Towards full abstraction for class-based, multithreaded OO

— Work in progress — München, February 2003

- introduction, full-abstraction
- object-based and class-based calculus
- issues for full abstraction
- conclusion

- class-based oo
 - mainstream of oo (C⁺⁺, Smalltalk, Java, ...)
 - class as unit of code/reuse (inheritance) and (often) as unit of abstraction (type)
- object-based
 - no classes, no (class)-inheritance
 - dynamic method update

- basically: comparison between 2 semantics, resp. 2 implied notions of equality
- given a reference semantics, the 2nd one is
 - neither too abstract = sound
 - nor too concrete = complete
- Milner [Mil77], Plotkin [Plo77] for λ -calculus/LCF
- various *variations* of the theme

- *reference* semantics:
 - must be natural
 - easy to define
 - non-compositional

contextual, observational

- context C[] = "program with a hole"
- filling the hole with a part of a program (component C): complete program C[C]
- what is a context/component?: depends on the language/syntax (sequential/parallel/functional ... contexts)

 \Rightarrow

- given a closed program $P: \mathcal{O}(P) = observations$
- \Rightarrow observational equivalence:

$$c_1 \equiv_{obs} c_2$$
 iff $\forall \mathcal{C}. \ \mathcal{O}(\mathcal{C}[c_1]) = \mathcal{O}(\mathcal{C}[c_2])$

- given a denotational semantics $[\![_]\!]_{\mathcal{D}}$, resp. the implied equality $\equiv_{\mathcal{D}}$
- $\Rightarrow \equiv_{\mathcal{D}}$ is fully abstract wrt. \equiv_{obs} :



- formal model(s) of oo languages
- in the tradition of the λ -calculi, process calculi ...
- more specifically:
 - object-calculi of Abadi/Cardelli [AC96]
 - π -calculus: processes, parallelism, name-passing [MPW92][SW01]
 - ν -calculus: λ -calc. with name creation (references) respectively its concurrent version [PS93][GH98]

а

- program = "set" of named threads and objects running in parallel: $n\langle t \rangle$ and $n[l_1 = m_1, \dots, l_k = m_k]$
- dynamic scoping of names
 - $\nu n: T. (C_1 \parallel C_2)$
 - ν acts as binder: α -equivalence
 - communication of names changes the scope ("scope extrusion")
- methods = functions with specific "self"-parameter
- active entities: threads
 - sequencing + local, static scoping: let x = e in t
 - thread creation

^aIn the presence of subtyping, the parameter would be late-bound abstraction/OO-p.8

Adding classes

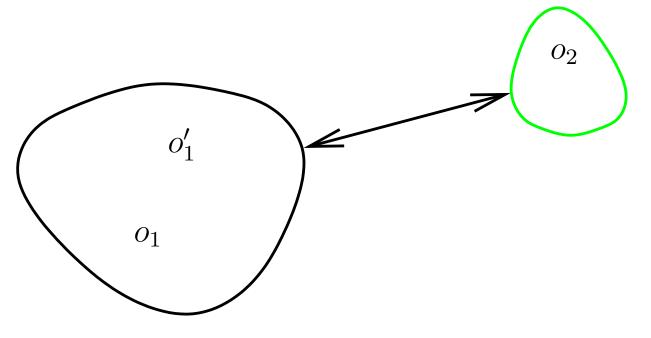
- class: just like objects:
 - named collection of methods $n[(l_1 = m_1, ...)]$
 - instantiated by name, not structure: new n !
 - class names are not first-order citizens, i.e., not subject to
 - ν -binding (= hiding)
 - storing, sending, receiving etc.
- method update not used in class-based setting

Adding classes

Semantics (1)

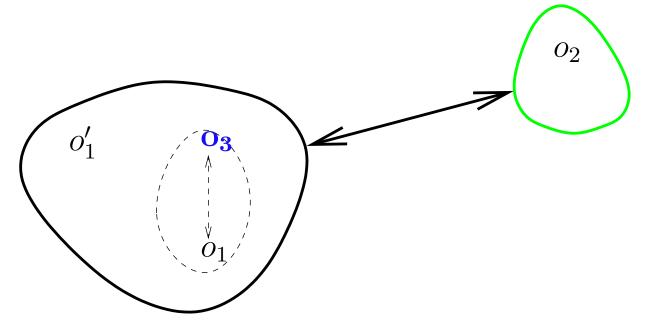
- given in various "stages"
 - internal (configuration-local) steps
 - external, global steps, interacting with the environment
 - computation steps modulo α -conversion
- typed operational semantics

