

# On the Decidability of Subtyping with Bounded Existential Types<sup>\*</sup>

Stefan Wehr and Peter Thiemann

Institut für Informatik, Universität Freiburg  
{wehr,thiemann}@informatik.uni-freiburg.de

**Abstract.** Bounded existential types are a powerful language feature for modeling partial data abstraction and information hiding. However, existentials do not mingle well with subtyping as found in current object-oriented languages: the subtyping relation is already undecidable for very restrictive settings.

This paper considers two subtyping relations defined by extracting the features specific to existentials from current language proposals (JavaGI, WildFJ, and Scala) and shows that both subtyping relations are undecidable. One of the two subtyping relations remains undecidable even if bounded existential types are removed.

With the goal of regaining decidable type checking for the JavaGI language, the paper also discusses various restrictions including the elimination of bounded existentials from the language as well as possible amendments to regain some of their features.

## 1 Introduction

Cardelli and Wegner [5] introduced bounded existential types to obtain a fine-grained modeling instrument for structured and partial data abstraction and information hiding, thus generalizing the concept of an abstract data type modeled with a plain existential type [12]. Language designers and type theorists rely on bounded existentials in diverse areas such as object-oriented languages [1], module systems [19], and functional languages [10].

In the realm of object-oriented languages, bounded existential types have found uses for modeling object-oriented languages in general [2], as well as for modeling specific features such as Java wildcards [3,20,21]. Only a few languages (e.g., Scala [13]) make bounded existential types in full generality available to the programmer. Also, the initial design of JavaGI [23] includes bounded existential types and provides interface types (i.e., the ability to form types from interface names) as a special case supported by syntactic sugar. Building directly on bounded existential types has several advantages compared to interface types: they properly generalize interface types, they encompass Java wildcards, and they have meaningful uses with interfaces abstracting over families of types. Thus, despite the complexity that they introduce in a type system, bounded existential types initially appear like a worthwhile feature.

---

<sup>\*</sup> A preliminary version of this work was presented at FTfJP 2008 [24].

Unfortunately, it turns out that subtyping —and hence type checking— for JavaGI is undecidable in the presence of general bounded existential types. Furthermore, subtyping for bounded existential types as used to encode Java wildcards is also undecidable. This article proves both of these undecidability results. Moreover, it also shows that replacing JavaGI’s existentials with plain interface types does not regain decidability of subtyping.

## Contributions and Overview

After refreshing some background on the JavaGI language in Sec. 2, Sec. 3 defines the calculus  $\mathcal{E}\mathcal{X}_{impl}$  that models the essential aspects of subtyping and bounded existential types in JavaGI. Subtyping in  $\mathcal{E}\mathcal{X}_{impl}$  is shown to be undecidable by reduction from Post’s Correspondence Problem [18]. Further, the section defines the calculus  $\mathcal{I}\mathcal{T}_{impl}$  by replacing existential types in  $\mathcal{E}\mathcal{X}_{impl}$  with plain interface types. Subtyping in  $\mathcal{I}\mathcal{T}_{impl}$  is also undecidable but various restrictions exist that ensure decidability.

Sec. 4 considers the calculus  $\mathcal{E}\mathcal{X}_{uplo}$  supporting existentials with lower and upper bounds. Subtyping in  $\mathcal{E}\mathcal{X}_{uplo}$  is also undecidable, as shown by reduction from subtyping in  $F_{\leq}^D$ , a restricted form of the polymorphic  $\lambda$ -calculus extended with subtyping [14]. The results in this section are relevant to Scala [13], formal systems for modeling Java wildcards [3, 4, 20], and JavaGI’s full type system.

Sec. 5 explores alternative design options for JavaGI that avoid bounded existential types but keep the remaining features. Finally, Sec. 6 reviews related work and Sec. 7 concludes. Detailed proofs may be found in an accompanying technical report [26].

## 2 Background

JavaGI [23, 25] is a conservative extension of Java 1.5. It generalizes Java’s interface concept to incorporate the essential features of Haskell type classes [8, 22]. The generalization allows for retroactive and type-conditional interface implementations, binary methods, static methods in interfaces, default implementations for interface methods, and multi-headed interfaces (interfaces over families of types). Furthermore, JavaGI’s initial design generalizes Java-like interface types to bounded existential types. This section only discusses the features relevant to this paper, namely retroactive interface implementations and existential types.

### 2.1 Retroactive Interface Implementations

Retroactive interface implementations allow programmers to implement an interface for some class without changing the source code of the class. For example, Java rejects the use of a `for`-loop to iterate over the characters of a string because the class `String` does not implement the interface `Iterable`:<sup>1</sup>

<sup>1</sup> Java’s enhanced `for`-loop allows to iterate over arrays and all types implementing the `Iterable<X>` interface, which contains a single method `Iterator<X> iterator()`.

```

implementation Iterable<Character> [String] {
  public Iterator<Character> iterator() {
    return new Iterator<Character>() {
      private int index = 0;
      public boolean hasNext() { return index < length(); }
      public Character next() { return charAt(index++); }
    };
  }
}

```

Fig. 1. Retroactive implementation of the `Iterable` interface.

```

for (Character c : "21 is only half the truth") { ... } // illegal in Java

```

As a class definition in Java must specify all interfaces that the class implements and the definition of `java.lang.String` is fixed, there is no hope of getting this code to work. A JavaGl programmer can overcome this restriction by adding implementations for interfaces to an existing class at any time, retroactively, without modifying the source code of the class.

For example, the *implementation definition* shown in Fig.1 specifies that the *implementing type* `String`, enclosed in square brackets `[]`, implements the interface `Iterable<Character>`.<sup>2</sup> The definition of the `iterator` method can use the methods `length` and `charAt` because they are part of `String`'s public interface.

## 2.2 Bounded Existential Types

Java uses the name of an interface as an *interface type* to denote the set of all types implementing the interface. Instead of interface types, the initial design of JavaGl features *bounded existential types* [5] and provides syntactic sugar for recovering interface types. For example, the interface type `List<String>` abbreviates the existential type  $\exists X \text{ where } X \text{ implements } \text{List}\langle\text{String}\rangle . X$ . The *implementation constraint* "`X implements List<String>`" restricts instantiations of the type variable `X` to types that implement the interface `List<String>`. Thus, the existential type denotes the set of all types implementing `List<String>`, exactly like the synonymous interface type. (The occurrence of "`List<String>`" in the implementation constraint does not abbreviate an existential type.)

Existentials are more general than interface types. For instance, the existential  $\exists X \text{ where } X \text{ implements } \text{List}\langle\text{String}\rangle, X \text{ implements } \text{Set}\langle\text{String}\rangle . X$  denotes the set of all types that implement both `List<String>` and `Set<String>`. Java supports such intersections of interface types only for specifying bounds of type variables. Existentials also encompass Java wildcards [3, 4, 20, 21]. For instance, the existential type  $\exists X \text{ where } X \text{ extends } \text{Number} . \text{List}\langle X \rangle$  corresponds to the wildcard type `List<? extends Number>`.<sup>3</sup>

<sup>2</sup> The implementation ignores the `remove` method of the `Iterator` interface.

