

Location Types for Safe Distributed Object-Oriented Programming*

Yannick Welsch and Jan Schäfer

University of Kaiserslautern, Germany
{welsch, jschaefer}@cs.uni-kl.de

Abstract. In distributed object-oriented systems, objects belong to different *locations*. For example, in Java RMI, objects can be distributed over different JVM instances. Accessing a reference in RMI has crucial different semantics depending on whether the referred object is local or remote. Nevertheless, such references are not statically distinguished by the type system.

This paper presents *location types*, which statically distinguish *far* from *near* references. We present a formal type system for a minimal core language. In addition, we present a type inference system that gives optimal solutions. We implemented location types as a pluggable type system for the ABS language, an object-oriented language with a concurrency model based on *concurrent object groups*. An important contribution of this paper is the combination of the type system with the flexible inference system and a novel integration into an Eclipse-based IDE by presenting the inference results as overlays. This drastically reduces the annotation overhead while providing full static type information to the user. The IDE integration is a general approach of its own and can be applied to many other type system extensions.

1 Introduction

In distributed object-oriented systems, objects belong to different *locations*. A location in this paper is regarded to be an abstract concept, but in practice it may, for example, refer to a physical computation node, some process (like a JVM instance in RMI [19]), or can even be a concept of a programming language. For example, in object-languages with concurrency models based on communicating groups of objects such as E [18], AmbientTalk/2 [24], JCoBox [22], or ABS [15], the location of an object can be considered as the group it belongs to. In these scenarios it often makes a difference whether a reference points to an object at the current location, i.e., the location of the current executing object (in the following called a *near* reference), or to an object at a different location (a *far* reference). For example, in the E programming language [18], a *far* reference can only be used for *eventual sends*, but not for *immediate* method calls. In Java RMI accessing a remote reference may throw a `RemoteException`, where accessing a normal reference cannot throw such an exception. It is thus desirable to be able to statically distinguish these two

* This research is funded by the EU project FP7-231620 HATS: Highly Adaptable and Trustworthy Software using Formal Models

kinds of references. This is useful for documentation purposes, to reason about the code, and to statically prevent runtime errors.

We present *location types* which statically distinguish far from near references. Location types can be considered as a lightweight form of ownership types [4, 21] with the following two characteristics. The first is that location types only describe a *flat set* of locations instead of a *hierarchy* of ownership contexts. The second is that ownership types typically define the ownership context of an object in a precise way. Location types abstract from these precise locations by only stating whether an object belongs to the current location or some other location. These two simplifications make location types very lightweight and easy to use, while still being expressive enough to guarantee their desired properties. Location types are *not* used to enforce encapsulation, which is the main goal of many ownership type systems.

As with any type system extension, writing down the extended types can become tiresome for programmers. Furthermore, these annotations may clutter up the code and reduce readability, especially when several of such pluggable type systems [2, 9] are used together. This reduces the acceptance of pluggable type systems in practice. The first issue can be solved by automatically inferring the type annotations and inserting them into the code. But this results again in cluttered code with potentially many annotations. Our solution is to leverage the power of an IDE and present the inferred types to the programmer by using unobtrusive *overlays*. They give the programmer full static type information without cluttering the code with annotations nor reducing readability. The overlays can be turned on and off according to the programmer's need. Type annotations are only needed to make the type checking and inference modular, where the degree of modularity just depends on the interfaces where type annotations appear. This way of integrating type inference into the IDE drastically simplifies the usage of the proposed type system and is applicable to similar type system extensions.

Contributions. The three main contributions of this paper are the following. (1) We give the formalization of a type system for location types in a core object-oriented language. (2) We describe a type inference system that gives optimal solutions and helpful error messages. (3) We present an implementation of the type and inference system for the ABS language and show how to integrate such a system into an IDE by using a novel way of visualizing inferred type information.

Outline. The remainder of this paper is structured as follows. In Sect. 2 we give an informal introduction to location types and illustrate their usage by an example. Section 3 presents the formalization of location types for a core object-oriented language and the inference system. In Sect. 4 we explain how we implemented and integrated location types into an IDE, and provide a short evaluation. Section 5 discusses location types in the context of related work. Section 6 concludes.

2 Location Types at Work

Location types statically distinguish *far* from *near* references. To do so, standard types are extended with additional type annotations, namely *location types*. There

are three different location types: Near, Far, and Somewhere. Location types are always interpreted *relatively* to the current object. A variable typed as Near means that it may only refer to objects that belong to the *same* location as the current object. Accordingly, a Far typed variable may only refer to objects that belong to a *different* location than the current object. Somewhere is the super-type of Far and Near and means that the referred object may either be Near or Far. Important to note is that only Near precisely describes a certain location. A Far annotation only states that the location of the referred object is *not* Near. This means that a Far typed variable may over time refer to different locations which are not further defined, except that they are not the location of the current object. What a location actually means is irrelevant to the type system. So whether the location of an object refers to a JVM instance or has some other form of object grouping does not matter. It is only important that an object belongs to a unique location for its entire lifetime.

We illustrate the location type system by applying it to a small implementation of a chat application. For the description we use the abstract behavioral specification language (ABS) [15], which we explain hand-in-hand with the example.

ABS is an object-oriented language with a Java-like syntax. It has a concurrency model that is based on so-called concurrent object groups (COGs). COGs can be regarded as the unit of concurrency and distribution in ABS. Every object in ABS belongs to exactly one unique COG for its entire lifetime. This is similar to the Java RMI setting where objects belong to certain JVM instances, which may run distributed on different machines. At creation time of an object it is specified whether the object is created in the current COG (using the standard **new** expression) or is created in a fresh COG (using the **new cog** expression). Communication in ABS between different COGs happen via *asynchronous method calls* which are indicated by an exclamation mark (!). A reference in ABS is *far* when it targets an object of a different COG, otherwise it is a *near* reference. Similar to the E programming language [18], ABS has the restriction that synchronous method calls (indicated by the standard dot notation) are only allowed on near references. Using a far reference for a synchronous method call results in a runtime exception. Our location type system can be used to statically guarantee the absence of these runtime exceptions.

The chat application is a simple IRC-like application, which consists of a single server and multiple clients. For simplicity, there is only a single chat room, so all clients actually broadcast their messages to all other clients. The basic interfaces of the chat application in the ABS language are given in Fig. 1. Note that only Server, Client, and Session are actually reference types, the types Unit, ClientMsg, and ServerMsg are *data types* and represent immutable data and not objects.

