in type-checking overloading methods for classes proposed by Aczel [Acz94] was a major influence to the idea of coercive subtyping. This idea goes back to the work on giving coercion semantics to lambda calculi with subtyping by Breazu-Tannen et al [BCGS91]. Bailey's development of coercions in Lego share the major practical motivation with this work. His implementation is based on the syntactical equality rather than computational equality. This allows him to implement certain forms of parameterised coercions as well. However, Bailey has not considered contravariant subtyping for Π-types in his implementation. It is not yet clear how the coercive definition rules may be directly related to such implementations of implicit coercions based on syntactical equality, though the equational theory presented in this paper provides a type-theoretic framework in which coercion-based implicit syntax can be understood. Work is in progress to study implementation issues based on the notion of typed reduction and more efficient reduction strategies.

Subtyping has also been studied extensively by many researchers in the context of typed functional programming, inspired by the notion of inheritance as found in object-oriented programming (cf., [CW85, Car88]). Semantic studies on subtyping in type systems (without dependent types) such as the second-order lambda calculus (ie, the system F) include the use of the PER model [BL90] and the coercion-based approach [BCGS91]. A more recent logical study of subtyping in system F can be found in [LMS95]. There is not much work on subtyping in dependent type systems in the context of proof development systems based on type theory. Pfenning's work on refinement types [Pfe93] and Aspinall and Compagnoni's work on the decidability of Edinburgh LF with subtyping [AC96] are among the recent research development. It is clear that coercive subtyping is strongly motivated by the need in proof development, where inductive types are essential. Traditional approaches based on overloading term constructors (such as λ) cannot be generalised in this context. However, some ideas in the research on subtyping in programming language context may be very useful for proof languages, examples of which include introduction of subtyping assumptions into contexts and bounded quantifiers, among others. Further research is needed to see whether they are useful and how they may be incorporated for coercive subtyping.

Acknowledgements I am very grateful to Peter Aczel, Anthony Bailey, Adriana Compagnoni, Healfdene Goguen, Paul Jackson, Alex Jones, Randy Pollack, Sergei Soloviev, and the anonymous referees for their useful comments on this work or the earlier drafts of this paper. Communication and discussion with Anthony Bailey on his implementation of coercions in Lego has been very useful. I want to thank Rod Burstall and Keith Bennett for their continuous support and encouragement.

References

[AC96] D. Aspinall and A. Compagnoni. Subtyping dependent types. *Proc. of LICS96*, 1996. To appear.

[Acz94] P. Aczel. Simple overloading for type theories. Draft, 1994.

[Bai96] A. Bailey. Lego with implicit coercions. 1996. Draft.

[BCGS91] V. Breazu-Tannen, T. Coquand, C. Gunter, and A. Scedrov. Inheritance and explicit coercion. *Information and Computation*, 93, 1991.

[Bee85] M.J. Beeson. *Foundations of Constructive Mathematics.* Springer-Verlag, 1985.

[BL90] K. Bruce and G. Longo. A modest model of records, inheritance, and bounded quantification. *Information and Computation*, 87, 1990.

[BM92] R. Burstall and J. McKinna. Deliverables: a categorical approach to program development in type theory. LFCS report ECS-LFCS-92-242, Dept of Computer Science, University of Edinburgh, 1992.

[Bur93] R. Burstall. Extended Calculus of Constructions as a specification language. In R. Bird and C. Morgan, editors, *Mathematics of Program Construction*, 1993. Invited talk.

[C$^+$86] R.L. Constable et al. *Implementing Mathematics with the NuPRL Proof Development System.* Pretice-Hall, 1986.

[Car88] L. Cardelli. Type-checking dependent types and subtypes. *Lecture Notes in Computer Science*, 306, 1988.

[Coq92] Th. Coquand. Pattern matching with dependent types. Talk given at the BRA workshop on Proofs and Types, Bastad, 1992.