Internal steps



- black: objects of the component
- green: objects of the environment

Internal steps



*o*₁ creates an internal object **o**₃ (assume: thread *n* visits *o*₁)

$$\mathbf{c}[(O)] \parallel n \langle let x:T = new \mathbf{c} in t \rangle \rightsquigarrow \\ c[(O)] \parallel \nu \mathbf{o_3}:\mathbf{T}. \ (n \langle let x:T = \mathbf{o_3} in t \rangle \parallel \mathbf{o_3}[\mathbf{O}]. \)$$

- 4 exemplary axioms
- confluent (\rightarrow) and non-confluent ($\xrightarrow{\tau}$) internal steps
- for CALL_i: O.l(o)(v) in t: parameter passing, and especially replacing the *ς*-bound self-parameter by o.

 $n\langle let x:T = \mathbf{v} in t \rangle \rightsquigarrow n\langle t[\mathbf{v}/\mathbf{x}] \rangle$ RED

$$n\langle let x_2:T_2 = (let \mathbf{x_1}:\mathbf{T_1} = \mathbf{e_1} in \mathbf{e}) in t \rangle \rightsquigarrow n\langle let \mathbf{x_1}:\mathbf{T_1} = \mathbf{e_1} in(let x_2:T_2 = \mathbf{e} in t) \rangle \qquad \mathsf{LET}$$

$$\mathbf{c}[(O)] \parallel n \langle let x: T = new \mathbf{c} in t \rangle \rightsquigarrow c[(O)] \parallel \nu \mathbf{o}: \mathbf{T}. \ (n \langle let x: T = \mathbf{o} in t \rangle \parallel \mathbf{o}[\mathbf{O}]. \) \qquad \mathsf{NEWO}_i$$

 $\mathbf{o}[O] \parallel n \langle let \, x:T = \mathbf{o}.\mathbf{l}(\mathbf{\vec{v}}) \, in \, t \rangle \xrightarrow{\tau} o[O] \parallel n \langle let \, x:T = \mathbf{O}.\mathbf{l}(\mathbf{o})(\mathbf{\vec{v}}) \, in \, t \rangle \qquad \mathsf{CALL}_i$

- "typed" operational semantics
- i.e., labeled steps between typing judgments: $\Delta \vdash P : \Theta$
 - $-\Delta$ = "assumptions"
 - names assumed present in the rest
 - $\Theta =$ "commitments"
 - names guaranteed to the rest
- steps labeled by
 - thread id
 - communicated values
 - kind of communication (!/?, call/return)

External steps (2)

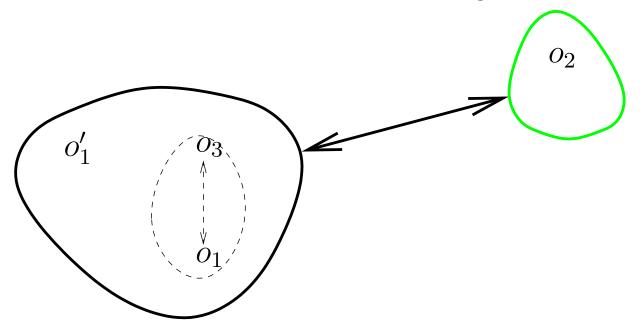
• e.g.: outgoing calls and incoming returns

$$\frac{\mathbf{o} \text{ in } \Delta}{\Delta \vdash C \parallel n \langle \text{let } x:T = o.l(\vec{v}) \text{ in } t \rangle : \Theta \xrightarrow{\mathbf{n} \langle \text{ call } \mathbf{o}.\mathbf{l}(\vec{v}) \rangle!} \Delta \vdash C \parallel n \langle \text{let } x:T = b \text{lock in } t \rangle : \Theta}$$