Figure 2 shows a possible runtime structure of the chat application. As the clients and the server run independently of each other, they live in their own COGs. This means that all references between clients and the server are far references. The Session objects that handle the different connections with the clients live in the same COG as the Server object. This means that references between Session and Server are near references. In a typical scenario, the client calls the connect method of the server and passes a reference to itself and a user name as arguments. The server then returns a reference to a Session object, which is used by the client to send messages

```

interface Server {
  [Near] Session connect(
    [Far] Client c, String name); }
interface Session {
  Unit receive(ClientMsg m);
  Unit close(); }
interface Client {
  Unit connectTo([Far] Server s);
  Unit receive(ServerMsg m); }

```

Fig. 1. The annotated interfaces of the chat application.

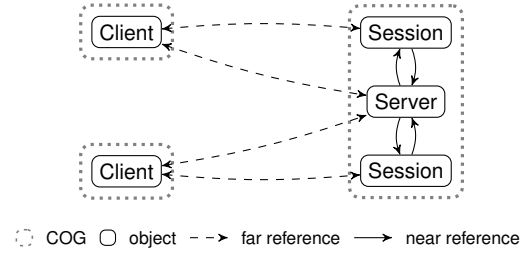


Fig. 2. Runtime structure of the chat application.

```

1 class ClientImpl(String name) implements Client {
2   [Far] Session session; ...
3   Unit connectTo([Far] Server server) {
4     Fut<[Far] Session> f = server!connect(this, name);
5     session = f.get; } }

```

Fig. 3. Fully annotated implementation of the ClientImpl class.

to the server. The interfaces of Fig. 1 are annotated accordingly, e.g., the connect method of the server returns a reference to a Session object that is Near to the server.

Figure 3 shows the ClientImpl class, an implementation of the Client interface. It has a field session which stores a reference to the Session object which is obtained by the client when it connects to the server. Lines 3-5 show the connectTo method. As specified in the interface, the Server parameter has type Far. In Line 4, the client asynchronously (using the ! operator) calls the connect method of the server. The declared result type of the connect method is [Near] Session (see Fig. 1). The crucial fact is that the type system now has to apply a *viewpoint adaptation* [7]. As the target of the call (server) has location type Far, the return type of connect (which is Near) is adapted to Far. Furthermore, as the call is an asynchronous one, the value is not directly returned, but a future instead (i.e. a placeholder for the value). In Line 5, the client waits for the future to be resolved.

Figure 4 shows the ServerImpl class, an implementation of the Server interface. It has an internal field sessions to hold the sessions of the connected clients. List is a polymorphic data type in ABS whose type parameter is instantiated with [Near] Session, which means that it holds a list of near references to Session objects. When a client connects to the server using the connect method, the server creates a new SessionImpl object in its current COG (using the standard **new** expression), which means that it is statically clear that this object is Near. It then stores the reference in its internal list, publishes that a new client has connected, and returns a reference to the session object. In the publish method at Line 16, the send method is synchronously called. As ABS requires that synchronous calls are only done on near objects, the type system guarantees that s always refers to a near object.

```

1  class ServerImpl implements Server {      10  Unit publish(ServerMsg m) {
2    List<[Near] Session> sessions = Nil;    11    List<[Near] Session> sess =
3    [Near] Session connect(                12    sessions;
4    [Far] Client c, String name) {        13    while (~isEmpty(sess)) {
5    [Near] Session s =                    14    [Near] Session s = head(sess);
6    new SessionImpl(this, c, name);      15    sess = tail(sess);
7    sessions = Cons(s,sessions);          16    s.send(m);
8    this.publish(Connected(name));        17    }} ...
9    return s; }                            18  }

```

Fig. 4. Fully annotated implementation of the ServerImpl class.

3 Formalization

This section presents the formalization of the location type system in a core calculus called LocJ. We first present the abstract syntax of the language and its dynamic semantics. In Sect. 3.1 we introduce the basic type system for location types as well as its soundness properties. In Sect. 3.2 we improve the precision of the basic type system by introducing named Far types. In Sect. 3.3 we present the location type inference system.

Notations. We use the overbar notation \bar{x} to denote a list. The empty list is denoted by \bullet and the concatenation of list \bar{x} and \bar{y} is denoted by $\bar{x} \cdot \bar{y}$. Single elements are implicitly treated as lists when needed. $\mathcal{M}[x \mapsto y]$ yields the map \mathcal{M} where the entry with key x is updated with the value y , or, if no such key exists, the entry is added. The empty map is denoted by $[\]$ and $\text{dom}(\mathcal{M})$ and $\text{rng}(\mathcal{M})$ denote the domain and range of the map \mathcal{M} .

Abstract Syntax. LocJ models a core sequential object-oriented Java-like language, formalized in a similar fashion to Welterweight Java [20]. The abstract syntax is shown in Fig. 5. The main difference is that objects in LocJ can be created at different *locations*. For this, the new-expression has an additional argument, given by the in part, that specifies the target location. The target can either be fresh to create the object in a new (fresh) location, or a variable x to create the object in the same location as the object that is referenced by x ¹. We do not introduce locations as first class citizens as they can be encoded using objects, i.e., objects can be simply used to denote locations. To keep the presentation short, LocJ does not include inheritance and subtyping. However, the formalization can be straightforwardly extended to support these features.

