# Linearity, Control Effects, and Behavioral Types

Luís Caires<sup>1</sup> and Jorge A. Pérez<sup>1,2</sup>

<sup>1</sup> NOVA LINCS and Departamento de Informática, FCT, Universidade Nova de Lisboa
<sup>2</sup> University of Groningen & CWI, Amsterdam

Abstract. Mainstream programming idioms intensively rely on state mutation, sharing, and concurrency. Designing type systems for handling and disciplining such idioms is challenging, due to long known conflicts between internal nondeterminism, linearity, and control effects such as exceptions. In this paper, we present the first type system that accommodates non-deterministic and abortable behaviors in the setting of session-based concurrent programs. Remarkably, our type system builds on a Curry-Howard correspondence with (classical) linear logic conservatively extended with two dual modalities capturing an additive (co)monad, and provides a first example of a Curry-Howard interpretation of a realistic programming language with built-in internal non-determinism. Thanks to its deep logical foundations, our system elegantly addresses several well-known tensions between control, linearity, and non-determinism: globally, it enforces progress and fidelity; locally, it allows the specification of non-deterministic and abortable computations. The expressivity of our system is illustrated by several examples, including a typed encoding of a higher-order functional language with threads, session channels, non-determinism, and exceptions.

# 1 Introduction

In this paper, we study a principled, typeful foundation to represent a relevant class of control effects within a behavioral type system for stateful concurrent programs. So-phisticated structural type systems have shaped mainstream static type checking for a long time now, and are fairly complete tools to discipline and effectively check programs that manipulate pure values. Unfortunately, the same cannot be said for most mainstream programming idioms, which intensively rely on state mutation, sharing, and often concurrency, about which "standard" type systems are quite silent.

Interactive concurrent systems need to manipulate stateful resources, ranging from basic memory references and passive objects (such as files, locks, and communication channels) to dynamic entities (such as threads or web references) typically subject to linearity constraints. To extend type-based verification techniques to this challenging setting, substructural type systems, based on various forms of linearity and affinity, have been increasingly investigated [27,41,43,42,31], and start to make their way towards practical adoption. Recent examples include Mozilla's Rust, but also embeddings of session types in target languages without linear types [28,37]. Some approaches use types to model states (cf. assertions); examples include typestate and several affine, linear, and stateful type systems, see, e.g., [16,27,29]. In other works, types are used to model behaviors (cf. processes); examples in this line include session types and usage types such as, e.g., [23,24,30], often referred to as *behavioral types* [25].

Linear types are in general very expressive, and the fine-grained specifications of usage types they typically support may simultaneously bring a benefit and a curse. In

LET $a = Ref()$ IN LET $b = Ref()$ IN	LET $a = Ref()$ IN LET $b = Ref()$ IN
LET $x1 = f()$ IN	LET $x1 = f()$ IN
LET $a' = write \ a \ x1$ IN	LET $x 2 = g()$ IN
LET $x 2 = g()$ IN	LET $a' = write \ a \ x1$ IN
LET $b' = write \ b \ x1$ IN (free $a'$ ; free $b'$ )	LET $b' = write \ b \ x2 \ IN \ (free \ a'; free \ b')$

#### Fig. 1. Two code snippets.

particular, a still open issue is how to seamlessly combine linearity with many other useful programming mechanisms—a prominent example being the interaction of linearity with non-determinism and control effects, such as exceptions and (linear) continuations. The general issue is that by their very essence control effects (and associated programming language constructs) conflict with the linear, stateful usage discipline of values manipulated by programs, which makes it difficult to statically check programs. For example, when a communication channel is aborted, any linear values held by a channel client continuation must be aborted as well (which may not be always possible), or passed away to some candidate consumer code in scope. Likewise, after an exception is raised, it is not always clear how to safely discard a continuation holding linear values, nor how to proceed after the exception is caught. This situation is already present in non-deterministic programs where subexpressions may return more than one result, or even no result at all (e.g., "fail"), as in the non-deterministic monad.

These underlying conflicts and challenges may be illustrated by the simple examples in Fig. 1, adapted from [42], which we express in an idealized linear functional language. The Ref function is assumed to return a stateful fresh value r that may either be initially discarded, or else subject to a strictly linear protocol consisting of a write r x operation followed by a free r operation. We use a common idiom when programming with usage types, in which a operation f acting on a linear value a that needs to be used according to a stateful protocol would be called as a' = (fa) where the returned a' refers to the new state of a (which gets "consumed" in the call (f a)). So, in our examples, failure to call free r' after r' = write r x may result in, say, a memory leak. Now, in Fig. 1 (left), if for some reason the call f() raises an exception, the continuation will be safely aborted, but if f() succeeds and the call g() raises an exception instead, the resulting behavior would be ill-defined, as the required execution of free a' will be discarded. On the other hand, consider the slightly different code snippet in Fig. 1 (right). Now, even assuming that f() or g() may raise an exception, there will be no value usage violations, since both a and b will still be in their initially discardable state at such stage. A suitable typing discipline should deem the left snippet unsafe but the right snippet safe, taking into account the interference between control effects and the linear usage behavior of values in scope.

For a further example, consider the slightly more involved scenario given in Fig. 2, which involves concurrent communication and explicit exception handling. Our idealized linear functional language is now assumed to include the ability to fork threads, and manipulate session typed communication channels. In the first line of Fig. 2, a thread representing a logging server is forked: the fork primitive FORK l.e spawns a thread with body e accessing one endpoint of session channel l, and returns the other endpoint (bound to log) to the caller. The logging server receives precisely two messages, 
$$\begin{split} & \texttt{LET} \log = \texttt{FORK} \, l.\texttt{RECV}(l, m_1); \texttt{RECV}(l, m_2); \texttt{CLOSE} ? \, l \, \texttt{IN} \\ & \texttt{LET} \, res = \texttt{FORK} \, f.\texttt{RECV}(f, code); \texttt{SEND}(f, qs(code)); \texttt{RECV}(f, bk); \texttt{CLOSE} ? \, f \, \texttt{IN} \\ & \texttt{SEND}(res, \texttt{QU2112}); \texttt{RECV}(res, qss); \\ & \texttt{TRY} \\ & \texttt{LET} \, log = Check \, log \, qss \, \texttt{IN} \\ & \texttt{SEND}(res, bkok); \texttt{SEND}(log, "ok"); \texttt{CLOSE} \, ! \, log; \texttt{CLOSE} \, ! \, res \\ & \texttt{CATCH} \, logc.\texttt{SEND}(logc, "fail"); \texttt{CLOSE} \, ! \, logc \end{split}$$

Fig. 2. Code snippet for the server example.

whose payload is a string, and closes the connection. Then another concurrent thread is spawned, mocking a resource allocation service: it receives a resource code, returns the quality of service constraints, and receives some reservation information bk.

The server at channel *res* is used by client code that first sends a resource code QU2112, receives a quality of service *qss* spec, and then calls a conformance checking operation *Check* which, crucially, may raise an exception:

LET  $Check = \lambda l.\lambda x.SEND(l, "checkin"); if Valid(x) then l else THROW(l) IN...$ 

All interactions between client and server are meant to be logged, so channel *log* is also passed to the *Check* function together with the quality of service specification.

The overall expected behavior would be as follows. If *Valid()* returns true, then *Check* succeeds and the client will proceed booking the resource; however, if *Check* raises an exception, the continuation will be discarded and the exception handler invoked. In this case, the continuation of the resource allocation thread at the other endpoint will also need to be aborted. If the overall ongoing (linear) session between client and server could be safely discarded at that particular stage of the protocol, then the overall behavior may be deemed safe, even if the log could not be safely aborted, since the linear outstanding interaction on the *loq* endpoint will be performed anyway by the exception handler. Indeed, notice that the log is linearly passed to the exception handler as *loqc*. Again, a suitable typing discipline combining effects and linearity should be able to express the assumptions underlying the reasoning above, and deem the code snippet safe, under the stated assumptions, but unsafe if the client server session is not safely abortable exactly before the RECV(f, bk); interaction (cf. Line 2 in Fig. 2). Moreover, any such typing discipline should also be compatible with internal non-determinism, in the sense that if the result of Valid() is non-deterministic, then the resulting computation must be soundly typed for any alternative result, including the degenerate situation in which no value at all is returned, and the rest of the computation needs to safely abort, including, in that extreme case, the exception handler code itself!