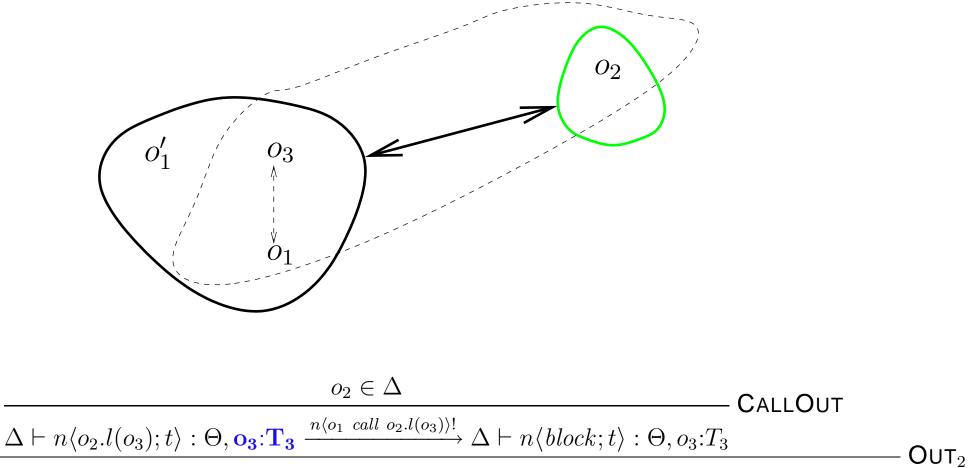
$$\frac{;\Delta, \Theta \vdash v : T}{\Delta \vdash C \parallel n \langle \text{let } x:T = b \text{lock in } t \rangle : \Theta \xrightarrow{\mathbf{n} \langle \text{ return}(\mathbf{v}) \rangle?} \Delta \vdash C \parallel n \langle \mathbf{t}[\mathbf{v}/\mathbf{x}] \rangle : \Theta}$$
RETURNIN

- names
 - for object and thread id's
 - can be generated freshly: "new"
 - valid within dynamic scopes
 - up-to renaming
- dynamic, i.e.,
 - names can be sent around: scope is extended
 - also: across component interface
 - bound exchange of names: " ν "

internal o_3 is sent outside, as argument to method call at o_2



External steps: Scoping



 $\Delta \vdash \nu \mathbf{o_3}.(n \langle o_2.l(o_3); t \rangle \parallel o_3[\dots]) : \Theta \xrightarrow{\nu o_3. n \langle o_1 \ call \ o_2.l(o_3) \rangle !} \Delta \vdash n \langle block; t \rangle \parallel o_3[\dots] : \Theta, \mathbf{o_3}: \mathbf{T_3}$

- instantiation of a class in the context
- external request for instantiation of component class
- scope of the new id: immediate scope extrusion
- \Rightarrow extension of Δ , resp. Θ

$$\frac{c \in \Delta}{\Delta \vdash n \langle let \ x:T = new \ c \ in \ t \rangle : \Theta \xrightarrow{\nu \ \mathbf{o}_3: \mathbf{T}. \ creates \ \mathbf{o}_3!} \Delta, \mathbf{o}_3: \mathbf{T} \vdash n \langle let \ x:T = o_3 \ in \ t \rangle : \Theta} \mathsf{NEWO}$$

$$\frac{C(c) = \llbracket O \rrbracket}{\Delta \vdash C : \Theta \xrightarrow{\nu \ \mathbf{o_3}: \mathbf{T}. creates \ \mathbf{o_3}?} \Delta \vdash C \parallel \mathbf{o_3}[\mathbf{O}] : \Theta, \mathbf{o_3}: \mathbf{T}} \mathsf{NEWI}$$

- [JR02]: for the concurrent ν -calculus
- notion of observation: may-testing equivalence.
 Formalized here: whether a specific context method ("o.success()") is called
- **component** = set of parallelly running objects + threads
- observable: message exchange at the boundary
- ⇒ fully abstract observable behavior = communication traces of the labels of the OS

actually: they use may-preorder.

- classes are the units of exchange: C[n[(O)]]!
- i.e., internal and external classes
- component objects can instantiate external classes

can one use these objects for "observations"?

