Dependency-Free Capture Tracking

Edward Lee
Computer Science
University of Waterloo
Waterloo, ON, Canada

Kavin Satheeskumar
Computer Science
University of Waterloo
Waterloo, ON, Canada

Ondřej Lhoták
Computer Science
University of Waterloo
Waterloo, ON, Canada

ABSTRACT
Type systems usually characterize the shapes of values but not usually their free variables. Many desirable safety properties could be guaranteed by the type system if it knew exactly which variables were free in values.

There has been much recent work investigating such systems, with an eventual goal of incorporating a capture tracking system into Scala. These systems are unfortunately complicated by advanced features in Scala’s type system, particularly dependent types. We explore what a capture tracking system could look like without the full complication of dependent types.

CCS CONCEPTS
- Software and its engineering → General programming languages; Compilers.

KEYWORDS
System \( F_{<:}\), Capture Tracking, Type Systems

ACM Reference Format:

1 INTRODUCTION
Type systems usually characterize the shapes of values but not their free variables. This is unfortunate, as many desirable safety properties could be guaranteed by the type system if it knew exactly which variables were free in values.

For example, using the effects-as-capabilities discipline, one can reduce the problem of tracking where potentially unsafe effects could happen to the problem of tracking where capabilities corresponding to those effects could flow in the program. There is an ongoing effort to incorporate such a system into Scala, and a working prototype has been implemented in the Dotty compiler – see Odersky et al. [5].

Formalizing such a system has been difficult. The problem is that a free variable tracking system is inherently dependently typed, in that types track (term-level) variables. Consider an example Scala definition:

```scala
def delayedPrint(s: {*} String) = () => print(s)
```

Here, the \* in the type of \( s \) specifies that capturing \( s \) is to be tracked by the type system. Thus, the return type of `delayedPrint` could be \( \{ s \} \text{Unit} \Rightarrow \text{Unit} \) to indicate that the closure that is returned captures \( s \). A type like `print (s : {*} String) => \{ s \} \text{Unit}` will unfortunately contain term variables.

One attempt to formalize a tracking system by Boruch-Gruszecki et al. [3] runs into complications concerning dependent types, which they solve by enforcing variance restrictions on where type variables can occur. More recently, Odersky et al. [6] sidestep these issues by working in administrative normal form (ANF), a technique previously used by Rapoport and Lhoták [7] and Amin et al. [1] to work around similar complications in formalizing type dependency in DOT, Scala’s core calculus. However, the indirectness of ANF makes it difficult to acquire an intuition for the properties of the system.

This does not need to be the case, though. As Brachthäuser et al. [4] show, a capture tracking system can be done with a simple calculus with a simple mechanized proof soundness even in the presence of dependent types. While their system lacks subtyping, their system differs in one key way compared to Odersky et al. [6] and Boruch-Gruszecki et al. [3] which enables a simple soundness proof to go through while still keeping track of captured variables in types. They annotate their function applications with the capture information which should be substituted instead of relying on the computed capture information on the actual value being substituted in.

Inspired by Brachthäuser et al. [4], we show that a similar technique also works in systems like Boruch-Gruszecki et al. [3] and Odersky et al. [6]. We present a simple calculus System \( F_{<:;<:}\) that extends System \( F_{<:}\) with capture tracking without resorting to ANF nor to variance restrictions on type variables, for which we have mechanized a soundness proof, and we sketch out how System \( F_{<:;<:}\) can be used as a basis of soundness proofs for a surface syntax in a practical programming language. Our development has been mechanized in Coq, and can be found at https://github.com/e45lee/simple-capture-proof.

We hope that with this contribution, capture tracking systems will be better understood, and that System \( F_{<:;<:}\) can serve as the basis of a proof of soundness for capture tracking systems – not only in Scala, but for other languages as well.

