Using Ambients to Control Resources (long version)

David Teller, Pascal Zimmer, Daniel Hirschkoff

Using Ambients to Control Resources*

David Teller[1], Pascal Zimmer[2], and Daniel Hirschkoff[1]

[1] LIP - ENS Lyon, France, {David.Teller,Daniel.Hirschkoff}@ens-lyon.fr
[2] INRIA Sophia Antipolis, France, Pascal.Zimmer@sophia.inria.fr

Abstract

Current software and hardware systems, being parallel and reconfigurable, raise new safety and reliability problems, and the resolution of these problems requires new methods. Numerous proposals attempt at reducing the threat of bugs and preventing several kinds of attacks. In this paper, we develop an extension of the calculus of Mobile Ambients, named *Controlled Ambients*, that is suited for expressing such issues, specifically Denial of Service attacks. We present a type system for Controlled Ambients, which makes static resource control possible in our setting.

Introduction

The latest generation of computer software and hardware makes use of numerous new technologies in order to enhance flexibility or performances. Most current systems may be dynamically reconfigured or extended, allow parallelism or use it, and can communicate with other systems. This flexibility, however, induces the multiplication of subsystems and protocols. In turn, this multiplication greatly increases the possibility of bugs, the feasibility of attacks and the sensitivity to possible breakdown of individual subsystems.

This paper presents a formalism for *resource control* in parallel, distributed, mobile systems, called *Controlled Ambients* (CA for short). The calculus of CA is based on Mobile Ambients [5], extends Safe Ambients [18], and is equipped with a type system to express and verify resource control policies.

In the first section, we present our point of view on the problem of resource control. We provide motivations for using ambient calculi to represent the notion of resource in a distributed setting, and claim that a specific calculus should be designed for the purpose of guaranteeing some control on the use of resources. In Sec. 2, we introduce our calculus of Controlled Ambients and explain why it matches our goals. We then develop in Sec. 3 a type system which uses the specifics of this language to make resource control possible; we prove its correctness (i.e. that it does indeed monitor the acquisition and release of resources), and use it to treat several examples. We then discuss some refinements of our type system, and, in the last section, we present possible extensions of this study as well as related works.

*Work supported by european project FET - Global Computing. This paper is an extended version of [26] – full proofs of the results presented here can be found in [25].

1 Resource Control

For the sake of the present study, we define a resource as an entity which may *at will* be acquired, used, then released. We thus work with a rather broad notion of resource, that encompasses ports, CPUs, computers or RAM, but not time, or (presumably) money. A *resource-controlled system* is a system in which no subsystem will ever require more resources than may be available.

In order to prevent problems such as Denial of Service attacks, we need a formalism making resource control possible. This formalism should in particular provide means to describe systems in terms of resource availability and resource requirement, and should also support the description of concurrent and mobile computations. Lastly, the model should provide some kind of entity that can be regarded as a resource. We now present Ambient calculi, and explain why they can be used for these purposes (see also Sec. 5 for a discussion of related works).

Ambient Calculi. Ambient Calculi are based on the notion of locality: each *ambient* is a site. In turn, any ambient may contain subambients, as well as *process*es, controlling its behaviour through the use of *capabilities*. Capabilities let the structure of ambients evolve: $\mathsf{in}\ m$ and $\mathsf{out}\ m$ let an ambient move (resp. entering ambient m or leaving ambient m), while $\mathsf{open}\ m$ opens ambient m and releases its contents in the current ambient. This is expressed by the following reduction rules of the Mobile Ambients calculus [5], that describe the basic evolution steps (captured by relation $\longrightarrow$) of terms:

$$\begin{array}{rcll} m[\mathsf{in}\ n.P \mid Q] \mid n[R] & \longrightarrow & n[m[P \mid Q] \mid R] & \textit{m entering ambient n} \\ n[m[\mathsf{out}\ n.P \mid Q] \mid R] & \longrightarrow & m[P \mid Q] \mid n[R] & \textit{m exiting ambient n} \\ \mathsf{open}\ m.P \mid m[Q] & \longrightarrow & P \mid Q & \textit{opening ambient m} \end{array}$$

In the terms above, $n[P]$ stands for process P running at site (or, equivalently, *ambient*) n, while $\mid$ denotes parallel composition of terms. Hence for instance $n[P] \mid n'[P']$ represents two adjacent sites named n and n', with their corresponding contents P and P'. A capability can be used to prefix a term (as for instance in $\mathsf{open}\ m.P$), which results in a process liable to execute this capability when appropriate, as defined by the rules for $\longrightarrow$. When a capability is triggered, it is consumed by the corresponding reduction step. A more precise, formal, definition of the syntax and semantics of Ambients will be provided below, when we present our calculus of Controlled Ambients.

To draw some analogies with real systems, the in and out primitives can represent the movement of data in a computer or in a network, while open could be used for cleaning memory, for reading data or for loading programs into memory. As for ambients, they could stand for computers, programs, data, components...

These correspondences open the way for a natural model of resource control, where each site may have a finite (or infinite) quantity of resources of a given category. Resources will be used for data, programs, ... In other words, each ambient has a given *capacity* and each subambient *uses* a part of this capacity. Basically, *controlling resources means checking the number of* **direct** *subambients (according to the amount of resources these are using) which may be present in one ambient at any time.*

Message emitted by client *client* at site *from* to call a cab

$$call\ from\ client \triangleq call[\mathsf{out}\ client.\mathsf{out}\ from.\mathsf{in}\ cab.\mathsf{in}\ from.loading[\mathsf{out}\ cab.\mathsf{in}\ client]]$$

Instructions given by client *client* going from site *from* to site *to*

$$trip\ from\ to\ c \triangleq trip[\mathsf{out}\ client.\mathsf{out}\ from.\mathsf{in}\ to.unloading[\mathsf{in}\ c]]$$

The client itself, willing to go from *from* to *to*

$$\begin{aligned} client\ from\ to \triangleq\ & (\boldsymbol{\nu}\, c)c[call\ from\ c \mid \mathsf{open}\ loading.\mathsf{in}\ cab.trip\ from\ to\ c \\ & \mid \mathsf{open}\ unloading.\mathsf{out}\ cab.bye[\mathsf{out}\ c.\mathsf{in}\ cab.\mathsf{out}\ to]] \end{aligned}$$

The cab and the city

$$cab \triangleq cab[\mathsf{rec}\ X.\mathsf{open}\ call.\mathsf{open}\ trip.\mathsf{open}\ bye.X]$$

$$city \triangleq city[cab \mid cab \mid \cdots \mid site_1[client\ site_1 site_i \mid client\ site_1\ site_j \mid \cdots] \mid \cdots \mid site_i[\cdots]]$$

Figure 1: Cab protocol - first attempt

Do note that we could have chosen different points of view and decided to take into account all subambients at all depths, or possibly only "leaf" ambients. We believe, however, that our approach is more general and flexible, which is the reason why we chose it.

An example. We shall use as our main running example a cab protocol: the system consists of one city, n sites, and several cabs and clients. Cabs may be either "anywhere in the city" or in a precise site. Each client may be either in a given site or in a cab. Any client may call a cab, asking for a trip from a site to another site.

In this scenario, several non-trivial properties concerning the interaction among participants and the managment of resources may be expressed. Typically, we impose that if a cab is available, one (and only one) cab must come fetch the client and bring her to her destination. Moreover, if we consider the unique passenger seat of a cab as a resource, the system will be resource-controlled if each cab contains at most one client at any time.

Fig. 1 presents the cab protocol as written in the calculus of Mobile Ambients[1]. The *city* itself is an ambient, which may contain *site*s and *cab*s. Each site s is in turn an ambient, which may contain *client*s, and ambient movements are used to simulate the movements in the protocol (client entering a cab, cab moving from site to site,...). In order for this protocol to work, there must be at least one *cab* and each "*client from to*" declaration must be coherent, i.e. *from* must be the name of the site which hosts the *client* and *to* must be the name of some site.