<sup>3</sup> Because `List` is an interface,  $\exists X \text{ where } X \text{ extends } \text{Number} . \text{List}\langle X \rangle$  stands for  $\exists X, L \text{ where } X \text{ extends } \text{Number}, L \text{ implements } \text{List}\langle X \rangle . L$ .

Syntax

$$\begin{aligned}
P, Q, R &::= X \text{ implements } I\langle\bar{T}\rangle \\
T, U, V, W &::= X \mid \exists X \text{ where } P. X \\
\text{def} &::= \text{interface } I\langle\bar{X}\rangle \mid \text{implementation}\langle\bar{X}\rangle I\langle\bar{T}\rangle [I\langle\bar{T}\rangle] \\
X, Y, Z &\in \text{TyvarName} \quad I, J \in \text{IfaceName}
\end{aligned}$$

$\Theta; \Delta \Vdash T \text{ implements } I\langle\bar{T}\rangle$

$$\begin{array}{c}
\text{E}_1\text{-IMPL} \\
\frac{\text{implementation}\langle\bar{X}\rangle I\langle\bar{T}\rangle [U] \in \Theta}{\Theta; \Delta \Vdash [\bar{V}/\bar{X}] (U \text{ implements } I\langle\bar{T}\rangle)} \\
\text{E}_1\text{-LOCAL} \\
\frac{P \in \Delta}{\Theta; \Delta \Vdash P}
\end{array}$$

$\Theta; \Delta \vdash T \leq T$

$$\begin{array}{c}
\text{S}_1\text{-REFL} \\
\Theta; \Delta \vdash T \leq T \\
\text{S}_1\text{-TRANS} \\
\frac{\Theta; \Delta \vdash T \leq U \quad \Theta; \Delta \vdash U \leq V}{\Theta; \Delta \vdash T \leq V} \\
\text{S}_1\text{-OPEN} \\
\frac{\Theta; \Delta, P \vdash X \leq T \quad X \notin \text{ftv}(\Theta, \Delta, T)}{\Theta; \Delta \vdash (\exists X \text{ where } P. X) \leq T} \\
\text{S}_1\text{-ABSTRACT} \\
\frac{\Theta; \Delta \Vdash [T/X]P}{\Theta; \Delta \vdash T \leq (\exists X \text{ where } P. X)}
\end{array}$$

**Fig. 2.** Syntax, entailment, and subtyping for  $\mathcal{EX}_{impl}$ .

The initial design of `JavaGl` allows implementation definitions for existentials. For example, given an interface `I`, a programmer may write an implementation definition to specify that all types implementing `List<X>` also implement `I`.

```
implementation<X> I [List<X>] { /* implement methods of I */ }
```

Such a definition is feasible only if all methods of `I` can be implemented using only methods of the `List` interface. The example also demonstrates that `JavaGl` supports *generic implementation definitions*, which are parameterized by type variables.

### 3 Subtyping Existential Types with Implementation Constraints

This section introduces  $\mathcal{EX}_{impl}$ , a subtyping calculus with existentials ( $\mathcal{EX}$ ) and implementation constraints (*impl*). The calculus is a subset of `Core-JavaGl` [23]. It does not model all aspects of `JavaGl`'s initial design, but contains only those features that make subtyping undecidable. In particular, the syntax of existentials is restricted such that all existentials are encodings of interface types. Consequently, undecidability of subtyping also holds for  $\mathcal{IT}_{impl}$ , a variant of  $\mathcal{EX}_{impl}$  where interface types ( $\mathcal{IT}$ ) replace existentials.

#### 3.1 Definition of $\mathcal{EX}_{impl}$

Fig. 2 defines the syntax along with the entailment and subtyping relations of  $\mathcal{EX}_{impl}$ . An implementation constraint  $P$  has the form  $X \text{ implements } I\langle\bar{T}\rangle$  and

constrains the type variable  $X$  to types that implement the interface  $I\langle\bar{T}\rangle$ . An interface without type parameters is written  $I$  instead of  $I\langle\bullet\rangle$ .<sup>4</sup>

A type  $T$  is either a type variable  $X$  or a bounded existential type of the form  $\exists X \text{ where } X \text{ implements } I\langle\bar{T}\rangle. X$ . For simplicity, there are no class types, existentials have a single quantified type variable  $X$ , they have exactly one constraint  $X \text{ implements } I\langle\bar{T}\rangle$ , and the body of an existential must be the quantified type variable. Existentials are considered equal up to renaming of bound type variables.

A definition  $def$  in  $\mathcal{E}\mathcal{X}_{impl}$  is either an interface or an implementation definition. Interface and implementation definitions do not have method signatures or bodies, because methods do not matter for the entailment and subtyping relations of  $\mathcal{E}\mathcal{X}_{impl}$ . Moreover,  $\mathcal{E}\mathcal{X}_{impl}$  does not support interface inheritance. An implementation definition  $\text{implementation}\langle\bar{X}\rangle I\langle\bar{T}\rangle [J\langle\bar{U}\rangle]$  implicitly assumes that  $\bar{X} = \text{ftv}(J\langle\bar{U}\rangle)$ .<sup>5</sup>

The entailment relation  $\Theta; \Delta \vdash T \text{ implements } I\langle\bar{T}\rangle$  expresses that type  $T$  implements interface  $I\langle\bar{T}\rangle$ . It relies on a program environment  $\Theta$ , which is a finite set of definitions  $def$ , and a type environment  $\Delta$ , which is a finite set of constraints  $P$ , where  $\Delta, P$  abbreviates  $\Delta \cup \{P\}$ . A type implements an interface either because it corresponds to an instance of a suitable implementation definition (rule  $E_1\text{-IMPL}$ ) or because the type environment contains the constraint (rule  $E_1\text{-LOCAL}$ ).<sup>6</sup>

The subtyping relation  $\Theta; \Delta \vdash T \leq U$  states that  $T$  is a subtype of  $U$ . It is reflexive and transitive as usual. Rule  $S_1\text{-OPEN}$  opens an existential on the left-hand side of the subtyping relation by moving its constraint into the type environment. The premise  $X \notin \text{ftv}(\Theta, \Delta, T)$  ensures that the existentially quantified type variable is sufficiently fresh and does not escape from its scope. Rule  $S_1\text{-ABSTRACT}$  deals with existentials on the right-hand side of the subtyping relation. It states that  $T$  is a subtype of some existential if the constraint of the existential holds after substituting  $T$  for the existentially quantified type variable.

As part of a type soundness proof for `Core-JavaGl`, we verified that the subtyping relation of  $\mathcal{E}\mathcal{X}_{impl}$  supports the usual principle of subsumption: we can always promote the type of an expression to some supertype without causing runtime errors.

### 3.2 Undecidability of Subtyping in $\mathcal{E}\mathcal{X}_{impl}$

The undecidability of subtyping in  $\mathcal{E}\mathcal{X}_{impl}$  follows by reduction from Post's Correspondence Problem (PCP). It is well known that PCP is undecidable [7, 18].

**Definition 1 (PCP).** *Let  $\{(u_1, v_1) \dots, (u_n, v_n)\}$  be a set of pairs of non-empty words over some finite alphabet  $\Sigma$  with at least two elements. A solution of PCP*

<sup>4</sup> The notation  $\bar{\xi}$  abbreviates a sequence  $\xi_1, \dots, \xi_n$  of syntactic entities with  $\bullet$  standing for the empty sequence. Sometimes, the sequence  $\bar{\xi}$  stands for the set  $\{\bar{\xi}\}$ .