The main goal of the paper is to investigate a principled foundation to express, reason and type-check a wide class of control effects in the context of a linear behavioral type system. Crucially, our approach builds on prior work on Curry-Howard correspondences between session types and various fragments of linear logic; our type system is a conservative extension of a standard system of classical linear logic. By approaching a Curry-Howard correspondence from the programming language perspective (in the spirit of, e.g., [15,4]), we introduce two new dual logical modalities—monadic & A and co-monadic  $\oplus A$ —and associated programming constructs and proof reductions. As is often the case for type systems motivated by Curry-Howard correspondences, our sys-

3

tem ensures global progress and usage / session protocol fidelity. Moreover, it is intrinsically compatible with all other logically motivated constructs and methods introduced in prior/related work, such as behavioral polymorphism [8,44], logical relations [33], dependent types [38], higher-order code mobility [40], and multiparty protocols [7,14].

It turns out that the new modalities &A and  $\oplus A$  suffice to express very general forms of internal non-determinism, and, importantly, include abortion—an explicitly typed form of affinity—as a special case. These two modalities can be technically seen as an additional pair of linear logic exponentials, and as such obey the basic monadic / co-monadic laws. However, while the standard linear logic modalities !A and ?A encapsulate contraction and weakening, our &A and  $\oplus A$  encapsulate non-determinism and abortion, in a sense to be made precise below. Although related to the non-deterministic monad and to well-known powerdomain models of non-determinism [35], a key novelty of our work is the perfect Curry-Howard match between proof reductions associated to the &A and  $\oplus A$  modalities and sensible operational rules. This correspondence allows us to state cut-elimination, and to naturally derive key properties with practical impact (e.g., lock freedom, fidelity, and strong normalization), while supporting natural effectful programming idioms and powerful reasoning techniques (such as logical relations).

We will illustrate, through a sequence of examples, how expressive linear usage protocols involving effects may be represented (compiled down) just to the basic linear logic system extended with these two primitives and associated programming constructs. We will present our basic results and examples for a canonical session-based  $\pi$ -calculus model realizing a Curry-Howard interpretation of session types as linear logic propositions. As it is well known, the  $\pi$ -calculus is a complete foundational model, able to represent general concurrent computation, encode higher-order data, object-oriented features and so on [36], so our development is carried on the setting of most higher-level programming languages. In particular, we will use this basic process model as target language in a typed encoding an effectful linear higher-order functional language with threads, session-typed channels, non-determinism, and exceptions, allowing us to show typings and discuss the examples shown above in a precise sense.

The rest of the paper is structured as follows. Section 2 presents the Curry-Howard interpretation of session types for concurrent processes via examples. Section 3 establishes basic meta-theoretical results for typed processes: cut elimination (Theorem 3.1), type preservation (Theorem 3.2), progress (Theorem 3.3). We also establish a postponing result (Theorem 3.5), which connects our process model with the (non confluent) non-determinism typical of process calculi. Section 4 discusses an embedding of  $\lambda^{\text{exc}}$  (a linear, higher-order functional language with concurrency and exceptions) into typed processes.  $\lambda^{\text{exc}}$  is the reference language for the motivating examples above. Theorem 4.2 ensures that our embedding preserves typing; therefore, all results in Section 3 will carry over to  $\lambda^{\text{exc}}$ . In Section 5 we discuss related works, and Section 6 concludes.

# 2 The Core Language and its Type System

The language we base our development on is a standard session-typed  $\pi$ -calculus, a core language in which general higher-order concurrent programs may be modeled and analyzed [36]. The (binary) session discipline [23,24] applies to pairs of name passing

processes that communicate through point-to-point channels. In this setting, interaction between processes always occur in matching pairs: when one partner sends, the other receives; when one partner offers a selection, the other chooses; when a partner closes the session, the other must acknowledge and no further interactions may be happen in the same channel. Sessions are initiated when a participant invokes a server, which acts as a shared provider, with the capability of unboundedly spawning fresh sessions between the invoking client and the newly created service instance process. A service name may be publicly shared by any clients in the environment. A session-based system exhibits concurrency and parallelism because many sessions may be executing simultaneously and independently. No races in communications within a session (or even between different sessions) can occur. Both session and server names may be passed around in communications. Session channels are subject to a linear usage discipline, conforming to a specific state dependent protocol, while server channels can be freely shared, and used by an arbitrary number of concurrent clients that can call on them for spawning new session instances.

Next we gradually introduce the ingredients of our typed process model (syntax, semantics, session types and their linear logic interpretation), which are summarized in Figure 3 (Page 14). This presentation allows us to better motivate and describe the the key novelty in this paper—the(dual) types for non-deterministic behaviors in sessions.

Caires and Pfenning [9] introduced a type system for  $\pi$ -calculus processes that corresponds to a linear logic proof system, revealing the first Curry-Howard interpretation of session types as linear logic propositions. Unlike traditional session type systems, Curry-Howard interpretations of behavioral types ensure global progress (meaning that well-typed processes never get stuck), livelock-freedom, and confluence (up to  $\equiv$ ), and may be developed within either intuitionistic [9,11] or classical linear logic [11,44], with certain subtle differences in expressiveness. The system presented in this paper is based on the presentation  $\Sigma_2$  of [1] of classical linear logic, here extended with mix principles and, crucially, with new exponential modalities  $\oplus A$  and  $\otimes A$ , which will be interpreted as (dual) types for non-deterministic sessions.

**Definition 2.1 (Types).** Types (A, B, C) are given by

 $A,B ::= \bot \mid \mathbf{1} \mid !A \mid ?A \mid A \otimes B \mid A \otimes B \mid A \oplus B \mid A \oplus B \mid A \otimes B \mid \oplus A \mid \&A$ 

In examples we will also assume given some basic (data) types, such as natural numbers, strings, etc, but will not elaborate the nature of such basic types (see, e.g., [38]).

For any type A, we define its dual  $\overline{A}$ , where  $(\cdot)$  corresponds to linear logic negation  $(\cdot)^{\perp}$ , following standard de Morgan-like laws. Intuitively, the type of a session endpoint is the dual of the type of the opposite end-point.

**Definition 2.2 (Duality).** *The duality relation on types is given by:* 

$\overline{1} = ot$	$\overline{!A} = ?\overline{A}$	$\overline{A\otimes B} = \overline{A} \otimes \overline{B}$	$\overline{A \oplus B} = \overline{A} \otimes \overline{B}$	$\overline{\oplus A} = \& \overline{A}$
$\overline{\perp} = 1$	$\overline{?A} = !\overline{A}$	$\overline{A \otimes B} = \overline{A} \otimes \overline{B}$	$\overline{A \otimes B} = \overline{A} \oplus \overline{B}$	$\overline{\&A} = \oplus\overline{A}$

Typing judgments have the form  $P \vdash \Delta; \Theta$ , where P is a program term,  $\Delta$  is the linear context and  $\Theta$  is the unrestricted context, along the lines of DILL [3] and  $\Sigma_2$  [1].

Both contexts are assignments of types to (channel) names x, y, z, ... We write '.' to denote empty typing environments.

After erasing the term P, our judgment corresponds exactly to a logical sequent in the presentation of classical linear logic  $\Sigma_2$  of [1]. Remarkably, this formulation naturally supports a Curry-Howard interpretation for the exponentials !A and ?A in terms of standard ( $\pi$ -calculus) semantics for lazy replication [9,11].

### 2.1 Reduction semantics

The operational semantics of our session calculus is defined by a relation of reduction (denoted  $P \rightarrow Q$ ) that expresses dynamic evolution, and a relation of structural congruence (denoted  $P \equiv Q$ ), which equates processes with the same spatial (or static) structure. This semantics exhibits a precise correspondence with cut elimination at the logic level. While most cut-reduction steps directly correspond to process reductions, other cut-reduction steps are better expressed in the process world as structural congruence principles or as behavioral equivalences; this applies similarly to the so-called commuting conversions, which are known to capture typed behavioral equivalences [33].

To describe reductions and conversions on proof trees (which correspond to typing derivations), it is convenient to introduce a simple algebraic notation. For each typing rule (T\*) with k premises  $d_1, \ldots, d_k$  we denote by  $T^*(p_1, p_2, \ldots)$  the derivation obtained by applying rule (T\*) to the derivations  $p_1, \ldots, p_k$ . In case the proof rule binds some names  $\tilde{x}$  in the conclusion, as in, e.g., cut, we would then write  $T^*(\tilde{x})(p_1, p_2, \ldots)$  to make this binding explicit.