Dynamic Semantics. The dynamic semantics of our language is defined as a small-step operational semantics. The main difference to standard object-oriented languages is that we explicitly model locations to partition the heap. The runtime entities are shown in Fig. 6. Runtime configurations ζ consist of a stack, which is a list of

¹ In ABS, **new cog** C() creates a new location (i.e., corresponds to "new c in fresh" in LocJ) whereas **new** C() creates a new object in the same location as the current object (i.e., corresponds to "new c in this" in LocJ).

$$\begin{array}{l}
P ::= \bar{C} \\
C ::= \text{class } c \{ \bar{V} \bar{M} \} \\
V ::= T \ x \\
M ::= T \ m(\bar{V}) \{ \bar{V} \bar{S} \} \\
S ::= x \leftarrow E \mid x.f \leftarrow y \\
E ::= \text{new } c \text{ in fresh} \\
\quad \mid \text{new } c \text{ in } x \mid x \\
\quad \mid x.m(\bar{y}) \mid x.f \\
T ::= c
\end{array}$$

Fig. 5. Abstract syntax of LocJ. c ranges over class names, m over method names and x, y, z, f over field and variable names (including this and result).

$$\begin{array}{l}
\zeta ::= \bar{F}, \mathcal{H} \quad \text{runtime config.} \\
\mathcal{H} ::= \iota \mapsto (l, c, \mathcal{D}) \quad \text{heap} \\
\mathcal{F} ::= (\bar{S}, \mathcal{D})^{c,m} \quad \text{stack frame} \\
\mathcal{D} ::= x \mapsto v \quad \text{variable-value map} \\
v ::= \iota \mid \text{null} \quad \text{value}
\end{array}$$

Fig. 6. Runtime entities of LocJ. ι ranges over object identifiers and l over locations.

stack frames, and a heap. The heap maps object identifiers to object states (l, c, \mathcal{D}) , consisting of a location l , a class name c , and a mapping from field names to values \mathcal{D} . A stack frame consists of a list of statements and a mapping from local variable names to values. Furthermore the stack frame records with which class c and method m it is associated, which we sometimes omit for brevity.

The reduction rules are shown in Fig. 7. They are of the form $\zeta \rightsquigarrow \zeta'$ and reduce runtime configurations. The rules use the helper functions `initO` and `initF` to initialize objects and stack frames. The function `initO`(l, c) creates a new heap entry (l, c, \mathcal{D}) where $\mathcal{D} = [] [f \mapsto \text{null}]$ and f are the field names of class c . Similarly, `initF`(m, c, ι, \bar{v}) creates a new stack frame $(\bar{S}, \mathcal{D})^{c,m}$ where \bar{S} are the statements in the method body of method m in class c and $\mathcal{D} = [] [\text{this} \mapsto \iota] [\text{result} \mapsto \text{null}] [\bar{x} \mapsto \bar{v}] [\bar{y} \mapsto \text{null}]$ and \bar{x} are the variable names of the formal parameters of method m in class c and \bar{y} are the local variable names.

$$\begin{array}{c}
\frac{\iota \notin \text{dom}(\mathcal{H}) \quad l \text{ is fresh}}{\mathcal{H}' = \mathcal{H}[\iota \mapsto \text{initO}(l, c)] \quad \mathcal{D}' = \mathcal{D}[x \mapsto \iota]} \\
\hline
(x \leftarrow \text{new } c \text{ in fresh} \cdot \bar{S}, \mathcal{D}) \cdot \bar{F}, \mathcal{H} \rightsquigarrow (\bar{S}, \mathcal{D}') \cdot \bar{F}, \mathcal{H}' \\
\\
\frac{\iota \notin \text{dom}(\mathcal{H}) \quad (l, _ , _) = \mathcal{H}(\mathcal{D}(y))}{\mathcal{H}' = \mathcal{H}[\iota \mapsto \text{initO}(l, c)] \quad \mathcal{D}' = \mathcal{D}[x \mapsto \iota]} \\
\hline
(x \leftarrow \text{new } c \text{ in } y \cdot \bar{S}, \mathcal{D}) \cdot \bar{F}, \mathcal{H} \rightsquigarrow (\bar{S}, \mathcal{D}') \cdot \bar{F}, \mathcal{H}' \\
\\
\frac{\iota = \mathcal{D}(x) \quad (l, c, \mathcal{D}') = \mathcal{H}(\iota)}{\mathcal{D}'' = \mathcal{D}'[f \mapsto \mathcal{D}(y)] \quad \mathcal{H}' = \mathcal{H}[\iota \mapsto (l, c, \mathcal{D}'')]} \\
\hline
(x.f \leftarrow y \cdot \bar{S}, \mathcal{D}) \cdot \bar{F}, \mathcal{H} \rightsquigarrow (\bar{S}, \mathcal{D}) \cdot \bar{F}, \mathcal{H}' \\
\\
\frac{(_ , _ , \mathcal{D}'') = \mathcal{H}(\mathcal{D}(y)) \quad \mathcal{D}' = \mathcal{D}[x \mapsto \mathcal{D}''(f)]}{(x \leftarrow y.f \cdot \bar{S}, \mathcal{D}) \cdot \bar{F}, \mathcal{H} \rightsquigarrow (\bar{S}, \mathcal{D}') \cdot \bar{F}, \mathcal{H}} \\
\\
\frac{\mathcal{D}' = \mathcal{D}[x \mapsto \mathcal{D}(y)]}{(x \leftarrow y \cdot \bar{S}, \mathcal{D}) \cdot \bar{F}, \mathcal{H} \rightsquigarrow (\bar{S}, \mathcal{D}') \cdot \bar{F}, \mathcal{H}} \\
\\
\frac{\mathcal{F} = (x \leftarrow y.m(\bar{z}) \cdot \bar{S}, \mathcal{D})}{(_ , c, _) = \mathcal{H}(\mathcal{D}(y))} \\
\frac{\mathcal{F}' = \text{initF}(c, m, \mathcal{D}(y), \mathcal{D}(\bar{z}))}{\mathcal{F} \cdot \bar{F}, \mathcal{H} \rightsquigarrow \mathcal{F}' \cdot \mathcal{F} \cdot \bar{F}, \mathcal{H}} \\
\\
\frac{\mathcal{F} = (x \leftarrow y.m(\bar{z}) \cdot \bar{S}, \mathcal{D}')}{\mathcal{D}'' = \mathcal{D}'[x \mapsto \mathcal{D}(\text{result})]} \\
\hline
(\bullet, \mathcal{D}) \cdot \mathcal{F} \cdot \bar{F}, \mathcal{H} \rightsquigarrow (\bar{S}, \mathcal{D}'') \cdot \bar{F}, \mathcal{H}
\end{array}$$

Fig. 7. Operational semantics of LocJ.

$$\begin{aligned}
T & ::= \dots \mid L \ c && \text{annotated type} \\
L & ::= \text{Near} \mid \text{Far} \mid \text{Somewhere} && \text{location type}
\end{aligned}$$

Fig. 8. Basic location types.