<sup>5</sup> The notation  $\text{ftv}(\xi)$  denotes the set of type variables free in  $\xi$ .

<sup>6</sup> The notation  $[T/X]$  stands for the capture-avoiding substitution replacing each  $X_i$  with  $T_i$ .

is a sequence of indices  $i_1 \dots i_r$  such that  $u_{i_1} \dots u_{i_r} = v_{i_1} \dots v_{i_r}$ . The decision problem asks whether such a solution exists.

**Theorem 2.** *Subtyping in  $\mathcal{EX}_{impl}$  is undecidable.*

*Proof.* Let  $\mathcal{P} = \{(u_1, v_1), \dots, (u_n, v_n)\}$  be a particular instance of PCP over the alphabet  $\Sigma$ . We can encode  $\mathcal{P}$  as an equivalent subtyping problem in  $\mathcal{EX}_{impl}$  as follows. First, words over  $\Sigma$  must be represented as types in  $\mathcal{EX}_{impl}$ .

```
interface E      // empty word  $\varepsilon$ 
interface L<X>  // letter, for every  $L \in \Sigma$ 
```

Words  $u \in \Sigma^*$  are formed with these interfaces through nested existentials. For example, the word AB is represented by

$$\exists X \text{ where } X \text{ implements } A < \exists Y \text{ where } Y \text{ implements } B < \exists Z \text{ where } Z \text{ implements } E . Z > . Y > . X$$

The abbreviation  $I < \bar{T} >$  stands for the type  $\exists X \text{ where } X \text{ implements } I < \bar{T} > . X$ . Using this notation, the word AB is represented by  $A < B < E > >$ .

Formally, we define the representation of a word  $u$  as  $\llbracket u \rrbracket := u \# E$ , where  $u \# T$  is the concatenation of a word  $u$  with a type  $T$ :

$$\varepsilon \# T := T \qquad Lu \# T := L < u \# T >$$

Two interfaces are required to model the search for a solution of PCP:

```
interface S<X,Y>  // search state
interface G      // search goal
```

The type  $S < \llbracket u \rrbracket, \llbracket v \rrbracket >$  represents a particular search state where we have already accumulated indices  $i_1, \dots, i_k$  such that  $u = u_{i_1} \dots u_{i_k}$  and  $v = v_{i_1} \dots v_{i_k}$ . To model valid transitions between search states, we define implementations of  $S$  for all  $i \in \{1, \dots, n\}$  as follows:

$$\text{implementation} < X, Y > \ S < u_i \# X, v_i \# Y > \ [S < X, Y >] \tag{1}$$

The type  $G$  represents the goal of a search, as expressed by the following implementation:

$$\text{implementation} < X > \ G \ [S < X, X >] \tag{2}$$

To get the search running we ask whether there exists some  $i \in \{1, \dots, n\}$  such that  $\Theta_{\mathcal{P}}; \emptyset \vdash S < \llbracket u_i \rrbracket, \llbracket v_i \rrbracket > \leq G$  is derivable. The program  $\Theta_{\mathcal{P}}$  consists of the interfaces and implementations just defined. In the extended version [26], we prove a lemma showing that  $\mathcal{P}$  has a solution if, and only if, there exists some  $i \in \{1, \dots, n\}$  such that  $\Theta_{\mathcal{P}}; \emptyset \vdash S < \llbracket u_i \rrbracket, \llbracket v_i \rrbracket > \leq G$  is derivable.  $\square$

*Example.* Suppose the PCP instance  $\mathcal{P} = \{(u_1, v_1), (u_2, v_2)\}$  with  $u_1 = A$ ,  $u_2 = ABA$ ,  $v_1 = AA$ , and  $v_2 = B$  is given. The instance has the solution 1, 2, 1 because  $u_1 u_2 u_1 = v_1 v_2 v_1 = AABAA$ . The program  $\Theta_{\mathcal{P}}$  for the  $\mathcal{EX}_{impl}$  encoding of this problem looks like this:

Syntax

$$\begin{aligned} \varphi &::= \forall \bar{X}. I \langle \bar{T} \rangle \leq I \langle \bar{T} \rangle \\ T, U, V, W &::= X \mid I \langle \bar{T} \rangle \\ X, Y, Z &\in \text{TyvarName} \quad I, J \in \text{IfaceName} \end{aligned}$$

$\Phi \vdash T \leq T$

$$\begin{array}{c} \text{S}_2\text{-REFL} \\ \Phi \vdash T \leq T \end{array} \quad \frac{\text{S}_2\text{-TRANS} \quad \Phi \vdash T \leq U \quad \Phi \vdash U \leq V}{\Phi \vdash T \leq V} \quad \frac{\text{S}_2\text{-IMPL} \quad (\forall \bar{X}. T \leq U) \in \Phi}{\Phi \vdash [V/X]T \leq [V/X]U}$$

Fig. 3. Syntax and subtyping for  $\mathcal{IT}_{impl}$ .

```

interface E           interface A<X>   interface B<X>
interface S<X,Y>     interface G
implementation<X,Y> S<A<X>,      A<A<Y>>>  [S<X,Y>]      // (1)
implementation<X,Y> S<A<B<A<X>>>, B<Y>>  [S<X,Y>]      // (2)
implementation<X>   G                [S<X,X>]      // (3)

```

We then need to ask whether there exists some  $i \in \{1, 2\}$  such that  $\Theta_{\mathcal{P}}; \emptyset \vdash \mathbf{S} \langle \llbracket u_i \rrbracket, \llbracket v_i \rrbracket \rangle \leq \mathbf{G}$  is derivable. Verifying that such a derivation exists for  $i = 1$  is left as an exercise to the reader.

### 3.3 Undecidability Without Existential Types

The proof of undecidability of subtyping in  $\mathcal{EX}_{impl}$  reveals that subtyping remains undecidable even if plain interface types replace existentials. To make this claim concrete, Fig. 3 defines the calculus  $\mathcal{IT}_{impl}$ , which essentially is a version of  $\mathcal{EX}_{impl}$  with plain interface types instead of existentials. To simplify the syntax,  $\mathcal{IT}_{impl}$  drops interface definitions altogether and uses subtyping schemes instead of implementation definitions: a *subtyping scheme*  $\varphi = \forall \bar{X}. J \langle \bar{T} \rangle \leq I \langle \bar{U} \rangle$  corresponds to an implementation definition **implementation**< $\bar{X}$ >  $I \langle \bar{U} \rangle$  [ $J \langle \bar{T} \rangle$ ]. Such a subtyping scheme implicitly assumes that  $\bar{X} = \text{ftv}(J \langle \bar{T} \rangle)$ .  $\mathcal{IT}_{impl}$  also replaces constraint entailment and the rules  $\text{S}_1\text{-OPEN}$  and  $\text{S}_1\text{-ABSTRACT}$  with a single rule  $\text{S}_2\text{-IMPL}$ . The symbol  $\Phi$  ranges over finite sets of subtyping schemes  $\varphi$ .

