Object Connectivity in a Concurrent Calculus of Classes

— Extended Abstract —

September 26, 2003

Erika Ábrahám^{1,2}, Marcello M. Bonsangue³, Frank S. de Boer⁴, and Martin $$\rm Steffen^1$$

¹ Christian-Albrechts-University Kiel, Germany

² University Freiburg, Germany

³ University Leiden, The Netherlands

⁴ CWI Amsterdam, The Netherlands

Abstract. The concurrent ν -calculus has been investigated as a core calculus for imperative, object-oriented languages with multithreading and heap-allocated objects. From an abstract point of view, the combination of this form of concurrency with objects corresponds to features known from the popular language Java. One distinctive feature, however, of the concurrent object calculus is that it is *object-based*, whereas the mainstream of object-oriented languages is *class-based*.

This work extends the concurrent ν -calculus by introducing classes and exploring some of the semantical consequences. The external behavior of a component, given as a set of threads and objects, can be described in a linear-time setting as the set of traces of interactions with the environment. Considering classes as part of the component makes instantiation a possible interaction between component and environment. A striking consequence of this new interaction is that we need to take into account *connectivity* information, i.e., the way objects may have knowledge of each other, in order to get a precise characterization of possible traces. We formulate an operational semantics that incoporates the connectivity information into the scoping mechanism of the ν -calculus. Since instantiation itself is to be considered as non-observable, we formulate a semantics where objects are instantiated only when they are accessed for the first time ("lazy instantiation").

1 Introduction

The concurrent ν -calculus has been proposed as core calculus for imperative, object-oriented languages with multithreading and heap-allocated objects. In the context of concurrent, *object-based* programs and starting from may-testing as a very simple notion of observation, Jeffrey and Rathke [4] provide a fully abstract trace semantics for the language. Their result roughly says that, given a *component* as a set of objects and threads, the fully abstract semantics consists

of the set of traces at the boundary of the component, where the traces record incoming and outgoing calls and returns. At this level, the result is as one would expect, since intuitively in the chosen setting, the only possible way to observe something about a set of objects and threads is by exchanging messages. It should be equally clear, however, that for the language featuring multithreading, object references with aliasing, and creation of new objects and threads, the details of defining the semantics and proving the full abstraction result are far from trivial.

The result in [4] is developed within the concurrent ν -calculus [3], an extension of the sequential ν -calculus [6] which stands in the tradition of various object calculi [1] and also of the π -calculus [5,7]. One distinctive feature of the ν -calculus is that it is *object-based*, which in particular means that there are no *classes* as templates for new objects. This is in contrast to the mainstream of object-oriented languages where the code is organized in classes. This report addresses therefore the following question:

What changes when switching from an object-based to a class-based setting?

Considering the observable behavior of a component, we have to take into account that in addition to objects, which are the passive entities containing the instance state and the methods, and threads, which are the active entities, *classes* come into play. Classes serve as a blueprint for their instances and can be conceptually understood as particular objects supporting just a method which allows to generate instances.

Important in our context is that now the division between the program fragment under observation and its environment also separates classes: There are classes internal to the component and those belonging to the environment. As a consequence, not only calls and returns are exchanged at the interface between component and environment, but instantiation requests, as well. This possibility of cross-border instantiation is absent in the object-based setting: Objects are created by directly providing the code of their implementation, not referring to the name of a class, which means that the component creates only componentobjects and dually the environment only environment objects. To understand the bearing of this change on the semantics, we must realize that the interesting part of the problem is not so much to just cover the possible behavior at the interface —there is little doubt that sequences of calls, returns, and instantiations with enough information at the labels would do— but to characterize it exactly, i.e., to exclude impossible environment interaction. As an obvious example, a trace with two consecutive calls from the same thread without outgoing communication in between cannot be part of the component behavior.

Let's concentrate on the issue of instantiation across the demarcation line between component and its environment, and imagine that the component creates an instance of an environment class. The first question is: does this yield a component object or an environment object? As the code of the object is provided by the external class which is in the hand of the observer, the interaction between the component and the newly created object can lead to observable ef-

Introduction

fects and must thus be traced. In other words, instances of environment classes belong to the environment, and those of internal classes to the component.

Whereas in the above situation, the object is instantiated to be part of the environment, the *reference* to it is kept at the creator, for the time being. So in case, an object of the program, say o_1 instantiates two objects o_2 and o_3 of the environment, the situation informally looks as shown in Figure 1, where the dotted bubbles indicate the scope of o_2 respectively o_3 .



Fig. 1. Instances of external classes

In this situation it is impossible, that there be an incoming call from the environment carrying both names o_2 and o_3 , as the only entity aware of both references is o_1 . Unless the component gives away the reference to the environment, o_2 and o_3 are completely separated.

Thus, in order to exclude impossible combinations of object reference in the communication labels, the component has to keep track which objects of the environment are connected. The component has, of course, by no means full informa-

tion about the complete system; after all it can at most trace what happens at the interface, and the objects of the environment can exchange information "behind the component's back". Therefore, the component must conservatively over-approximate the potential knowledge of objects in the environment, i.e., it must make *worst-case assumptions* concerning the proliferation of knowledge, which means it must assume that

- 1. once a name is out, it is never forgotten, and
- 2. if there is a possibility that a name is leaked from one environment object to another, this will happen.

Sets of environment objects which can possibly be in contact with each other form therefore equivalence classes of names —we call them *cliques*— and the formulation of the semantics must include a representation of them. New cliques can be created, as new objects can be instantiated without contact to others, and furthermore cliques can merge, if the component leaks the identity of a member of one clique to a member of another.

This extended abstract investigates a class-based variant of the concurrent ν -calculus, formalizing the ideas sketched above about cliques of objects. Instantiation itself, even across the environment-program boundary, is unobservable, since the calculus does not have constructor methods. In the semantics, an externally instantiated object is created only at the point, when it is actually accessed the for the first time, which we call "lazy instantiation". For want of space, we concentrate here on the intuition and, in the formal part, the operational semantics describing the external behavior of a component. For deeper coverage we refer to the technical report [2]; the derivation rules for typing and for the operational semantics are included also in the appendix for reference.

2 A concurrent class calculus

In this section, we present the calculus used in our development. As we concentrate on the semantical issues of connectivity of objects and the interface behavior of a component, we gloss over the exact syntax, ignore typing issues and also omit structural equivalence rules, as they are rather standard. As mentioned, the reader will find details in the appendix and the accompanying technical report.