3.1 Basic Location Type System

In this subsection, we present the basic location type system and its soundness properties. To incorporate location types into LocJ programs, we extend types T with location types L (see Fig. 8), where a location type can either be Near, Far, or Somewhere. We assume that a given program is already well-typed using a standard Java-like type system and we only provide the typing rules for typing the location type extension. The typing rules are shown in Fig. 9. Statements and expressions are typed under a type environment \bar{V} , which defines the types of local variables. The typing judgment for expressions is of the form $\bar{V} \vdash e : L$ to denote that expression e has location type L . The helper functions $\text{anno}(c, f)$ and $\text{anno}(c, m, x)$ return the declared location type of field f or variable x of method m in class c and $\text{params}(c, m)$ returns the formal parameter variables of method m in class c .

The crucial parts of the type system are the subtyping ($L <: L'$) and the viewpoint adaptation ($L \triangleright_K L'$) relations which are shown in Fig. 10. The location types Near and Far are both subtypes of Somewhere but are unrelated otherwise. Viewpoint adaptation is always applied when a type is used in a different context. There are two different directions ($K \in \{\text{From}, \text{To}\}$) to consider. (1) Adapting a type L from another viewpoint L' to the current viewpoint, written as $L \triangleright_{\text{From}} L'$. (2) Adapting a type L from the current viewpoint to another viewpoint L' , written as $L \triangleright_{\text{To}} L'$.² In typing rule WF-FIELDGET we adapt the type of the field from the viewpoint of y to the current viewpoint, whereas in rule WF-FIELDSET we adapt the type of y from the current viewpoint to the viewpoint of x .

As an example for the viewpoint adaptation, assume a method is called on a Far target and the argument is of type Near. Then the adapted type is Far, because the parameter is Near in relation to the caller, but from the perspective of the callee, it is actually Far in that case. Important is also the case where we pass a Far typed variable x to a Far target. In that case we have to take Somewhere as the adapted type, because it is not statically clear whether the object referred to by x is in a location that is different from the location of the target object.

Type Soundness. The location type system guarantees that variables of type Near only reference objects that are in the same location as the current object and that variables of type Far only reference objects that are in a different location to the current object. We formalize this by defining a well-formed runtime configuration. As helper functions, we define the location of a heap entry as $\text{loc}((l, c, \mathcal{D})) = l$ and the dynamically computed location type as $\text{dtype}(l, l') = \text{Near}$ if $l = l'$, and Far otherwise.

² Whereas in other ownership type systems (e.g. [7]), only one direction is considered, we chose to explicitly state the direction in order to achieve a simple and intuitive encoding.

$$\begin{array}{c}
\text{(WF-P)} \\
\frac{P = \overline{C} \quad \vdash C_i}{\vdash P} \\
\\
\text{(WF-M)} \\
\frac{\text{Near } c \text{ this } \cdot T \text{ result } \cdot \overline{V} \cdot \overline{V}' \vdash S_i}{c \vdash T \ m(\overline{V}) \{ \overline{V}' \ \overline{S} \}} \\
\\
\text{(WF-ASSIGN)} \\
\frac{\overline{V} \vdash E : L \quad L' _ x \in \overline{V} \quad L <: L'}{\overline{V} \vdash x \leftarrow E} \\
\\
\text{(WF-FIELDSET)} \\
\frac{L \ c \ x \in \overline{V} \quad L' = \text{anno}(c, f) \quad L'' _ y \in \overline{V} \quad (L'' \triangleright_{\text{To}} L) <: L'}{\overline{V} \vdash x.f \leftarrow y} \\
\\
\text{(WF-C)} \\
\frac{c \vdash M_i}{\vdash \text{class } c \{ \overline{V} \ \overline{M} \}} \\
\\
\text{(WF-NEWFRESH)} \\
\frac{}{\overline{V} \vdash \text{new } c \text{ in fresh} : \text{Far}} \\
\\
\text{(WF-NEWSAME)} \quad \text{(WF-VAR)} \\
\frac{L _ x \in \overline{V}}{\overline{V} \vdash \text{new } c \text{ in } x : L} \quad \frac{L _ x \in \overline{V}}{\overline{V} \vdash x : L} \\
\\
\text{(WF-FIELDGET)} \\
\frac{L \ c \ y \in \overline{V} \quad L' = \text{anno}(c, f)}{\overline{V} \vdash y.f : L' \triangleright_{\text{From}} L} \\
\\
\text{(WF-CALL)} \\
\frac{L \ c \ y \in \overline{V} \quad L_i _ z_i \in \overline{V} \quad \overline{x} = \text{params}(c, m) \quad (L_i \triangleright_{\text{To}} L) <: \text{anno}(c, m, x_i)}{\overline{V} \vdash y.m(\overline{z}) : \text{anno}(c, m, \text{result}) \triangleright_{\text{From}} L}
\end{array}$$

Fig. 9. Typing rules of LocJ. Note that indices are implicitly all-quantified.

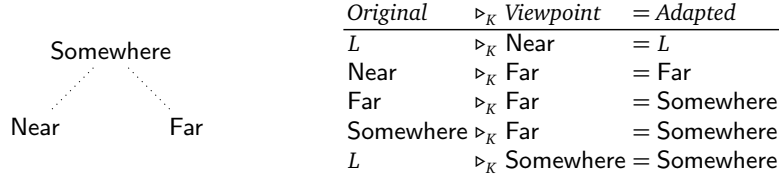


Fig. 10. Subtyping and viewpoint adaptation (where $K \in \{\text{From}, \text{To}\}$). Note that the direction K does not influence basic location types, but is important for our extension in Sect. 3.2.

Definition 1 (Well-formed runtime configuration). Let $\zeta = \overline{\mathcal{F}}, \mathcal{H}$ be a runtime configuration. ζ is well-formed iff all heap entries $(l, c, \mathcal{D}) \in \text{rng}(\mathcal{H})$ and all stack frames $\mathcal{F} \in \overline{\mathcal{F}}$ are well-formed under \mathcal{H} and the configuration satisfies all the standard conditions of a class-based language.

Definition 2 (Well-formed heap entry). $(l, _ , \mathcal{D})$ is well-formed under \mathcal{H} iff for all f with $\mathcal{D}(f) = \iota$ and $(l', c, _) = \mathcal{H}(\iota)$, we have $\text{dtype}(l, l') <: \text{anno}(c, f)$.

