A -calculus with limited resources, garbage-collection and guarantees

David Teller

Abstract.

this study, we attempt to go further, by taking into account notions of allocation, deallocation, reallocation of previously deallocated resources and garbage-collection.

This document presents a process algebra based on the -calculus, the *controlled calculus*, or c , built upon ideas previously expressed in the previous incarnation of the controlled

A -calculus with limited resources, garbage-collection and guarantees

3

ProcessesP, Q ::= (a:r)P / P/Q / iInstructionsi, j ::= 0 / new a:r in i / spawn i in j / a(b).i/ a b .i / !i / i fnull a then i else jContexts $C[\cdot] ::= (a:r)C[\cdot] / C[\cdot]/P / P/C[\cdot] / [\cdot]$ Figure 1. Syntax of c .

hence releasing resources. Complete enough garbage-collection schemes may also free resources held by deadlocks or livelocks. As there are many algorithms which may produce a garbage-collector and as, as we will see, complete garbage-collection is an undecidable problem, we use a parametric relation $_{GC}$ to determine when a channel may be removed.

This enriched -calculus is both simpler and more generic than the original c as well as more adapted to reasoning and programming with resources than the standard -calculus. Well-chosen garbage-collectors permit dynamic handling of resources and exception-like error mechanisms. In turn, this allows the writing of processes which enforce resource usage policies. We complete the language byano1(amic)-388hmolRot isoli1eroved.

their -calculus counterpart. As a syntactical shortcut, we will write \overline{a} for a_1, a_2, \ldots, a_n and $(\overline{a} : \overline{r})$ for $(a_1 : r_1) \ldots (a_n : r_n)$. The sets fn/bn of free/bound names are defined as in the -calculus – note that $fn((a : r)i) = fn(new \ a : r \ in \ i) = fn(i) \quad \{a\}$ and that _ is always free.

R-Par
$$\frac{P - pre P'}{P/Q - pre P'/Q}$$

R-Label $\frac{P - pre P'}{P/Q - pre P'/Q}$
R-Comm $\frac{P \stackrel{a(b)}{-} pre P' - Q \stackrel{a(b)}{-} pre Q'}{P/Q - pre P'/Q}$
R-Entity $\frac{P - pre P'}{(c:r)P - pre (c:r)P'}$
R-Hide $\frac{P - pre P'}{(a:r)P - pre (x:r)P'}$ a /
R-Equiv $\stackrel{P - P' - Q' - Q - P' - pre}{P}$

_

Push p

destruction of names whenever they are not referenced anymore, in a manner similar to that of the traditional -calculus. Such a mechanism is commonly found in programming languages, implemented as a referencecounter in Python, Visual Basic or C++ frameworks such as Microsoft's Com or Mozilla's XPCom. Note that this aspect of garbage-collection is orthogonal to the management of printers themselves.

The Garbage-Collection relation is the smallest $_2$ verifying *x*, *P*, *Q*, {*a* : _} $_2$ *delete a* .*i* / *Q*

Figure 6. A print spooler without dangling pointers

The process *BRMDriver*₂ and the garbage-collection scheme $_2$, presented on figure 6, provide a more robust management of resources. Relation $_2$ mimmicks a generic manual deallocator: any name *a* can be destroyed by calling *delete a*. Since the destruction is handled by the garbage-collector, the semantics of c guarantee that name *a* e ectively disappears.

The manager takes advantage of this deallocator to improve safety. Instead of giving full control to the client, it transmits a (dynamically created) handler, which can be revoked at any time by calling *delete handler*. For this example, revokation only takes place when it is explicitly requested through *destructor*. The printer can then safely be put back onto the pool, without any risk of being reused by the client and without any dangling pointers.

Although this strategy makes deallocation safer, it still does not work whenever a client fails to call the destructor. As in modern programming languages, such problems can be avoided using garbage-collection and finalisation, as shown on figure 7.

Relation ₃ defines a garbage-collection scheme, which supports a mechanism similar to reference-counting, in which names can be removed whenever they only appear in receptions or finalisations, as well as manual deallocation of handlers using *delete*. The notion of finalisation, as encountered in many garbage-collected programming languages such as



Java, C# or OCaml, and as defined here by *F* inalize x.i, triggers a function/method/process (here, *i*) in response to the deallocation of an entity (here, *x*). Note that, as in our previous works [11], and by opposition to these languages, finalisation is safe, insofar as *resurrection* of an entity [1] is impossible. Also note that, by opposition to the first version of c, finalisation is a macro rather than a primitive of the language.