2 CAPTURE TRACKING
At its core, capture tracking is concerned with:

1. Tracking what free variables are present in values, and,
2. For functions, tracking how arguments flow into results.
For example, in our earlier example:

```java
def delayedPrint(s: {*} String) = () => print(s)
```

The above term would receive the following type:

```java
(print) (s : {*} String) => (s) Unit => Unit
```

This type describes that delayedPrint is a function that returns a thunk which captures the parameter s passed to delayedPrint.

Applying delayedPrint to a string returns a value whose type reflects that it captures that string. For example, in the following listing, delayedPrint(hello) has type (hello) Unit => Unit.

```scala
val hello: {*} String = "Hello_World".
delayedPrint(hello) // returns thunk with type (hello)
```

### 2.1 Dependent Types

Dependent types pose a problem. Consider the following example.

```scala
val goodbye: {*} String = "Goodbye_World".
def wantsEqual = {s: {*} () => String} =>
  {t: (s) => String} => ( ... )
wantsEqual
  (if true then () => hello else () => goodbye)
  () => goodbye

Here, we declare that goodbye is a String which should be tracked, and wantsEqual is a curried function that expects a value s that can capture anything and a value t that captures the same variables as s does. We then apply wantsEqual to a conditional expression that returns two thunks: one returning hello and one returning goodbye:

```scala
if true then () => hello else () => goodbye
```

Now, what should the type of this conditional term be? Both branches return a thunk returning a String, but one captures hello and the other goodbye. All we can say is that the conditional could possibly capture hello or goodbye, giving the conditional the type (hello, goodbye) () => String. Thus, wantsEqual(if true then () => hello else () => goodbye) has type ((hello, goodbye) () => String) => T for some T. Accordingly, we widen the type of the second argument of wantsEqual from (goodbye) () => String to (hello, goodbye) () => String. The overall program is well typed.

Now, consider what happens after a single step of reduction, which reduces the conditional term to just the thunk () => hello. The overall program reduces to:

```
wantsEqual(() => hello)() => goodbye // what type does this function have?
```

The next step will be to reduce the application wantsEqual(() => hello). As is standard for function application, we substitute the argument () => hello for the parameter s in the body of wantsEqual. But s occurs not only in the body of the function, but also in the type of the second parameter t, which is (s) () => String. What should be substituted for s in the type of t?

There are a few options. One option is to substitute with the actual set of variables that are captured by the value, namely the free variables of () => hello. This is given by the substitution rule:

\[
(\lambda(x : T). t)[v] \rightarrow t[x \mapsto \text{fv}(o)] \quad (\text{BETA-V-BAD})
\]

However, this cannot be done so easily. In our example, if we replaced s with hello in the types of the body of wantsEqual, we end up with the term:

\[
((t: \text{hello}) () \Rightarrow \text{String}) \Rightarrow \text{...})() \Rightarrow \text{goodbye}
\]

This is obviously ill-typed, since the argument captures goodbye but is required by the parameter type to capture hello. The reduction in this example violates type preservation. What went wrong? In this example two things went wrong for this to transpire:

1. The capture set that we had "assigned" to s shrank under reduction from (hello, goodbye) to just (hello) and, and,
2. We used s in a negative, contravariant position – namely, in the type annotation for what we expect to be passed in the second parameter t.

This breaks preservation under reduction, as contravariance flips the subtyping relation. It would be unsound for s to shrink in this manner. It must either stay static or expand. In order for preservation to hold, we need to prevent this from happening...but how?

One possibility is to simply disallow contravariant occurrences of term variables in types, which is the approach taken by Boruch-Gruszecki et al. [3]. This is restrictive, however, and complicates the proof of soundness, which needs to enforce and check that term variables occur only covariantly in types, even under reduction.

Another possibility is to make what we “assign” to s static, by putting programs in administrative normal form, in effect assigning a name to every (sub)expression in a program:

```scala
val tmp =
  if true then () => hello else () => goodbye
wantsEqual(tmp) // expects (tmp) String
```

Now, since we have named the conditional term tmp, we can give wantsEqual(tmp) the type ((tmp) () => String) => T. Since tmp is named, we can use the capture set (tmp) to represent exactly what tmp captures, even as what is stored in tmp shrinks under reduction from something capturing (hello, goodbye) to just hello.