The calculus is a syntactic extension of the concurrent object calculus or concurrent ν -calculus from [3, 4]. The basic change is the introduction of *classes*, where a class is a named collection of methods. In contrast to object references, class names are literals introduced when defining the class; they may be hidden using the ν -binder but unlike object names, the scopes for class names are *static*. Object names, on the other hand, are first-order citizens of the calculus in that they can be stored in variables, passed to other objects as method parameters, making the scoping *dynamic*, and especially they can be created freshly by instantiating a class.

A program is given by a collection of classes. A class c[O] carries a name c and defines the implementation of its methods, and analogously for objects. A method $\varsigma(n:T).\lambda(x_1:T_1,\ldots,x_n:T_k).t$ provides the definition of the method body abstracted over the formal parameters of the method and the ς -bound "self" parameter [1]. Besides named objects and classes, the dynamic configuration of a program can contain as active entities named threads $n\langle t \rangle$, which, like objects, can be dynamically created. Unlike objects, threads are not instantiated by some statically named entity (a "thread class"), but directly created by providing the code. A thread, whose exact syntax is of no particular interest here, basically is either a value (especially a reference to another named entity) or a sequence of expressions, notably method calls (written $o.l(\vec{v})$) and creation of new objects and new threads (*new c* and *new* $\langle t \rangle$ where c is a class name and t a thread). We will generally use n and its syntactic variants as name for threads (or just in general for names), o for objects, and c for classes. Furthermore we will use f specifically for instance variables or fields, we use f = v for field variable declaration, field access is written as x.f, and field update as $x := v.^5$

Concerning the *operational semantics* of the calculus, the basic steps are given in two levels: *internal* steps whose effect is completely confined within a configuration, and those with external effect. Interested mainly in the external behavior we elide the definition of the internal steps (but see Appendix A.3).

The *external* behavior of a component is given in terms of labeled transitions describing the communication at the interface of an *open* program. For the completeness of the semantics, it is crucial ultimately to consider only communication traces realizable by an actual program context which, together with the component, yields a well-typed closed program.

 $^{^5}$ In the class-based setting we don't use general method update as supported by the object-based $\nu\text{-calculus.}$

A concurrent class calculus

The concentration on actually realizable traces has various aspects, e.g., the transmitted values needs to adhere to the static typing assumptions, only publicly known objects can be called from the outside, and the like. Being concerned in the dynamic relationship among objects, we omit also these aspects here. Besides that, this part is rather standard and also quite similar to the one in [4].

2.1 Connectivity contexts and cliques

The informal discussion in the introduction argued that in the presence of internal and external classes and cross-border instantiation, the component must *keep track* of which identities it gives away to which objects in order to exclude impossible as described for instance in connection with Figure 1. The external semantics is formalized as labeled transitions between judgments of the from

$$\Delta; E_{\Delta} \vdash C : \Theta; E_{\Theta} , \qquad (1)$$

where $\Delta; E_{\Delta}$ are the assumptions about the environment of the component Cand $\Theta; E_{\Theta}$ the commitments. The assumptions consists of a part Δ concerning the existence (plus static typing information) of named entities in the environment. For the book-keeping of which objects of the environment have been told which identities, a well-typed component must take into account the relation of object names from the assumption context Δ amongst each other, and the knowledge of objects from Δ about those exported by the component, i.e., those from Θ . ⁶ In analogy to the name contexts Δ and Θ , E_{Δ} expresses assumptions about the environment, and E_{Θ} commitments of the component.

$$E_{\Delta} \subseteq \Delta \times (\Delta + \Theta) . \tag{2}$$

and dually $E_{\Theta} \subseteq \Theta \times (\Theta + \Delta)$. We will write $o_1 \hookrightarrow o_2$ (" o_1 may know o_2 ") for pairs from these relations. The component has by no means full information about the complete system; after all it can at most trace what happens at the interface, and the objects of the environment can exchange information "behind the component's back". Thus it must conservatively overapproximate the potential knowledge of objects in the environment, i.e., it must make *worst-case assumptions* concerning the proliferation of knowledge by assuming:

- 1. once a name is out, it is never forgotten, and
- 2. if there is a possibility that a name is leaked from one environment object to another, this will happen.

More technically, the worst case assumptions about the actual situation are represented as the *reflexive*, *transitive*, and *symmetric* closure of the \hookrightarrow -pairs of

⁶ Besides the relationships amongst objects, we need to keep one piece of information concerning the "connectivity" of *threads*. To exclude situations where a known thread leaves the component into one clique of objects but later returns to the component coming from a different clique without connection to the first, we remember for each thread that has left the component the object from Δ it has left into.

objects from Δ the component maintains. Given Δ , Θ , and E_{Δ} , we write \rightleftharpoons for this closure, i.e.,

$$\Leftrightarrow \triangleq (\hookrightarrow \downarrow_{\Delta} \cup \longleftrightarrow \downarrow_{\Delta})^* \subseteq \Delta \times \Delta . \tag{3}$$

Note that we close the part of \hookrightarrow concerning only environment objects from Δ , but not wrt. objects at the *interface*, i.e., the part of $\hookrightarrow \subseteq \Delta \times \Theta$. We will also need the union of $= \cup =; \hookrightarrow \subseteq \Delta \times (\Delta + \Theta)$, for which we will also write =:. As judgment, we use $\Delta; E_{\Delta} \vdash v_1 = v_2 : \Theta$ respectively $\Delta; E_{\Delta} \vdash v_1 = \hookrightarrow v_2 : \Theta$. For Θ, E_{Θ} , and Δ , the definitions are applied dually.

The relation \rightleftharpoons is an equivalence relation on the objects from Δ and partitions them in equivalence classes. As a manner of speaking, we call a set of object names from Δ (or dually from Θ) such as for all objects o_1 and o_2 from that set, Δ ; $E_{\Delta} \vdash o_1 \rightleftharpoons o_2 : \Theta$, a *clique*, and if we speak of *the* clique of an object we mean the whole equivalence class.