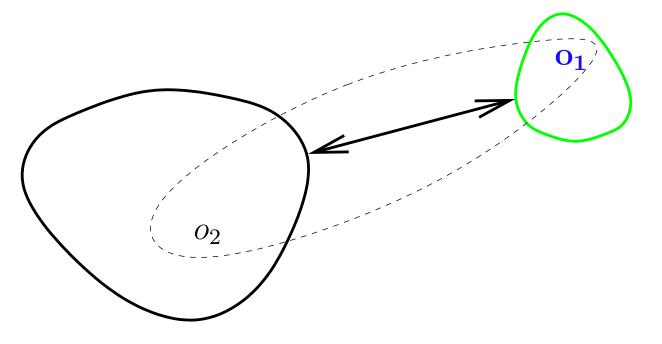
- instances of external classes,
 - instantiation itself is unobservable
 - comm. between component and object observable
 - but:
 - their existence is (principally) unknown to the rest of environment (≠ OC),
 - unless the component gives away their identity!

Consequences/Completeness: Idea

- starting point: component's semantics = set of traces
- Expressibility
- 2 problems for completeness (apart from many technicalities)
 - 1. expressibility \Rightarrow : what are legal traces?
 - 2. what can be observed/distinguished?

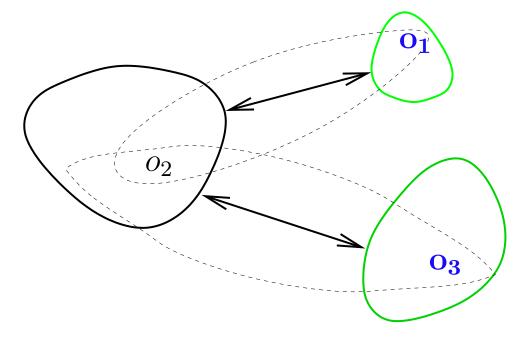
- For completeness: component must realize all potential traces but not more!
- various aspects
 - "global": call-return discipline = balanced/"parenthetic" (per thread)
 - "local"
 - no name clashes: scoping/renaming
 - well-typedness
 - impossible name communication (input)

 Assume: component instantiates two external classes (into o₁ and o₃)



• can o_1 call the component with o_3 as argument?

Impossible incoming names?

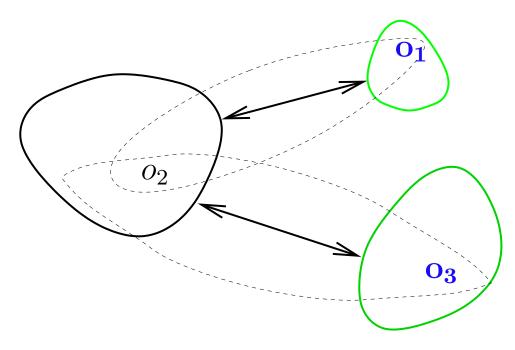


• trace labelled

 $\nu o_1.o_2 creates \mathbf{o_1}!$. $\nu o_3.o_2 creates \mathbf{o_3}!$. $n \langle \mathbf{o_1} \ call \ o_2.l(\mathbf{o_3}) \rangle$?

impossible!

- situation as before:
 - o_1 and o_3 created externally by component



- instantiation itself is not observable
- communication with the 2 objects is observable
- but!: existence of o_1 unknown to o_3 , and vice versa
- \Rightarrow observable are communication traces from/to o_1 and from/to o_3

but not their mutual order!

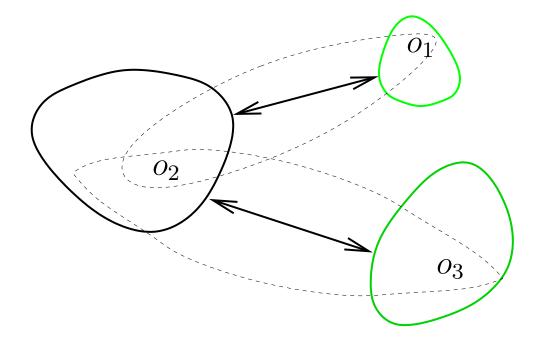
 \Rightarrow separated trace sets

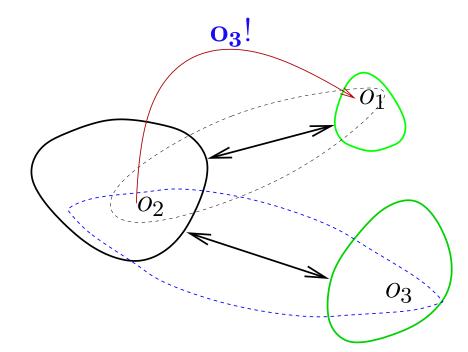
- o_1 and o_3
 - cannot occur in the same label and
 - cannot determine the order of events mutually,