[CPM90] Th. Coquand and Ch. Paulin-Mohring. Inductively defined types. *Lecture Notes in Computer Science*, 417, 1990.

[CW85] L. Cardelli and P. Wegner. On understanding types, data abstraction and polymorphism. *Computing Surveys*, 17, 1985.

[D$^+$91] G. Dowek et al. *The Coq Proof Assistent: User's Guide (version 5.6).* INRIA-Rocquencourt and CNRS-ENS Lyon, 1991.

[Dyb91] P. Dybjer. Inductive sets and families in Martin-Löf's type theory and their set-theoretic semantics. In G. Huet and G. Plotkin, editors, *Logical Frameworks.* Cambridge University Press, 1991.

[Gog94] H. Goguen. *A Typed Operational Semantics for Type Theory.* PhD thesis, University of Edinburgh, 1994.

[HHP87] R. Harper, F. Honsell, and G. Plotkin. A framework for defining logics. *Proc. 2nd Ann. Symp. on Logic in Computer Science. IEEE*, 1987.

[LMS95] G. Longo, K. Milsted, and S. Soloviev. A logic of subtyping. In *Proc. of LICS'95*, 1995.

[LP92] Z. Luo and R. Pollack. LEGO Proof Development System: User's Manual. LFCS Report ECS-LFCS-92-211, Department of Computer Science, University of Edinburgh, 1992.

[Luo89] Z. Luo. **ECC**, an extended calculus of constructions. In *Proc. of the Fourth Ann. Symp. on Logic in Computer Science*, Asilomar, California, U.S.A., June 1989. IEEE.

[Luo90] Z. Luo. *An Extended Calculus of Constructions.* PhD thesis, University of Edinburgh, 1990. Also as Report CST-65-90/ECS-LFCS-90-118, Department of Computer Science, University of Edinburgh.

[Luo91] Z. Luo. A higher-order calculus and theory abstraction. *Information and Computation*, 90(1):107–137, 1991.

[Luo92] Z. Luo. A unifying theory of dependent types: the schematic approach. *Proc. of Symp. on Logical Foundations of Computer Science (Logic at Tver'92), LNCS 620*, 1992. Also as LFCS Report ECS-LFCS-92-202, Dept. of Computer Science, University of Edinburgh.

[Luo93] Z. Luo. Program specification and data refinement in type theory. *Mathematical Structures in Computer Science*, 3(3), 1993.

[Luo94] Z. Luo. *Computation and Reasoning: A Type Theory for Computer Science.* Oxford University Press, 1994.

[Luo95] Z. Luo. Developing reuse technology in proof engineering. *Proceedings of AISB95, Workshop on Automated Reasoning: bridging the gap between theory and practice, Sheffield, U.K.*, April 1995.

[ML84] P. Martin-Löf. *Intuitionistic Type Theory.* Bibliopolis, 1984.

[ML94] P. Martin-Löf. Analytic and synthetic judgements in type theory. In P. Parrini, editor, *Kant and Contemporary Epistemology.* Kluwer Accademic Publishers, 1994.

[MN94] L. Magnusson and B. Nordström. The ALF proof editor and its proof engine. In *Types for Proof and Programs, LNCS*, 1994.

[NPS90] B. Nordström, K. Petersson, and J. Smith. *Programming in Martin-Löf's Type Theory: An Introduction.* Oxford University Press, 1990.

[Pfe93] F. Pfenning. Refinement types for logical frameworks. *Preliminary Proceedings of BRA Workshop on Types and Proofs*, 1993.

[Pol90] R. Pollack. Implicit syntax. In the preliminary Proceedings of the 1st Workshop on Logical Frameworks, 1990.

[Pol94] R. Pollack. *The Theory of LEGO: A Proof Checker for the Extended Calculus of Constructions.* PhD thesis, Edinburgh University, 1994.

[Sai96] A. Saibi. Typing algorithm in type theory with inheritance. Draft, 1996.