2.2 External steps

As mentioned before, the external semantics is given by transitions between $\Delta; E_{\Delta} \vdash C : \Theta; E_{\Theta}$ judgments. Table 1 for the exchange of free names and in Table 2 dealing with bound names. Beside internal steps a component exchanges information with the environment by transitions; the core labels γ are of the form $n\langle [o] call \ o.l(\vec{v}) \rangle$ and $n\langle return(v) \rangle$; names may occur bound in a label $\nu(n:T).\gamma$, and sending and receiving labels are written as γ ? and γ !.

The component exchanges information with the environment via *calls* and *returns*. Using a lazy instantiation scheme for cross-border object creation, there are no separate external labels for *new*-steps. In the extended abstract, we omit the static typing premises in the operational rules as they are straightforward and we concentrate in the discussion on the novel aspects, namely the connectivity information. The external steps are given in Tables 1 and 2.

As for the relationship of communicated values, incoming and outgoing communication play dual roles: E_{Θ} overapproximates the actual connectivity of the component, while the assumption context E_{Δ} is consulted to exclude impossible combinations of incoming values. For incoming communication (cf. rule CALLI₂ and RETI)⁷ we require that the sender be acquainted with the transmitted arguments. In case of a call, the caller o_1 must additionally be acquainted with the callee o_2 and furthermore the calling thread must originate from a clique of objects in connection with the one to which the thread had left the component the last time: $\Delta; E_{\Delta} \vdash n \rightleftharpoons [o_1] : \Theta$.⁸ To assure these connectivity conditions, the identity of the callee has been remembered as part of the block-syntax when the call was issued. It is worth mentioning that in rule RETI the proviso that the callee o_2 knows indirectly the caller o_1 , i.e., $\Delta; E_{\Delta} \vdash o_2 \rightleftharpoons o_1 : \Theta$ is not needed. Neither is it necessary to require in analogy to the situation for the incoming call

⁷ We omit rule CALLI₁ which deals with a thread that entering for the first time. The rule is identical to CALLI₂ as far as the treatment of E_{Δ} and E_{Θ} is concerned.

⁸ Since the caller o_1 is in the domain of Δ , we can write $n \rightleftharpoons [o_1]$ instead of $n \rightleftharpoons \bigcirc [o_1]$.

$\begin{split} E_{\Delta} \vdash [o_1] &\leftrightarrows \hookrightarrow \vec{v} : \Theta \qquad E_{\Delta} \vdash [o_1] \leftrightarrows \hookrightarrow o_2 : \Theta \qquad E_{\Delta} \vdash n \leftrightarrows [o_1] : \Theta \\ \dot{E}_{\Delta} &= E_{\Delta} \setminus n \qquad \dot{E}_{\Theta} = E_{\Theta} + (o_2 \hookrightarrow \vec{v}, n \hookrightarrow o_2) \end{split}$	CALL
$\begin{split} &\Delta; E_{\Delta} \vdash C \parallel n \langle \det x' : T' = [o_2] \ blocks \ for \ o'_2 \ in \ t \rangle : \Theta; E_{\Theta} \xrightarrow{n \langle [o_1] call \ o_2 . l(\vec{v}) \rangle;} \\ &\Delta; \acute{E}_{\Delta} \vdash C \parallel n \langle \det x : T = o_2 . l(\vec{v}) \ in \ return [o_1] \ x; \det x' : T' = [o_2] \ blocks \ for \ o'_2 \ in \ t \rangle : \Theta; \acute{E}_{\Theta} \end{split}$	CALLI2
$ \acute{E}_{\Delta} = E_{\Delta} + ([o_1] \hookrightarrow v, n \hookrightarrow [o_1]) \qquad \acute{E}_{\Theta} = E_{\Theta} \setminus n $	BETO
$\Delta; E_{\Delta} \vdash C \parallel n \langle let x: T = return[o_1] v in t \rangle : \Theta; E_{\Theta} \xrightarrow{n \langle return(v) \rangle!} \Delta; \acute{E}_{\Delta} \vdash C \parallel n \langle t \rangle : \Theta; \acute{E}_{\Theta}$	Itero
$o_2 \in \Delta$ $\acute{E}_{\Delta} = E_{\Delta} + (o_2 \hookrightarrow \vec{v}, n \hookrightarrow o_2)$ $\acute{E}_{\Theta} = E_{\Theta} \setminus n$	
$\begin{aligned} \Delta; E_{\Delta} \vdash C \parallel n \langle let x:T = [o_1] \ o_2.l(\vec{v}) \ in \ t \rangle : \Theta; E_{\Theta} \xrightarrow{n \langle [o_1] \ call \ o_2.l(\vec{v}) \rangle!} \\ \Delta; \dot{E}_{\Delta}; \vdash C \parallel n \langle let x:T = [o_1] \ blocks \ for \ o_2 \ in \ t \rangle : \Theta; E_{\Theta} \end{aligned}$	
$ \begin{array}{ccc} ; \Delta, \Theta \vdash v : T & \Delta; E_{\Delta} \vdash o_{2} \leftrightarrows v : \Theta & \Delta; E_{\Delta} \vdash n \hookrightarrow o_{2} : \Theta \\ \dot{E}_{\Delta} = E_{\Delta} \setminus n & \dot{E}_{\Theta} = E_{\Theta} + (o_{1} \hookrightarrow v, n \hookrightarrow [o_{1}]) \end{array} $	
$ \begin{array}{l} \Delta; E_{\Delta} \vdash C \parallel n \langle let x:T = [o_1] \ blocks \ for \ o_2 \ in \ t \rangle : \Theta; E_{\Theta} \xrightarrow{-n \langle return(v) \rangle?} \\ \Delta; \dot{E}_{\Delta} \vdash C \parallel n \langle t[v/x] \rangle : \Theta; \dot{E}_{\Theta} \end{array} $	

Table 1. External steps (free core labels)

that the thread is acquainted with the callee. If fact, both requirements will be automatically assured for traces where calls and return occur in correct manner. A commonality for incoming communication from a thread n is that the (only) pair $n \hookrightarrow o$ for some object reference o is removed from E_{Δ} , for which we write $E_{\Delta} \setminus n$. While E_{Δ} imposes restrictions for incoming communication, the commitment context E_{Θ} is updated when receiving new information. For instance in CALLI₂, the commitment \acute{E}_{Θ} after reception marks that now the callee o_2 is acquainted with the received arguments and furthermore that the thread n is visiting (for the time being) the callee o_2 . For outgoing communication, the E_{Δ} and E_{Θ} play dual roles. Note further that E_{Θ} is not mentioned (except for the connectivity of the thread identity n).