because

they don't "know" of each other

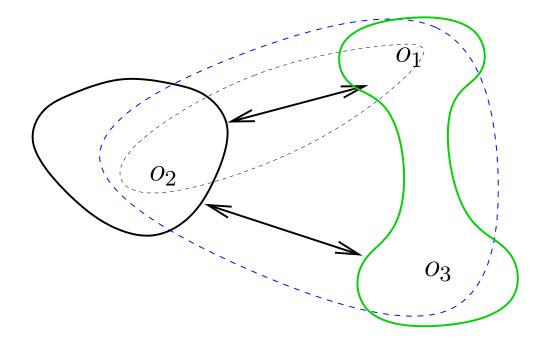
- if "connected", they
 - could occur in the same label and
 - could (in principle) cooperate to observe the order
- connectivity or "acquaintance" is dynamic
- the only one to make o_2 and o_3 acquainted: the component





 $\Delta \vdash n \langle o_1.l(o_3); t \rangle \parallel o_2[\dots] : \Theta, o_2: T_2 \xrightarrow{n \langle o_2 \ call \ o_1.l(o_3) \rangle!} \Delta \vdash n \langle block; t \rangle \parallel o_2[\dots] : \Theta, o_2: T_2$

• no scope extrusion from perspective of the component



- scope enlarged
- o_1 knows o_3
- $\Rightarrow o_3$ could know now o_1 , too

- observers: not just one static outside context but
- dynamic cliques of acquainted objects
 - existing cliques only grow larger: merging
 - new ones can be created by the component
- for full-abstraction:
 - traces per clique, partial-order semantics
 - worst-case: "conspiracy" of environment
- acquaintance = equivalence relation on object id's
- \Rightarrow component keeps track of (the worst-case) of cliques \Rightarrow set of equations; clique: implied equational theory
 - e.g., sending o_1 to o_2 , adds $o_1 \hookrightarrow o_2$ to the equations

- component keeps book about "whom it told what"
- transitions

$$\mathbf{E}; \Delta' \vdash C : \Theta \xrightarrow{a} \mathbf{E'}; \Delta' \vdash C' : \Theta'$$

- $E \subseteq \Delta \times (\Delta + \Theta)$ = pairs of objects
- written $o_1 \hookrightarrow o_2$:
- worst case: equational theory implied by E (on Δ):

 $\mathbf{E} \vdash \mathbf{o_1} \leftrightarrows \mathbf{o_2}$

(for $o_2 \in \Theta$: $E \vdash o_1 \rightleftharpoons; \hookrightarrow o_2$)

- outgoing call to o_2 , \Rightarrow callee now may know the arguments
- \Rightarrow extend *E*

$$c_{2} \in \Delta \qquad E' = E + (\mathbf{o}_{2} \hookrightarrow \vec{\mathbf{v}})$$
$$E; \Delta \vdash C \parallel n \langle let \ x:T = o_{1} \ o_{2}.l(\vec{v}) \ in \ t \rangle : \Theta \xrightarrow{n \langle o_{1} \ call \ o_{2}.l(\vec{v}) \rangle!} \mathbf{E}'; \Delta \vdash C \parallel n \langle let \ x:T = block?o_{1} \ in \ t \rangle : \Theta$$

 same when an argument is sent ν-bound (where Θ is extended, as well)

- *E* unchanged (in first approx.)
- checking for legality:

is, according to E, the incoming label possible?

$$;\Delta, n: thread, \Theta \vdash o_2.l(\vec{v}) : T \qquad o_1 \in \Delta \qquad o_2 \in \Theta$$
$$E \vdash \mathbf{o_2} \longleftrightarrow; \Leftrightarrow \mathbf{o_1} \qquad E \vdash \mathbf{v} \hookleftarrow; \Leftrightarrow \mathbf{o_1} \lor E \vdash \mathbf{v} \Leftrightarrow \mathbf{o_1}$$
$$E; \Delta, n: thread \vdash C : \Theta \xrightarrow{n\langle o_1 \ call \ o_2.l(\vec{v}) \rangle^2} E; \Delta \vdash C \parallel n\langle let \ x:T = \ldots \rangle : n: thread, \Theta$$