Definition 3 (Well-formed stack frame). $(\overline{\mathcal{S}}, \mathcal{D})^{c, m}$ is well-formed under \mathcal{H} iff for all x with $\mathcal{D}(x) = \iota$, we have $\text{dtype}(\text{loc}(\mathcal{H}(\mathcal{D}(\text{this}))), \text{loc}(\mathcal{H}(\iota))) <: \text{anno}(c, m, x)$.

Theorem 1 (Preservation for location types). Let ζ be a well-formed runtime configuration. If $\zeta \rightsquigarrow \zeta'$, then ζ' is well-formed as well.

Proof. The proof proceeds by a standard case analysis on the reduction rule used and is available in the accompanying report [25].


```

1 [Far] Server server = new cog ServerImpl();
2 [Far] Client client1 = new cog ClientImpl("Alice");
3 [Far] Client client2 = new cog ClientImpl("Bob");
4 client1 ! connectTo(server);
5 client2 ! connectTo(server);

```

Fig. 11. The code of the main block of the chat application, annotated with location types.

3.2 Named Far Location Types

The location type system so far can only distinguish near from far references. The type system knows that a near reference always points to a different location than a far reference. But whether two far references point to the same location or different ones is not statically known. This makes the type system often too weak in practice. As an example, let us consider the *main block*³ of the ABS chat application in Fig. 11, annotated with location types. The server and both clients are created by using the **new cog** expression. This means that all these objects live in their own, fresh COG and thus they can be typed to *Far*, because these locations are different to the current COG (the *Main* COG). However, for the method call `client1!connectTo(server)` to successfully type-check, the formal parameter of the `connectTo` method would need to be typed as *Somewhere* because the actual (adapted) parameter type is of type *Somewhere* ($= \text{Far} \triangleright_{\text{To}} \text{Far}$). This issue arises because the type system cannot distinguish that `client1` and `server` point to different locations. The example shows that in its basic form, the location type system often has to conservatively use the *Somewhere* type to remain sound, which in fact means that the type system cannot say anything about the location.

To improve the precision of the location type system we introduce *named* far types:

$$L ::= \dots \mid \text{Far}(i)$$

A named far type is a far type parametrized with an arbitrary name⁴. Far types with different names represent disjoint sets of far locations and are incompatible to each other. The following typing rule WF-NEWFRESHP is added, which allows new locations to be more precisely described.

$$\frac{(\text{WF-NEWFRESHP})}{\overline{V} \vdash \text{new } c \text{ in fresh} : \text{Far}(i)}$$

The subtyping and viewpoint adaptation relations are extended accordingly in Fig. 12. Adapting a $\text{Far}(i)$ to a $\text{Far}(j)$ for $i \neq j$ yields a $\text{Far}(i)$, as they denote different sets of locations. Adapting a $\text{Far}(i)$ to a $\text{Far}(i)$ does not yield *Near*, however, as two variables with the same $\text{Far}(i)$ type can refer to objects of different locations.

In practice the user does not explicitly provide the names. Instead the inference system automatically infers them when possible. These refined far types are then

³ A main block in ABS corresponds to a main method in Java.

⁴ Note that these are *not* object identifiers.

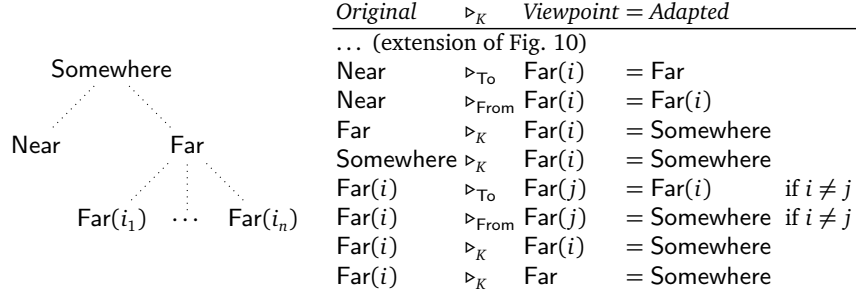


Fig. 12. Subtyping and viewpoint adaptation for extended location types.

used to improve the viewpoint adaptation. In the chat example our type system is now able to infer that the server and the client variables actually refer to different far locations. This means that the argument of the `connectTo` method call can be typed to `Far` instead of `Somewhere`.

Our experience with case studies shows that this extension is expressive enough for our purposes (cf. Sect. 4). However, other extensions to improve the expressiveness and precision of the location type system are imaginable, e.g. location type polymorphism similar to owner polymorphism in ownership type systems [4, 3, 17].

Type Soundness. Similar as for Thm. 1, a proof of type soundness for the named far location type system extension is available in the accompanying report [25].

3.3 Location Type Inference

The type system presented in the previous section requires the programmer to annotate all type occurrences with location types. In this subsection we present an inference system for location types. We first present a sound and complete inference system, which makes it possible to use the location type system without writing any type annotations and only use type annotations for achieving modular type checking. The second part then presents an inference system that can deal with type-incorrect programs and that finds not only *some* solution but an *optimal* solution.

Sound and Complete Inference. The formal model for inferring location types follows the formalization of other type system extensions [8]. The idea is to introduce location type variables at places in the program where location types occur in our typing rules. Type inference then consists of two steps. First, generating constraints for the location type variables. Second, checking whether a substitution for the location type variables exists such that all constraints are satisfied.

To introduce location type variables into programs we extend the syntax of location types accordingly:

$$L ::= \dots \mid \alpha \quad \text{location type variables (also } \beta, \gamma, \text{ and } \delta)$$

$$\begin{aligned}
 \mathcal{Q} ::= & \alpha \triangleright_K \beta = \gamma && \text{adaptation constraint} \\
 & | \alpha <: \beta && \text{subtype constraint} \\
 & | \alpha = L \mid \alpha \neq L && \text{constant constraint}
 \end{aligned}$$

Fig. 13. Location type constraints.