2.3 Scoping, lazy instantiation, and guessing of connectivity

Next we discuss the exchange of new names by means of scope extrusion and intrusion and related to that instantiation across the component boundary (cf. Table 2). Besides bound names in the form $\nu(n:T)$, the labels contain also *connectivity* information which is exchanged. This means the labels are of the form $\nu(n:T, E_{\Theta}).\gamma!$ of outgoing and $\nu(n:T, E_{\Delta}).\gamma!$ for incoming communication.

As expected, the ν -binder influences only names occurring freely in the label (cf. rule COMM). In case of a *bound input* with label $\nu(n:T, \tilde{E}_{\Delta}).\gamma$? in rule BIN, the name's scope is extruded into the component. The treatment of the knowledge base E_{Δ} merits a closer look. As specified in $\tilde{E}_{\Delta} \subseteq \Delta' + (\Delta' \times \Theta)$, a new

$n \notin fn(a) \qquad \Delta; E_{\Delta} \vdash C : \Theta, n:T \xrightarrow{a} \Delta; \acute{E}_{\Delta} \vdash \acute{C} : \acute{\Theta}, n:T; \acute{E}_{\Theta} $	
$\Delta; E_{\Delta} \vdash \nu(n:T).C: \Theta \xrightarrow{a} \Delta'; E_{\Delta}' \vdash \nu(n:T).C': \acute{\Theta}; \acute{E}_{\Theta} \setminus n$	
$n\in fn(\gamma) \qquad \varDelta'=\varDelta, n: T E_{\varDelta}'=E_{\varDelta}+\tilde{E}_{\varDelta} \varDelta; E_{\varDelta}\vdash \varDelta'; E_{\varDelta}'\downarrow_{\varDelta\times (\varDelta+\varTheta)}: \varTheta$	
$\Delta'; E'_{\Delta} \vdash C : \Theta \xrightarrow{\gamma?} \acute{\Delta}; \acute{E}_{\Delta} \vdash \acute{C} : \acute{\Theta}; \acute{E}_{\Theta}$	BIN
$\Delta; E_{\Delta} \vdash C : \Theta \xrightarrow{\nu(n:T; \tilde{E}_{\Delta}) \cdot \gamma^{?}} \acute{\Delta}; \acute{E}_{\Delta} \vdash \acute{C} : \acute{\Theta}; \acute{E}_{\Theta}$	BIN
$n \in fn(\gamma)$ $ extstyle Z; E_{ extstyle } eq T: \llbracket \dots rbrace$	
$\underline{\Delta}; E_{\Delta} \vdash (C \setminus n) : \Theta, n:T; E_{\Theta} + E_{\Theta}(C, n) \xrightarrow{\gamma!} \underline{\Delta}; \underline{\acute{E}}_{\Delta} \vdash \underline{\acute{C}} : \underline{\acute{\Theta}}; \underline{\acute{E}}_{\Theta} $ BOUT	
$\Delta; E_{\Delta} \vdash \nu(n:T).C: \Theta; E_{\Theta} \xrightarrow{\nu(n:T; E_{\Theta}^{i=i} \hookrightarrow (C,n)).\gamma!} \acute{\Delta}; \acute{E}_{\Delta} \vdash \acute{C}: \acute{\Theta}; \acute{E}_{\Theta}$	
$o \in fn(\gamma)$ $\Theta \vdash c : \llbracket \dots rbrace$ $C(c) = \llbracket O rbrace$	
$\Theta' = \Theta, o:c \qquad E'_{\varDelta} = E_{\varDelta} + \tilde{E}_{\varDelta} \qquad \varDelta; E_{\varDelta} \vdash \varDelta; E'_{\varDelta} \downarrow_{\varDelta \times (\varDelta + \Theta)}: \Theta$	
$\Delta; E'_{\Delta} \vdash C \parallel o[O] : \Theta'; E_{\Theta} \xrightarrow{\gamma^{?}} \acute{\Delta}; \acute{E}_{\Delta} \vdash \acute{C} : \acute{\Theta}; \acute{E}_{\Theta} $ BINnew	
$\Delta; E_{\Delta} \vdash C : \Theta \xrightarrow{\nu(o:c; \tilde{E}_{\Delta}) \cdot \gamma^{?}} \acute{\Delta}; \acute{E}_{\Delta} \vdash \acute{C} : \acute{\Theta}; \acute{E}_{\Theta}$	
$o\in fn(\gamma)$ $ extstyle extstyle c: \llbracket \ldots rbrace$	
$\underline{\Delta, o:c; E_{\Delta} \vdash C: \Theta; E_{\Theta} + E_{\Theta}^{\overleftarrow{\leftarrow}}(C, n) \xrightarrow{\gamma!} \underline{\Delta}; \underline{\acute{E}_{\Delta}} \vdash \underline{\acute{C}: \Theta; \acute{E}_{\Theta}}}_{\text{BOUT}}$	
$\Delta; E_{\Delta} \vdash \nu(o:c).C:\Theta; E_{\Theta} \xrightarrow{\nu(o:c; E_{\Theta}^{\underline{\leftarrow}} \hookrightarrow (C,n)).\gamma!} \Delta; \acute{E}_{\Delta} \vdash \acute{C}:\acute{\Theta}; \acute{E}_{\Theta}$	

 Table 2. External steps (scoping)

name of an object is related to *some* other objects. Now the gist is to understand that while the rule guesses which acquaintances the new object has, it is not completely free to do so! Since E_{Δ} is maintained as a worst-case assumption about the connectivity of the known external objects, learning about of a fresh object must not contradict this assumption. Intuitively, by creating new objects, which are initially unknown to the component, the environment cannot contact objects it could not contact otherwise. This restriction is captured in the proviso

$$\Delta; E_{\Delta} \vdash \dot{\Delta}; \dot{E}_{\Delta} \downarrow_{\Delta \times (\Delta + \Theta)} : \Theta , \qquad (4)$$

where $\Delta' = \Delta$, n:T in the rule, which requires that the addition of connectivity of the new identity n added to Δ may not lead to new derivable equations for the objects previously known. The requirement, Δ ; $E_{\Delta} \vdash E'_{\Delta} \downarrow_{\Delta \times (\Delta + \Theta)}$: Θ thus stands for the implication: If Δ' ; $E'_{\Delta} \vdash o_1 \rightleftharpoons o_2 : \Theta$, then Δ ; $E_{\Delta} \vdash o_1 \rightleftharpoons o_2 : \Theta$, for all o_1 from Δ and o_2 from Δ, Θ . In other words, E'_{Δ} is a *conservative* extension of E_{Δ} wrt. the old objects. More colloquially, incoming communication brings "no news about old objects" as we assume the worst already.