Incoming bound values

- incoming ν -bound value
- \Rightarrow value new to the component (i.e., not (yet) in Δ)

 $E; \Delta \vdash C: \Theta \xrightarrow{\nu \mathbf{o}_3: \mathbf{T}_3 \nu o_1: T_1.n \langle o_1 \ call \ o_2.l(o_3) \rangle?}$

- incoming ν -bound value
- \Rightarrow value new to the component (i.e., not (yet) in Δ)
 - $\Delta' = \Delta, \mathbf{o_3}: \mathbf{T_3}$

$$E; \Delta' \vdash C: \Theta \xrightarrow{\nu o_1: \mathbf{T}_1.n \langle o_1 \ call \ o_2.l(o_3) \rangle?} \\E; \Delta \vdash C: \Theta \xrightarrow{\nu o_3: T_3 \nu o_1: T_1.n \langle o_1 \ call \ o_2.l(o_3) \rangle?}$$

- incoming ν -bound value
- \Rightarrow value new to the component (i.e., not (yet) in Δ)
 - $\Delta'' = \Delta', \mathbf{o_1}: \mathbf{T_1}, \mathbf{E}'' = E, \mathbf{o_1} \hookrightarrow \mathbf{o_2}, \mathbf{o_1} \hookrightarrow \mathbf{o_3}$

$$\underbrace{\mathbf{E}''; \Delta'' \vdash C: \Theta}_{E; \Delta' \vdash C: \Theta} \xrightarrow{n\langle o_1 \ call \ o_2.l(o_3) \rangle?} \\
 \underbrace{E; \Delta' \vdash C: \Theta}_{E; \Delta \vdash C: \Theta} \xrightarrow{\nu o_1: T_1.n \langle o_1 \ call \ o_2.l(o_3) \rangle?} \\
 \underbrace{E; \Delta \vdash C: \Theta}_{E; \Delta \vdash C: \Theta} \xrightarrow{\nu o_3: T_3 \nu o_1: T_1.n \langle o_1 \ call \ o_2.l(o_3) \rangle?} \\$$

- incoming ν -bound value
- \Rightarrow value new to the component (i.e., not (yet) in Δ)

$$c_{2} \in \Theta \qquad o_{1} \in \Delta''$$

$$E'' \vdash \mathbf{o_{1}} \leftrightarrows \mathbf{o_{3}} \qquad E'' \vdash \mathbf{o_{2}} \longleftrightarrow; \leftrightarrows \mathbf{o_{1}}$$

$$\overline{E''; \Delta'' \vdash C: \Theta \xrightarrow{n\langle o_{1} \ call \ o_{2}.l(o_{3})\rangle?}}$$

$$E; \Delta' \vdash C: \Theta \xrightarrow{\nu o_{1}:T_{1}.n\langle o_{1} \ call \ o_{2}.l(o_{3})\rangle?}$$

$$E; \Delta \vdash C: \Theta \xrightarrow{\nu o_{3}:T_{3}\nu o_{1}:T_{1}.n\langle o_{1} \ call \ o_{2}.l(o_{3})\rangle?}$$

- incoming ν -bound value
- \Rightarrow value new to the component (i.e., not (yet) in Δ)

$$c_{2} \in \Theta \qquad o_{1} \in \Delta''$$

$$E'' \vdash o_{1} \rightleftharpoons o_{3} \qquad E'' \vdash o_{2} \hookleftarrow; \rightleftharpoons o_{1}$$

$$\overline{E''; \Delta'' \vdash C: \Theta} \xrightarrow{n\langle o_{1} \ call \ o_{2}.l(o_{3})\rangle?} E''; \Delta'' \vdash C': \Theta'}$$

$$\overline{E; \Delta' \vdash C: \Theta} \xrightarrow{\nu o_{1}:T_{1}.n\langle o_{1} \ call \ o_{2}.l(o_{3})\rangle?} E''; \Delta'' \vdash C': \Theta'}$$

$$\overline{E; \Delta \vdash C: \Theta} \xrightarrow{\nu o_{3}:T_{3}\nu o_{1}:T_{1}.n\langle o_{1} \ call \ o_{2}.l(o_{3})\rangle?} E''; \Delta'' \vdash C': \Theta'}$$