Term *BRMDriver*₃ takes advantage of the automatic garbage-collection and finalisation: *destructor* and *handler* are automatically destroyed, while finalisation premits returning the printer to the pool after the dallocation of *handler*. This behaviour is more robust than that of either *BRMDriver*₁ or *BRMDriver*₂ and could be rendered even more robust by more complete garbage-collectors.

3.2 Error-handling

Let us consider the following scenario: a client has acquired a printer but has started misbehaving, possibly by sending a stream of incorrect instructions to that printer. Assuming that the spooler can detect such a situation, it should stop the printing transaction and return the printer to the pool. A number of other external reasons may require stopping the printing transaction, such as lack of memory or prioritization of a specific client.

These behaviours can be modelled easily, as shown on figure 8, by modifying the garbage-collector to send signals representing the error/ exception. A signal *ERR* is sent to represent a non-deterministic client

BRMDriver₄ = !alloc(request).Pop (x).new destructor in new *handler* in new *sigmem* : *MEM* in new *prio* : *PRIO* in new err: ERR in new flag: Flag in ($| \overline{request} handler, destructor .!handler(y).x y$ / destructor().0 Finalize handler.i fnul | prio then prioritize(c).... el se Push x | Finalize err. delete handler | Finalize prio.delete handler | Finalize mem. delete handler) The Garbage-Collection relation is the smallest 4 verifying ${X}_{3}P$ ${X}_{4}P$ $\{x : ERR\}$ 4 *P* non-deterministically res(P) memory_limit {signal: MEM} 4 P {signal : PRIO, flag : Flag} _ P | prioritize client | flag

Figure 8. A garbage-collected print spooler with signal- and error-handling

error, a signal *PRIO* to represent a reprioritization, requested on channel *prioritize* (*flag* serves to guarantee that only one transaction will be cancelled), and a signal *MEM* is triggered whenever processes use too much memory. In all three cases, the spooler destroys the handler, hence terminating the authorization of the client. If the request was a prioritization, the prioritized client receives a new handler, without going through the queue. Otherwise, the printer is returned to the pool.

The process $BRMDriver_4$ defines the responses of the spooler to these signals. From the point of view of programming languages, *Finalize err*, *Finalize prio* and *Finalize mem* are exception-handlers, comparable to try…catch blocks, although in a concurrent setting.

4 Behaviours and properties

4.1 Properties of the language

Proposition 1 (c can contain).

There is a "good" encoding of the -calculus to an instance of c .

We produce a simple encoding of a monadic synchronous -calculus with structural equivalence and guarded replication, without choice, with a set of names not containing _, to an instance of c with the trivial set of resources and a garbage-collector of unused names. This encoding preserves termination, reduction, structure, distribution, structural equivalence and

A *-calculus with limited resources, garbage-collection and guarantees* 11

barbs.

Proposition 2 (More resources give more freedom). If *S* is a set of resources and if *r* and *s* are elements of *S* such that *r* s then, for any garbage-collection scheme $GC_r - \frac{GC}{r} - \frac{GC}{r}$.

The inclusion derives directly from the definition of $-\frac{GC}{r}$. The nonequality can be proved by examining process (a:s)($a(x)/\overline{a}$), as this process has no reduction in $-\frac{GC}{r}$ and one step of reduction in $-\frac{GC}{s}$.

As in the -calculus, we may observe behaviours of terms in c using barbs and simulations.

4.2 Behaviours

Definition 4 (Barbs).

If *P* and *P* are processes such that $P \stackrel{x(-)}{=}_{pre} P$ (respectively $P \stackrel{\overline{x}}{=}_{pre} P$), we say that *P* has a barb *x*() (respectively \overline{x}). Whenever *P* has a barb , we write *P*.

Definition 5 (Weak barbed simulation). For a resource-aware instance of c on the set of resources S and with a limit n, a relation R is a weak barbed simulation if, whenever (P, Q) = R,

- if P , then Q
- if $P \frac{GC}{n}P$ then, for some Q, $Q \frac{GC}{n}Q$ and)e3910.9ifreedom 9230-301(a)-293(r)1(e)-1(l)1(ati)]TJ/F

Definition 8 (Complete).