The rule BOUT for bound output is similar, except that here we can consult the component C and especially the " \hookrightarrow "-parts of it to determine the connectivity part of the label: We add the *all* names from $\Theta' \times (\Theta' + \Delta)$ where $\Theta' = \Theta, n:T$ which according to C are acquainted with n. The set of acquainted objects are

Conclusion

those in the reflexive, transitive, and symmetric closure in the sense of Equation 3 from. We write $E^{\rightleftharpoons \hookrightarrow}(C,n)$ for that set of names. With the scope opened, we remove from C all \hookrightarrow -pairs mentioning n.

Object creation across component boundary is not immediately visible. Instead, new objects are actually created only when first communicated to or used by the other side. We call the mechanism *lazy instantiation*. If the component creates an instance of an external class $c \in \Delta$, a new reference o_3 is generated locally. Additionally it is remembered in a separate "parallel component" $o_1 \hookrightarrow o_3$ that the creator o_1 may now know o_3 . The \sim -step itself is not externally visible:

$c\in\varDelta$	NEWO,
$\varDelta; E_{\varDelta} \vdash n \langle let x : c = [o_1] \; new c in t \rangle : \Theta; E_{\Theta} \rightsquigarrow$	I I I I O lazy
$\Delta; E_{\Delta} \vdash \nu(o_3:c).o_1 \hookrightarrow o_3 \parallel n \langle let x:c = o_3 in t \rangle : \Theta; E_{\Theta}$	

If now the accesses the provisionally created object, it is an external step where the new object is mentioned in the label (cf. rule BOUT_{new}). In contrast to ordinary scope extrusion in BOUT, the one for lazy instantiation extends the *assumption* context Δ . Since furthermore the object is *unknown* in the environment, the assumption context E_{Δ} is used unchanged in the premise of the rule before the output step. The rule BIN_{new} works symmetric, where the component incorporates the component class instance upon request from the environment.

3 Conclusion

In this report we presented, inspired by the work of [4], an operational semantics of a class-based, object-oriented calculus with multithreading. The seemingly innocent step from an *object-based* setting as in [4] to a framework with classes requires quite some extension in the operational semantics to characterize the possible behavior of a component. In particular it is necessary to keep track of the potential *connectivity* of objects of the environment to exclude impossible communication labels.

It is therefore instructive, to review the differences in this conclusion, especially to try to understand how the calculus of [4] can be understood as a special case of the framework explored here.

The fundamental dichotomy underlying the observational definition of equivalence is the one between the inside and the outside: program or component vs. environment or observer. This leads to the crucial difference between objectbased languages, instantiating from objects, and class-based language, instantiating from classes: In the class-based setting, instantiation may *cross the demarcation line between component and environment*, while in the object-based setting, this is not possible: the program only instantiates program objects, and the environment only objects belonging to the environment. All other complications, expounded here at some great length, follow from this difference. The most visible complication is that it is necessary to represent the dynamic object structure into the semantics, or rather an approximation of the connectivity of the environment objects. Another way to see it is, that in the setting of [4], there is only *one clique* in the environment, i.e., in the worst case, which is the relevant one, all environment objects are connected with each other. Since the component cannot create environment objects (or vice versa), never new isolated cliques are created. The object-based case can therefore be understood by invariantly (and trivially) taking $E_{\Delta} = \Delta \times (\Delta + \Theta)$, while in our setting, E_{Δ} may be more specific.

We see this study of the semantics as a step towards a full-abstraction result for the class-based calculus. Other future work is to extend the language and the semantics in a number of ways. One inherent feature of the calculus is that objects are *input enabled*. This disallows to model directly *synchronized methods* as in Java. Another generalization is to consider *cloning* of objects, i.e., to create a replica of an object In a certain way, *instantiation* of a class is just like cloning with the restriction that only objects in their initial state can be obtained by instantiation, while cloning can be applied to an object in mid-life. The ability to create an object in a state different from the initial one makes new observations possible, most notably the branching structured gets exposed. One therefore has to generalize the *linear-time* framework of traces to a *branching-time* view. More challenging is to take seriously the notion of classes in that they are not only considered as generator of new objects by instantiation, but also as template for new classes, i.e., to consider inheritance and subtyping. This makes new "observations" on classes possible, namely by subclassing.

References

- 1. M. Abadi and L. Cardelli. A Theory of Objects. Monographs in Computer Science. Springer, 1996.
- E. Abrahám, M. M. Bonsangue, F. S. de Boer, and M. Steffen. A structural operational semantics for a concurrent class calculus. Technical Report 0307, Institut für Informatik und Praktische Mathematik, Christian-Albrechts-Universität zu Kiel, Aug. 2003.
- A. D. Gordon and P. D. Hankin. A concurrent object calculus: Reduction and typing. In U. Nestmann and B. C. Pierce, editors, *Proceedings of HLCL '98*, volume 16.3 of *Electronic Notes in Theoretical Computer Science*. Elsevier Science Publishers, 1998.
- A. Jeffrey and J. Rathke. A fully abstract may testing semantics for concurrent objects. In *Proceedings of LICS '02.* IEEE, Computer Society Press, July 2002.
- R. Milner, J. Parrow, and D. Walker. A calculus of mobile processes, part I/II. Information and Computation, 100:1–77, Sept. 1992.
- A. M. Pitts and D. B. Stark. Observable properties of higher-order functions that dynamically create local names, or: What's new. In A. M. Borzyszkowski and S. Sokołowski, editors, *Proceedings of MFCS '93*, volume 711 of *Lecture Notes in Computer Science*, pages 122–141. Springer-Verlag, Sept. 1993.
- 7. D. Sangiorgi and D. Walker. The π -calculus: a Theory of Mobile Processes. Cambridge University Press, 2001.