$$\begin{array}{c}
 \frac{P = \bar{C} \quad \vdash C_i : \bar{Q}_i}{\vdash P : \bar{Q}_1 \cdot \dots \cdot \bar{Q}_n} \\
 \\
 \frac{c \vdash M_i : \bar{Q}_i}{\vdash \text{class } c \{ \bar{V} \bar{M} \} : \bar{Q}_1 \cdot \dots \cdot \bar{Q}_n} \\
 \\
 \frac{\bar{V} \vdash E : \beta, \bar{Q} \quad \alpha _ x \in \bar{V}}{\bar{V} \vdash x \leftarrow E : \beta <: \alpha \cdot \bar{Q}} \\
 \\
 \frac{\delta \text{ is fresh}}{\bar{V} \vdash \text{new } c \text{ in fresh } : \delta, \delta \neq \text{Near}} \\
 \\
 \frac{\alpha _ y \in \bar{V} \quad \alpha _ x \in \bar{V}}{\bar{V} \vdash \text{new } c \text{ in } y : \alpha, \bullet \quad \bar{V} \vdash x : \alpha, \bullet} \\
 \\
 \frac{\delta \text{ c this} \cdot T \text{ result} \cdot \bar{V} \cdot \bar{V}' \vdash S_i : \bar{Q}_i \quad \delta \text{ is fresh}}{c \vdash T \ m(\bar{V}) \{ \bar{V}' \bar{S} \} : \delta = \text{Near} \cdot \bar{Q}_1 \cdot \dots \cdot \bar{Q}_n} \\
 \\
 \frac{\alpha \text{ c } x \in \bar{V} \quad \beta = \text{anno}(c, f) \quad \gamma _ y \in \bar{V} \quad \delta \text{ is fresh}}{\bar{V} \vdash x.f \leftarrow y : \delta <: \beta \cdot \gamma \triangleright_{\text{To}} \alpha = \delta} \\
 \\
 \frac{\alpha \text{ c } y \in \bar{V} \quad \beta = \text{anno}(c, f) \quad \gamma \text{ is fresh}}{\bar{V} \vdash y.f : \gamma, \beta \triangleright_{\text{From}} \alpha = \gamma} \\
 \\
 \frac{\alpha \text{ c } y \in \bar{V} \quad \alpha_i _ z_i \in \bar{V} \quad \bar{x} = \text{params}(c, m) \quad \beta_i = \text{anno}(c, m, x_i)}{\beta = \text{anno}(c, m, \text{result}) \quad \mathcal{Q}_i = \alpha_i \triangleright_{\text{To}} \alpha = \gamma_i \cdot \gamma_i <: \beta_i} \\
 \\
 \frac{\gamma_i \text{ is fresh} \quad \gamma \text{ is fresh}}{\bar{V} \vdash y.m(\bar{z}) : \gamma, \beta \triangleright_{\text{From}} \alpha = \gamma \cdot \bar{Q}_1 \cdot \dots \cdot \bar{Q}_n}
 \end{array}$$

Fig. 14. Constraint generation rules.

In the following we consider P as a program which is fully annotated with pairwise distinct location type variables. The constraints which are generated by the inference system are shown in Fig. 13. We use the judgment $\vdash P : \bar{Q}$, defined in Fig. 14, to denote the generation of the constraints \bar{Q} from program P . Note that additional *fresh* location type variables are introduced during the constraint generation.

Soundness and Completeness. Let σ be a mapping function from location type variables to location types, i.e., α to $\{\text{Near}, \text{Far}, \text{Somewhere}, \text{Far}(i_1), \dots, \text{Far}(i_n)\}$. Then $\sigma \models \bar{Q}$ if the constraints \bar{Q} are satisfiable under σ . We write σP to denote that all location type variables in P have been replaced by location types according to the substitution function σ .

Conjecture 1 (Soundness and Completeness of the Inference). The inference is sound and complete in the sense that every typing inferred can be successfully type-checked and every typing which type-checks can also be inferred.

- *Soundness:* If $\vdash P : \bar{Q}$ and $\sigma \models \bar{Q}$, then $\vdash \sigma P$.
- *Completeness:* If $\vdash \sigma P$ for some minimal σ , then $\exists \bar{Q}$ such that $\vdash P : \bar{Q}$ and $\exists \sigma'$ such that σ' is an extension of σ and $\sigma' \models \bar{Q}$. Note that σ' is an extension of σ iff $\sigma'(\alpha) = \sigma(\alpha)$ for all $\alpha \in \text{dom}(\sigma)$.

Optimal and Partial Inference. Whereas soundness and completeness is important, it is not sufficient for an inference system to be usable in practice. Two additional properties are required, namely:

1. If multiple inference solutions exist, an *optimal* solution should be taken. This is important, because the user in general wants to have the most precise solution, i.e., with the least amount of Somewhere annotations.
2. If no typable solution can be inferred, at least a partially typable solution should be provided. It is otherwise nearly impossible to use the inference system if one only gets a “No solution can be found” result. In addition, this partially typable solution should lead to the least amount of type errors.

To support these two properties, we extend our formal model in the following way. We introduce three constraint categories: *must-have*, *should-have*, and *nice-to-have*. The *must-have* constraints must always be satisfied. These are for example in Fig. 14 the adaptation constraints ($\alpha \triangleright_K \beta = \gamma$) and the constant constraints ($\alpha = L$, $\alpha \neq L$), characterizing the types of subexpressions. They also encompass the constant constraints which result from user annotations (not considered in the formalization of Fig. 14, but present in the implementation). Note that there is always a solution to these constraints in our inference system as they are based on freshly allocated location type variables. The *should-have* constraints, e.g. the subtype constraints ($\alpha <: \beta$) in Fig. 14, should always be satisfied in order to get a valid typing, but can be unsatisfied for partially correct solutions. The *nice-to-have* constraints are those that give us a *nice* (optimal) solution, i.e., with the least amount of Somewhere annotations or with Far types at the places where the precision of $\text{Far}(i)$ types is not needed.

Inferring an optimal solution consists of solving the following problem. First, all *must-have* constraints, then the most amount of *should-have* constraints, and finally the most amount of *nice-to-have* constraints should be satisfied. The problem can be encoded as a partially weighted MaxSAT problem by assigning appropriate weights to the constraints. This means that *must-have* constraints are hard clauses (maximum weight) and *should-have* constraints correspond to soft clauses whose weight is greater than the sum of all weighted *nice-to-have* clauses. Solving such a problem can be efficiently done using specialized SAT solvers.

As an example for partial inference, consider the `ServerImpl` class in Fig. 4. Assume that there are no annotations on the signature and the body of the `connect` method except for the return type which has been wrongly annotated by the programmer as `Far`. The inference system then still gives a solution where all constraints are satisfied except one *should-have* constraint, namely $s <: \text{result}$ which is generated at the last line of the `connect` method. The inference system assigns the type `Near` to variable `s` because if it were to assign `Far` to `s`, more *should-have* constraints would be unsatisfied (i.e. those resulting from lines 5 to 7).