A garbage-collection scheme GC is complete if and only if it contains all sound garbage-collection schemes.

Proposition 3 (Perfect garbage-collection). Sound and complete garbage-collection is undecidable.

We prove this by examining process $P = (a:r)(\overline{a} / a().M_b)$ where M_b encodes a Turing machine and emits a message on channel *b* after termination. As a sound and complete garbage-collector must decide whether P = Q, it must also decide whether M_b terminates, hence solve the halting problem.

4.4 Properties of garbage-collectors

Proposition 4 (Print spoolers).

From the garbage-collectors presented in section 3, $_1$ is sound, while $_2$, $_3$ and $_4$ are unsound. None is complete.

Soundness By definition, if $\{a\}_1 P$, *a* is not free in *P*, therefore $P\{a_{a}\} = P$. We also have $(a:r)P P/(a:r)\mathbf{0}$. We can prove easily that $P/(a:r)\mathbf{0} P$.

Unsoundness Let us write

 $P = (handler : Handler)\overline{delete} handler / \overline{handler} a / handler(x).x b$ and

 $Q = \overline{delete} / a / (x).x b$.

We have {handler : Handler} $_2 P$ and P - Q by garbage-collection. Since $P - \frac{1}{a}$ and $Q - \frac{1}{a}$, we conclude that P - Q, hence $_2$ is unsound. The proof is identical for $_3$ and $_4$.

Uncompleteness None of these schemes will garbage collect (a) a b.

Proposition 5 (Actual garbage-collection).

Informally, the Garbage-Collection of Jvm, .Net's Cl i or OCaml is unsound and incomplete.

Unsoundness All three platforms have unsafe weak references, which can be dereferenced even when they point to null. Therefore, assuming that weak is a weak reference, let us consider an extract such as

• Java/Jvm

```
String s = weak.get().toString();
out.println("Action");
```

```
• C#/Cli
```

```
A -calculus with limited resources, garbage-collection and guarantees 13
string s = weak.get().target;
Consol e.WriteLine("Action");
```

• OCaml

```
match Weak.get weak 0 with
   Some x -> print\_endline "Action";;
```

If the garbage-collector has removed the object referenced by ref, a null-pointer or match-failure exception will prevent the observable output "Acti on" from being performed.

Incompleteness As garbage-collection relies purely on the analysis of stack and heap, in the following example, the value of s is never recovered:

```
boolean value = true;
final String s = "useless";
while(value) ;
System.out.println(s);
```

5 A type system for resource guarantees

5.1 The system

The semantics of c are parametrized on a notion of resources. The mechanism of parametric garbage-collection combined with the use of terms such as *Finalize* permit to write systems which take into account allocation of resources as well as deallocations. We now introduce a type system to provide guarantees on the usage of such resources.

$$T ::= Bound(t,) r S, : N - r N ::= Name(C, r) e S C ::= Chan(N, g,) g S, : N - r / Ssh$$

Judgement P: Bound(t,) states that, under environment , P

Figure 9 presents the rules of this type system. For the sake of readability, we slighly alter the syntax to allow writing new a: N in \cdots and (a: N). When necessary, we will write 0 for the function defined on N whose value is uniformly and a r for the function defined on Nwhose value is r for a and for everything else.

Properties

Lemma 1 (Weakening).

If is an environment and P a process such that P: Bound(t,), then, for any t t and any , we have P: Bound(t,).

The proof of this lemma is trivial, as each rule of the type system allows growing t and $\ .$

Theorem 1 (Subject Reduction).

If P is a	process, if	P : Bound	d(r,) and	Ρ –	Ρ	then	there is	s a r
and a	such that	Bound(r ,) and r	x N	(<i>x</i>)	r	x N	(X).

To understand this, let us first consider the case where = 0. This case corresponds to a system closed as far as resource deallocation is concerned, as it does not reuse resources held by free names. In this case, the property becomes r r: the guaranteed bound on resources cannot increase.

The more general case where is not necessarily 0 also covers transitory states between the deallocation of a name and the reuse of resources previously held by that name.

The proof is detailed in the annex.

Theorem 2 (Resource control).