**Theorem 3.** *Subtyping in  $\mathcal{IT}_{impl}$  is undecidable.*

*Proof.* Similar to the proof of Theorem 2.

The rest of this section investigates decidable fragments of  $\mathcal{IT}_{impl}$ . It starts with the observation that the undecidability proofs of subtyping in  $\mathcal{EX}_{impl}$  and  $\mathcal{IT}_{impl}$  rely on two main ingredients:

**Cyclic interface subtyping.** Implementation definitions in  $\mathcal{EX}_{impl}$  (or subtyping schemes in  $\mathcal{IT}_{impl}$ ) allow the introduction of cycles in the subtyping graph of interfaces. Consider one of the implementations defined by Equation (1) on page 6: it states that  $\mathbf{S} \langle u_i \# \mathbf{X}, v_i \# \mathbf{Y} \rangle$  is a supertype of  $\mathbf{S} \langle \mathbf{X}, \mathbf{Y} \rangle$ . In the reduction from PCP, such cycles are used to encode the individual steps in the search for a solution.

**Multiple instantiation subtyping.** Implementation definitions in  $\mathcal{E}\mathcal{X}_{impl}$  (or subtyping schemes in  $\mathcal{I}\mathcal{T}_{impl}$ ) allow to introduce two different instantiations of the same interface as supertypes of some other interface. Consider again the implementations defined by Equation (1): for  $u_i \neq u_j$  or  $v_i \neq v_j$  the implementations state that  $\mathbf{S}\langle u_i \# \mathbf{X}, v_i \# \mathbf{Y} \rangle \neq \mathbf{S}\langle u_j \# \mathbf{X}, v_j \# \mathbf{Y} \rangle$  are both supertypes of  $\mathbf{S}\langle \mathbf{X}, \mathbf{Y} \rangle$ . In the reduction from PCP, multiple instantiation subtyping encodes the choice between different pairs  $(u_i, v_i)$  and  $(u_j, v_j)$ .

An obvious way to obtain decidable subtyping for  $\mathcal{I}\mathcal{T}_{impl}$  is to restrict the set of subtyping schemes  $\Phi$  such that, for all types  $T$ , only a finite set of  $T$ -supertypes is derivable from  $\Phi$ .

**Definition 4.** The set of  $T$ -supertypes derivable from  $\Phi$ , written  $\mathcal{S}_{T,\Phi}$ , is defined as the smallest set closed under the following rules:

$$T \in \mathcal{S}_{T,\Phi} \quad \frac{(\forall \bar{X}. V \leq U) \in \Phi \quad [\bar{W}/\bar{X}]V \in \mathcal{S}_{T,\Phi}}{[\bar{W}/\bar{X}]U \in \mathcal{S}_{T,\Phi}}$$

**Restriction 1.** The set  $\mathcal{S}_{T,\Phi}$  must be finite for all types  $T$ .

**Theorem 5.** Under Restriction 1, subtyping in  $\mathcal{I}\mathcal{T}_{impl}$  is decidable.

*Proof.* See the extended version [26]. □

Here is a restriction that eliminates cyclic interface subtyping.

**Definition 6.** A finite set of subtyping schemes  $\Phi$  is contractive if, and only if, there exists no sequence  $\varphi_1, \dots, \varphi_n \in \Phi$  such that  $\varphi_i = \forall \bar{X}_i. I_i \langle \bar{T}_i \rangle \leq J_i \langle \bar{U}_i \rangle$  for all  $i = 1, \dots, n$  and  $J_i = I_{i+1}$  for all  $i = 1, \dots, n-1$  and  $J_n = I_1$ .

**Restriction 2.** The set  $\Phi$  must be contractive.

**Lemma 7.** Restriction 2 implies Restriction 1.

*Proof.* See the extended version [26]. □

*Remark.* Restriction 1 does not imply Restriction 2. Consider the set  $\Phi = \{\forall \bullet. I \leq I\}$ , which obviously meets Restriction 1 but is not contractive.

The next restriction is strictly stronger than Restriction 2.

**Restriction 3.** For all  $\varphi_1, \varphi_2 \in \Phi$  it must hold that  $\varphi_1 = \forall \bar{X}. I_1 \langle \bar{T} \rangle \leq J_1 \langle \bar{U} \rangle$  and  $\varphi_2 = \forall \bar{Y}. I_2 \langle \bar{V} \rangle \leq J_2 \langle \bar{W} \rangle$  imply  $J_1 \neq I_2$ .

The last restriction considered eliminates multiple instantiation subtyping.

**Restriction 4.** If  $\Phi \vdash I \langle \bar{T} \rangle \leq J \langle \bar{U} \rangle$  and  $\Phi \vdash I \langle \bar{T} \rangle \leq J \langle \bar{V} \rangle$  then it must hold that  $\bar{U} = \bar{V}$ .

**Lemma 8.** Restriction 4 implies Restriction 1.

*Proof.* See the extended version [26]. □

*Remark.* Neither Restriction 1 nor Restriction 2 implies Restriction 4 as demonstrated by  $\Phi = \{\forall \bullet. I \leq J \langle A \rangle, \forall \bullet. I \leq J \langle B \rangle\}$ . Moreover, Restriction 4 does not imply Restriction 2 as demonstrated by  $\Phi' = \{\forall \bullet. I \leq J, \forall \bullet. J \leq I\}$ .

Syntax

$$\begin{aligned}
N, M &::= C\langle\bar{X}\rangle \mid \mathbf{Object} \\
T, U, V, W &::= X \mid N \mid \exists\bar{X} \mathbf{where} \bar{P}. N \\
P, Q, R &::= X \mathbf{extends} T \mid X \mathbf{super} T \\
X, Y, Z &\in \mathit{TyvarName} \quad C, D \in \mathit{ClassName}
\end{aligned}$$

$\Delta \Vdash T \mathbf{extends} T \quad \Delta \Vdash T \mathbf{super} T$

$$\frac{\text{E}_3\text{-EXTENDS} \quad \Delta \vdash T \leq U}{\Delta \Vdash T \mathbf{extends} U}$$

$$\frac{\text{E}_3\text{-SUPER} \quad \Delta \vdash U \leq T}{\Delta \Vdash T \mathbf{super} U}$$

$\Delta \vdash T \leq T$

$$\frac{\text{S}_3\text{-REFL}}{\Delta \vdash T \leq T}$$

$$\frac{\text{S}_3\text{-TRANS} \quad \Delta \vdash T \leq U \quad \Delta \vdash U \leq V}{\Delta \vdash T \leq V}$$

$$\frac{\text{S}_3\text{-OBJECT}}{\Delta \vdash T \leq \mathbf{Object}}$$

$$\frac{\text{S}_3\text{-EXTENDS} \quad X \mathbf{extends} T \in \Delta}{\Delta \vdash X \leq T}$$

$$\frac{\text{S}_3\text{-SUPER} \quad X \mathbf{super} T \in \Delta}{\Delta \vdash T \leq X}$$

$$\frac{\text{S}_3\text{-OPEN} \quad \Delta, \bar{P} \vdash N \leq T \quad \bar{X} \cap \mathit{fv}(\Delta, T) = \emptyset}{\Delta \vdash \exists\bar{X} \mathbf{where} \bar{P}. N \leq T}$$

$$\frac{\text{S}_3\text{-ABSTRACT} \quad T = [\bar{U}/\bar{X}]N \quad (\forall i) \Delta \vdash [\bar{U}/\bar{X}]P_i}{\Delta \vdash T \leq \exists\bar{X} \mathbf{where} \bar{P}. N}$$

Fig. 4. Syntax, entailment, and subtyping for  $\mathcal{E}\mathcal{X}_{uplo}$ .