4 Implementation and IDE Integration

We have implemented the type and inference system for location types, including named far location types and optimal and partial inference, as an extension of the

ABS compiler suite. The type and inference system is integrated into an Eclipse-based IDE, but can also be used from the command line.

Inference System. The inference system internally uses the Max-SAT solver SAT4J [16] to solve the generated inference constraints. As the inference system may return a solution that is not fully typable, we use the type checker for location types to give user-friendly error messages.

The alias analysis for named Far locations (cf. Sect. 3.2) can be configured to use scopes of different granularity: basic (no alias analysis), method-local, class-local, module-local, and global analysis. This allows the user to choose the best tradeoff between precision and modularity. For the inference, an upper bound on the number of possible named $\text{Far}(i)$ locations is needed. This is calculated based on the number of new c in fresh expressions in the current scope.

IDE Integration. ABS features an Eclipse-based IDE⁵ for developing ABS projects. The interesting part of the IDE for this paper is that we have incorporated visual overlays which display the location type inference results. For each location type there is a small overlay symbol, e.g., **N** for Near and **F** for Far, which are shown as superscripts of the type name. For example, a Far Client appears as Client^{F} . Whenever the user saves a changed program, the inference is triggered and the overlays are updated. They give the user complete location type information of all reference types, without cluttering the code. In addition, the overlays can easily be toggled on or off. It is also possible to write the inference results back as annotations into the source code, with user-specified levels of granularity, e.g., method signatures in interfaces.

Evaluation. We evaluated the location type system by applying it to three case studies. The Trading System (1164 LOC, 150 types to annotate) and Replication System (702 LOC, 62 types to annotate) case studies are ABS programs developed as parts of the case studies in the HATS project. The Chat Application (251 LOC, 55 types to annotate) is an extended version of the one presented in Sect. 2.

The evaluation results are presented in Fig. 15. They show how precise the case studies can be typed and how fast the inference works. We also restricted the alias analysis by various scopes to see the impact on performance and precision. First of all, all case studies can be fully typed using our type system. The chart on the left shows the precision (percentage of near and far annotations) of the type inference. As can be seen, the basic type system already has a good precision (> 60%) in all three case studies. As expected, the precision increased with a broader analysis scope. Using a global aliasing analysis, the inference achieved a precision of 100% in the Chat as well as the Trading System case studies. In the Replication Server case study, the best precision was already achieved with a method-local scope.

The chart on the right shows the performance results of the inference. It shows that the performance of the inference is fast enough for the inference system to be used interactively. It also shows that the performance depends on the chosen scope for the aliasing analysis. Note that the examples were completely unannotated, so that

⁵ <http://tools.hats-project.eu/eclipseplugin/installation.html>

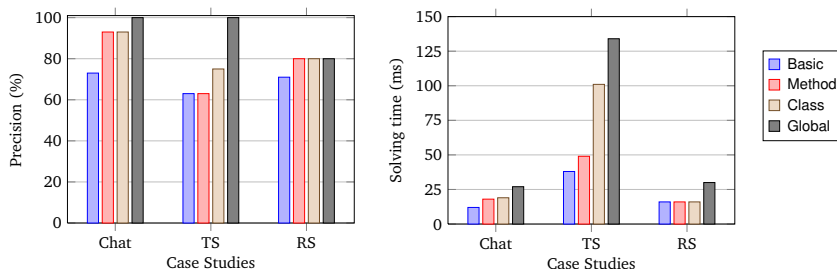


Fig. 15. Precision and solving time of the location type inference for the three case studies, using four different scopes for the aliasing analysis. The measurements were done on a MacBook Pro laptop (Intel Core 2 Duo T7400 2.16GHz CPU, 2GB RAM, Ubuntu 10.04, Sun JDK 1.6.16). We used the `-Xms1024` parameter to avoid garbage collection. As working in an IDE usually consists of an edit-compile-run cycle, we provide the performance results (the mean of 20 complete runs) after warming up the JVM with 5 dry runs. We measured the time that the SAT-solver required for finding a solution using the `System.nanoTime()` method.

all types had to be inferred. In practice, programs are often partially annotated, which additionally improves the performance of the type inference. Our implementation of the inference focused more on correctness than performance, which means that many improvements in the encoding, and thus in the solving time are still possible.

5 Discussion and Related Work

Location types are a variant of ownership types that concentrate on *flat* ownership contexts. We presented the type system in the context of distributed object-oriented systems, but it can be applied to any context where flat ownership contexts are sufficient. Ownership types [4, 3, 17] and similar type systems [1, 6] typically describe a hierarchical heap structure. On one hand this makes these systems more general than location types, because ownership types could be used for the same purpose as location types; on the other hand this makes these systems more complex. An ownership type system which is close to location types in nature is that by Clarke et al. [5], which applies ownership types to active objects. In their system ownership contexts are also flat, but ownership is used to ensure encapsulation of objects with support for a safe object transfer using unique references and cloning. Haller and Odersky [13] use a capability-based type system to restrict aliasing in concurrent programs and achieve full encapsulation. As these systems are based on encapsulation they do not have the concept of *far* references. Places [12] also partition the heap. However, the set of places is fixed at the time the program is started. Similar, but less expressive than our type system, is Loci [26], which only distinguishes references to be either thread-local or shared. Loci only uses defaults to reduce the annotation overhead. Loci is also realized as an Eclipse plug-in. Regions are also considered in region-based memory management [23], but for another purpose. They give the guarantee that objects inside a region do not refer to objects inside another region to ensure safe deallocation.

Using a Max-SAT solver with weighted constraints was also used in [11] to infer types that prevent data-races and in [8] to find good inference solutions for universe types. A crucial aspect of our work is the integration of type inference results into the IDE by using overlays. To the best of our knowledge there is no comparable approach. A widely used type system extension is the non-null type system [10]. For variations of this type system, there exist built-in inference mechanisms in Eclipse⁶ and IntelliJ IDEA⁷ as well as additional plug-ins such as [14]. None of these IDE integrations provide type information by using overlays, but only give warnings in cases of type errors, which makes it difficult for the user to find the root of the problem.