### 2.2 Basic Typing Rules, Congruence Rules, and Reduction Rules

The parallel composition of processes is typed in our system by rules corresponding to the cut (dependent composition) and mix (independent composition) principles.

$$\frac{P \vdash \Delta; \Theta \quad Q \vdash \Delta'; \Theta}{P \mid Q \vdash \Delta, \Delta'; \Theta} (\mathsf{T} \mid ) \quad \frac{P \vdash \Delta, x:\overline{A}; \Theta \quad Q \vdash \Delta', x:A; \Theta}{(\nu x)(P \mid Q) \vdash \Delta, \Delta'; \Theta} (\mathsf{Tcut}) \quad \frac{\mathbf{0} \vdash; \Theta}{\mathbf{0} \vdash; \Theta} (\mathsf{T} \cdot)$$

The mix rule  $(T \mid )$  types the composition of two processes P and Q that do not share any linear name; thus, P and Q run in parallel, but do not interact. The cut rule (Tcut) types the composition of processes P and Q, while establishing a session between both using a single linear channel x; each process holds one of the two (dually typed) endpoints x of a session of type A and  $\overline{A}$ . Such communication channel is kept private to the composition by the restriction operator  $(\nu x)(...)$  so that the newly established binary session will not be affected by interference. Rule  $(T \cdot)$  allows the empty process **0** to be introduced. Neutrality of **0** is expressed by the conversion  $T|(T, D) \cong D$  at the level of proofs, whicih corresponds exactly to the familiar structural congruence principle  $\mathbf{0} \mid P \equiv P$  (we consider here a conversion, not a computational reduction, since it does not involve any process interaction). We take process terms up to basic structural congruence principles, namely we assume that -| – is commutative and associative with unit 0, etc. This way, e.g.,  $P \mid Q$  and  $Q \mid P$  denote the same process, i.e., the (unique) parallel composition of P and Q. Thus, Rule (Tcut) is symmetric w.r.t. its premises: if  $Tcut(D_1, D_2)$  is a derivation then  $Tcut(D_2, D_1)$  is the same derivation; we then consider the conversion  $\operatorname{Tcut}(D_1, D_2) \cong \operatorname{Tcut}(D_2, D_1)$  as well.

Session Communication, Send and Receive Session-typed processes communicate by sending and receiving messages according to some session discipline. The message payload can be a value of some primitive data type or a session channel; we focus here on the general case of session passing. Type  $A \otimes B$  is the type of a session that first sends a session of type A and then continues as a session of type B. As such, it corresponds to the session type !A.B of [23]. Dually,  $A \otimes B$  is the type of a session that first receives a session of type A and then continues as a session of type B; it thus corresponds to the session type  $!\overline{A.B}$ . So, the session type ?A.B corresponds to the linear type  $A \longrightarrow B$ . We have the following typing rules for send  $A \otimes B$  and receive  $A \otimes B$ .

$$\frac{P \vdash \Delta, y:A; \Theta \quad Q \vdash \Delta', x:B; \Theta}{\overline{x}(y).(P \mid Q) \vdash \Delta, \Delta', x:A \otimes B; \Theta} \text{ (T \otimes)} \qquad \frac{R \vdash \Gamma, y:C, x:D; \Theta}{x(y).R \vdash \Gamma, x:C \otimes D; \Theta} \text{ (T \otimes)}$$

An input process has the form x(y).R, a process that receives on session x a session n, passed in parameter y, and then proceeds as specified by R. The continuation R will use the received session but also any other open sessions (including x). An output process has the form  $\overline{x}(y).M$ , where y is a freshly created name. The behavior of such a output process is to send session y on x and then proceed as defined by M. In our typed language, the output continuation M has the form  $P \mid Q$ , where P defines the behavior of the session y being sent and Q the behavior of the continuation session on x. Notice that y is bound both in  $\overline{x}(y).M$  and in x(y).R, and so only fresh names can be sent in output processes; this corresponds to the *internal mobility* discipline investigated in [5], which does not restrict expressiveness. The associated principal cut reduction corresponds to process communication, where  $C = \overline{A}$ ,  $D = \overline{B}$ , expressed by

 $\operatorname{Tcut}(x)(\operatorname{T}\otimes(y)(D_1, D_2), \operatorname{T}\otimes(y)(D_3)) \to \operatorname{Tcut}(x)(\operatorname{Tcut}(y)(D_1, D_3), D_2)$ 

This reduction exactly captures (bound output) communication in the  $\pi$ -calculus

$$(\nu x)(\overline{x}(y).M \mid x(y).R) \to (\nu x)(\nu y)(M \mid R)$$

where we write  $M \equiv P \mid Q$ . Although  $- \mid -$  is commutative there is no ambiguity in  $(T\otimes)$ ; as P and Q are the split of M typed by  $P \vdash \Delta, y:A; \Theta$  and  $Q \vdash \Delta', x:B; \Theta$ , respectively, as  $\Delta$  and  $\Delta'$  are the split of the linear context in the conclusion. The multiplicative units  $\perp$  and 1 type session termination actions as seem from each one of the two end-points; no partner can make further use of a closed session.

$$\frac{P \vdash \Delta; \Theta}{x. close; P \vdash x: 1; \Theta} (T1) \qquad \frac{P \vdash \Delta; \Theta}{x. close; P \vdash x: \bot, \Delta; \Theta} (T\bot)$$

The associated principal cut reduction corresponds to session termination, which we define respectively at the level of processes and proof trees by the rules

$$(\nu x)(x.\overline{\mathtt{close}} \mid x.\mathtt{close}; P) \to P \qquad \mathrm{Tcut}(x)(\mathrm{T}\mathbf{1}, \mathrm{T}\bot(D)) \to D$$

Types 1 and  $\perp$  correspond in traditional session types to a single type, denoted end, and are usually silent at the programming language level. In the presence of mix principles, as we consider here, the propositions  $\perp -\infty 1$  and  $1 - \infty \perp$  are valid. Morally, we could consider  $\perp = 1$ , and define a single type end for completed sessions as standing for "both" 1 or  $\perp$ , where end = end.

Session Offer and Choice The linear type  $A \oplus B$  types a session that first *chooses* (from the dual partner menu) either "left" or "right", and then continues as a session of type A or B, depending on the choice. This type is the binary version of the session type  $\bigoplus_{i \in I} \{l_i:A_i\}$  (labeled internal choice). The linear type  $A \otimes B$  types a session that first *offers* both "left" or "right" menu options (for a partner to choose) and then continues as a session of type A or B, depending on the choice made. Thus,  $A \otimes B$  corresponds to the binary version of the session type  $\bigotimes_{i \in I} \{l_i:A_i\}$  (labeled external choice). Offers and choices are typed by the additive linear conjunction and disjunction  $\bigotimes$  and  $\bigoplus$ , as defined by the rules:

$$\frac{R \vdash \Delta, x:A; \Theta}{x.\mathtt{inl}; R \vdash \Delta, x:A \oplus B; \Theta} (\mathsf{T} \oplus_1) \qquad \frac{R \vdash \Delta, x:B; \Theta}{x.\mathtt{inr}; R \vdash \Delta, x:A \oplus B; \Theta} (\mathsf{T} \oplus_2) \\ \frac{P \vdash \Delta, x:A; \Theta \quad Q \vdash \Delta, x:B; \Theta}{x.\mathtt{case}(P, Q) \vdash \Delta, x:A \otimes B; \Theta} (\mathsf{T} \otimes)$$

The associated principal cut reductions correspond to the process and proof reductions

$$\begin{array}{l} (\nu x)(x. \mathtt{case}(P,Q) \mid x. \mathtt{inl}; R) \to (\nu x)(P \mid R) \\ (\nu x)(x. \mathtt{case}(P,Q) \mid x. \mathtt{inr}; R) \to (\nu x)(Q \mid R) \\ \mathrm{Tcut}(x)(\mathrm{T} \& (D_1, D_2), \mathrm{T} \oplus_1(D_3)) \to \mathrm{Tcut}(x)(D_1, D_3) \\ \mathrm{Tcut}(x)(\mathrm{T} \& (D_1, D_2), \mathrm{T} \oplus_2(D_3)) \to \mathrm{Tcut}(x)(D_2, D_3) \end{array}$$