## 4 Subtyping Existential Types with Upper and Lower Bounds

This section considers the calculus  $\mathcal{E}\mathcal{X}_{uplo}$ , which is similar in spirit to  $\mathcal{E}\mathcal{X}_{impl}$ , but supports upper and lower bounds (*uplo*) for type variables but no implementation constraints. Other researchers [3, 4, 20] use formal systems very similar to  $\mathcal{E}\mathcal{X}_{uplo}$  for modeling Java wildcards [21]. It is not the intention of  $\mathcal{E}\mathcal{X}_{uplo}$  to provide another formalization of wildcards, but rather to expose the essential ingredients that make subtyping undecidable in a calculus as simple as possible. Scala [13] as well as the initial design of the full JavaGl language employ existential types with upper and lower bounds as a replacement for Java wildcards.

### 4.1 Definition of $\mathcal{E}\mathcal{X}_{uplo}$

Fig. 4 defines the syntax and the entailment and subtyping relations of  $\mathcal{E}\mathcal{X}_{uplo}$ . A class type  $N$  is either  $\mathbf{Object}$  or an instantiated generic class  $C\langle\bar{X}\rangle$ , where the type arguments must be type variables. A type  $T$  is a type variable, a class type, or an existential. Unlike in  $\mathcal{E}\mathcal{X}_{impl}$ , existentials in  $\mathcal{E}\mathcal{X}_{uplo}$  may quantify over several type variables, they support multiple constraints, and the body of an existential must be a class type. A constraint  $P$  places either an upper bound ( $X \mathbf{extends} T$ ) or a lower bound ( $X \mathbf{super} T$ ) on a type variable  $X$ . Type environments  $\Delta$  are finite set of constraints  $P$  with  $\Delta, P$  standing for  $\Delta \cup \{P\}$ .

Class definitions and inheritance are omitted from  $\mathcal{E}\mathcal{X}_{uplo}$ . The only assumption is that every class name  $C$  comes with a fixed arity that is respected when applying  $C$  to type arguments. There are some further (implicit) restrictions:

Syntax

$$\begin{aligned}\tau^+ &::= \mathbf{Top} \mid \forall \alpha_0 \leq \tau_0^- \dots \alpha_n \leq \tau_n^- . \neg \tau^- \\ \tau^- &::= \alpha \mid \forall \alpha_0 \dots \alpha_n . \neg \tau^+ \\ \Gamma^- &::= \emptyset \mid \Gamma^-, \alpha \leq \tau^-\end{aligned}$$

$\Gamma^- \vdash \sigma^- \leq \tau^+$

$$\begin{array}{c} \text{D-TOP} \\ \Gamma \vdash \tau \leq \mathbf{Top} \end{array} \quad \begin{array}{c} \text{D-VAR} \\ \tau \neq \mathbf{Top} \\ \Gamma \vdash \Gamma(\alpha) \leq \tau \\ \hline \Gamma \vdash \alpha \leq \tau \end{array} \quad \begin{array}{c} \text{D-ALL-NEG} \\ \Gamma, \alpha_0 \leq \phi_0 \dots \alpha_n \leq \phi_n \vdash \tau \leq \sigma \\ \hline \Gamma \vdash \forall \alpha_0 \dots \alpha_n . \neg \sigma \leq \forall \alpha_0 \leq \phi_0 \dots \alpha_n \leq \phi_n . \neg \tau \end{array}$$

**Fig. 5.** Syntax and subtyping for  $F_{\leq}^D$ .

- (1) If  $T = \exists \bar{X} \mathbf{where} \bar{P} . N$ , then  $\bar{X} \neq \bullet$  and  $\bar{X} \subseteq \text{ftv}(N)$ . That is, an existential must abstract over at least one type variable and all its bounded type variables must appear in the body type  $N$ .
- (2) If  $T = \exists \bar{X} \mathbf{where} \bar{P} . N$  and  $P \in \bar{P}$ , then  $P = Y \mathbf{extends} T$  or  $P = Y \mathbf{super} T$  with  $Y \in \bar{X}$ . That is, only bound variables may be constrained.
- (3) A type variable must not have both upper and lower bounds.<sup>7</sup>

These three restrictions simplify the formulation of a variant of  $\mathcal{E}\mathcal{X}_{uplo}$ 's subtyping relation without an explicit rule for transitivity (see the extended version [26]).

Constraint entailment ( $\Delta \Vdash T \mathbf{extends} U$  and  $\Delta \Vdash U \mathbf{super} T$ ) uses subtyping ( $\Delta \vdash T \leq U$ ) to check that the constraint given holds. The subtyping rules for  $\mathcal{E}\mathcal{X}_{uplo}$  are similar to those for  $\mathcal{E}\mathcal{X}_{impl}$ , except that **Object** is now a supertype of every type and that rules  $s_3\text{-EXTENDS}$  and  $s_3\text{-SUPER}$  use assumptions from  $\Delta$ . Moreover, rule  $s_3\text{-ABSTRACT}$  possibly needs to “guess” some of the types  $\bar{U}$  because, unlike in  $\mathcal{E}\mathcal{X}_{impl}$ , existentials in  $\mathcal{E}\mathcal{X}_{uplo}$  may quantify over more than one type variable.

## 4.2 Undecidability of Subtyping in $\mathcal{E}\mathcal{X}_{uplo}$

To get a feeling how subtyping derivations in  $\mathcal{E}\mathcal{X}_{uplo}$  may lead to infinite regress, consider the goal  $\{X \mathbf{extends} \neg U\} \vdash X \leq \neg C \langle X \rangle$ , where  $U$  is defined as  $\exists X \mathbf{where} X \mathbf{extends} \neg C \langle X \rangle . C \langle X \rangle$  and the notation  $\neg T$  is an abbreviation for  $\exists X \mathbf{where} X \mathbf{super} T . D \langle X \rangle$  such that  $X$  is fresh. Searching for a derivation of this goal quickly leads to a subgoal of the form  $\{X \mathbf{extends} \neg U, Z \mathbf{super} U\} \vdash X \leq \neg C \langle X \rangle$ , where  $Z$  is a fresh type variable introduced by rule  $s_3\text{-OPEN}$ . The details are left as an exercise to the reader.