In order to call a cab, the client sends a *call* ambient. This ambient then enters a cab, where it gets opened. Opening ambient *call* unleashes process

$$\mathsf{in}\ from.loading[\mathsf{out}\ cab.\mathsf{in}\ client]\,.$$

Therefore, after opening, the cab goes in *from*, to meet its client, and releases ambient *loading*. Once *loading* has been released, it enters *client*. As soon as the client opens *loading*, she knows that the cab is present, and therefore that she may enter it. Consequently, the client enters the cab and releases

[1]As a matter of fact, we are not exactly using the original MA calculus, since we work with a recursion operator (rec) instead of replication, which suits better our purposes.

ambient *trip*, which the cab, in turn, receives and opens. Once again, a process is unleashed: $\mathsf{out}\ from.\mathsf{in}\ to.unloading[\mathsf{in}\ c]$. This process moves the cab to its destination and releases another synchronization ambient, *unloading*, to tell the client she may get out. When the client receives this ambient, she opens it, leaves, and sends the last synchronization ambient *bye* to the *cab*, to tell it it may leave.

Limitations. By examining the code of Fig. 1, one may see that several aspects of this implementation may lead to unwanted behaviors. The most visible flaw is the sending of ambient *bye*: if, for any reason, there are several *cab*s in the site, nothing guarantees that *bye* will reach the right *cab*. And if it does not, it may completely break the system by making one *cab* wait forever for its client to exit, although it already has left, while making the other *cab* leave its destination site with its unwilling *client*. In turn, the *client* may then get out of the cab about anywhere.

Although this problem is partly due to the way this implementation has been designed, its roots are deeply nested within the calculus of Mobile Ambients itself. One may notice that any malicious ambient may, at any time, enter the cab: in the calculus of Mobile Ambients, there is no such thing as a filtering of entries/exits. This lack of filtering and accounting is a security threat as well as an obstacle for resource control: for security, since it prevents modeling a system which could check and refuse entry to unwanted mobile code, and for control, since one cannot maintain any information about who is using which resources in a given ambient.

Towards a better control. Difficulties with security and control are due, for the greatest part, to the nature of capabilities in, out and open. Actually, the way these capabilities are used seems too simplistic: in any real system, arrival or departure of data cannot happen without the consent of the acting subsystem, much less go unnoticed, not to mention the opening of a program. In practice, if a program wishes to receive network information, it must first "listen" on some communication port. If a binary file is to be loaded and executed, it must have some executable structure and some given entry point.

A calculus derived from Mobile Ambients is presented in [18]; in this calculus of *Safe Ambients*, three *cocapabilities* are introduced, which we will note $\overline{\mathsf{SAin}}$, $\overline{\mathsf{SAout}}$ and $\overline{\mathsf{SAopen}}$. When executed in m, capability $\overline{\mathsf{SAin}}\ m$ allows an ambient to enter m (by execution of capability $\mathsf{in}\ m$). Similarly, $\overline{\mathsf{SAout}}\ m$ allows an ambient to leave m using $\mathsf{out}\ m$, while $\overline{\mathsf{SAopen}}\ m$ allows m's parent to open m using $\mathsf{open}\ m$. These cocapabilities make synchronizations more explicit and considerably decrease the risk of security breaches. Getting back to the example above, a rewritten cab may thus easily refuse entry right to parasites as long as it is not in any site, or while it contains a client. Moreover, a form of resource control is indeed possible, since an ambient having no more available resource may refuse entrance of new subambients.

However, in this model, ambients are not always warned when they receive or lose subambients by some kind of side effect: in Safe Ambients, when the process $h[m[n[\mathsf{out}\ m] \mid \overline{\mathsf{SAout}}\ m]]$ evolves to $h[m[\mathbf{0}] \mid n[\mathbf{0}]]$, h receives n from m but is not made aware of this. Moreover, while $\overline{\mathsf{SAin}}\ m$ serves as a warning for m that it will receive a new subambient, m does not know which one. Since a subambient representing static data and another one modeling some internal message will not occupy the same amount of resources, this model is probably not sufficient for our purposes.

[13] offers an alternative to these cocapabilities, in order to further enhance systems' robustness: in this formalism, $\overline{\mathsf{in}}\ m$ does not allow *entering* m but rather *m to enter*. This approach solves one of our problems: identifying incoming data. Controlled Ambients, that shall be presented in the next section, may be considered as a development of [13] towards even more robustness as well as resource control. Let us also mention [19], where a different mechanism for the $\overline{\mathsf{SAout}}$ cocapability w.r.t. [18] is introduced. Our proposal subsumes the solutions of [19] and [18].

Embedding resource control. In Sec. 3, we equip our langage with a type system for resource control. Basically, the type of an ambient carries two informations:

– its *capacity* - how many resources the ambient offers to its subambients;
– its *weight* - how many resources it requires from its parent ambients.

The type system allows one to statically divide the available resources between parallel processes, and check that resources will be controlled along movements and openings of ambients.

2 The Language of Controlled Ambients

2.1 Syntax and Semantics

In CA, each movement is subject to a 3-way synchronization between the moving ambient, the ambient welcoming a new subambient and the ambient letting a subambient go. As for the opening of an ambient, it is triggered by a synchronization between the opener and the ambient being opened. These forms of synchronization are somewhat reminiscent of early versions of Seal [28]. Interaction is handled using *cocapabilities*: $\overline{\mathsf{in}_\uparrow}$, $\overline{\mathsf{out}_\uparrow}$, $\overline{\mathsf{in}_\downarrow}$, $\overline{\mathsf{out}_\downarrow}$ and $\overline{\mathsf{open}}$.

$\overline{\mathbf{in}_\uparrow}\ m$ the *up coentry*, welcomes m coming from a subambient;

$\overline{\mathbf{in}_\downarrow}\ m$ the *down coentry*, welcomes m coming from the parent ambient;

$\overline{\mathbf{out}_\uparrow}\ m$ the *up coexit*, allows m to leave the current ambient by exiting it;

$\overline{\mathbf{out}_\downarrow}\ m$ the *down coexit*, allows m to leave by entering a subambient;

$\overline{\mathbf{open}}\ \{m, h\}$ the *coopening*, allows the parent ambient h to open the current ambient m.

Do note that the direction tags $\uparrow$ and $\downarrow$ are not strictly necessary for resource control. We added them since we found they ease the task of specification in mobile ambients. We will return on the use of these annotations in Sec. 2.3.

The syntax of Controlled Ambients is presented in Fig. 2. We suppose we have two infinite sets of term variables, ranged over with capital letters (X, Y), and of names, ranged over with small letters $(m, n, h, x, \ldots)$. Name binders (input and restriction) are decorated with some type information, that shall be made explicit in the next section. While several proposals for Mobile Ambient

P	$::=$	$\mathbf{0}$	null process	M	$::=$	$\mathsf{in}\ m$	enter m
	$\mid$	$M.P$	capability		$\mid$	$\mathsf{out}\ m$	leave m
	$\mid$	$m[P]$	ambient		$\mid$	$\mathsf{open}\ m$	open m
	$\mid$	$P_1 \mid P_2$	parallel composition		$\mid$	$\overline{\mathsf{in}_\uparrow}\ m$	m may enter upwards
	$\mid$	$(\boldsymbol{\nu}\, n : A)P$	restriction		$\mid$	$\overline{\mathsf{in}_\downarrow}\ m$	m may enter downwards
	$\mid$	$\mathsf{rec}\ X.P$	recursion		$\mid$	$\overline{\mathsf{out}_\uparrow}\ m$	m may leave upwards
	$\mid$	X	process variable		$\mid$	$\overline{\mathsf{out}_\downarrow}\ m$	m may leave downwards
	$\mid$	$(n : A)P$	abstraction		$\mid$	$\overline{\mathsf{open}}\ \{m, h\}$	h may open m
	$\mid$	$\langle m \rangle$	message emission				

Figure 2: Controlled Ambients – Syntax

$$P \equiv P \,|\, \mathbf{0} \qquad P \,|\, Q \equiv Q \,|\, P \qquad P \,|\, (Q \,|\, R) \equiv (P \,|\, Q) \,|\, R$$

$$(\boldsymbol{\nu}\, n : A)\, \mathbf{0} \equiv \mathbf{0} \qquad (\boldsymbol{\nu}\, n : A)(\boldsymbol{\nu}\, m : B)\, P \equiv (\boldsymbol{\nu}\, m : B)(\boldsymbol{\nu}\, n : A)\, P$$

$$(\boldsymbol{\nu}\, n : A)\, (P \,|\, Q) \equiv ((\boldsymbol{\nu}\, n : A)\, P) \,|\, Q \quad \text{if } n \notin fn(Q)$$

$$(\boldsymbol{\nu}\, n : A)\, m[P] \equiv m[(\boldsymbol{\nu}\, n : A)\, P] \quad \text{if } n \neq m$$

Figure 3: Controlled Ambients – Structural Congruence

calculi use replication, infinite behaviour is represented using recursion in CA. This is mostly due to the fact that recursion allows for an easier specification of loops, especially in the context of resource consumption. Note also that, compared to the original calculus of Mobile Ambients, we restrict ourselves to communication of ambient names only, and we do not handle communicated capabilities.