If *S* is a set of resources, if *GC* is a garbage-collection scheme, if *P* is a process, if $P - {}^{GC}P$ and if P : Bound(r, 0) then, for all r r, we also have $P - {}^{GC}P$.

The proof (detailed in the annex) is straight-28(e)(sour)1(c)-7120Td[(:)]TJ/F119.963Tf5.70Td[

Proposition 7 (Print spool er). Typing the print spooler permits us to determine the following properties:

- The spooler uses at most n printers.
- Each incoming call causes the allocation of at most one handler.
- There can be at most *n* handlers running at any time.
- The spooler allocates handlers only on demand.
- The spooler sends messages to the printer only when requested to do so by a client.

The main idea is to use the set of resources N^4 where P: Bound((p, h, k, m),) means that P uses resources to allocate at most p printers, h handlers, k handlers and m messages. For this example, we use both h and k to count handlers, respectively from the point of view of the client and from that of the spooler – creating a handler uses resource (0, 1, 1, 0).

Channel *alloc* serves to transfer resource (0, 1, 0, 0) from the client to the spooler, while channel *handler* serves to transfer resource (0, 0, 0, 1) from the client to the spooler and each printing channel p_1, \dots, p_n serves to transfer resource (0, 0, 0, 1) from the spooler to the printer. Channel *printer* transfers one resource (0, 0, 1, 0) from the pool to the spooler, for allocation to a handler.

It is thus su cient to check that

$BRMDriver_4 | Pool : Bound((n, 0, n, 0), 0)$

to prove the proposition. Conversely, a client will have type

Bound((0, h, 0, m), 0)

if it requests at most *h* printers/handlers and sends at most *m* messages. Depending on the actual type of *request*, *h* can measure either the total number of handlers allocated during the execution of the client or the maximal number of handlers held at any time by the client, assuming that the client uses finalization to recover the resources held by the handler. By using a slightly more complicated set of resources, it is possible to measure both properties at once

The typing derications themselves are long but straightforward.

5.3 Extending the type system

This version of the type system permits transferring resources from an agent to another using a communication channel. This situation, however, fails to take into account the fact that a process may charge for some

 $C ::= Chan(N, g, g, p, p) g, p \quad S, g, p : N - r$ $a : Name(Chan(N, g, g, p, p), ..), b : N \quad i : Bound(t_i g, i g)$ $\underbrace{t_i \quad t_i \quad p \quad i \quad p}_{a(b).i : Bound(t_i, i)} \text{T-Rcv-Exchange}$ $a : Name(Chan(N, g, g, p, p), ..), j : Bound(t_j p, j p) \quad b : N$ $t_j \quad t$ T-Snd-Exchange

language, it starts to deal with error-handling and it adds the notions of transfer of resources.

Rel ated works Other approaches of resource management have been proposed. The BoCa [2] calculus is a variant of Mobile Ambients with a notion of resources which can be dynamically transferred, acquired or released. Our notion of resources held during the execution of a process, in particular, is close to the corresponding notion of weight of a process in that language, although that notion is part of the well-formedness of a BoCa term and is central to the semantics of the calculus.

The Mobile Resource Guarantees [6] project builds on a linear type system to provide guarantees of safe memory deallocation and reuse as well as memory bounds in a single-threaded ML dialect. The Vault project [3] uses in a multithreaded yet safe subset of C and a complex type system to guarantee that resources are in a correct state whenever they are used. TyPiCal [8] has comparable aims with the -calculus. None of these works, however, takes into account garbage-collection.

Several other, mostly dynamic, solutions have been o ered, from Guardians for Mobile Ambients [4] to JML or Spec#'s design-by-contract. These works, however, fail to provide static guarantees, behavioral observation of resources or to take into account deallocation and reuse.

Future developments As we mentioned, instructions such as spawn ··· in ··· and new ··· in ··· instanciate processes or resources and, in an implementation of c , would be accompanied by constructors. Although we have not dealt with constructors for processes, a number of processes such as the print spooler can be seen as constructors for resources, which brings a number of question – firstly, if it is possible to write a constructor in c , how such a constructor should be defined, invoked, and what properties it should have.