In examples we may consider labeled sums, close to traditional session types constructs.

$$\frac{R \vdash \Delta, x:A; \Theta}{x.\mathbf{l}_{\mathbf{i}}; R \vdash \Delta, x: \oplus_{i \in I} \{\mathbf{l}_{\mathbf{i}}:A_i\}; \Theta} \quad \frac{P_i \vdash \Delta, x:A_i; \Theta \quad (\text{all } i \in I)}{x.\mathsf{case}_{i \in I}(\mathbf{l}_{\mathbf{i}}.P_i) \vdash \Delta, x: \otimes_{i \in I} \{\mathbf{l}_{\mathbf{i}}:A_i\}; \Theta}$$

with associated principal cut reduction expressed by

$$(\nu x)(x.\mathtt{case}_{i \in I}(\mathtt{l}_{\mathtt{i}}.P_i) \mid x.\mathtt{l}_i; R) \rightarrow (\nu x)(P_i \mid R)$$

*Example 2.3 (Movie Server (1)).* Consider a toy scenario involving a movie server and some clients. We first model a single session (on channel s) established between client Alice(s) and server instance SBody(s). The server session offers two options: "buy movie" (in1), and "preview trailer" (inr). Alice selects the "preview" option from the server menu, and plays the corresponding protocol. Consider now the following terms:

$$\begin{split} SBody(s) &\triangleq s.\texttt{case}(s(title).s(card).s\langle movie \rangle.\texttt{close}, s(title).s\langle trailer \rangle.\texttt{close}) \\ Alice(s) &\triangleq s.\texttt{inr}; s\langle "mullholanddrive" \rangle.s(preview).\texttt{close}; \mathbf{0} \\ System_1 &\triangleq (\nu s)(SBody(s) \mid Alice(s)) \end{split}$$

Assuming some given basic types for movie titles (T), credit card data (C) and movie files M (and for which we may assume to be self-dual; since they do not type communication capabilities but values of basic types), we can then provide the following types and derivable type assignments for the various system components as follows:

$$SBT \triangleq (T \multimap C \multimap M \otimes \mathbf{1}) \& (T \multimap M \otimes \mathbf{1})$$
  

$$SBody(s) \vdash s : SBT ; \cdot \qquad Alice(s) \vdash s : \overline{SBT} ; \cdot$$

We would then have  $System_1 \vdash \cdot ; \cdot$ . While the type of the server endpoint is SBT, the type of a client endpoint would be  $\overline{SBT} = (T \otimes C \otimes M \multimap \bot) \oplus (T \otimes M \multimap \bot)$ .  $\Box$ 

Shared Service Definition and Invocation Shared service definition and invocation are typed by the linear logic exponentials ! and ?. Type !A types a shared channel that persistently offers a replicated service which whenever invoked spawns a fresh session of type A (from the server's perspective). Dually, type ?A types a shared channel on which requests to a persistently replicated service of type A can be unboundedly issued (from the client's perspective). We consider the following typing rules:

$$\frac{P \vdash \Delta; x:A, \Theta}{P \vdash \Delta, x:?A; \Theta} \text{ (T?)} \quad \frac{Q \vdash y:A; \Theta}{!x(y).Q \vdash x:!A; \Theta} \text{ (T!)} \quad \frac{P \vdash \Delta, y:A; x:A, \Theta}{\overline{x}?(y).P \vdash \Delta; x:A, \Theta} \text{ (Tcopy)}$$

The associated principal cut reduction corresponds to shared service invocation

 $(\nu x)(!x(y).Q \mid \overline{x}?(y).P) \to (\nu x)(!x(y).Q \mid (\nu y)(P \mid Q))$ 

This operational interpretation of the rules for !A and ?A (cf. [1,34,3], implementing "lazy" contraction) exactly coincides with the usual interpretation of lazy replication. Notice that Rule (T?) is silent on the term assignment: it implements a bookkeeping device to move the typed channel x:? A to the unrestricted context, and does not induce a computational effect (e.g., as exchange is also implicitly handled).

As our typing judgments have two different regions, linear and exponential, two cut rules are required [3], one for cutting a linear (session) channel in the linear context (Tcut), already presented in Section 2.2, and the following rule, for cutting a unrestricted (shared) channel in the exponential context (Tcut<sup>?</sup>) [34,3]:

$$\frac{P \vdash y:\overline{A}; \Theta \quad Q \vdash \Delta; x:A, \Theta}{(\nu x)(!x(y).P \mid Q) \vdash \Delta; \Theta} (\mathsf{Tcut}^?)$$

For typing "source programs" only the linear (Tcut) is required, but (Tcut?) is required for cut-elimination; hence, Rule (Tcut?) is a "runtime" typing rule. The principal reduction above is expressed at the level of proofs by

 $\operatorname{Tcut}(x)(\operatorname{T}!(y)(D_1), \operatorname{Tcopy}(y)(D_2)) \to \operatorname{Tcut}_{?}(xy)(D_1, \operatorname{Tcut}(y)(D_1, D_2))$ 

*Example 2.4 (Movie Server (2)).* We illustrate the usage of !A and ?A types using a shared movie server, which may answer requests from an unbounded number of clients; here we use just two concurrent clients, *SAlice* and *SBob*. Alice still selects the "preview trailer" option as in Example 2.3, but Bob selects the "buy movie" option. Recall the definitions of type *SBT* and processes *SBody(s)* and *Alice(s)* from Example 2.3.

$$\begin{array}{l} MOVIES(srv) \triangleq !srv(s).SBody(s) \\ Bob(s) \triangleq s.\texttt{inl}; \overline{s}(``inception").\overline{s}(bobscard).s(mpeg).s.\texttt{close}; \textbf{0} \\ SAlice(srv) \triangleq \overline{srv}(s).Alice(s) \qquad SBob(srv) \triangleq \overline{srv}(s).Bob(s) \\ System_2 \triangleq (\nu srv)(MOVIES(srv) \mid SAlice(srv) \mid SBob(srv)) \end{array}$$

The following typing judgments are derivable:

$$\begin{array}{ll} MOVIES(srv) \vdash srv : !SBT ; \cdot & SAlice(srv) \vdash \cdot ; srv : \overline{SBT} \\ SBob(srv) \vdash \cdot ; srv : \overline{SBT} & Alice(srv) \mid Bob(srv) \vdash srv : ?\overline{SBT} ; \cdot \end{array}$$

We can obtain  $System_2 \vdash \cdot; \cdot$  as follows: we first use the (mix) Rule (T | ) to compose the two clients; then, Rule (T?) is used to merge the shared endpoints under the explicit type  $?\overline{SBT}$ ; finally, the clients are composed with the server using Rule (Tcut).

**Identity** We interpret the identity axiom by the *forwarder process*  $[x \leftrightarrow y]$  [10,44], which denotes a bidirectional (linear) link between sessions x and y, giving a logical justification to a known concept in  $\pi$ -calculi (cf. [20]). The forwarder at type A is typed

$$[x \leftrightarrow y] \vdash x:A, y:\overline{A}; \Theta$$
 (Tid)

The associated cut reduction  $(\nu x)(P | [x \leftrightarrow y]) \rightarrow P\{y/x\}$  (where y is not free in P) is akin to the application of an explicit substitution. It is known since [26] that linear forwarders can simulate substitution in the sense of the above reduction rule. We also introduce  $[x \leftrightarrow y] \equiv [y \leftrightarrow x]$  as a structural congruence axiom, as a direct consequence of (implicit) exchange in the typing context. While a well-typed copycat forwarder process  $F_A$  without forwarder links can be easily constructed for any concrete type A by  $\eta$ -expansion (see [11]) the primitive forwarder is particularly important when considering polymorphism [8]. It also allows us to represent the "free" output construct  $x\langle y\rangle \cdot P$  (where y is a free channel name in scope) by  $\overline{x}(z) \cdot ([y \leftrightarrow z] | P)$  (cf. [5]).

### 2.3 Non-Determinism and Abortion

The developments of this paper focus on the challenge of expressing fundamental primitives for non-deterministic behavior—including the special important case of abortable behavior—in the setting of our Curry-Howard correspondence for session types.