Appendix

A Appendix

The appendix contains further material, especially the formalization of the type system and the operational semantics for internal component behavior. Further material and discussions can be found in the technical report [2].

A.1 Syntax

Table 3 and 4 give the definitions of the types and the abstract syntax. For the types $[(l_1:U_1,\ldots,l_k:U_k)]$ is the type of a class, whose instances support methods labeled l_1 to l_k of the indicated types, and $[l_1:U_1,\ldots,U_k]$ the type of corresponding instances.

 $\begin{array}{l} T ::= B \mid none \mid thread \mid [l:U, \ldots, l:U] \mid n \mid [(l:U, \ldots, l:U)] \\ U ::= T \times \ldots \times T \rightarrow T \end{array}$

Table 3. Types

$C ::= 0 \mid C \parallel C \mid \nu(n:T).C \mid n[(O)] \mid n[O] \mid n\langle t \rangle$	program
$O ::= l = m, \dots, l = m$	object
$m ::= \varsigma(n:T).\lambda(x:T,\ldots,x:T).t$	method
$t ::= v \mid stop \mid let x:T = e in t$	thread
$e ::= t \mid if v = v then e else e$	expr.
$ v.l(v, \dots, v) n.l \leftarrow m currentthread$	
$\mid new n \mid new \langle t angle$	
$v ::= x \mid n$	values

 Table 4. Abstract syntax

A.2 Type system

The type system or static semantics presented next characterizes the well-typed programs. The derivation rules are given in Table 5 and 6.

A.3 Internal steps

For the component-internal steps from Table 7 we distinguish, as in [4], confluent \rightsquigarrow -step and competing τ -steps, i.e., steps that may lead to race conditions. Both kind of steps are invisible to the outside.

$$\frac{\Delta \vdash \mathbf{0}:()}{\Delta \vdash \mathbf{0}:()} \operatorname{T-EMPTY} \qquad \frac{\Delta, \Theta_2 \vdash C_1 : \Theta_1 \qquad \Delta, \Theta_1 \vdash C_2 : \Theta_2}{\Delta \vdash C_1 \parallel C_2 : \Theta_1, \Theta_2} \operatorname{T-PAR} \\
\frac{\Delta \vdash C: \Theta, n:T}{\Delta \vdash \nu(n:T).C: \Theta} \operatorname{T-NU} \\
\frac{; \Delta, c:T \vdash \llbracket O \rrbracket : T}{\Delta \vdash c \llbracket O \rrbracket : (c:T)} \operatorname{T-NCLASS} \qquad \frac{; \Delta, o:c \vdash [O] : [T] \qquad ; \Delta \vdash c : \llbracket T \rrbracket}{\Delta \vdash o[O] : (o:c)} \operatorname{T-NOBJ} \\
\frac{; \Delta, n: thread \vdash t : none}{\Delta \vdash n \langle t \rangle : (n: thread)} \operatorname{T-NTHREAD}$$

 Table 5. Static semantics (components)

A.4 External steps

In contrast to the abridged presentation in Section 2.2 we include here the full rules for external steps. In particular, the static typing premises are listed. The full rules are shown in Table 9.

$\Gamma; \Delta \vdash m_1:T_1 \dots \Gamma; \Delta \vdash m_k:T_k \qquad T = [[l_1:T_1, \dots, l_k:T_k]]$
$\Gamma; \Delta \vdash [(l_1 = m_1, \dots, l_k = m_k)]: T$
$\Gamma; \Delta \vdash m_1:T_1 \dots \Gamma; \Delta \vdash m_k:T_k \qquad T = [l_1:T_1, \dots, l_k:T_k]$
$\Gamma; \Delta \vdash [l_1 = m_1, \dots, l_k = m_k] : T$
$\Gamma, x_1:T_1, \dots, x_k:T_k; \Delta, n:c \vdash t: T' \Gamma; \Delta \vdash c: T T = [(\dots, l:T_1 \times \dots \times T_k \to T', \dots)]$
$\Gamma; \Delta \vdash \varsigma(n:c).\lambda(x_1:T_1, \dots, x_k:T_k).t:T.l$
$\Gamma; \Delta \vdash v: c \Gamma; \Delta \vdash c: \llbracket \dots, l: T_1 \times \dots \times T_k \to T, \dots \rrbracket \Gamma; \Delta \vdash v_1: T_1 \ \dots \ \Gamma; \Delta \vdash v_k: T_k $
$\Gamma; \Delta \vdash v.l(v_1, \dots, v_k) : T$
$\Gamma; \Delta \vdash v: c \qquad \Gamma; \Delta \vdash c: T \qquad \Gamma; \Delta \vdash v': T.f$
$\Gamma; \Delta \vdash v.f := v': c$
$\Gamma; \Delta \vdash c : [(T)] \qquad \qquad \Gamma; \Delta \vdash t : T \qquad \qquad$
$\frac{1}{\Gamma; \Delta \vdash new c: c} \qquad \frac{1}{\Gamma; \Delta \vdash new \langle t \rangle: thread}$
${\Gamma: \Delta \vdash current thread : thread}$
$\Gamma; \Delta \vdash e: T_1 \qquad \Gamma, x:T_1; \Delta \vdash t: T_2$ T_LET
$\Gamma; \Delta \vdash let x: T_1 = e in t : T_2$
$\Gamma; \Delta \vdash v_1 : T_1 \qquad \Gamma; \Delta \vdash v_2 : T_1 \qquad \Gamma; \Delta \vdash e_1 : T_2 \qquad \Gamma; \Delta \vdash e_2 : T_2 \qquad T \text{COND}$
$\Gamma; \Delta \vdash if v_1 = v_2 then e_1 else e_2 : T_2$
${\Gamma; \Delta \vdash stop : T} \text{T-Stop}$
$\Gamma(x) = T \qquad \Delta(n) = T T-VAR \qquad T-NAME$
$\Gamma; \Delta \vdash x: T$ $\Gamma; \Delta \vdash n: T$
$\Gamma; \Delta \vdash block: T \qquad \qquad \Gamma; \Delta \vdash return[o_1] \ v: T'$