Closely related is the question of transformation and composition of resources. While some resources, such as hard drive space and perhaps some authorizations, can be composed into bigger resources, and while we can take this into account at the level of typing, at the level of the language, we have no way of express such behaviour. Similarely, while some resources can be transformed by operations – such as a *file* becoming an *opened file*, our definition of resources is insu cient to model this.

We have started working on all these problems. Preliminary results seem to indicate that the controlled -calculus and its type system may be adapted to take into account constructors, composition and transformations and to provide static guarantees based on the state of resources.

We have also started to investigate whether the notion of static re-

source exchange could be generalized to more than two participants, perhaps using some form of n-ary communication as seen in the Join-Calculus [5] or in the Kell-Calculus [9].

Garbage-collection schemes raise another series of questions. As we have seen, our definitions of soundness and completeness of a garbage-collector are too restrictive for common garbage-collectors such as those found in Java, C# or OCaml. We thus hope to better criteria to classify such services.

More importantly, we have observed that nearly all the garbage-collection schemes we have been using in our examples, both in this document and during our research, could be classified as simple cases of pattern-matching. We wonder whether this observation can be generalized and if a "useful" set of garbage-collectors can be easily defined. In particular, we have attempted to define stack-based as well as regions-based techniques as instances of and preliminary results lead us to believe in the feasibility of the task.

References

- [1] K. Arnold and J. Gosling. *The Java Programming Language*. Addison-Wesley, 1998.
- [2] F. Barbanera, M. Bugliesi, M. Dezani, and V. Sassone. A calculus of bounded capacities. In *Proceedings of Advances in Computing Science, 9th Asian Computing Science Conference, ASIAN'03*, volume 2896 of *Lecture Notes in Computer Science.* Springer, 2003.
- [3] R. DeLine and M. Fahndrich. Enforcing high-level protocols in low-level software. In SIGPLAN Conference on Programming Language Design and Implementation, 2001.
- [4] G. Ferrari, E. Moggi, and R. Pugliese. Guardians for ambient-based monitoring. In V. Sassone, editor, *F-WAN: Foundations of Wide Area Network Computing*, number 66 in ENTCS. Elsevier Science, 2002.
- [5] C. Fournet and G. Gonthier. The reflexive cham and the join-calculus. In Proceedings of the 23rd ACM SIGPLAN-SIGACT symposium on Principles of programming languages. ACM Press, 1996.
- [6] M. Hofmann. A type system for bounded space and functional in-place updateextended abstract. *Nordic Journal of Computing*, 7(4), Autumn 2000. An earlier version appeared in ESOP2000.
- [7] H. P. Hofstee. Power e cient processor architecture and the cell processor. In HPCA, pages 258–262. IEEE Computer Society, 2005.
- [8] N. Kobayashi. TyPiCal: Type-based static analyzer for the pi-calculus.
- [9] J.-B. Stefani. A calculus of kells. In V. Sassone, editor,

A -calculus with limited resources, garbage-collection and guarantees 21 Garbage-collection If $P = (a)\mathbf{0}$ and $Q = \mathbf{0}$ then $[\![P]\!]_p = [\![Q]\!]_p$. Lemma 5 (Null values). For any P and p such that $[\![P]\!]_p - p, p$ does not contain any occurrence of _. Trivial.

Proposition 8 (Soundness). For all processes P and Q of the -calculus such that P - Q, we have $[\![P]\!]_p - [\![Q]\!]_p$.

By induction hypothesis, we also have

 $i\{x \ a\}$: Bound(t r, x^{a}) $j\{x \ a\}$: Bound(t, x^{a})

Let us prove that the relation between , i and j still holds after substitution. Let us write = i (x r).

For any *z* distinct of *x* and *a*, we have

$$x a(Z) = (Z) j(Z) = x a$$

and

$$^{X} a(Z) = (Z) (Z) = ^{X} a(Z)$$
.

We also have

$$x^{X} = a(X) = x^{X} = a(X) = x^{X$$

Also,

$$x = a(a) = (x) (a)$$
.

Z) .

A -calculus with limited resources, garbage-collection and guarantees 23 Trivially, we have weight(Bound(t,)) Bound(t, r, i). Which proves the case.