It is often believed that a Curry-Howard interpretation of a programming language is hard to reconcile with true (so-called internal) non-determinism in computation, since reduction steps should express at most behavioral equivalences on processes, via proof identities, which are inherently confluent from an operational viewpoint. However, it is clear, at least from work on denotational semantics and functional programming, that non-determinism can be handled equationally by working on the powerdomain of computation results. In the logical setting, developments on differential linear logic [18] also require the interpretation domain for proofs to be closed under a (formal) notion of "sum", which may possibly be interpreted as non-deterministic choice. Although partially inspired by such approaches, our proposal picks a fairly different road, which turns out to lead to the first example of a Curry-Howard interpretation of a realistic programming language with built-in internal non-determinism.

It is well known after Girard that the linear logic exponential modalities !A and ?A, which have been used above to model the type of shared channel names, are not uniquely defined by their standard proof rules: not surprisingly, if one adds additional operators defined by the same rules, we obtain independent monad / comonad pairs. We exploit this fact to our advantage, noting that it allows us to modularly add new "exponential" modalities to the base logical system, defined by identical proof rules (in Girard's original formulation), without semantically interfering with the existing ones. Any such pair of connectives (say, &A and  $\oplus A$ ) will yield a dual monad / comonad pair defined by the fundamental principles (in a simplified form):

$$\frac{\vdash \Delta, A}{\vdash \Delta, \& A} \quad \frac{\vdash \& \Delta, A}{\vdash \& \Delta, \oplus A}$$

In the particular case of the usual modalities !A and ?A, additional specific rules for ?A define the intended semantics of the linear logic exponentials, which encapsulate the

structural rules of weakening and contraction:

$$\frac{\vdash \Delta}{\vdash \Delta, ?A} \quad \frac{\vdash \Delta, ?A, ?A}{\vdash \Delta, ?A}$$

These observations suggest a methodology for the addition of new monadic operators to the basic linear logic framework in a logically justified way, by means of independent monad / comonad pairs, in which the monad semantics is defined by specific additional logical principles. In this work, we develop our type system on top of (classical) linear logic, conservatively extended with two operators capturing a (co)monad defined by (a refined version of the) following principles, which can be verified to be sound for an (additive) monad &- and comonad  $\oplus$ -.

$$\frac{\vdash \varDelta, A}{\vdash \varDelta, \& A} \qquad \frac{\vdash \& \varDelta, A}{\vdash \& \varDelta, \oplus A} \qquad \frac{\vdash}{\vdash \& A} \qquad \frac{\vdash \& \varDelta}{\vdash \& \varDelta}$$

The resulting proof (and type) system provides a Curry-Howard interpretation of a realistic programming language with built-in internal non-determinism and abortion.

So, getting back to the presentation of our type system, we capture non-deterministic behavior in the type structure by operators &A and  $\oplus A$  related by duality ( $\overline{\&A} = \oplus \overline{A}$ ) and defined by the following rules:

$$\frac{P \vdash \Delta, x:A; \Theta}{x.\overline{\mathtt{some}}; P \vdash \Delta, x:\&A \Theta} (\mathsf{T} \otimes^{x}_{d}) \qquad \frac{x.\overline{\mathtt{none}} \vdash x:\&A \Theta}{x.\overline{\mathtt{none}} \vdash x:\&A \Theta} (\mathsf{T} \otimes^{x}_{d}) \\ \frac{P \vdash \overline{w}:\&\Delta, x:A; \Theta}{x.\mathtt{some}_{\overline{w}}; P \vdash \overline{w}:\&\Delta, x:\oplus A; \Theta} (\mathsf{T} \oplus^{x}_{\overline{w}}) \qquad \frac{P \vdash \&\Delta; \Theta \ Q \vdash \&\Delta; \Theta}{P \oplus Q \vdash \&\Delta; \Theta} (\mathsf{T} \otimes)$$

Intuitively, &A is the type of a session that *may produce* a behavior of type A: this potential is made concrete in Rule  $(T \otimes_d^x)$  where the behavior x:A is indeed available (some), whereas Rule  $(T \otimes_d^x)$  describes the case in which x:A is not available (none). Dually, the type  $\oplus A$  is the type of a session that *may consume* a behavior of type A. Rule  $(T \oplus_w^x)$  accounts for the possibility of not being able to consume an A by considering sessions different from x as potentially not available (i.e., abortable - cf.  $\overline{w}: \& \Delta$  in the rule, where  $\overline{w}$  denotes a sequence  $w_1, \ldots, w_n$  of names). Rule  $(T \otimes)$  expresses non-deterministic sum, it may be seen to a formal sum of proofs (cf. [18]), but in our case exactly corresponds to non-deterministic choice  $P \oplus Q$  of processes, and can only be used inside the monad &A. The principal cut reductions are:

$$\begin{aligned} \operatorname{Tcut}(x)(\operatorname{T}\otimes^x,\operatorname{T}\oplus^x_{\overline{w}}(D_2)) &\to \operatorname{T} \mid (\operatorname{T}\otimes^{w_1},\cdots,\operatorname{T}\otimes^{w_i}) \\ \operatorname{Tcut}(x)(\operatorname{T}\otimes^x_d(D_1),\operatorname{T}\oplus^x_{\overline{w}}(D_2)) &\to \operatorname{Tcut}(x)(D_1,D_2) \end{aligned}$$

At the level of the process interpretation, these reduction rules are expressed by

$$(\nu x)(x.\overline{\mathtt{some}}; P \mid x.\mathtt{some}_{\overline{w}}; Q) \to (\nu x)(P \mid Q)$$
$$(\nu x)(x.\overline{\mathtt{none}} \mid x.\mathtt{some}_{\overline{w}}; Q) \to w_1.\overline{\mathtt{none}} \mid \cdots \mid w_n.\overline{\mathtt{none}}$$

Notice how the reduction for  $\overline{none}$  safely discards the continuation Q. We also consider the following proof conversion and corresponding process congruence

$$\begin{aligned} \operatorname{Tcut}(x)(\operatorname{T}\&(D_1,D_2),D_3) &\equiv \operatorname{T}\&(\operatorname{Tcut}(x)(D_1,D_3),\operatorname{Tcut}(x)(D_2,D_3))\\ (\nu x)(P \mid (Q \oplus R)) &\equiv (\nu x)(P \mid Q) \oplus (\nu x)(P \mid R) \end{aligned}$$

that expresses the distribution of parallel composition over internal choice. Notice that the two computational reduction rules above could, in principle, be formally used to express the reduction rules for the "sharing" exponentials (cf. [44]) in presentations of linear logic with explicit weakening and dereliction rules, instead of the DILL-style presentation we prefer, as it more tightly express the behavior of sharing present in traditional session types. On the other hand, together with the conversion principle just shown, the primitives and reduction rules just presented turn out to be quite adequate to express the behavior of non-determinism and abortion.

Before closing the section, we discuss initial examples using  $\oplus A$  and & A types.

*Example 2.5 (Movie Server (3)).* Getting back to our movie server scenario we illustrate how to model a system with a client Randy(s) that non-deterministically decides between either buying a movie or just seeing its trailer. Recalling process definitions for SBody(s), Alice(s), and Bob(s) from Examples 2.3 and 2.4, we would have:

$$Randy(s) \triangleq s.\overline{\texttt{some}}; Alice(s) \oplus s.\overline{\texttt{some}}; Bob(s)$$
$$USystem \triangleq (\nu s)(s.\texttt{some}_{\emptyset}; SBody(s) \mid Randy(s))$$

where the suitable types and type assignments are now given by

$$Randy(s) \vdash s : \otimes \overline{SBT} ; \cdot s.some_{\emptyset}; SBody(s) \vdash s : \oplus SBT ;$$

To type process Randy(s), one would first use Rule  $(T \otimes_d^s)$  on each individual client; then, using Rule  $(T \otimes)$  one would obtained a typeful non-deterministic choice between them. To type the server, one uses Rule  $(T \oplus_{\overline{w}}^s)$ : in this case,  $\overline{w} = \emptyset$  because there are no other sessions (besides s) in the linear context (recall that  $SBody(s) \vdash s : SBT$ ;  $\cdot$ ). This way, we can derive  $USystem \vdash \cdot$ ;  $\cdot$ .