The null process $\mathbf{0}$ does nothing. Process $M.P$ is ready to execute M, then to proceed with P. $P \,|\, Q$ is the parallel composition of P and Q. $m[P]$ is the definition of an ambient with name m and contents P. The process $(\boldsymbol{\nu}\, n : A)P$ creates a new, private name n, then behaves as P. The recursive construct $\mathsf{rec}\ X.P$ behaves like P in which occurences of X have been replaced by $\mathsf{rec}\ X.P$. Process $(n : A)Q$ is ready to accept a message, then to proceed with Q with the actual message replacing the formal parameter n. $\langle m \rangle$ is the asynchronous emission of a message m. In most cases, we omit the terminal $\mathbf{0}$ process. We say that a process is *prefixed* if it is of the form $M.P$, $\mathsf{rec}\ X.P$ or $(x : A)P$.

The operational semantics of CA is defined in two steps. Structural congruence, written $\equiv$, is defined as the least congruence relation that contains α-equivalence (capture-free renaming of bound names) and satisfying the laws of Fig. 3. Two processes are deemed equal by $\equiv$ when they only differ by some elementary syntactical manipulations. Reduction ($\longrightarrow$) is defined by the rules of Fig. 4. The first three rules specify movement and opening in CA as described informally above: note the three-way synchronisation for the movement rules, and the role of the direction tags in cocapabilities. The other reduction rules are standard: they describe communication in Ambients, recursion unfolding, and express the fact that reduction can occur anywhere in non-prefixed contexts, and that $\longrightarrow$ is defined modulo $\equiv$. We let $\longrightarrow^*$ stand for the reflexive transitive closure of $\longrightarrow$.

$$m[\mathsf{in}\ n.P \mid Q] \mid n[\overline{\mathsf{in}_\downarrow}\ m.R \mid S] \mid \overline{\mathsf{out}_\downarrow}\ m.T \ \longrightarrow\ n[m[P \mid Q] \mid R \mid S] \mid T$$
$$n[m[\mathsf{out}\ n.P \mid Q] \mid \overline{\mathsf{out}_\uparrow}\ m.R \mid S] \mid \overline{\mathsf{in}_\uparrow}\ m.T \ \longrightarrow\ m[P \mid Q] \mid n[R \mid S] \mid T$$
$$h[\mathsf{open}\ m.P \mid Q \mid m[\overline{\mathsf{open}}\ \{m,h\}.R \mid S]] \ \longrightarrow\ h[P \mid Q \mid R \mid S]$$
$$\langle n\rangle \mid (x:A)P \ \longrightarrow\ P\{x \leftarrow n\}$$
$$\mathsf{rec}\ X.P \ \longrightarrow\ P\{X \leftarrow \mathsf{rec}\ X.P\}$$
$$\frac{P \longrightarrow Q}{(\boldsymbol{\nu}\, n:A)\,P \longrightarrow (\boldsymbol{\nu}\, n:A)\,Q} \qquad \frac{P \longrightarrow Q}{R \,|\, P \longrightarrow R \,|\, Q} \qquad \frac{P \longrightarrow Q}{n[P] \longrightarrow n[Q]}$$
$$\frac{P \equiv Q \qquad Q \longrightarrow R \qquad R \equiv S}{P \longrightarrow S}$$

Figure 4: Controlled Ambients – Reduction

2.2 Examples of CA Programming

We now provide a few examples to illustrate the use of Controlled Ambients. We omit in the examples given below type annotations in restrictions; these will be made explicit in the next section.

Renaming. Since movements in Controlled Ambients require full knowledge about the name of moving ambients (also in cocapabilites, which is not the case in Safe Ambients), renaming turns out to be often useful in order to comply with some protocols. One may write the renaming of ambient a to b as follows:

$$a\ \mathit{be}\ b.P \ \triangleq\ b[\mathsf{out}\ a.\overline{\mathsf{in}_\downarrow}\ a.\mathsf{open}\ a] \mid \overline{\mathsf{out}_\uparrow}\ b.\mathsf{in}\ b.\overline{\mathsf{open}}\ \{a,b\}.P\,.$$

We then have $\overline{\mathsf{in}_\uparrow}\ b.\overline{\mathsf{out}_\downarrow}\ a \mid a[a\ \mathit{be}\ b.P] \ \longrightarrow^*\ b[P]$. This important example is also characteristic of Controlled Ambients, since $\overline{\mathsf{in}_\uparrow}\ b.\overline{\mathsf{out}_\downarrow}\ a$ illustrates a particular programming discipline: a's parent ambient must accept the replacement of a by b. This means that, at any time, the father ambient knows its own contents, that is both the number of subambients and their names.

Safe Ambients Cocapabilities. As mentioned above, Safe Ambients [18] introduce another kind of cocapabilities, similar to ours, though weaker. We concentrate here on the $\overline{\mathsf{SAin}}$ cocapability (the case of $\overline{\mathsf{SAout}}$ being symmetrical). Its semantics is defined by

$$a[\mathsf{in}\ b.P \mid Q] \mid b[\overline{\mathsf{SAin}}\ b.R \mid S] \ \longrightarrow\ b[R \mid S \mid a[P \mid Q]]\,.$$

By carrying on the idea behind renaming, we can approximate the specifics of this cocapability in CA. In other words, $a[\mathsf{in}\ b.P \mid Q] \mid b[\overline{\mathsf{SAin}}\ b.R \mid S]$ may be written

$$\begin{aligned}(\boldsymbol{\nu}\, m,n)\quad \big(a\big[\,&\overline{\mathsf{out}_\uparrow}\ m.\mathsf{in}\ b.(P \mid n[\mathsf{out}\ a.\overline{\mathsf{open}}\ \{n,b\}] \mid \overline{\mathsf{out}_\uparrow}\ n) \mid Q\\ &\mid m[\mathsf{out}\ a.\mathsf{in}\ b.\overline{\mathsf{open}}\ \{m,b\}.\overline{\mathsf{in}_\downarrow}\ a]\,\big]\\ \mid b[&\overline{\mathsf{in}_\downarrow}\ m.\mathsf{open}\ m.\overline{\mathsf{in}_\uparrow}\ n.\mathsf{open}\ n.R \mid S] \mid \overline{\mathsf{in}_\uparrow}\ m.\overline{\mathsf{out}_\downarrow}\ m.\overline{\mathsf{out}_\downarrow}\ a\big)\,.\end{aligned}$$

As specified, this expression reduces to $b[R \mid S \mid a[P \mid Q]]$. We use here two auxiliary ambients m and n to simulate the $\overline{\mathsf{SAin}}$ cocapability. At start, ambient b does not know name a, so the role of m is to bring this knowledge into b, in order for it to be able to execute the CA cocapability $\overline{\mathsf{in}_\downarrow}\ a$ (which is carried in m). Ambient n is used as a synchronisation device, in order to block the execution of R as long as a is not inside b. As was the case for renaming, the father must accept the transaction with $\overline{\mathsf{in}_\uparrow}\ m.\overline{\mathsf{out}_\downarrow}\ m.\overline{\mathsf{out}_\downarrow}\ a$. This entails in particular that the father ambient must be aware of the presence of a.