Communication Let us write

, x : N i : Bound(t r, _a) a : Name(Chan(N, r, _a), ₋) j : Bound(t_j, _j) b : N

		Typage de <i>P</i>	
		Typage de <i>a</i> (x).i	
, x : N	<i>i</i> :	Bound(t r, _a)	Par hypothse
	<i>a</i> :	Name(Chan(N, r, a), _)	Par hypothse
	a(x).i:	$Bound(t_1, 1)$	Par T-Rc∨
	Avec	$t_1 t$	
	1		
		Typage de ā b .j	
	<i>j</i> :	Bound(tj, j)	Par hypothse
	<i>a</i> :	Name(Chan(N, r, a), _)	Par hypothse
	<i>b</i> :	N	Par hypothse
	а <i>b.j</i> :	Bound(t_2 , ₂)	Par T-Snd
	Avec	t ₂ t _j r	
	2 j a		
		Typage de P	
	a(x).i:	$Bound(t_1, 1)$	Cf. plus haut
	:	$Bound(t_2, 2)$	Cf. plus haut
	P :	$Bound(t_3, 3)$	Par T-Par
	Avec	t_3 t_1 t_2	
	3 1 2		

Typage de <i>Q</i>						
Typage de <i>i{x b}</i>						
, x : N	i{x	b}:	Bound(t	r,	a)	Par hypothse
	<i>i</i> :		Bound(t	r,	a)	Par Substitution
Typage de <i>Q</i>						
	i{x	b}:	Bound(t	r,	a)	Cf. plus haut
	j :		Bound(tj	, j)		Par hypothse
Comme t_3 t_1 t_2						

David Teller t_2 t_j r t_1 tComme 3 1 2 1 2 j aQ: Bound(t_3 , 3) Par T-Par

The case is proved.

Structure Proof of the various structural cases are identical to the corresponding proofs in our previous works [13].

Garbage-collection Cases GC-Receive, GC-Send, GC-RReceive and GC-RSend are trivial as *Q* is **0**, which can always be typed, with any type.

Case GC-Deallocate derives directly from the substitution lemma (lemma 6).

The induction is thus proved. Hence the subject-reduction property.

C Resource control

Lemma 8 (Resource total). If P : Bound(r, 0) then res(P) r. Trivial.

C.1 Main proof

From the Resource total lemma and subject-reduction, we conclude the resource control theorem.

D Typing finalization

We have

 $Loop = Name(Chan(_, r,),)$ i: Bound(r r_n,) r r x r_n

Typage de *Toop*

	0	: <i>Bound</i> (,0)	Par <i>T-Nil</i>				
	loop	: Name(Chan(_, r', _'), _)	Par hypothse				
	Τοορ	: <i>Bound</i> (<i>r'</i> , ')	Par T-Snd				
	Typage de i f	null <i>x</i> then <i>i</i> else loop					
	Тоор	: <i>Bound</i> (<i>r</i> ′, ′)	Cf. plus haut				
	i	: Bound(r' r _n ,)	Par hypothse				
	ifnull <i>x</i> then <i>i</i> else <i>loop</i>	: Bound(r', ')	Par T-Test-Nil				
	Typage de <i>loop</i> ().i fnul I x then i el se loop						
-	ifnull <i>x</i> then <i>i</i> el se <i>loop</i>	: <i>Bound</i> (<i>r</i> ′, ′)	Cf. plus haut				
	Гоор	: Name(Chan(_, r', '), _)	Par hypothse				
	loop()	: <i>Bound</i> (,0)	Par T-Rcv				
	Typage de ! <i>loop</i> ().i fnul I x then i el se loop						
<u>.</u>	loop()	: <i>Bound</i> (,0)	Cf. plus haut				
	! <i>loop</i> ()	: <i>Bound</i> (,0)	Par T-Bang				
Typage de ! <i>loop</i> ().i fnul I x then i el se loop / loop							
	! <i>loop</i> ()	: <i>Bound</i> (,0)	Cf. plus haut				
	Тоор	: Bound(r', ')	Cf. plus haut				
	! <i>loop</i> () · · · / <i>loop</i>	: $Bound(r', \prime)$	Par T-Par				
Typage de <i>F inalize x.i</i>							
	! <i>loop</i> () · · · / <i>loop</i>	: Bound(r', ')	Cf. plus haut				
	$(loop: Loop)(\cdots)$: <i>Bound</i> (<i>r</i> ′, ′)					

A -calculus with limited resources, garbage-collection and guarantees 25