6 Conclusion and Future Work

We have presented a type system for distributed object-oriented programming languages to distinguish near from far references. We applied the type system to the context of the ABS language to guarantee that far references are not used as targets for synchronous method calls. A complete type inference implementation allows the programmer to make use of the type system without making any annotations. The type inference results are visualized as overlay annotations directly in the development environment. Application of the type system to several case studies shows that the type system is expressive enough to type realistic code. The type inference implementation is fast enough to provide inference results within fractions of a second, so that interactive use of the system is possible.

We see two directions for future work. First, the type system could be applied to other settings where the location of an object is important, e.g., Java RMI [19]. Second, it would be interesting to investigate the visual overlay technique for other (pluggable) type systems, e.g., the nullness type system [14].

Acknowledgements. We thank our Master's students Thomas Fischer, Christian Seise, Florian Strauß, and Mathias Weber for implementing most parts of the ABS-Eclipse integration. We also thank Peter Wong for applying the location type system to the Replication System case study, and Arnd Poetzsch-Heffter and Patrick Michel for proof-reading earlier drafts of this paper. Finally, we thank the anonymous reviewers of TOOLS Europe 2011 for their constructive feedback.

References

1. Aldrich, J., Chambers, C.: Ownership Domains: Separating Aliasing Policy from Mechanism. In: ECOOP. Ed. by Odersky, M. Vol. 3086. LNCS. Springer, 2004, pp. 1–25.
2. Andrae, C., Noble, J., Markstrum, S., Millstein, T.: A framework for implementing pluggable type systems. In: OOPSLA. Ed. by Tarr, P. L., Cook, W. R. ACM, 2006, pp. 57–74.

⁶ http://wiki.eclipse.org/JDT_Core/Null_Analysis

⁷ <http://www.jetbrains.com/idea/webhelp/inferring-nullity.html>

3. Boyapati, C., Liskov, B., Shriram, L.: Ownership types for object encapsulation. In: POPL. ACM, Jan. 2003, pp. 213–223.
4. Clarke, D., Potter, J., Noble, J.: Ownership Types for Flexible Alias Protection. In: OOPSLA. ACM Press, Oct. 1998, pp. 48–64.
5. Clarke, D., Wrigstad, T., Östlund, J., Johnsen, E. B.: Minimal Ownership for Active Objects. In: APLAS. Ed. by Ramalingam, G. Vol. 5356. LNCS. Springer, 2008.
6. Dietl, W.: Universe Types: Topology, Encapsulation, Genericity, and Tools. PhD thesis. ETH Zurich, Switzerland, 2009.
7. Dietl, W., Drossopoulou, S., Müller, P.: Generic Universe Types. In: ECOOP. Ed. by Ernst, E. Vol. 4609. LNCS. Springer, July 2007.
8. Dietl, W., Ernst, M., Müller, P.: Tunable Universe Type Inference. Tech. rep. 659. Department of Computer Science, ETH Zurich, Dec. 2009.
9. Ernst, M. D.: Type Annotations Specification (JSR 308) and The Checker Framework: Custom pluggable types for Java. URL: <http://types.cs.washington.edu/jsr308/>.
10. Fähndrich, M., Leino, K. R. M.: Declaring and checking non-null types in an object-oriented language. In: OOPSLA '03. ACM, 2003, pp. 302–312.
11. Flanagan, C., Freund, S. N.: Type inference against races. In: Sci. Comput. Program. 64 (1 2007), pp. 140–165.
12. Grothoff, C.: Expressive Type Systems for Object-Oriented Languages. PhD thesis. University of California, Los Angeles, 2006.
13. Haller, P., Odersky, M.: Capabilities for Uniqueness and Borrowing. In: ECOOP. Ed. by D'Hondt, T. LNCS. Springer, June 2010, pp. 354–378.
14. Hubert, L., Jensen, T., Pichardie, D.: Semantic Foundations and Inference of Non-null Annotations. In: FMOODS '08. Springer, 2008, pp. 132–149.
15. Hähnle, R. et al.: Report on the Core ABS Language and Methodology: Part A. Report. The HATS Consortium, Mar. 2010. URL: <http://www.hats-project.eu/>.
16. Le Berre, D., Parrain, A.: The SAT4J library, Release 2.2, System Description. In: Journal on Satisfiability, Boolean Modeling and Computation 7 (2010), pp. 59–64.
17. Lu, Y., Potter, J.: On Ownership and Accessibility. In: ECOOP. Ed. by Thomas, D. Vol. 4067. LNCS. Springer, July 2006, pp. 99–123.
18. Miller, M. S., Tribble, E. D., Shapiro, J. S.: Concurrency Among Strangers. In: TGC. Ed. by Nicola, R. D., Sangiorgi, D. Vol. 3705. LNCS. Springer, 2005, pp. 195–229.
19. Oracle Corporation: Java SE 6 RMI documentation. <http://download.oracle.com/javase/6/docs/technotes/guides/rmi/index.html>.
20. Östlund, J., Wrigstad, T.: Welterweight Java. In: TOOLS (48). Ed. by Vitek, J. Vol. 6141. LNCS. Springer, 2010, pp. 97–116.
21. Potanin, A., Noble, J., Clarke, D., Biddle, R.: Generic Ownership for Generic Java. In: OOPSLA. Ed. by Tarr, P. L., Cook, W. R. ACM, 2006.
22. Schäfer, J., Poetzsch-Heffter, A.: JCoBox: Generalizing Active Objects to Concurrent Components. In: ECOOP. Ed. by D'Hondt, T. LNCS. Springer, June 2010, pp. 275–299.
23. Tofte, M., Talpin, J.-P.: Region-based Memory Management. In: Inf. Comput. 132.2 (1997), pp. 109–176.
24. Van Cutsem, T., Mostinckx, S., Boix, E. G., Dedeker, J., Meuter, W. D.: AmbientTalk: Object-oriented Event-driven Programming in Mobile Ad hoc Networks. In: SCC. IEEE Computer Society, 2007, pp. 3–12.
25. Welsch, Y., Schäfer, J.: Location Types for Safe Distributed Object-Oriented Programming. Tech. rep. 383/11. University of Kaiserslautern, Apr. 2011.
26. Wrigstad, T., Pizlo, F., Meawad, F., Zhao, L., Vitek, J.: Loci: Simple Thread-Locality for Java. In: ECOOP. Ed. by Drossopoulou, S. LNCS. Springer, 2009, pp. 445–469.