Firewall. We revisit the firewall example of [5], and consider a system f, protected by a firewall. Only agents knowing the password g are allowed in f. This may be modeled as:

$$\begin{array}{lcl}
\mathit{Agent}\ P\ Q & \triangleq & \mathit{agent}[\mathsf{in}\ g.\overline{\mathsf{in}_\downarrow}\ \mathit{entered}.\mathsf{open}\ \mathit{entered}.P \mid Q] \\
\mathit{System} & \triangleq & (\boldsymbol{\nu} f)f\big[\mathsf{rec}\ X.\big(\,g[\mathsf{out}\ f.\overline{\mathsf{in}_\downarrow}\ \mathit{agent}.\mathsf{in}\ f.\overline{\mathsf{open}}\ \{g,f\}] \\
 & & \quad \mid \overline{\mathsf{out}_\uparrow}\ g.\overline{\mathsf{in}_\downarrow}\ g.\mathsf{open}\ g.(\mathit{entered}[\mathsf{in}\ \mathit{agent}.\overline{\mathsf{open}}\ \{\mathit{entered},\mathit{agent}\}] \\
 & & \qquad\qquad\qquad\qquad \mid \overline{\mathsf{out}_\uparrow}\ \mathit{entered}.X)\,\big)\big] \\
 & & \mid\ \mathsf{rec}\ Y.\overline{\mathsf{in}_\uparrow}\ g.\overline{\mathsf{out}_\downarrow}\ \mathit{agent}.\overline{\mathsf{out}_\downarrow}\ g.Y
\end{array}$$

This specification behaves as follows: *System* receives *agent* and then recovers its original structure thanks to rec . The structure of g guarantees that, at any time, g may only contain one *agent*. On the other hand, *System* may contain any number of *agent*s. This system implements two authentifications: in the first place, the Agent must be named *agent* - it will not enter f by accident. In the second place, it must know the password. Note that this is not the Firewall described in the original paper on Mobile Ambients [5], which relied on the secrecy of three keys. This version uses only one key and takes advantage of the synchronization mechanism to execute correctly.

Cab. Fig. 5 presents a CA version of the cab protocol from Sec. 1. We do not give definitions for the city or for the sites, which only need to contain all movement authorizations, in addition to *client*s and *cab*s. Using cocapabilities, synchronizations in CA are both easier than in Mobile Ambients and atomic. Additionally, the system is not subject to the interferences we have presented: only *client*s may enter the cab, not just any "parasite" ambient which happens to contain capability in *cab*. Similarly, sites only welcome *client*s, *cab*s and *call*s.

Note that in this version, all clients must be named *client* in order to enter a *cab*. One could use renaming or the approximation of $\overline{\mathsf{SAin}}$ to relax this constraint (see above).

Additionally, Controlled Ambients permit the control of resources such as available space in cabs. As opposed to the Mobile Ambients version, we can easily check that the cab may contain at most *only one passenger* and possibly an auxiliary ambient *call*, *trip*, *arrived* or *end*. These properties will be expressed formally using our type system in Sec. 3.

2.3 Benefits

We believe that the formalism of Controlled Ambients is more reasonable than Mobile Ambients or Safe Ambients. More reasonable insofar as the implemen-

$$
\begin{array}{lcl}
\mathit{call\ from} & \triangleq & \mathit{call}[\mathsf{out}\ \mathit{client}.\mathsf{out}\ \mathit{from}. \\
& & \quad \mathsf{in}\ \mathit{cab}.\overline{\mathsf{open}}\ \{\mathit{call},\mathit{cab}\}.\mathsf{in}\ \mathit{from}.\overline{\mathsf{in}_{\downarrow}}\ \mathit{client}] \\
\mathit{trip\ from\ to} & \triangleq & \mathit{trip}[\mathsf{out}\ \mathit{client}.\overline{\mathsf{open}}\ \{\mathit{trip},\mathit{cab}\}.\mathsf{out}\ \mathit{from}. \\
& & \quad \mathsf{in}\ \mathit{to}.\mathit{arrived}[\overline{\mathsf{open}}\ \{\mathit{arrived},\mathit{cab}\}.\mathit{end}[\overline{\mathsf{open}}\ \{\mathit{end},\mathit{cab}\}.\mathsf{out}\ \mathit{to}]]] \\
\mathit{client\ from\ to} & \triangleq & \mathit{client}[\mathit{call\ from} \mid \overline{\mathsf{out}_{\uparrow}}\ \mathit{call}.\mathsf{in}\ \mathit{cab}.\mathit{trip\ from\ to} \\
& & \quad \mid \overline{\mathsf{out}_{\uparrow}}\ \mathit{trip}.\mathsf{out}\ \mathit{cab}] \\
\mathit{cab} & \triangleq & \mathit{cab}[\mathsf{rec}\ X.\overline{\mathsf{in}_{\downarrow}}\ \mathit{call}.\mathsf{open}\ \mathit{call}.\overline{\mathsf{in}_{\uparrow}}\ \mathit{trip}.\mathsf{open}\ \mathit{trip}.\mathsf{open}\ \mathit{arrived}. \\
& & \qquad \overline{\mathsf{out}_{\uparrow}}\ \mathit{client}.\mathsf{open}\ \mathit{end}.X]
\end{array}
$$

Figure 5: Cab protocol – CA-style (see Fig. 1)

tation of movements in ambient calculi suggests this kind of three-way synchronization. To illustrate our claim, let us consider the following transition in Mobile Ambients:

$$h[m[\mathsf{in}\ n] \mid n[\mathbf{0}]] \longrightarrow h[n[m[\mathbf{0}]]]\,.$$

As shown in [10, 22], a practical implementation of this rule requires that h must be aware of the presence of n, no matter how n may have entered h. More generally, the execution of this movement will involve a synchronization between n (who is actually present), m (who looks for n) and h (who knows about the presence of m and n). Similarly, the opening of ambient m by ambient h requires some complex synchronization between m and h in order to recover all processes and subambients of m within h and update presence registers of h. A prototype implementation has been developped [11] in order to experiment with CA-like synchronisation.

Controlled Ambients are also more realistic as modeling tools. When a system receives informations, it must be by some action of his: the operating system “listens” on a device, the configuration server waits for a request by “listening” on some given TCP/IP port... Unfortunately, this listening behaviour is not rendered at all by Mobile Ambients and only in half of the cases by Safe Ambients. Similarly, a system is liable to request several kinds of informations and to sort them according to their origin: the OS is able to differentiate data read on a disk from data read on the network or on the keyboard, while software may listen on several communication ports, for example. We can easily model such phenomena in CA, and if necessary take into account situations where some part of the system (like the network connexion itself) accepts data without listening explicitely for it, using renaming and infinite loops of cocapacities.

3 Typing Controlled Ambients

This section is devoted to the presentation of a type system for resource control in Controlled Ambients. We first describe the system and its properties, and then show the kind of information it is liable to check on some examples.

A	$::=$	$\text{CA}\textsc{am}(s,e)[T]$	$s \in \overline{\mathbb{N}}, e \in \mathbb{N}$	ambient types
U	$::=$	$\text{CA}\textsc{pr}(t)[T]$	$t \in \overline{\mathbb{N}}$	process types
T	$::=$	Ssh		message types
	$\mid$	t, A	$t \in \overline{\mathbb{N}}$	

Figure 6: Types

3.1 The Type System

Type Judgments. The grammar for types is given in Fig. 6, and includes entries for the types of ambients, processes and messages ($\overline{\mathbb{N}}$ stands for $\mathbb{N} \cup \{\infty\}$).

Typing environments, ranged over with Γ, are lists of associations of the form $x : A$ (for ambient names) or $X : U$ (for process variables). We write $\Gamma(x) = A$ (resp. $\Gamma(X) = U$) to represent the fact that environment Γ associates A (resp. U) to x (resp. X). $\Gamma, x : A$ stands for the extension of Γ with the association $x : A$, possibly hiding some previous binding for x (and similarly for $\Gamma, X : U$).