Interestingly, the non-deterministic choices enabled at the level of types by &A and  $\oplus A$  and at the process level by  $\oplus$  is completely orthogonal to the usual deterministic choices enabled by labeled internal and external choices. The following example illustrates the pleasant interaction between the two kinds of choices:

*Example 2.6 (Movie Server (4)).* Consider now a variant of the movie server that logs the request made by the client on a log service l of type  $\mathbf{B} = \mathbf{1} \oplus \mathbf{1}$ . We extend the process SBody(s) from Example 2.3 as follows:

 $SBodyL(s) \triangleq s.case(s(title).s(card).s(movie).l.inl; close, s(title).s(trailer).l.inr; close)$ 

We may provide a typing  $SBodyL(s) \vdash s:SBT$ ,  $l:\mathbf{B}$ ;  $\cdot$  which cannot be composed with the non-deterministic client Randy(s) from Example 2.5. However, consider

 $s.\texttt{some}_l; l.\overline{\texttt{some}}; SBodyL(s)$ 

This process may now be composed with client process Randy(s) as

 $ULSystem \triangleq (\nu s)(s.\texttt{some}; l.\texttt{some}; SBodyL(s) \mid Randy(s))$ 

Now we may derive:

 $l.\overline{\mathtt{some}}; SBodyL(s) \vdash s: SBT, l: \& \mathbf{B}; \cdot \qquad ULSystem \vdash l: \& \mathbf{B}; \cdot \\ s.\mathtt{some}_l; l.\overline{\mathtt{some}}; SBodyL(s) \vdash s: \oplus SBT, l: \& \mathbf{B}; \cdot \end{cases}$ 

We obtain the reduction sequence (we denote  $P \Rightarrow Q$  the reflexive-transitive closure of  $P \rightarrow Q$ ) ULSystem  $\Rightarrow (l.\overline{\texttt{inr}}; \mathbf{0} \oplus l.\overline{\texttt{inl}}; \mathbf{0})$ .

Notice that the visible behavior of log channel l in ULSystem must be given the non-deterministic type  $\& \mathbf{B}$ : there is no typing  $ULSystem \vdash l:\mathbf{B}$ , since the resulting interaction is essentially non-deterministic. Alternatively, we could directly express the server body with the intended non-deterministic typings, as follows:

$$\begin{split} SBodyNDL(s) &\triangleq s.\texttt{some}_l; s.\texttt{case}(s(title).s(card).s\langle movie \rangle.l.\texttt{some}; l.\texttt{inl};\texttt{close}, \\ s(title).s\langle trailer \rangle.l.\texttt{some}; l.\texttt{inr};\texttt{close}) \\ SBodyNDL(s) \vdash s: \oplus SBT, l: \& \mathbf{B}; \cdot \\ (\nu s)(SBodyNDL(s) \mid Randy(s)) \vdash l: \& \mathbf{B}; \cdot \end{split}$$

In our system, the ability of representing (internal) non-determinism is intrinsically tied to that of describing, in a completely logically motivated manner, abortable behaviors as typical of programming constructs such as, e.g., (structured) exceptions and compensations. Our following example illustrates this distinctive aspect of our model.

*Example 2.7 (Movie Server (5)).* To consider the possibility of modeling abortion, we introduce the code for a "faulty" client, that non-deterministically behaves like Bob(s) (cf. Example 2.4) or does not produce any behavior at all. Consider the non-deterministic server SBodyNDL(s) from Example 2.6; we may now have:

$$\begin{array}{ll} Buzz(s) \triangleq \overline{\texttt{some}}; Bob(s) \oplus \overline{\texttt{none}} & Buzz(s) \vdash s : \& \overline{SBT} ; \cdot \\ (\nu s)(SBodyNDL(s) \mid Buzz(s)) \vdash l : \& \mathbf{B} ; \cdot \end{array}$$

Notice how abortion / failure of sub computations propagates inside the monad &-, encapsulated in an hereditarily safe way. Here, we have the reduction sequence

 $(\nu s)(SBodyNDL(s) \mid Buzz(s)) \Rightarrow (l.\overline{\texttt{none}} \oplus l.\overline{\texttt{inl}}; 0)$ 

reflecting that the composed system either aborts or chooses l.inl on the log.

Our next example illustrates, in a simplest setting, how systems encapsulating nondeterministic behavior can nevertheless be given a globally deterministic type, thus showing that the internal non-determinism and abortion is not observable at public channels as long as they are typed by "plain" deterministic types.

Example 2.8. Consider the following processes and typings:

 $\begin{array}{ll} Some(y) \triangleq y.\overline{\texttt{some}}; y.\texttt{in1}; \mathbf{0} \oplus y.\overline{\texttt{some}}; y.\texttt{inr}; \mathbf{0} & Some(y) \vdash y : \& \mathbf{B} \\ Prod \triangleq \overline{x}(y).(Some(y) \mid b\langle ``done" \rangle.\mathbf{0}) & Prod \vdash x : (\& \mathbf{B}) \otimes \bullet, b : Str \otimes \bullet \\ Cons \triangleq x.\texttt{some}; x.\texttt{case}(\mathbf{0}, \mathbf{0}) & Cons \vdash x: (\oplus \mathbf{B}) \otimes \bullet \\ Blob \triangleq (\nu x)(Prod \mid Cons) & Blob \vdash b : String \otimes \bullet \end{array}$ 

(Processes)

$$\begin{split} P &::= \begin{bmatrix} x \leftrightarrow y \end{bmatrix} \mid P \mid Q \mid (\nu y)P \mid \mathbf{0} \\ \mid \overline{x}(y).P \mid x(y).P \mid \overline{x}?(y).P \mid !x(y).P \\ \mid x.\texttt{case}(P,Q) \mid x.\texttt{inr}; P \mid x.\texttt{inl}; P \mid x.\overline{\texttt{close}} \mid x.\texttt{close}; P \\ \mid P \oplus Q \mid x.\overline{\texttt{some}}; P \mid x.\overline{\texttt{none}}; P \mid x.\texttt{some}_{\overline{w}}; P \end{split}$$

(Reduction - Contextual Congruence Rules omitted)

$$\begin{split} \overline{x}(y).Q \mid x(y).P &\to (\nu y)(Q \mid P) & \overline{x}?(y).Q \mid !x(y).P \to (\nu y)(Q \mid P) \mid !x(y).P \\ x.\operatorname{inr}; P \mid x.\operatorname{case}(Q,R) \to P \mid R & x.\operatorname{inl}; P \mid x.\operatorname{case}(Q,R) \to P \mid Q \\ (\nu x)([x \leftrightarrow y] \mid P) \to P\{y/x\} & x.\overline{\operatorname{some}}; P \mid x.\operatorname{some}_{\overline{w}}; Q \to P \mid Q \\ x.\overline{\operatorname{close}} \mid x.\operatorname{close}; P \to P & x.\overline{\operatorname{none}} \mid x.\operatorname{some}_{\overline{w}}; Q \to w_1.\overline{\operatorname{none}} \mid \dots \mid w_n.\overline{\operatorname{none}} \\ P \to Q \Rightarrow (\nu y)P \to (\nu y)Q & P \equiv P', P' \to Q', Q' \equiv Q \Rightarrow P \to Q \\ Q \to Q' \Rightarrow P \mid Q \to P \mid Q' & Q \to Q' \Rightarrow P \oplus Q \to P \oplus Q' \end{split}$$

(Structural Congruence - Contextual Congruence Rules omitted)

$$P \mid \mathbf{0} \equiv P \qquad P \equiv_{\alpha} Q \Rightarrow P \equiv Q \qquad (\nu x)\mathbf{0} \equiv \mathbf{0} \qquad P \mid Q \equiv Q \mid P$$
$$P \mid (Q \mid R) \equiv (P \mid Q) \mid R \qquad x \notin fn(P) \Rightarrow P \mid (\nu x)Q \equiv (\nu x)(P \mid Q)$$
$$[x \leftrightarrow y] \equiv [y \leftrightarrow x] \qquad (\nu x)(\nu y)P \equiv (\nu y)(\nu x)P \qquad \mathbf{0} \oplus \mathbf{0} \equiv \mathbf{0} \qquad P \oplus Q \equiv Q \oplus P$$
$$P \oplus (Q \oplus R) \equiv (P \oplus Q) \oplus R \qquad (\nu x)(P \mid (Q \oplus R)) \equiv (\nu x)(P \mid Q) \oplus (\nu x)(P \mid R)$$

Fig. 3. The Process Language.

Notice that the although the producer *Prod* sends (on *x*) a non-deterministic boolean Some(y) (on *y*) to the consumer *Cons*, which actually uses it, the type of the composed system *Blob* is  $b : String \otimes \bullet$ , a deterministic type.