Table 6. Static semantics (2)

$$\begin{split} n\langle let \, x:T = v \ in \, t \rangle & \rightarrow n\langle t[v/x] \rangle & \text{Red} \\ n\langle let \, x_2:T_2 = (let \, x_1:T_1 = e_1 \ in \, e) \ in \, t \rangle & \rightarrow n\langle let \, x_1:T_1 = e_1 \ in \ (let \, x_2:T_2 = e \ in \, t) \rangle & \text{Let} \\ n\langle let \, x:T = (\text{if } v = v \ \text{then} \ e_1 \ \text{else} \ e_2) \ in \, t \rangle & \rightarrow n\langle let \, x:T = e_1 \ in \, t \rangle & \text{COND}_1 \\ n\langle let \, x:T = (\text{if } v_1 = v_2 \ \text{then} \ e_1 \ \text{else} \ e_2) \ in \, t \rangle & \rightarrow n\langle let \, x:T = e_2 \ in \, t \rangle & \text{COND}_2 \\ n\langle let \, x:T = (\text{if } v_1 = v_2 \ \text{then} \ e_1 \ \text{else} \ e_2) \ in \, t \rangle & \rightarrow n\langle let \, x:T = e_2 \ in \, t \rangle & \text{COND}_2 \\ n\langle let \, x:T = currenthread \ in \, t \rangle & \rightarrow n\langle let \, x:T = n \ in \, t \rangle & \text{CURRENTTHREAD} \\ c[[O]] \parallel n\langle let \, x:c = [o_1] \ new \ c \ in \, t \rangle & \rightarrow n\langle let \, x:T = n \ in \, t \rangle & \text{NEWO}_i \\ c[[O]] \parallel n\langle let \, x:c = [o_1] \ new \ c \ in \, t \rangle & \rightarrow \nu(n_2:T).(n\langle let \, x:T = n_2 \ in \, t_1 \rangle \parallel n_2\langle t \rangle) & \text{NEWT} \\ n\langle let \, x:T = stop \ in \, t \rangle & \rightarrow n\langle stop \rangle & \text{STOP} \\ o[O] \parallel n\langle let \, x:T = o.l(\vec{v}) \ in \, t \rangle & \stackrel{\tau}{\rightarrow} o[O] \parallel n\langle let \, x:T = O.l(o)(\vec{v}) \ in \, t \rangle & \text{CALL}_i \\ o[O] \parallel n\langle let \, x:T = o.f \ := v \ in \, t \rangle & \stackrel{\tau}{\rightarrow} o[O.f \ := v] \parallel n\langle let \, x:T = o \ in \, t \rangle & \text{FUPDATE} \end{split}$$

 Table 7. Internal steps

 $\begin{array}{l} \gamma ::= n \langle [o] \, call \, \, o.l(\vec{v}) \rangle \mid n \langle return(v) \rangle \mid \nu(n:T).\gamma \\ a ::= \gamma? \mid \gamma! \end{array}$

basic labels receive and send labels

Table 8. Labels

$\begin{array}{ll} ;\Delta\vdash o_{1}:[\ldots] & ;\Delta,\Theta\vdash o_{2}.l(\vec{v}):T o_{2}\in\Theta \Delta\vdash n: thread n\notin dom(\Theta) \\ \Delta;E_{\Delta}\vdash [o_{1}]=\hookrightarrow \vec{v}:\Theta \Delta;E_{\Delta}\vdash [o_{1}]=\hookrightarrow o_{2}:\Theta \Delta;E_{\Delta}\vdash n\coloneqq [o_{1}]:\Theta \\ \dot{E}_{\Delta}=E_{\Delta}\setminus n \dot{\Theta}=\Theta,n: thread \dot{E}_{\Theta}=E_{\Theta}+(o_{2}\hookrightarrow \vec{v},n\hookrightarrow o_{2}) \end{array}$
$\Delta; E_{\Delta} \vdash C : \Theta; E_{\Theta} \xrightarrow{n \langle [o_1] call \ o_2 . l(\vec{v}) \rangle^2} \\ \Delta; E_{\Delta} \vdash C \parallel n \langle let \ x:T = o_2 . l(\vec{v}) \ in \ return[o_1] \ x \rangle : \acute{\Theta}; \acute{E}_{\Theta}$
$\begin{array}{ll} ;\Delta\vdash o_{1}:[\ldots] & ;\Delta,\Theta\vdash o_{2}.l(\vec{v}):T o_{2}\in\Theta \Delta\vdash n:thread\\ E_{\Delta}\vdash [o_{1}]\leftrightarrows \hookrightarrow \vec{v}:\Theta E_{\Delta}\vdash [o_{1}]\leftrightarrows \hookrightarrow o_{2}:\Theta E_{\Delta}\vdash n\leftrightarrows [o_{1}]:\Theta\\ & \acute{E}_{\Delta}=E_{\Delta}\setminus n \acute{E}_{\Theta}=E_{\Theta}+(o_{2}\hookrightarrow \vec{v},n\hookrightarrow o_{2}) \end{array}$
$\begin{aligned} \Delta; E_{\Delta} \vdash C \parallel n \langle let x':T' = [o_2] \ blocks \ for \ o'_2 \ in t \rangle : \Theta; E_{\Theta} \xrightarrow{n \langle [o_1] \ call \ o_2 . l(\vec{v}) \rangle?} \\ \Delta; \acute{E}_{\Delta} \vdash C \parallel n \langle let x:T = o_2 . l(\vec{v}) \ in \ return [o_1] \ x; let \ x':T' = [o_2] \ blocks \ for \ o'_2 \ in \ t \rangle : \Theta; \acute{E}_{\Theta} \end{aligned}$
$E_{\Delta} = E_{\Delta} + ([o_1] \hookrightarrow v, n \hookrightarrow [o_1]) \qquad E_{\Theta} = E_{\Theta} \setminus n$ $\Delta; E_{\Delta} \vdash C \parallel n \langle let x:T = return[o_1] v in t \rangle : \Theta; E_{\Theta} \xrightarrow{n \langle return(v) \rangle!} \Delta; \acute{E}_{\Delta} \vdash C \parallel n \langle t \rangle : \Theta; \acute{E}_{\Theta}$ $c_{\Delta} \in \Delta \qquad \acute{E}_{\Delta} = E_{\Delta} + (o_2 \hookrightarrow \vec{v}, n \hookrightarrow o_2) \qquad \acute{E}_{\Theta} = E_{\Theta} \setminus n$ $E_{\Delta} = E_{\Phi} \setminus n$ $E_{\Delta} = E_{\Phi} \setminus n$ $E_{\Delta} = E_{\Phi} \setminus n$
$ \begin{array}{c} \begin{array}{c} \begin{array}{c} \begin{array}{c} \begin{array}{c} \begin{array}{c} \begin{array}{c} \end{array}\\ \end{array}\\ \end{array} \\ \Delta; E_{\Delta} \vdash C \parallel n \langle let x:T = [o_1] \ o_2.l(\vec{v}) \ in \ t \rangle : \Theta; E_{\Theta} \end{array} \\ \hline \begin{array}{c} \begin{array}{c} \end{array}\\ \begin{array}{c} \end{array}\\ \begin{array}{c} \end{array}\\ \begin{array}{c} \end{array}\\ \begin{array}{c} \end{array}\\ \end{array} \\ CALLO \\ \hline \begin{array}{c} \end{array}\\ \end{array} \\ \Delta; E_{\Delta}; \vdash C \parallel n \langle let x:T = [o_1] \ blocks \ for \ o_2 \ in \ t \rangle : \Theta; E_{\Theta} \end{array} $
$\begin{array}{ccc} ; \Delta, \Theta \vdash v : T & \Delta; E_{\Delta} \vdash o_{2} \leftrightarrows v : \Theta & \Delta; E_{\Delta} \vdash n \hookrightarrow o_{2} : \Theta \\ \hline \dot{E}_{\Delta} = E_{\Delta} \setminus n & \dot{E}_{\Theta} = E_{\Theta} + (o_{1} \hookrightarrow v, n \hookrightarrow [o_{1}]) \\ \hline \end{array} $ RETI
$\begin{aligned} \Delta; E_{\Delta} \vdash C \parallel n \langle let x:T = [o_1] \ blocks \ for \ o_2 \ in \ t \rangle : \Theta; E_{\Theta} \xrightarrow{n \langle return(v) \rangle?} \\ \Delta; E_{\Delta} \vdash C \parallel n \langle t[v/x] \rangle : \Theta; E_{\Theta} \end{aligned}$