The typing judgment for ambient names is of the form

$$\Gamma \vdash n : \text{CA}\textsc{am}(s,e)[T]\,,$$

and expresses the fact that under assumptions Γ, n is the name of an ambient of *capacity* s, *weight* e, and within which messages carrying information of type T may be exchanged. The capacity s represents the amount of space (or of *resources*) available for subambients within n, while e is the number of resources this ambient is occupying in its surrounding ambient. Note that while an ambient may have an infinite capacity ($s = \infty$), it cannot manipulate infinitely many resources ($e < \infty$). Moreover, if we decide to impose $e \geq s$ in ambient types, we may develop an analysis close to what is done in [7], where the weight of an ambient takes into account the weight of all its subambients, *at any depth.* The type T for messages captures the kind of names being exchanged within n, similarly to Cardelli and Gordon's *topics of conversation* [6], augmented with an information t which represents a higher bound on the effect of exchanging messages within n (we shall come back to this below).

The typing judgment for processes is written

$$\Gamma \vdash P : \text{CA}\textsc{pr}(t)[T]\,,$$

meaning that according to Γ, P is a process that may use up to t resources, and take part in conversations (that is, emit and receive messages) having type T.

Typing Rules. The rules defining the typing judgments are given on Fig. 7. We now comment on them. While typing (subjective) movements has no effect from the point of view of resources (rules T-IN and T-OUT), the rules T-COIN and T-COOUT, for the co-capabilities (where δ ranges over a direction tag, which can be $\uparrow$ or $\downarrow$), express the meaning of t in $\text{CA}\textsc{pr}(t)[T]$, according to the weight e of the moving ambient. Note that the number t of resources allocated to the process must remain positive after decreasing (rule T-COOUT). This is made possible by the subtyping property of the system (Lemma 1), together with

$$\text{T-NAME}\ \frac{\Gamma(n) = A}{\Gamma \vdash n : A} \qquad \text{T-VAR}\ \frac{\Gamma(X) = \mathrm{CA}\textsc{pr}(t)[T]}{\Gamma \vdash X : \mathrm{CA}\textsc{pr}(t')[T]}\ t' \geq t$$

$$\text{T-REC}\ \frac{\Gamma, X : \mathrm{CA}\textsc{pr}(t)[T] \vdash P : \mathrm{CA}\textsc{pr}(t)[T]}{\Gamma \vdash \mathsf{rec}\ X.P : \mathrm{CA}\textsc{pr}(t')[T]}\ t' \geq t$$

$$\text{T-IN}\ \frac{\Gamma \vdash P : \mathrm{CA}\textsc{pr}(t)[T]}{\Gamma \vdash \mathsf{in}\ m.P : \mathrm{CA}\textsc{pr}(t)[T]} \qquad \text{T-OUT}\ \frac{\Gamma \vdash P : \mathrm{CA}\textsc{pr}(t)[T]}{\Gamma \vdash \mathsf{out}\ m.P : \mathrm{CA}\textsc{pr}(t)[T]}$$

$$\text{T-COIN}\ \frac{\Gamma \vdash P : \mathrm{CA}\textsc{pr}(t)[T] \qquad \Gamma \vdash m : \mathrm{CA}\textsc{am}(s,e)[T']}{\Gamma \vdash \overline{\mathsf{in}_\delta}\ m.P : \mathrm{CA}\textsc{pr}(t+e)[T]}$$

$$\text{T-COOUT}\ \frac{\Gamma \vdash P : \mathrm{CA}\textsc{pr}(t)[T] \qquad \Gamma \vdash m : \mathrm{CA}\textsc{am}(s,e)[T']}{\Gamma \vdash \overline{\mathsf{out}_\delta}\ m.P : \mathrm{CA}\textsc{pr}(t-e)[T]}\ t \geq e$$

$$\text{T-OPEN}\ \frac{\Gamma \vdash m : \mathrm{CA}\textsc{am}(s,e)[T] \qquad \Gamma \vdash P : \mathrm{CA}\textsc{pr}(t)[T]}{\Gamma \vdash \mathsf{open}\ m.P : \mathrm{CA}\textsc{pr}(t-e+s)[T]}\ t-e+s \geq 0$$

$$\text{T-COOPEN}\ \frac{\Gamma \vdash m : \mathrm{CA}\textsc{am}(s,e)[T] \qquad \Gamma \vdash R : \mathrm{CA}\textsc{pr}(t)[T]}{\Gamma \vdash \overline{\mathsf{open}}\ \{m,h\}.R : \mathrm{CA}\textsc{pr}(t)[T]}$$

$$\text{T-NIL}\ \Gamma \vdash \mathbf{0} : \mathrm{U} \qquad \text{T-AMB}\ \frac{\begin{array}{c}\Gamma \vdash m : \mathrm{CA}\textsc{am}(s,e)[T] \\ \Gamma \vdash P : \mathrm{CA}\textsc{pr}(a)[T]\end{array}}{\Gamma \vdash m[P] : \mathrm{CA}\textsc{pr}(t)[T']}\ \begin{cases} a \leq s \\ e \leq t \end{cases}$$

$$\text{T-RES}\ \frac{\Gamma, n : A \vdash P : U}{\Gamma \vdash (\boldsymbol{\nu}\, n : A)P : U} \qquad \text{T-PAR}\ \frac{\Gamma \vdash P : \mathrm{CA}\textsc{pr}(t)[T] \qquad \Gamma \vdash Q : \mathrm{CA}\textsc{pr}(t')[T]}{\Gamma \vdash P \mid Q : \mathrm{CA}\textsc{pr}(t+t')[T]}$$

$$\text{T-SND}\ \frac{\Gamma \vdash m : A}{\Gamma \vdash \langle m \rangle : \mathrm{CA}\textsc{pr}(t')[t,A]}\ t' \geq t \qquad \text{T-RCV}\ \frac{\Gamma, x : A \vdash P : \mathrm{CA}\textsc{pr}(t)[t,A]}{\Gamma \vdash (x : A)P : \mathrm{CA}\textsc{pr}(t')[t,A]}$$

Figure 7: Typing rules

rules T-NIL, T-AMB, ..., which allow one to allocate any number of resources to an inert process (inert from the point of view of the current ambient). This mechanism can be used for example to derive a typing for a process of the form $\overline{\mathsf{out}_\uparrow}\ n.\mathbf{0}$. Note also that the side condition $a \leq s$ in rule T-AMB expresses conformity with the capacity of the ambient.

When opening an ambient, we release the resources it had acquired (e), but at the same time we have to provide at least as many resources as its original capacity (s). The $\overline{\mathsf{open}}$ capability plays no role from the point of view of resource control, as illustrated by rule T-COOPEN (note, still, that message types in the opening ambient and in the type of R are unified using this rule). We shall present in Sec. 4 a richer system where a more precise typing of opening (and co-opening) permits a better control.

We now explain the typing rules for communication. Since reception of a message can trigger a process which will necessitate a certain amount of resources, we attach to the type of an ambient the maximum amount of resources needed by a receiving process running within it: this is information t in an ambient's topic of conversation. Put differently, messages are decorated with an integer representing at least as many resources as needed by the processes they are liable to trigger: we are thus somehow measuring an effect in this case. Note that our approach is based on the idea that one emission typically corresponds to several receptions. The dual point of view could have been adopted, by putting in correspondence one reception and several concurrent emissions. Our experience in writing examples suggests that the first choice is more useful.

Finally, rule T-REC expresses the fact that a recursively defined process should run "in constant space": as required by the premise, each time a recursive call is triggered (by X), the number t of allocated resources is the same as the number of resources allocated to the whole recursive process P.

3.2 Static Resource Control

We now present the main properties of the type system. Proofs are not given, but can be found in [25].We start by some technical properties of typing derivations.