```scala
val tmp = () => hello // reduced
wantsEqual(tmp) // still expects (tmp) String.
```

This is the approach taken by Odersky et al. [6]. Similar approaches have been used to handle variance issues in other contexts – for example, Rapoport and Lhoták [7] and Amin et al. [1]. However, administrative normal form adds complexity to the reduction rules. Moreover, it does not even improve expressiveness for this example. While wantsEqual can be typed, the only variable that can be passed to wantsEqual(tmp) is either tmp itself or a variable that does not capture any other tracked variables. In particular, the ANF-converted version of the original term wantsEqual(tmp)() => goodbye remains ill-typed.

In this paper, in Section 3, we explore a third approach, which is to annotate each term application in the original program with the capture set that will be substituted in place of that term variable in capture set position. That set is fixed in the original program and does not shrink with reduction. Formally, we replace (BETA-V-BAD) with the following:

\[
(\lambda(x : T). t)[C] (v) \rightarrow t[x \mapsto \text{type } C] \quad (\text{BETA-V})
\]
This neatly sidesteps the issue of shrinking variables and variance, at the cost of annotating each term application with a capture set. The resulting system satisfies type preservation and allows us to type the `wantsEqual` example with the capture set annotation `{hello, goodbye}:

```plaintext
wantsEqual
  (hello, goodbye)
  if true then () => hello else () => goodbye
  (hello, goodbye)() => goodbye
  =>

wantsEqual
  (hello, goodbye)() => hello
  (hello, goodbye)() => goodbye
  =>
  (t: (hello, goodbye) () => String)
  (hello, goodbye)() => goodbye
```

The System C calculus of Brachthäuser et al. [4] also uses this approach, but in the context of a rather different calculus.

### 2.2 Types, Shapes, and Polymorphism

Another issue that arises is that of parametric polymorphism. What should a type variable \( X \) quantify over, now that types contain both information describing the shape of a value, as well as what that value captures? The seemingly obvious answer is that \( X \) should stand for both the shape and the capture information of a value, but this imposes some restrictions. For one, if \( X \) is a type variable that stands for a full type, how do we interpret a type like \( \text{hello} \ X \)? What does the type \( \text{hello} \) even mean? Do we somehow combine the two capture sets (\( \text{hello} \)) and (\( \text{goodbye} \))?

One solution is to prevent type variables from further being annotated by capture sets. This has the advantage of being simple conceptually at first glance. However, it comes at the cost of complexity in the calculus: Boruch-Gruszeczi et al. [3] require two very distinct typing rules for term variables depending on what sort of type they are bound to. Furthermore, complications arise in the presence of generic data types. Consider an arbitrary list type `List[X]` describing a list containing elements of type \( X \), and the generic map function over said lists. What should the type of map be - in particular, what should the return type of map be?

```plaintext
def map(X,Y)(l: {*} List[X], f: {*} X => Y) : {??} List[Y]
```

Assuming map does not capture anything else, the values that could be captured in the resulting list must come from \( f \) and the elements of \( l \). Naming the former is easy in a capture set, but not the latter, as the result does not capture \( 1 \) directly but the elements of \( l \), which do not have any name in this scope. We could have the capture set \( 1 \) stand in for all its elements when \( l \) is given a `List` type, but that requires special support in the calculus for every data structure.

To sidestep this issue, Odersky et al. [6] require generic data structures to hold only pure elements with empty capture sets. This would be impractically restrictive, however, so they add a special mechanism called `boxing` to "tunnel" captures, by temporarily making impure values pure so they can be stored in a generic data structure, but marking their type to ensure that the necessary captures are covered by the type when the element is later retrieved from the data structure.

In this paper, as we discuss in Section 3, we use a more familiar mechanism, for which soundness proofs are well understood: an additional binding form for variables that stand for capture sets. Thus, we separate the binders for the shape part and the capture part of a type. In this setting, we could give map the following signature:

```plaintext
def map[C,X,Y](l: {*} List[{C} X], f: {*} X => Y) : {C, f} List[{C, f} Y]
```

Here, the capture set binder allows us to introduce a name \( C \) for the capture set of the elements of \( l \).

### 3 A SIMPLER CALCULUS

To this end, we develop a simpler capture tracking calculus System \( F_{\text{C}} \) inspired by these observations. As discussed in Section 2.3, we explicitly annotate term applications with the capture sets which should be substituted in. Specifically, we define application reduction to be as below:

\[(\lambda (x : T).t) [C] (v) \rightarrow t [x \mapsto \text{type } C] [x \mapsto \text{term } v] \text{ (BETA-V)}\]

Obviously, though, \( C \) needs to bound to what could be carried in \( v \) - this is handled through the typing rule `APP`.

\[\Gamma \vdash t : \{D\} (x : \{C\} S) \rightarrow T \quad \Gamma \vdash s : \{C\} S \quad \Gamma \vdash t : \{C\} S \text{ (APP)}\]

To handle parametric polymorphism, as discussed in Section 2.2, we adopt a split shape-type system with polymorphism on shape variables. We support capture polymorphism directly using a third binder \( \Lambda (Y < C) t \) for capture set variables. This allows for straightforward mechanized proofs of Progress and Preservation, without requiring variance restrictions or administrative normal form.

**Theorem 3.1 (Preservation).** Suppose \( \Gamma \vdash s : T \), and \( s \rightarrow t \).

Then \( \Gamma \vdash t : T \) as well.

**Theorem 3.2 (Progress).** Suppose \( \emptyset \vdash s : T \). Either \( s \) is a value, or \( s \rightarrow t \) for some term \( t \).

Our development has been mechanized in Coq, based on the mechanization System \( F_{C} \) of Aydemir et al. [2]. While there are many lemmas in our mechanization, due to our three binding forms, all of those proofs were straightforward and were not difficult to discharge. The only proofs that were difficult were those that dealt with sets, which is mainly due to lack of good automation for working with sets in Coq.

### 3.1 The Burden of Annotations?

One might object to the extra annotation required on term application. After all, one has to find a \( C \) - a capture set - to fill in there. However, this is not such an onerous requirement. As Brachthäuser et al. [4] observe, one infers the appropriate annotation anyways while type checking. Concretely, with the same typing rules, we could devise a calculus System \( F_{C,0} \) without reduction rules, but with almost the same syntax and typing rules, with the following two changes to syntax and typing:
\[
s, t ::= \begin{array}{ll}
\lambda(x : T).t & \text{term abstraction} \\
\Lambda(X << S).t & \text{shape abstraction} \\
\Lambda(Y << C).t & \text{capture abstraction} \\
x & \text{term variable} \\
s[C](t) & \text{application} \\
s[S] & \text{shape application} \\
s[C] & \text{capture application} \\
\end{array}
\]

\[
\Gamma ::= \begin{array}{ll}
\text{Environment} & \text{empty} \\
\Gamma, x : T & \text{term binding} \\
\Gamma, X << S & \text{shape binding} \\
\Gamma, Y << C & \text{capture binding} \\
\end{array}
\]

\[
S ::= \begin{array}{ll}
X & \text{function variable} \\
(x : T_1) \rightarrow T_2 & \text{function shapes} \\
\forall(X << S).T & \text{for-all shapes} \\
\forall(Y << C).T & \text{capture for-all shapes} \\
\end{array}
\]

\[
T ::= \begin{array}{ll}
\{C, S\} & \text{capturing type} \\
\end{array}
\]

\[
C, D ::= \begin{array}{ll}
\{x_1, x_2, \ldots, Y_1, Y_2, \ldots\} & \text{proper set} \\
\top & \text{universal set} \\
\text{where } Y & \text{capture variable} \\
\end{array}
\]

\[\Gamma \vdash t : T\]

\[\Gamma \vdash \lambda(x : T).t : \{f_0(x) = x\} (x : T_1) \rightarrow T_2\]

\[\Gamma \vdash \Lambda(X << S).T : (f_0(x)) \forall(X << S).T\]