The undecidability proof of subtyping in  $\mathcal{E}\mathcal{X}_{uplo}$  is by reduction from  $F_{\leq}^D$  [14], a restricted version of  $F_{\leq}$  [5]. Pierce defines  $F_{\leq}^D$  for his undecidability proof of  $F_{\leq}$  subtyping [14]. Fig. 5 recapitulates  $F_{\leq}^D$ 's syntax and subtyping relation. Let  $n$  be a fixed natural number. A type  $\tau$  is either an  $n$ -positive type,  $\tau^+$ , or an  $n$ -negative

<sup>7</sup> Modeling Java wildcards requires upper and lower bounds for the same type variable in certain situations.

$$\begin{aligned}
\llbracket \text{Top} \rrbracket^+ &= \text{Object} \\
\llbracket \forall \alpha_0 \leq \tau_0^- \dots \alpha_n \leq \tau_n^- . \neg \tau^- \rrbracket^+ &= \neg \exists Y, \overline{X^{\alpha_i}} \text{ where } X^{\alpha_0} \text{ extends } \llbracket \tau_0 \rrbracket^- \dots \\
&\quad X^{\alpha_n} \text{ extends } \llbracket \tau_n \rrbracket^-, Y \text{ extends } \llbracket \tau \rrbracket^- \\
&\quad . C^{n+2} \langle Y, \overline{X^{\alpha_i}} \rangle \\
\llbracket \alpha \rrbracket^- &= X^\alpha \\
\llbracket \forall \alpha_0 \dots \alpha_n . \neg \tau^+ \rrbracket^- &= \neg \exists Y, \overline{X^{\alpha_i}} \text{ where } Y \text{ extends } \llbracket \tau \rrbracket^+ . C^{n+2} \langle Y, \overline{X^{\alpha_i}} \rangle \\
\llbracket \emptyset \rrbracket^- &= \emptyset \\
\llbracket \Gamma, \alpha \leq \tau^- \rrbracket^- &= \llbracket \Gamma \rrbracket^-, X^\alpha \text{ extends } \llbracket \tau \rrbracket^- \\
\llbracket \Gamma^- \vdash \tau^- \leq \sigma^+ \rrbracket &= \llbracket \Gamma \rrbracket^- \vdash \llbracket \tau \rrbracket^- \leq \llbracket \sigma \rrbracket^+
\end{aligned}$$

**Fig. 6.** Reduction from  $F_{\leq}^D$  to  $\mathcal{E}\mathcal{X}_{uplo}$ .

type,  $\tau^-$ , where  $n$  stands for the number of type variables (minus one) bound at the top-level of the type. An  $n$ -negative type environment  $\Gamma^-$  associates type variables  $\alpha$  with upper bounds  $\tau^-$ . The polarity (+ or  $-$ ) characterizes at which positions of a subtyping judgment a type or type environment may appear. For readability, the polarity is often omitted and  $n$  is left implicit.

An  $n$ -ary subtyping judgment in  $F_{\leq}^D$  has the form  $\Gamma^- \vdash \sigma^- \leq \tau^+$ , where  $\Gamma^-$  is an  $n$ -negative type environment,  $\sigma^-$  is an  $n$ -negative type, and  $\tau^+$  is an  $n$ -positive type. Only  $n$ -negative types appear to the left and only  $n$ -positive types appear to the right of the  $\leq$  symbol. The subtyping rule  $D$ -ALL-NEG compares two quantified types  $\sigma = \forall \alpha_0 \dots \alpha_n . \neg \sigma'$  and  $\tau = \forall \alpha_0 \leq \tau_0 \dots \alpha_n \leq \tau_n . \neg \tau'$  by swapping the left- and right-hand sides of the subtyping judgment and checking  $\tau' \leq \sigma'$  under the extended environment  $\Gamma, \alpha_0 \leq \tau_0 \dots \alpha_n \leq \tau_n$ . The rule is correct with respect to  $F_{\leq}$  because we may interpret every  $F_{\leq}^D$  type as an  $F_{\leq}$  type:

$$\begin{aligned}
\forall \alpha_0 \dots \alpha_n . \neg \sigma' &\equiv \forall \alpha_0 \leq \text{Top} \dots \forall \alpha_n \leq \text{Top} . \forall \beta \leq \sigma' . \beta \quad (\beta \text{ fresh}) \\
\forall \alpha_0 \leq \tau_0 \dots \alpha_n \leq \tau_n . \neg \tau' &\equiv \forall \alpha_0 \leq \tau_0 \dots \forall \alpha_n \leq \tau_n . \forall \beta \leq \tau' . \beta \quad (\beta \text{ fresh})
\end{aligned}$$

Using these abbreviations, every  $F_{\leq}^D$  subtyping judgment can be read as an  $F_{\leq}$  subtyping judgment. The subtyping relations in  $F_{\leq}^D$  and  $F_{\leq}$  coincide for judgments in their common domain [14].

It is sufficient to consider only *closed judgments*. A type  $\tau$  is closed under  $\Gamma$  if  $\text{ftv}(\tau) \subseteq \text{dom}(\Gamma)$  (where  $\text{dom}(\alpha_1 \leq \tau_1, \dots, \alpha_n \leq \tau_n) = \{\alpha_1, \dots, \alpha_n\}$ ) and, if  $\tau = \forall \alpha_0 \leq \tau_0 \dots \alpha_n \leq \tau_n . \neg \sigma$ , then no  $\alpha_i$  appears free in any  $\tau_j$ . A type environment  $\Gamma$  is closed if  $\Gamma = \emptyset$  or  $\Gamma = \Gamma', \alpha \leq \tau$  with  $\Gamma'$  closed and  $\tau$  closed under  $\Gamma'$ . A judgment  $\Gamma \vdash \tau \leq \sigma$  is closed if  $\Gamma$  is closed and  $\tau, \sigma$  are closed under  $\Gamma$ .

These notions are sufficient to state the central theorem of this section and sketch its proof.

**Theorem 9.** *Subtyping in  $\mathcal{E}\mathcal{X}_{uplo}$  is undecidable.*

*Proof.* The proof is by reduction from  $F_{\leq}^D$ . Fig. 6 defines a translation from  $F_{\leq}^D$  types, type environments, and subtyping judgments to their corresponding  $\mathcal{E}\mathcal{X}_{uplo}$  forms. The translation of an  $n$ -ary subtyping judgment assumes the

existence of two  $\mathcal{E}\mathcal{X}_{uplo}$  classes:  $C^{n+2}$  accepts  $n + 2$  type arguments, and  $D^1$  takes one type argument. The superscripts in  $[\![\cdot]\!]^+$  and  $[\![\cdot]\!]^-$  indicate whether the translation acts on positive or negative entities.

As before, a negated type, written  $\neg T$ , is an abbreviation for an existential with a single **super** constraint:  $\neg T \equiv \exists X \text{ where } X \text{ super } T . D^1 \langle X \rangle$ , where  $X$  is fresh. The **super** constraint simulates the behavior of the  $F_{\leq}^D$  subtyping rule D-ALL-NEG, which swaps the left- and right-hand sides of subtyping judgments.

An  $n$ -positive type  $\forall \alpha_0 \leq \tau_0^- \dots \alpha_n \leq \tau_n^- . \neg \tau^-$  is translated into a negated existential. The existentially quantified type variables  $X^{\alpha_0}, \dots, X^{\alpha_n}$  correspond to the universally quantified type variables  $\alpha_0, \dots, \alpha_n$ . The bound  $[\![\tau]\!]^-$  of the fresh type variable  $Y$  represents the body  $\neg \tau^-$  of the original type. We cannot use  $[\![\tau]\!]^-$  directly as the body because existentials in  $\mathcal{E}\mathcal{X}_{uplo}$  have only class types as their bodies. The translation for  $n$ -negative types is similar to the one for  $n$ -positive types. It is easy to see that the  $\mathcal{E}\mathcal{X}_{uplo}$  types in the image of the translation meet the restrictions defined in Sec. 4.1. Type environments and subtyping judgments are translated in the obvious way.