Lemma 1 (Subtyping) *Let P be a process and Γ an environment such that $\Gamma \vdash P : \mathrm{CA{\scriptstyle PR}}(t)[T]$ for some t. Then for any $t' \geq t$, $\Gamma \vdash P : \mathrm{CA{\scriptstyle PR}}(t')[T]$.*

Corollary 2 (Minimal typing) *If a process P is typeable in Γ with a conversation topic type T, then there is a minimal $t \in \overline{\mathbb{N}}$ such that $\Gamma \vdash P : \mathrm{CA{\scriptstyle PR}}(t)[T]$.*

Note that the minimal parameter t can be different for each possible value T (see for example rule T-SND).

Let us now examine resource control. In order to be able to state the properties we are interested in, we extend the notion of weight, which has been used for ambients, to processes, by introducing the notion of resource usage, together with a natural terminology:

Definition 3 (Resource policy and resource usage) *We call* resource policy *a typing context. Given a resource policy Γ, we define the* resource usage of a process P according to Γ, *written $Res_\Gamma(P)$, as follows:*

- *if $\Gamma(a) = \mathrm{CA{\scriptstyle AM}}(s,e)[T]$, then $Res_\Gamma(a[P]) = e$;*
- *$Res_\Gamma(P_1 \mid P_2) = Res_\Gamma(P_1) + Res_\Gamma(P_2)$;*
- *$Res_\Gamma((\boldsymbol{\nu}\, n : A)\, P) = Res_{\Gamma,n:A}(P)$.*
- *in all other cases, $Res_\Gamma(P) = 0$;*

Note in particular that according to this definition, prefixed terms (capabilities, reception, recursion) do not contribute to a process' *current* resource usage (accordingly, their resource usage is equal to 0).

We now define formally what it means for a process to respect a given resource policy.

Definition 4 (Resource policy compliance) *Given a resource policy Γ, we define the judgment $\Gamma \models P$ (pronounced "P complies with Γ"), as follows:*

- *$\Gamma \models n[P]$ iff $\Gamma \models P$ and $Res_\Gamma(P) \leq s$, where capacity s is given by $\Gamma(n) = \mathrm{CA{\scriptstyle AM}}(s,e)[T]$;*

- $\Gamma \models P_1|P_2$ *iff* $\Gamma \models P_1$ *and* $\Gamma \models P_2$*;*
- $\Gamma \models (\boldsymbol{\nu}\, n : A)\, P$ *iff* $\Gamma, n : A \models P$*;*
- *in all other cases,* $\Gamma \models P$*.*

Intuitively, the judgment $\Gamma \models P$ means that any ambient occurring in P contains no more subambients (in relation to the corresponding weights) than what its capacity allows. The typing rules we have introduced ensure that a typeable term complies with a resource policy:

Lemma 5 (Typeable terms comply with resource policies) *For any process* P*, resource policy* Γ *and process type* U*, if* $\Gamma \vdash P : U$*, then* $\Gamma \models P$*.*

The following theorem states that typability is preserved by the operational semantics of Controlled Ambients:

Theorem 6 (Subject Reduction) *For any processes* P, Q*, resource policy* Γ *and type* U*, if* $\Gamma \vdash P : U$ *and* $P \longrightarrow Q$*, then* $\Gamma \vdash Q : U$*.*

As a direct consequence, we obtain our main result:

Theorem 7 (Resource control) *Consider a resource policy* Γ *and a process* P *such that* $\Gamma \vdash P : U$ *for some* U*. Then for any* Q *such that* $P \longrightarrow^* Q$*, it holds that* $\Gamma \models Q$*.*

3.3 Examples

We now revisit some examples of Sec. 2.2, and explain how they can be typed. In each case, we exhibit a resource policy (i.e., a typing context Γ) that captures a property we wish to guarantee, and describe the weight and capacity associated to every ambient in order to do so.

Renaming. The expression of renaming given in Sec. 2.2 is typeable as soon as there exists a typing environment Γ and a conversation type T such that

$$\Gamma(a) = \mathrm{CA\textsc{am}}(s, e)[T]\,, \quad \Gamma(b) = \mathrm{CA\textsc{am}}(s, e)[T] \text{ with } s \geq e\,, \\ \text{and} \quad \Gamma \vdash P : \mathrm{CA\textsc{pr}}(s)[T]\,.$$

We can actually slightly relax the conditions on types. One can show that the least set of conditions to type the renaming is

$$t_P \leq s_a, \quad e_b \leq s_a, \quad s_a \leq s_b, \quad \text{and} \quad e_a \leq s_b\,,$$

where $\Gamma(a) = \mathrm{CA\textsc{am}}(s_a, e_a)[T]$, $\Gamma(b) = \mathrm{CA\textsc{am}}(s_b, e_b)[T]$ and we have $\Gamma \vdash P : \mathrm{CA\textsc{pr}}(t_P)[T]$.

Firewall. Similarly, the firewall in Controlled Ambients, as defined in subsection 2.2, can be typed in a context Γ such that:

$$\Gamma(agent) = \mathrm{CA\textsc{am}}(a_P + a_Q, 1)[T], \quad \Gamma(entered) = \mathrm{CA\textsc{am}}(0, 0)[T], \\ \Gamma(f) = \mathrm{CA\textsc{am}}(\infty, 0)[T], \quad \text{and} \quad \Gamma(g) = \mathrm{CA\textsc{am}}(1, 0)[T]\,.$$

In particular, the typing of the recursive process $\mathsf{rec}\ X.\ldots$ in $System$ entails a constraint of the form $\mathrm{CA{\scriptstyle PR}}(t)[T] = \mathrm{CA{\scriptstyle PR}}(t+1)[T]$. This is possible if and only if $t = \infty$, and as a consequence the capacity of f should also be ∞, so that the firewall is supposed to have infinite size. This is no surprise, since it may actually receive any number of external ambients. However, these ambients are contained in the firewall. Hence, one may still integrate this firewall as a component in a system with limited resources.

Cab. Let us consider an environment Γ such that:

$$\left\{\begin{array}{lcl} \Gamma(client) & = & \mathrm{CA{\scriptstyle AM}}(0,1)[T] \\ \Gamma(call) & = & \mathrm{CA{\scriptstyle AM}}(1,0)[T] \\ \Gamma(trip) & = & \mathrm{CA{\scriptstyle AM}}(0,0)[T] \\ \Gamma(arrived) & = & \mathrm{CA{\scriptstyle AM}}(0,0)[T] \end{array}\right. \qquad \left\{\begin{array}{lcl} \Gamma(end) & = & \mathrm{CA{\scriptstyle AM}}(0,0)[T] \\ \Gamma(cab) & = & \mathrm{CA{\scriptstyle AM}}(1,0)[T] \\ \Gamma(site_i) & = & \mathrm{CA{\scriptstyle AM}}(\infty,0)[T] \\ \Gamma(city) & = & \mathrm{CA{\scriptstyle AM}}(0,0)[T] \end{array}\right.$$

Note in particular that this resource policy specifies that among the ambients that may enter the cab, only those named *client* are actually "controlled": this corresponds to the property we focus on when analyzing the cab. With these assumptions, the complete cab system is typeable. This means that resources are statically controlled in cabs: *at any step of its execution, the cab may contain at most one client.*

Moreover, we may adopt a different resource policy, defined as follows:

$$\left\{\begin{array}{lcl} \Gamma(client) & = & \mathrm{CA{\scriptstyle AM}}(0,0)[T] \\ \Gamma(call) & = & \mathrm{CA{\scriptstyle AM}}(0,1)[T] \\ \Gamma(trip) & = & \mathrm{CA{\scriptstyle AM}}(1,1)[T] \\ \Gamma(arrived) & = & \mathrm{CA{\scriptstyle AM}}(1,1)[T] \end{array}\right. \qquad \left\{\begin{array}{lcl} \Gamma(end) & = & \mathrm{CA{\scriptstyle AM}}(0,1)[T] \\ \Gamma(cab) & = & \mathrm{CA{\scriptstyle AM}}(1,0)[T] \\ \Gamma(site_i) & = & \mathrm{CA{\scriptstyle AM}}(\infty,0)[T] \\ \Gamma(city) & = & \mathrm{CA{\scriptstyle AM}}(0,0)[T] \end{array}\right.$$

The system is also typeable with this choice for Γ, which allows us to control the number of "auxiliary" ambients: at any time, at most one of those may be present in *cab*.