In fact, we may easily verify that  $Plug \Rightarrow b\langle ("done" \rangle .0$ . In Section 4, we will show how the extremely simple constructs associated to our non-deterministic types are expressive enough to support the encoding of a higher-order functional language with threads, session channels, non-determinism, and exceptions.

# 3 Main Results

We collect in this section main sanity results for our non-deterministic linear logicbased type system for session process behavior. First, our system enjoys the cut-elimination property. Cut elimination may be derived given a suitable congruence  $\cong_s$  on processes consisting of reduction (computational conversions), structural congruence (structural conversions), and some key commuting conversions (cf.[9,33,11]).

**Theorem 3.1 (Cut Elimination).** If  $P \vdash \Delta; \Theta$  then there is a process Q such that  $P \cong_s Q$  and  $Q \vdash \Delta; \Theta$  is derivable without using the rules (Tcut) and (Tcut?).

The proof is an extension of the proof for classical linear logic with mix, but considering the new reductions and conversions introduced above for revealing and reducing principal cuts involving the &A and &A modalities.

Then, we may state type safety, witnessed by theorems of type preservation and global progress for closed systems. Type preservation states that the observable interface of a system is invariant under reduction.

# **Theorem 3.2** (Type Preservation). If $P \vdash \Delta$ ; $\Theta$ and $P \rightarrow Q$ then $Q \vdash \Delta$ ; $\Theta$ .

*Proof.* (Sketch) By induction on typing derivations, and case analysis on reduction steps. In each case, the result easily follows, given that reductions come from well-defined proof conversions, which by construction preserve typing.

Unlike standard type systems for session types, our logical interpretation satisfies global process, meaning that well-typed processes never get stuck on pending linear communications. More precisely, we say that a process P is live, noted live(P), if and only if  $P \equiv C[\pi.Q]$  where C[-] is a static context (e.g. a process term context in which the hole is not behind an action prefix, but only under parallel composition - | -, name restriction  $(\nu x)$ -, or sum  $-\oplus$ - operators) and  $\pi.Q$  is not a non-replicated action prefixed process (i.e.,  $\pi$  is a simple session input, output, offer, choice, or non-deterministic action prefix). We then have:

### **Theorem 3.3 (Progress).** *If* $P \vdash ; \Theta$ *and* live(P) *then there is* Q *such that* $P \rightarrow Q$ *.*

*Proof.* (Sketch) By induction on the typing derivation. Our proof relies on a contextual progress lemma, which uses a labeled transition system for processes, compatible with reduction (cf. [11]). This lemma yields a more general progress property for processes with free linear channels that transition by means of immediate external interactions. It extends Lemma 4.3 in [11] (which holds for a language without non-determinism) as follows: If  $P \vdash \Delta; \Theta$  and live(P) then either (1) there is Q such that  $P \rightarrow Q$  or (2) there are  $P_i$  (i = 1..n) such that  $P \equiv \oplus P_i$  and for all  $P_i$  there exist  $Q_i$  and  $\alpha$  such that  $P_i \xrightarrow{\alpha} Q_i$ . The proof of this extended lemma is by induction on derivations.

We now discuss additional results that clarify some key features of the our type system. We say that a process P is *prime* if it is not structurally congruent to a process of the form  $Q \oplus R$  with non-trivial (i.e., equivalent to 0) Q and R. We can then prove:

**Proposition 3.4.** Let  $P \vdash \Delta$ ;  $\Theta$  where types in  $\Delta$ ;  $\Theta$  are deterministic (do not contain & A or  $\oplus A$  types at the top level), and let  $P \Rightarrow Q \not\rightarrow$ . Then Q is prime.

Proof. (Sketch) By induction on the typing derivation.

Based on a logical system in which reduction deeply matches cut-elimination, it turns out that typing in our system enforces the confluence property, and also strong normalization. These results can be established using (linear) logical relations, as developed in [33]. Intuitively, confluence holds because non-determinism is captured equationally without losing information, by means of delaying choice in processes  $P \oplus Q$ , which express sets of alternative states. Still, it is interesting to relate our system with standard process calculi which explicitly commit non-deterministic states into alternative components. For that purpose, we investigate the extension of the reduction relation in Figure 3 with non-confluent rules for internal choice, standard in process calculi but clearly incompatible with any Curry-Howard interpretation, namely  $P \oplus Q \rightarrow P$  and  $P \oplus Q \rightarrow Q$ . We denote by  $P \rightarrow_c Q$  the extended reduction relation, which can be proven to still satisfy preservation and progress in the sense of Theorems 3.2 and 3.3. We may then show the following property, expressing postponing of internal non-deterministic collapse of non-deterministic states into prime states.

**Theorem 3.5** (Postponing). Let  $P \vdash \Delta; \Theta$ . We have

1. If  $P \Rightarrow P_1 \oplus \ldots \oplus P_n \not\rightarrow$  with  $P_i$  prime for all i, then  $P \Rightarrow_c P_i$  for all  $0 < i \le n$ . 2. Let  $C = \{P_i \mid P \Rightarrow_c P_i \not\rightarrow_c \text{ and } P_i \text{ is prime } \}$ . Then C is finite up to  $\equiv$ , with #C = n, and for all  $0 < i \le n$ ,  $P \Rightarrow P_1 \oplus \ldots \oplus P_n \rightarrow_c P_i$ .

*Proof.* 1. Trivial by definition. 2. By induction on the reduction sequence, using the fact that we may commute  $\Rightarrow$  reduction steps backwards with  $\Rightarrow_c$  reduction steps.

Theorem 3.5(2) shows that no information is lost by  $\Rightarrow$  with respect to the (standard) non-deterministic (and non-confluent) semantics of internal choice  $P \oplus Q$  expressed by  $\Rightarrow_c$ . We may therefore tightly relate our system, based on a logically motivated reduction relation, with a standard non-confluent reduction relation including rules for internal choice, in the sense that the former precisely captures the multiset of observable alternatives defined by the latter, while preserving compositional and equational reasoning about system behavior as expected from a Curry-Howard interpretation.

## 4 Higher-Order Concurrency, Non Determinism, and Exceptions

We illustrate the expressive power of our typed process model by embedding  $\lambda^{\text{exc}}$ , a linear higher-order functional, concurrent programming language with concurrency, nondeterminism—including abortion—, and exceptions. Defined by a typed compositional encoding, this embedding allows us to showcase the generality of our developments and the relevance of our Curry-Howard correspondence in a broader setting; it will also enable us to give a rigorous footing to our motivating examples (cf. Figures 1 and 2).

The Target Calculus.  $\lambda^{\text{exc}}$  is a typed call-by-value functional calculus, defined by the grammar below. We use  $e, e', \ldots$  to range over expressions;  $v, v', \ldots$  to range over values;  $x, y, z, \ldots$  to range over variables;  $c, c', \ldots$  to range over channels, and T, U, A, B to range over types. The syntax of values, expressions, and types (T) is as follows:

 $\begin{array}{l} v ::= x \mid \ast \mid \lambda z.e \mid \langle \! \langle v \rangle \rangle \\ e ::= v \mid (f x) \mid \text{LET } a = e_1 \text{ IN } e_2 \\ \mid \text{ TRY } e_1 \text{ CATCH } z.e_2 \mid \text{THROW } z \\ \mid \text{ LIFT } e \mid \text{SOME } ! z;e \mid \text{SOME } ? z;e \mid \text{NONE } ! z;e \mid e_1 \oplus e_2 \\ \mid \text{ FORK } c.e \mid \text{SEND}(c,e_1);e_2 \mid \text{RECV}(c,z);e \mid \text{CLOSE } ! c;e \mid \text{CLOSE } ? c \\ T ::= \textbf{unit} \mid A \xrightarrow{T} B \mid A \xrightarrow{0} B \mid !T.T' \mid ?T.T' \mid \text{end}_! \mid \text{end}_? \mid \oplus T \mid \& T \end{array}$ 

Types for  $\lambda^{\text{exc}}$  integrate functional and communication features. Besides the unit type, we have (linear) arrow types of two forms:  $A \xrightarrow{0} B$  is the type of "pure" functions (that do not raise exceptions), whereas  $A \xrightarrow{T} B$  is the type of functions that may raise an exception of type T. We also have session types for output and input channel-based communication, denoted !T.T' and ?T.T'. Types for labeled selection and choice are not included, but can be easily accommodated. Types end<sub>1</sub> and end<sub>2</sub> denote the dual

views of terminated endpoints. Furthermore, we have types  $\oplus T$  and & T for expressions that may produce and consume values of a type T, respectively. We use write S, S' to denote the session fragment of the type structure (i.e., no unit nor arrow types). On this fragment, we assume a duality relation, denoted  $\overline{S}$ , defined as expected. The type syntax does not include general (non-linear) functional values nor shared sessions; the integration of these constructs is orthogonal and unsurprising.