We now need to verify that  $\Gamma \vdash \tau \leq \sigma$  is derivable in  $F_{\leq}^D$  if and only if  $[\![\Gamma \vdash \tau \leq \sigma]\!]^+$  is derivable in  $\mathcal{E}\mathcal{X}_{uplo}$ . The “ $\Rightarrow$ ” direction is an easy induction on the derivation of  $\Gamma \vdash \tau \leq \sigma$ . The “ $\Leftarrow$ ” direction requires more work because the transitivity rule  $s_3$ -TRANS (Fig. 4) involves an intermediate type which is not necessarily in the image of the translation. Hence, a direct proof by induction on the derivation of  $[\![\Gamma \vdash \tau \leq \sigma]\!]^+$  fails. To solve this problem, we give an equivalent definition of the  $\mathcal{E}\mathcal{X}_{uplo}$  subtyping relation that does not include an explicit transitivity rule. See the extended version [26] for details and the full proof.  $\square$

## 5 Lessons Learned

What are the consequences of this investigation for the design of JavaGI? While bounded existential types are powerful and unify several diverse concepts, they complicate the metatheory of JavaGI’s initial design considerably. Also, subtyping with existential types is undecidable in the general case, as demonstrated in the two preceding sections.

There are three alternatives for dealing with this problem: (1) Accept that the subtyping relation is undecidable. (2) Restrict existentials such that subtyping becomes decidable. (3) Remove existentials from JavaGI altogether.

The first alternative (opted for by the Scala compiler) is pragmatic and readily implementable by imposing a resource limit on the subtype checker to avoid divergence. Consequently, the subtyping algorithm would become incomplete with respect to its specification.

The second alternative turns out not to be viable. It is feasible to come up with a set of restrictions that keep the subtyping relation defined in Sec. 3 decidable. (Sec. 3.3 investigated such restrictions for plain interface types.) For the subtyping relation defined in Sec. 4, however, we were not able to identify sensible restrictions without giving up either lower or upper bounds (which are essential for encoding Java wildcards). Moreover, restricting existential types so

that subtyping becomes decidable would make an already complex type system even more complicated. In addition, such restrictions tend to be difficult to communicate to users of the language.

The third alternative comes with the realization that existentials may not be worth the trouble. `JavaGI`'s main contribution is its very general and powerful interface concept (which this paper does not explore, but see [23, 25]). While existential types are related to this concept, they are not at the heart of it. In fact, we conducted several real-world case studies using our implementation of `JavaGI` without a need for full-blown existential types arising [25].

Under these circumstances, it appears that the price of having bounded existential types at the core of `JavaGI` is too high. Hence, the current, revised version of `JavaGI` elides bounded existential types because of their poor power/cost ratio but retains all other features of the previous design. Thus, it gives up some of the power in favor of simplicity.

Several already existing features make up for the lack of existentials. More specifically, the revised design copes with most uses of existentials from Sec. 1: parametric polymorphism in combination with multiple subtyping constraints (as already present in Java) allows to emulate the composition of interface types in most situations; direct support for wildcards avoids their encoding through existential types;<sup>8</sup> `JavaGI`'s `implements` constraints in combination with parametric polymorphism allow the specification of meaningful types for interfaces over families of types.

Even the revised design has to accept compromises to avoid undecidability. Sec. 3.3 reveals that subtyping remains undecidable even if plain interface types replace existentials. The real culprit for undecidability is the ability to provide implementation definitions with interface types acting as implementing types.

Disallowing such implementation definitions completely is a rather severe restriction because it prevents useful implementation definitions such as the one given for `List<X>` in Sec. 2.2. Instead, the revised design allows interfaces as implementing types but it imposes the equivalent of Restriction 3 from Sec. 3.3: if an interface is used as an implementing type then no retroactive implementation can be provided for this interface.<sup>9</sup>

Sometimes, even this restriction is too strict. For example, the use of `List<X>` as an implementing type in Sec. 2.2 prevents retroactive implementations of the `List` interface. *Abstract implementation definitions* are a potential cure. They look similar to non-abstract implementation definitions but do not contribute to constraint entailment and subtyping, so the restriction just explained does

---

<sup>8</sup> It is an open question whether subtyping for Java wildcards is decidable (see Sec. 6 for details). Of course, the inclusion of wildcards in `JavaGI` is a concession to ensure backwards compatibility with Java 1.5. An embedding of `JavaGI`'s generalized interface concept in other languages such as C# could easily drop support for wildcards. Thus, the decidability question for wildcards is not intrinsic to decidability of subtyping in `JavaGI`.

<sup>9</sup> Restriction 2 is more flexible than Restriction 3 but the latter simplifies the detection of ambiguities arising through conflicting implementation definitions and it allows for an efficient implementation of dynamic method lookup.

not apply. The details of abstract implementation definitions are explained elsewhere [25].

## 6 Related Work

Kennedy and Pierce [9] investigate undecidability of subtyping under multiple instantiation inheritance and declaration-site variance. They prove that the general case is undecidable and present three decidable fragments. The proof in Sec. 3 is similar to theirs, although undecidability has different causes: Kennedy and Pierce’s system is undecidable because of contravariant generic types, expansive class tables, and multiple instantiation inheritance, whereas undecidability of our system is due to implementation definitions for existentials (or interface types), which cause cyclic interface and multiple instantiation subtyping.

Pierce [14] proves undecidability of subtyping in  $F_{\leq}$  by a chain of reductions from the halting problem for two-counter Turing machines. An intermediate link in this chain is the subtyping relation of  $F_{\leq}^D$ , which is also undecidable. Our proof in Sec. 4 works by reduction from  $F_{\leq}^D$  and is inspired by a reduction given by Ghelli and Pierce [6], who study bounded existential types in the context of  $F_{<}^D$  and show undecidability of subtyping. Crucial to the undecidability proof of  $F_{\leq}^D$  is rule D-ALL-NEG: it extends the typing context and essentially swaps the sides of a subtyping judgment. In  $\mathcal{E}\mathcal{X}_{uplo}$ , rule  $s_3$ -OPEN and rule  $s_3$ -ABSTRACT together with lower bounds on type variables play a similar role.

Torgersen and coworkers [20] present WildFJ as a model for Java wildcards using existential types. The authors do not prove WildFJ sound. Cameron and coworkers [4] define a similar calculus  $\exists J$  and prove soundness. However,  $\exists J$  is not a full model for Java wildcards because it does not support lower bounds for type variables. The same authors present with TameFJ [3] a sound calculus supporting all essential features of Java wildcards. WildFJ’s and TameFJ’s subtyping rules are similar to the ones of  $\mathcal{E}\mathcal{X}_{uplo}$  defined in Sec. 4, so the conjecture is that subtyping in WildFJ and TameFJ is also undecidable. The rule XS-ENV of TameFJ is roughly equivalent to the rules  $s_3$ -OPEN and  $s_3$ -ABSTRACT of  $\mathcal{E}\mathcal{X}_{uplo}$ .