4 More Accurate Analyses of Opening

In this section, we present several refinements of type system of Section 3, that we call systems R, Z and RZ. While the basic system we have presented so far allows one to type many interesting processes, some relatively simple examples show its limitations. For instance, let us define

$$P_1 \ \triangleq\ a[\overline{\mathsf{open}}\ \{a,b\}.\mathsf{rec}\ X.(X \mid b[\mathbf{0}])] \ \mid\ \mathsf{open}\ a\,,$$

and suppose that the weight of b is not 0. The construction $\mathsf{rec}\ X.(X \mid b[\mathbf{0}])$ then requires infinite resources. Although the execution would not use any resource inside a, our type system cannot capture this property: the typing will require a to have an infinite capacity.

Similarly, let us define

$$\begin{array}{rcl} P_2 & \triangleq & h[\mathsf{rec}\ X.(m[\overline{\mathsf{in}_\downarrow}\ n.\overline{\mathsf{out}_\uparrow}\ n.\overline{\mathsf{open}}\ \{m,h\}] \mid \overline{\mathsf{out}_\downarrow}\ n.\overline{\mathsf{in}_\uparrow}\ n.\mathsf{open}\ m.X) \\ & & \qquad\qquad\qquad\qquad\qquad\qquad \mid n[\mathsf{rec}\ Y.\mathsf{in}\ m.\mathsf{out}\ m.Y]]\,, \end{array}$$

and suppose that the weight of n is not 0. By following the evolution of this term, one may easily notice that a finite capacity for h should be sufficient.

However, when deriving a typing for P_2, we conclude that the capacity of h must be infinite.

In both cases, the typing system is not refined enough to express a resource control property. More specifically, the opening primitive is associated to a resource control that is too strict. For the discussion that follows, we shall use the following notations for the rule R-OPEN:

$$h[\mathsf{open}\ m.P \mid Q \mid m[\overline{\mathsf{open}}\ \{m,h\}.R \mid S]] \longrightarrow h[P \mid Q \mid R \mid S]$$

In order to try and refine the typing of opening, one may want to make the control on P, Q, R or S more precise. For technical reasons, we have chosen to concentrate on R and S.

System R In System R, we introduce a third parameter in ambient types, named r. In $\mathrm{CA}\textsc{am}(s,e,r)[T]$, $r \in \overline{\mathbb{N}}$ is an upper bound for the number of resources allocated to R in the opening ambient. Typing rules for open and $\overline{\mathsf{open}}$ become:

$$\frac{\Gamma \vdash m : \mathrm{CA}\textsc{am}(s,e,r)[T] \quad \Gamma \vdash P : \mathrm{CA}\textsc{pr}(t)[T]}{\Gamma \vdash \mathsf{open}\ m.P : \mathrm{CA}\textsc{pr}(t-e+s+r)[T]} \quad \begin{array}{c} CA-open \\ t-e+s+r \geq 0 \end{array}$$

$$\frac{\Gamma \vdash m : \mathrm{CA}\textsc{am}(s,e,r)[T] \quad \Gamma \vdash R : \mathrm{CA}\textsc{pr}(t)[T]}{\Gamma \vdash \overline{\mathsf{open}}\ \{m,h\}.R : \mathrm{CA}\textsc{pr}(t')[T]} \quad \begin{array}{c} CA-coopen \\ t \leq r \end{array}$$

Using these alternative rules, term P_1 may be satisfactorily typed (i.e. with a finite capacity for a), taking $r = \infty$. Additionally, all results of Section 3.2 still remain valid. However, System R does not help with term P_2.

System Z System Z, on the other hand, improves the control on S. This is particularly important, for processes such as

$$M_1.\cdots.M_n.\overline{\mathsf{open}}\ \{m,h\}.R\ :$$

although $M_1.\cdots.M_n$ might acquire as many as, say, s resources, it might also release some or all of them before the actual opening. By taking these releases into account, we may get a better approximation of resource consumption. To do so, we can introduce a parameter z which is compelled to satisfy $z \leq s$. In System Z, ambient types become $\mathrm{CA}\textsc{am}(s,e,z)[T]$ with $z \in \mathbb{N}$ and $z \leq s$, and the typing rules are:

$$\frac{\Gamma \vdash m : \mathrm{CA}\textsc{am}(s,e,z)[T] \quad \Gamma \vdash P : \mathrm{CA}\textsc{pr}(t)[T]}{\Gamma \vdash \mathsf{open}\ m.P : \mathrm{CA}\textsc{pr}(t-e+z)[T]} \quad \begin{array}{c} CA-open \\ t-e+z \geq 0 \end{array}$$

$$\frac{\Gamma \vdash m : \mathrm{CA}\textsc{am}(s,e,z)[T] \quad \Gamma \vdash R : \mathrm{CA}\textsc{pr}(t)[T]}{\Gamma \vdash \overline{\mathsf{open}}\ \{m,h\}.R : \mathrm{CA}\textsc{pr}(t+s-z)[T]} \quad CA-coopen$$

Results from Section 3.2 also remain valid on System Z. System Z permits a good analysis of term P_2, but cannot handle term P_1 any better than the basic system.

System RZ System R and System Z may be naturally merged into System RZ, which yields a more accurate analysis of resources, with ambient types of the form $\mathrm{CA}\textsc{am}(s,e,r,z)[T]$, $r \in \overline{\mathbb{N}}, z \in \mathbb{N}$ and $z \leq s$ and the following rules:

$$\frac{\Gamma \vdash m : \mathrm{CA}\textsc{am}(s,e,r,z)[T] \quad \Gamma \vdash P : \mathrm{CA}\textsc{pr}(t)[T]}{\Gamma \vdash \mathsf{open}\ m.P : \mathrm{CA}\textsc{pr}(t-e+z+r)[T]} \quad \begin{array}{c} CA-open \\ t-e+z+r \geq 0 \end{array}$$

$$\frac{\Gamma \vdash m : \mathrm{CA}\textsc{am}(s,e,r,z)[T] \quad \Gamma \vdash R : \mathrm{CA}\textsc{pr}(t)[T]}{\Gamma \vdash \overline{\mathsf{open}}\ \{m,h\}.R : \mathrm{CA}\textsc{pr}(t')[T]} \quad \begin{array}{c} CA-coopen \\ t \leq r, t' \geq s-z \end{array}$$

As expected, System RZ correctly handles both terms P_1 and P_2, and results from Section 3.2 also remain valid. Hence, System RZ is a more refined although more complicated type system.

5 Conclusion

The language of Controlled Ambients has been introduced to analyze resource control in a distributed and mobile setting through an accurate programming of movements and synchronisations. We have enhanced our formalism with a type system for the static control of resources, and extensions of the basic type system have also been presented. Further, examples show that indications on the maximal amount of resources needed by a process match rather closely the actual amount of resources which may be reached in the worst case, which suggests that the solution we propose could serve as the basis for a study of resource control properties on a larger scale.

Among extensions of the present work, we are currently enriching the language and type system to include communication of capabilities, as in the original Mobile Ambients calculus [5]. We are also studying type inference for our system, which would enhance (untyped) Controlled Ambients with a procedure for the automatic guess of resource needs. It seems that by requiring the recursion variables to be explicitely typed, type inference is decidable, and a rather natural algorithm can compute a minimal type for a given process, if it exists. In particular, the "message" component of terms leads to a classical unification problem. The question becomes more problematic if no information is given for recursion variables: one can compute a set of inequalities (resembling those given for the example of renaming in Sec. 3), but solving it in the general case would require more work.