\[\Gamma \vdash \Lambda(Y << C).T : (f_0(x)) \forall(Y << C).T\]

\[\Gamma \vdash t : \{D\} (x : \{C\} S) \rightarrow T\]

\[\Gamma \vdash t : \{D\} \forall(Y << C).T\]

\[\Gamma \vdash t[C'] : T[Y \mapsto C']\]

(1) Term application omits the capture set annotation \(-f_C(v)\) instead of \(f[C](v)\).

(2) The typing judgment is updated to reflect this:
Evaluation

\[
\begin{align*}
\lambda (x : T).t \triangleright [C](v) & \rightarrow t \{ x \mapsto \text{type } C \} [x \mapsto \text{term } v] \quad \text{(BETA-V)} \\
\Lambda (X : S_b).t[S] & \rightarrow t[X \mapsto S] \quad \text{(BETA-S)} \\
\Lambda (Y : C_b).t[C] & \rightarrow t[Y \mapsto C] \quad \text{(BETA-C)} \\
E[s] & \rightarrow E[t] \quad \text{(CONTEXT)}
\end{align*}
\]

\[
E ::= [\vphantom{\lambda}] | E[C](t) | s[C](E) \quad \text{Evaluation Context}
| E[S] \\
| E[C]
\]

Figure 5: Reduction rules for System \( F_{<:C} \)

At the surface level, System \( F_{<:C} \) looks almost like a dependently-typed calculus without capture annotations in the same style as Boruch-Gruszecki et al. \[3\]. This allows System \( F_{<:C} \) to serve as the basis for the surface-level syntax of a capture tracking system in a programming language, like Scala. It does not contain annotations like System \( F_{<:C} \). But System \( F_{<:C} \) can be used to show that well-typed System \( F_{<:C} \) programs behave properly as well. A typing derivation for a System \( F_{<:C} \) program is in one-to-one correspondence with a typing derivation for a System \( F_{<:C} \) program – simply lift out the Cs that appear in the (APP-OMIT) rule into the program itself. Progress and preservation in System \( F_{<:C} \) guarantee that the resulting System \( F_{<:C} \) program will reduce properly and not get stuck. Note, however, that the program will reduce according to the inferred capture sets on term application, unlike previous dependently-type capture tracking calculi. This avoids the restrictions that Odersky et al. \[6\] and Boruch-Gruszecki et al. \[3\] imposed on their systems.

4 FUTURE WORK

System \( F_{<:C} \) contributes a simple and sound basis for an expressive capture tracking system in a System \( F_{<:C} \)-based programming language. It does so by adding capture set annotations at every term application. Some existing systems share a similar property, in that the capture set to be substituted when the term is applied can be inferred by simply looking at the term. Notably, this is true of Odersky et al. \[6\]’s System \( CC_{<:C} \). Term application in their ANF-based calculus reduces as follows:

\[
f \text{is bound to } \lambda x.t \text{ in the enclosing ANF context} \\
f(y) \rightarrow t[x \mapsto \text{type } y] [x \mapsto \text{term } y] \quad \text{(BETA-V-VAR)}
\]

It would be interesting to explore whether our simpler System \( F_{<:C} \) could serve as a basis for showing the soundness of System \( CC_{<:C} \). More generally, it would be interesting to see if System \( F_{<:C} \) could serve as the basis of soundness proofs for other capture tracking systems which may come about in the future.

5 CONCLUSION

We contributed a simple and sound treatment of capture tracking as an extension of System \( F_{<:C} \), without the complications involving dependent types that arise in existing attempts. We show that simply annotating term applications with static capture sets allows for a simpler treatment of soundness in a core calculus, without the need for variance restrictions or ANF. We hope that this will contribute to better intuitive understanding of capture tracking systems and that System \( F_{<:C} \) can serve as the basis of a proof of soundness for capture tracking systems – not only for Scala, but for other languages as well.

REFERENCES


Received 2023-05-26; accepted 2023-06-23