- in the setting of [JR02] = may-testing equivalence
 - exactly one kind of observation (e.g., "success")
 - terminal i.e., not repeated observation
- ⇒ trace semantics gets weakened into a partial order semantics, relative to
 - dynamic cliques of connectivity of objects
 - note: we don't allow to observe (e.g.) divergence!
 - note: if we allowed
 - different, repeated observations (for instance success-method + divergence), or
 - if we had a global shared variables (e.g., stdout)
 we are back in linear trace semantics



- operational semantics clear, a generalization of the concurrent ν -calc.
- type system formalized
- exact formulation of the partial-order trace semantics (and proofs ...)

- are classes good composition units?
- what about cloning?
 - cloning means: obtaining an identical copy (up-to the object identity) of an object "on the run"
 - tree semantics
 - bisimulation equivalence instead of traces
- lock-sychronization
- subtype polymorphism & subclassing
- technology transfer to the proof systems, compositionality

References

- [AC96] Martín Abadi and Luca Cardelli. *A Theory of Objects*. Monographs in Computer Science. Springer-Verlag, 1996.
- [ÁMdB00] Erika Ábrahám-Mumm and Frank S. de Boer. Proof-outlines for threads in Java. In Catuscia Palamidessi, editor, *Proceedings of CONCUR 2000*, volume 1877 of *Lecture Notes in Computer Science*. Springer-Verlag, August 2000.
- [ÁMdBdRS02a] Erika Ábrahám-Mumm, Frank S. de Boer, Willem-Paul de Roever, and Martin Steffen. A compositional operational semantics for Java $_{MT}$. Technical Report TR-ST-02-2, Lehrstuhl für Software-Technologie, Institut für Informatik und Praktische Mathematik, Christian-Albrechts-Universität zu Kiel, May 2002.
- [ÁMdBdRS02b] Erika Ábrahám-Mumm, Frank S. de Boer, Willem-Paul de Roever, and Martin Steffen. Verification for Java's reentrant multithreading concept. In Mogens Nielsen and Uffe H. Engberg, editors, *Proceedings of FoSSaCS 2002*, volume 2303 of *Lecture Notes in Computer Science*, pages 4–20. Springer-Verlag, April 2002. A longer version, including the proofs for soundness and completeness, appeared as Technical Report TR-ST-02-1, March 2002.
- [GH98] Andrew D. Gordon and Paul D. Hankin. A concurrent object calculus: Reduction and typing. In Uwe Nestmann and Benjamin C. Pierce, editors, *Proceedings of HLCL '98*, volume 16.3 of *Electronic Notes in Theoretical Computer Science*. Elsevier Science Publishers, 1998.
- [JR02] Alan Jeffrey and Julian Rathke. A fully abstract may testing semantics for concurrent objects. In *Proceedings of LICS '02*. IEEE, Computer Society Press, July 2002.

[Mil77]	Robin Milner. Fully abstract models of typed λ -calculi. <i>Theoretical Computer Science</i> , 4:1–22, 1977.
[MPW92]	Robin Milner, Joachim Parrow, and David Walker. A calculus of mobile processes, part I/II. <i>Information and Computation</i> , 100:1–77, September 1992.
[Plo77]	Gordon Plotkin. LCF considered as a programming language. <i>Theoretical Computer Science</i> , 5:223–255, 1977.
[PS93]	A. M. Pitts and D. B. Stark. Observable properties of higher-order functions that dynamically create local names, or: What's new. In Andrzej M. Borzyszkowski and Stefan Sokołowski, editors, <i>Proceedings of MFCS '93</i> , volume 711 of <i>Lecture Notes in Computer Science</i> , pages 122–141. Springer-Verlag, September 1993.
[SW01]	Davide Sangiorgi and David Walker. The π -calculus: a The-

ory of Mobile Processes. Cambridge University Press, 2001.

33-1