As reported in [24], our approach can be adapted to other formalisms for mobile and distributed computation that provide a primitive notion of location, such as Seals [28], Boxed Ambients [2, 3], Nomadic π [27] and Kells [23]. In π-calculus-like languages, a natural notion of resource is given by *channels,* which represents a slightly different point of view w.r.t. the present work. Introducing resource control in calculi like the π-calculus or the distributed π-calculus [21] represents a challenging direction for future work.

We could also consider combining our type system for resource control with other typing disciplines, adapted from the Single Threadness types of [18], or the Mandatory Access Control of [2]. It seems that Controlled Ambients could also be used to approximate some of the analyses done in [9, 14], where, in a context where *security levels* are associated with processes, types are used to check that no agent can access an information having a security level higher than its own. For instance, in the simple case where we have two security levels, we could attach weight 0 to agents of high level, and 1 to low-level agents, and store high-level information in ambients of size 0: in such a framework, our type system can guarantee that only high-level processes can access high-level data. Of course this is a very rough approximation. We are currently working on a generalisation of our type system that would enhance its flexibility, making it possible to handle more complex kind of resources and related properties (such as security levels, non-releasable resources, and a form of movement typing along the lines of [4]).

We have not addressed the issue of behavioural equivalences for CA. A possible outcome of such a study could be to validate a more elaborate treatment of resources involving operations like garbage collection, which would allow one to make available uselessly occupied resources. An example is the perfect firewall equation of [12]: when $c \notin \mathrm{fn}(P)$, process $(\boldsymbol{\nu}\, c)\, c[P]$ may manipulate some resources while being actually equivalent to $\mathbf{0}$.

Other Related Works. Other projects aim at controlling resources in possibly mobile systems without resorting to mobile process algebras. [17] presents a modified ML language with sized types in which bounds may be given to stack consumption. Like in our framework, resources are releasable entities; however, this approach seems more specialized than ours, and moreover concentrates on a sequential model. Similarly, [8] introduces a variant of the Typed Assembly Language *"augmenting TAL's very low-level safety certification with running-time guarantees"*, while Quantum [20] may be used to describe distributed systems from the point of view of their resource consumption. In contrast to our work, both these approaches consider non-releasable resources. Another programming language, PLAN [15], has been designed specifically for active networks, and also handles some form of resource bounds. Although PLAN accounts for both releasable (space, bandwidth) and non-releasable (time) resources, it handles neither recursion nor concurrency on one node. A related line of research is followed in [16, 1], where means to guarantee bounds on the time or space consumption required for the execution of (sequential) functions are proposed.

These works all focus on resource control; however, none of these approaches can be directly compared to ours. It might be interesting to study if and how our methods could be integrated to these works, in order to combine several forms of resource control.

Another form of accounting on mobile ambients is introduced in [7]. In a calculus with a slightly different form of recursion than in CA (and without co-capabilities), the authors introduce a type system to count the number of active outputs and ambients (at any depth) in a process. This analysis, however, is not aimed at modelling resources: it tries and isolate a finite-control fragment of mobile ambients on which model checking w.r.t. the Ambient Logic is decidable through state-space exploration.

Acknowledgments We would like to thank Davide Sangiorgi for suggesting the original idea behind CA and providing insightful suggestions along this work.

References

[1] R. Amadio. Max-plus quasi-interpretations. In *Proc. of TLCA'03*, LNCS. Springer Verlag, 2003. to appear.

[2] M. Bugliesi, G. Castagna, and S. Crafa. Boxed ambients. In *Proc. TACS 2001*, LNCS 2215, pages 38–63. Springer Verlag, 2001.

[3] M. Bugliesi, S. Crafa, M. Merro, and V. Sassone. Communication Interference in Mobile Boxed Ambients. In *Proc. of FST-TCS'02*, LNCS. Springer Verlag, 2002.

[4] L. Cardelli, G. Ghelli, and A. Gordon. Types for the Ambient Calculus. *Information and Computation*, 200? to appear.

[5] L. Cardelli and A. D. Gordon. Mobile ambients. In *Proc. of FOSSACS'98*, volume 1378, pages 140–155. Springer Verlag, 1998.

[6] L. Cardelli and A. D. Gordon. Types for mobile ambients. In *Symposium on Principles of Programming Languages (POPL'99)*, pages 79–92. ACM Press, 1999.

[7] W. Charatonik, A. D. Gordon, and J.-M. Talbot. Finite-control mobile ambients. In *Proc. of ESOP'02*, volume 2305 of *LNCS*, pages 295–313, 2002.

[8] K. Crary and S. Weirich. Resource bound certification. In *Symposium on Principles of Programming Languages (POPL'00)*, pages 184–198. ACM Press, 2000.

[9] M. Dezani-Ciancaglini and I. Salvo. Security types for mobile safe ambients. In *Proc. of ASIAN'00*, LNCS 1961, pages 215–236. Springer Verlag, 2000.

[10] C. Fournet, J.-J. Lévy, and A. Schmitt. A distributed implementation of mobile ambients. In *Proc. of IFIP TCS'00*, pages 348–364. Springer Verlag, 1872.

[11] T. Gazagnaire and D. Pous. Implémentation des Controlled Ambients en JoCaml. Students project – Magistère d'Informatique ENS Lyon, 2002.

[12] A. D. Gordon and L. Cardelli. Equational properties of mobile ambients. In *Proc. of FOSSACS'99*, volume 1578 of *LNCS*, pages 212–226. Springer Verlag, 1999.

[13] X. Guan, Y. Yang, and J. You. Making ambients more robust. In *Proc. of the International Conference on Software: Theory and Practice*, pages 377–384, 2000.

[14] M. Hennessy and J. Riely. Resource access control in systems of mobile agents. In *Proceedings of HLCL '98*, number 16.3 in ENTCS, pages 3–17. Elsevier, 1998.

[15] M. Hicks, P. Kakkar, J. T. Moore, C. A. Gunter, and S. Nettles. PLAN: A Packet Language for Active Networks. In *Proc. ICFP'99*, pages 86–93. ACM Press, 1999.

[16] M. Hofmann. The strength of non-size increasing computation. In *Proc. 29th ACM Symp. on Principles of Programming Languages (POPL'02)*, pages 260–269. ACM Press, 2002.

[17] J. Hughes and L. Pareto. Recursion and dynamic data-structures in bounded space: Towards embedded ML programming. In *Proc. of ICFP'99*, pages 70–81. ACM Press, 1999.

[18] F. Levi and D. Sangiorgi. Controlling interference in ambients. In *Symposium on Principles of Programming Languages*, pages 352–364. ACM Press, 2000.

[19] M. Merro and M. Hennessy. Bisimulation congruences in safe ambients. In *Proc. of POPL'02*, pages 71–80. ACM Press, 2002.

[20] L. Moreau. A distributed garbage collector with diffusion tree reorganisation and mobile objects. In *Proc. of ICFP'98*, pages 204–215. ACM Press, 1998.

[21] J. Riely and M. Hennessy. A typed language for distributed mobile processes. In *Proc. of POPL'98*, pages 378–390. ACM Press, 1998.

[22] D. Sangiorgi and A. Valente. A distributed abstract machine for Safe Ambients. In *Proc. of ICALP'01*, 2001.

[23] J.-B. Stefani. A calculus of Higher-Order Distributed Components. Technical Report 4692, INRIA, 2003.

[24] D. Teller. Formalisms for Mobile Resource Control. submitted, 2003.

[25] D. Teller, P. Zimmer, and D. Hirschkoff. Using Ambients to Control Resources – technical annex. available at
http://www.ens-lyon.fr/~dtelle/recherche/Publications/annex_ca.ps.gz.

[26] D. Teller, P. Zimmer, and D. Hirschkoff. Using Ambients to Control Resources. In *Proc. of CONCUR'02*, LNCS. Springer Verlag, 2002.

[27] A. Unyapoth. *Nomadic Pi Calculi: Expressing and Verifying Infrastructure for Mobile Computation*. PhD thesis, Computer Laboratory, University of Cambridge, june 2001.

[28] J. Vitek and G. Castagna. Seal: A Framework for Secure Mobile Computations. In *Internet Programming Languages*, volume 1686 of *LNCS*. Springer Verlag, 1999.