Decidability of subtyping for Java wildcards is still an open question [11]. One step in the right direction might be the work of Plümicke, who solves the problem of finding a substitution  $s$  such that  $sT \leq sU$  for Java types  $T, U$  with wildcards [16, 17]. Our undecidability result for  $\mathcal{E}\mathcal{X}_{uplo}$  does not imply undecidability for Java subtyping with wildcards. The proof of this claim would require a translation from subtyping derivations in  $\mathcal{E}\mathcal{X}_{uplo}$  to subtyping derivations in Java with wildcards, which is not addressed in this paper. In general, existentials in  $\mathcal{E}\mathcal{X}_{uplo}$  are strictly more powerful than Java wildcards. For example, the existential  $\exists X.C\langle X, X \rangle$  cannot be encoded as the wildcard type  $C\langle ?, ? \rangle$  because the two occurrences of  $?$  denote two distinct types.

The programming language Scala [13] supports existential types in its latest release to provide better interoperability with Java libraries using wildcards and to address the avoidance problem [15, Chapter 8]. The subtyping rules for existentials (Sec. 3.2.10 and Sec. 3.5.2 of the specification [13]) are very similar to

the ones for  $\mathcal{E}\mathcal{X}_{uplo}$ . This raises the question whether Scala's subtyping relation with existentials is decidable.

A recent article [25] reports on practical experience with the revised design of JavaGI. It discusses several case studies and describes the implementation of a compiler and a runtime system for JavaGI, which employs the restrictions discussed in Sec. 5.

## 7 Conclusion

The paper investigated decidability of subtyping with bounded existential types in the context of JavaGI, Java wildcards, and Scala. It defined two calculi  $\mathcal{E}\mathcal{X}_{impl}$  and  $\mathcal{E}\mathcal{X}_{uplo}$  featuring bounded existential types in two variations and proved undecidability of subtyping for both calculi. Subtyping is also undecidable for  $\mathcal{T}\mathcal{T}_{impl}$ , a simplified version of  $\mathcal{E}\mathcal{X}_{impl}$  without existentials. The paper also suggested a revised version of JavaGI that avoids fully general existentials without giving up much expressivity. The revised version of JavaGI is available at <http://www.informatik.uni-freiburg.de/~wehr/javagi/>.

## Acknowledgments

We thank the anonymous FTfJP 2008 and APLAS 2009 reviewers for feedback on earlier versions of this article.

## References

1. M. Abadi and L. Cardelli. *A Theory of Objects*. Springer, 1996.
2. K. B. Bruce, L. Cardelli, and B. C. Pierce. Comparing object encodings. *Information and Computation*, 155(1-2):108–133, 1999.
3. N. Cameron, S. Drossopoulou, and E. Ernst. A model for Java with wildcards. In J. Vitek, editor, *22nd ECOOP*, volume 5142 of *LNCS*, Paphos, Cyprus, 2008. Springer.
4. N. Cameron, E. Ernst, and S. Drossopoulou. Towards an existential types model for Java wildcards. In *FTfJP, informal proceedings*, 2007. [http://www.doc.ic.ac.uk/~ncameron/papers/cameron\\_ftfjp07\\_full.pdf](http://www.doc.ic.ac.uk/~ncameron/papers/cameron_ftfjp07_full.pdf).
5. L. Cardelli and P. Wegner. On understanding types, data abstraction, and polymorphism. *ACM Comput. Surv.*, 17:471–522, Dec. 1985.
6. G. Ghelli and B. Pierce. Bounded existentials and minimal typing. *Theoretical Computer Science*, 193(1-2):75–96, 1998.
7. J. E. Hopcroft, R. Motwani, and J. D. Ullman. *Introduction to Automata Theory, Languages, and Computation*. Addison Wesley, third edition, 2006.
8. S. Kaes. Parametric overloading in polymorphic programming languages. In H. Ganzinger, editor, *Proc. 2nd ESOP*, volume 300 of *LNCS*, pages 131–144. Springer, 1988.
9. A. J. Kennedy and B. C. Pierce. On decidability of nominal subtyping with variance. In *FOOL/WOOD, informal proceedings*, Jan. 2007. <http://foolwood07.cs.uchicago.edu/program/kennedy-abstract.html>.

10. K. Läuffer. Type classes with existential types. *J. Funct. Program.*, 6(3):485–517, May 1996.
11. K. Mazurak and S. Zdancewic. Type inference for Java 5: Wildcards, F-bounds, and undecidability. <http://www.cis.upenn.edu/~stevez/note.html>, 2006.
12. J. C. Mitchell and G. D. Plotkin. Abstract types have existential types. *ACM TOPLAS*, 10(3):470–502, July 1988.
13. M. Odersky. The Scala language specification version 2.7, Apr. 2008. Draft, <http://www.scala-lang.org/docu/files/ScalaReference.pdf>.
14. B. C. Pierce. Bounded quantification is undecidable. *Information and Computation*, 112(1):131–165, 1994.
15. B. C. Pierce, editor. *Advanced Topics in Types and Programming Languages*. MIT Press, 2005.
16. M. Plümicke. Typeless programming in Java 5.0 with wildcards. In *5th PPPJ*. ACM, Sept. 2007.
17. M. Plümicke. Java type unification with wildcards. In *Applications of Declarative Programming and Knowledge Management*, volume 5437 of *LNCS*, pages 223–240. Springer, 2009.
18. E. L. Post. A variant of a recursively unsolvable problem. *Bulletin of the American Mathematical Society*, 53:264–268, 1946.
19. C. V. Russo. *Types for Modules*. PhD thesis, Edinburgh University, Edinburgh, Scotland, 1998. LFCS Thesis ECS–LFCS–98–389.
20. M. Torgersen, E. Ernst, and C. P. Hansen. Wild FJ. In *FOOL, informal proceedings*, 2005. <http://homepages.inf.ed.ac.uk/wadler/fool/program/14.html>.
21. M. Torgersen, E. Ernst, C. P. Hansen, P. von der Ahé, G. Bracha, and N. Gafter. Adding wildcards to the Java programming language. *Journal of Object Technology*, 3(11):97–116, Dec. 2004.
22. P. Wadler and S. Blott. How to make ad-hoc polymorphism less ad-hoc. In *Proc. 16th ACM Symp. POPL*, pages 60–76, Austin, Texas, USA, Jan. 1989. ACM Press.
23. S. Wehr, R. Lämmel, and P. Thiemann. JavaGI: Generalized interfaces for Java. In E. Ernst, editor, *21st ECOOP*, volume 4609 of *LNCS*, pages 347–372, Berlin, Germany, July 2007. Springer.
24. S. Wehr and P. Thiemann. Subtyping existential types. In *10th FTfJP, informal proceedings*, 2008. <http://www.informatik.uni-freiburg.de/~wehr/publications/subex.pdf>.
25. S. Wehr and P. Thiemann. JavaGI in the battlefield: Practical experience with generalized interfaces. In *Proc. 8th GPCE*, Denver, Colorado, USA, 2009. ACM.
26. S. Wehr and P. Thiemann. On the decidability of subtyping with bounded existential types (extended edition). Technical report, Universität Freiburg, Sept. 2009. <ftp://ftp.informatik.uni-freiburg.de/documents/reports/report250/report00250.ps.gz>.