 Table 9. External steps

$$\begin{array}{l} \begin{array}{l} n \notin fn(a) \qquad \Delta; E_{\Delta} \vdash C : \Theta, n:T \xrightarrow{a} \Delta; \dot{E}_{\Delta} \vdash \dot{C} : \dot{\Theta}, n:T; \dot{E}_{\Theta} \\ \hline \Delta; E_{\Delta} \vdash \nu(n:T).C : \Theta \xrightarrow{a} \Delta'; E'_{\Delta} \vdash \nu(n:T).C' : \dot{\Theta}; \dot{E}_{\Theta} \setminus n \end{array} \\ \begin{array}{l} n \in fn(\gamma) \qquad \Delta' = \Delta, n:T \quad E'_{\Delta} = E_{\Delta} + \tilde{E}_{\Delta} \quad \Delta; E_{\Delta} \vdash \Delta'; E'_{\Delta} \downarrow_{\Delta \times (\Delta + \Theta)} : \Theta \\ \hline \Delta'; E'_{\Delta} \vdash C : \Theta \xrightarrow{\gamma?} \dot{\Delta}; \dot{E}_{\Delta} \vdash \dot{C} : \dot{\Theta}; \dot{E}_{\Theta} \end{array} \\ \hline D; E_{\Delta} \vdash C : \Theta \xrightarrow{\nu(n:T; E_{\Delta}).\gamma?} \dot{\Delta}; \dot{E}_{\Delta} \vdash \dot{C} : \dot{\Theta}; \dot{E}_{\Theta} \end{array} \\ \begin{array}{l} BIN \\ \hline \Delta; E_{\Delta} \vdash (C \setminus n) : \Theta, n:T; E_{\Theta} + E_{\Theta}(C, n) \xrightarrow{\gamma!} \dot{\Delta}; \dot{E}_{\Delta} \vdash \dot{C} : \dot{\Theta}; \dot{E}_{\Theta} \\ \hline D; E_{\Delta} \vdash \nu(n:T).C : \Theta; E_{\Theta} \xrightarrow{\nu(n:T; E_{\Theta}^{\ominus \leftarrow (C,n)).\gamma!} \dot{\Delta}; \dot{E}_{\Delta} \vdash \dot{C} : \dot{\Theta}; \dot{E}_{\Theta} \\ \hline O \in fn(\gamma) \qquad \Theta \vdash c : [[\dots]] \\ \hline O' = \Theta, o:c \qquad E'_{\Delta} = E_{\Delta} + \tilde{E}_{\Delta} \qquad \Delta; E_{\Delta} \vdash \Delta; E'_{\Delta} \downarrow_{\Delta \times (\Delta + \Theta)} : \Theta \\ \hline \Delta; E'_{\Delta} \vdash C \parallel o[O] : \Theta'; E_{\Theta} \xrightarrow{\gamma?} \dot{\Delta}; \dot{E}_{\Delta} \vdash \dot{C} : \dot{\Theta}; \dot{E}_{\Theta} \\ \hline D \in fn(\gamma) \qquad \Delta \vdash c : [[\dots]] \\ \hline \Delta, o:c; E_{\Delta} \vdash C : \Theta; E_{\Theta} + E_{\Theta}^{\ominus \leftarrow (C,n) .\gamma!} \dot{\Delta}; \dot{E}_{\Delta} \vdash \dot{C} : \dot{\Theta}; \dot{E}_{\Theta} \\ \hline D \in fn(\gamma) \qquad \Delta \vdash c : [[\dots]] \\ \hline \Delta, o:c; E_{\Delta} \vdash C : \Theta; E_{\Theta} + E_{\Theta}^{\ominus \leftarrow (C,n) .\gamma!} \dot{\Delta}; \dot{E}_{\Delta} \vdash \dot{C} : \dot{\Theta}; \dot{E}_{\Theta} \\ \hline D \in fn(\gamma) \qquad \Delta \vdash c : [[\dots]] \\ \hline \Delta, e_{\Delta} \vdash \nu(o:c).C : \Theta; E_{\Theta} \xrightarrow{\nu(o:c; E_{\Theta}^{\ominus \leftarrow (C,n) .\gamma!} \dot{\Delta}; \dot{E}_{\Delta} \vdash \dot{C} : \dot{\Theta}; \dot{E}_{\Theta}} \\ \end{array}$$

 Table 10. External steps (scoping)