As values, we consider variables, abstractions, and the unit value \*; we also have the *abortable value*  $\langle\!\langle v \rangle\!\rangle$ , which represents discardable (affine) values: given a value v of type T, value  $\langle\!\langle v \rangle\!\rangle$  will be of type  $\oplus T$ . For convenience, the language is let-expanded; as a result, application is of the form (f x), for variables f and x. Expressions also include a try-catch construct for scoped exceptions, with the expected meaning, and a construct for raising/throwing exceptions with an explicit value. The key features of the non-deterministic process model in Section 2 appear as expressions that may produce a value of a certain type and one construct that may consume a value of a certain type. Non-deterministic choices between two expressions are also supported. Concurrency is enabled by spawning threads, using a forking construct. Moreover,  $\lambda^{\text{exc}}$  includes expressions for channel-based communication, including the possibility of creating, opening and closing a (session) channel to exchange values (including channels) of any type.

As mentioned above, the intended operational model for  $\lambda^{\text{exc}}$  is call-by-value; rather than directly giving the operational semantics for the language, we first delineate its behavior via a type system and then give its semantics indirectly, via a type respecting encoding into the basic type system introduced in Section 2.

The type system we consider here is actually a type-and-effect system in which the effect represents the type of the exception that can be raised by the typed expression. Judgments are then the form  $D \vdash^U e : T$ : under an environment D (a set of typing assignments), the expression *e* has return type *T*, while the effect type *U* is either 0 (the expression raises no exception - it is effect free) or *T* (the expected type of exceptions).

The typing rules for  $\lambda^{\text{exc}}$  are shown in Figure 4. Rule (ABS) types abstractions; it decrees that the type of the exception possibly raised by the abstraction body will be used as the effect associated to the arrow type. Rule (PRO) types abortable values  $\langle\!\langle v \rangle\!\rangle$ , as motivated earlier: it closely follows the principles of Rule  $(T \oplus \frac{x}{w})$  for session-based processes; in particular, it requires all free variables in v to be abortable (cf. the premise  $\oplus D$ ). Rule (LIFT) allows to cast an (trivially) effectual expression from a pure one.

We define three typing rules for let expressions LET  $a = e_1 \text{ IN } e_2$ ; the actual rule used depends on the exceptions possibly raised by its constituent expressions  $e_1$  and  $e_2$ . Of the three rules, Rule (LET1) is arguably the most interesting one: both  $e_1$  and  $e_2$ are effectful and can raise exceptions. Observe that  $e_2$  must be typable in an abortable environment, in order to safely account for an exception raised in  $e_1$ . Rule (LET2) handles the case in which both  $e_1$  and  $e_2$  are pure, while Rule (LET3) covers the case in which only  $e_1$  is pure. Notice that the three forms of the (LET) typing rules are crucial to isolate effects and exploit the intermingled combination of linear effect-free computations with effectful computations.

Rule (TRY) types the construct TRY  $e_1$  CATCH x.  $e_2$ ; the type of the exception possibly raised by  $e_1$  must match with the type of x in  $e_2$ . Notice that  $e_1$  and  $e_2$  must be of the same type (T in the rule). Rule (THROW) ensures that the type of the thrown

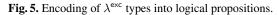
$$\begin{array}{c} \begin{array}{c} \operatorname{VAR} \\ \overline{x:T\vdash^{U}x:T} \\ \end{array} \\ \begin{array}{c} \operatorname{VAR} \\ \overline{x:T\vdash^{U}x:T} \\ \end{array} \\ \begin{array}{c} \operatorname{D}, x:A\vdash^{U}e:B \\ \overline{D\vdash^{U'}\lambda_{x,e}:A\stackrel{U}{\rightarrow}B} \\ \end{array} \\ \begin{array}{c} \operatorname{D}, x:A\vdash^{U}e:B \\ \overline{D\vdash^{U'}\lambda_{x,e}:A\stackrel{U}{\rightarrow}B} \\ \end{array} \\ \begin{array}{c} \operatorname{D}, D'\vdash^{U}(fa):B \\ \end{array} \\ \begin{array}{c} \operatorname{D}, D'\vdash^{U}(fa):B \\ \end{array} \\ \begin{array}{c} \operatorname{Lift} \\ \overline{D\vdash^{U}*:unit} \\ \end{array} \\ \begin{array}{c} \operatorname{PRO} \\ \overline{\oplus D\vdash^{0}v:T} \\ \overline{\oplus D\vdash^{0}\langle v\rangle} \\ \end{array} \\ \begin{array}{c} \operatorname{Lift} \\ \overline{D\vdash^{U}e:B} \\ \end{array} \\ \begin{array}{c} \operatorname{Lift} \\ \overline{D\vdash^{U}*:unit} \\ \end{array} \\ \begin{array}{c} \operatorname{Dides} \\ \overline{\oplus D\vdash^{0}\langle v\rangle} \\ \overline{\oplus D\vdash^{0}\langle v\rangle} \\ \end{array} \\ \begin{array}{c} \operatorname{Lift} \\ \overline{D\vdash^{U}e:B} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \operatorname{Lift} \\ \overline{D\vdash^{U}e:B} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \operatorname{Lift} \\ \overline{D\vdash^{U}e:B} \\ \end{array} \\ \begin{array}{c} \operatorname{Lift} \\ \overline{D\vdash^{U}e:B} \end{array} \\ \begin{array}{c} \operatorname{Lift} \\ \overline{D\vdash^{U}e:B} \end{array} \\ \begin{array}{c} \operatorname{Lift} \\ \overline{D\vdash^{U}e:B} \end{array} \\ \end{array} \\ \begin{array}{c} \operatorname{Lift} \\ \overline{D\vdash^{U}e:B} \end{array} \\ \begin{array}{c} \operatorname{Lift} \\ \overline{D\vdash^{U}e:B} \end{array} \\ \end{array} \\ \begin{array}{c} \operatorname{Lift} \\ \overline{D\vdash^{U}e:B} \end{array} \\ \end{array} \\ \begin{array}{c} \operatorname{Lift} \\ \overline{D\vdash^{U}e:B} \end{array} \\ \end{array} \\ \begin{array}{c} \operatorname{Lift} \\ \overline{D\vdash^{U}e:B} \end{array} \\$$

**Fig. 4.** Typing rules for  $\lambda^{\text{exc}}$ .

value is propagated as an effect. Rule (FORK) captures the essence of thread spawning for communication types, creating a new (linear) session channel where one endpoint is handed to the thread body and the other endpoint returned by the fork operation. Rules (CLOSE1) and (CLOSE1) type session channel closing operations; Rules (SEND) and (RECV) type operations for sending and receiving values along session channels.

Finally, Rule (SOME1) types the production of a non-deterministic value as z, whereas Rule (SOME2) types the consumption of a non-deterministic value as z. In particular, this rule applies to expressions that do not return values, since that may simply be discarded (abort) (but that may interact with expressions that do return values via channel-based communication). Notice the similarities between Rules (SOME1) and (SOME2) (for functional expressions) and Rules  $(T \otimes_d^x)$  and  $(T \oplus_w^x)$  (for process terms), respectively. In this line, Rule (NONE) can be seen as the analogue of Rule  $(T \otimes_x^x)$  but for abortable expressions in our functional language. Rule (NONE) enables the non-

$$\begin{bmatrix} \mathbf{unit} \end{bmatrix} = \mathbf{1}$$
$$\begin{bmatrix} A \xrightarrow{0} & B \end{bmatrix} = \left( \overline{\llbracket A \rrbracket} \otimes \left( \left( \uparrow \llbracket B \rrbracket \right) \otimes \bot \right) \right)$$
$$\begin{bmatrix} A \xrightarrow{U} & B \end{bmatrix} = \left( \overline{\llbracket A \rrbracket} \otimes \left( \left( \otimes \left( \llbracket B \rrbracket \div \overline{\llbracket U \rrbracket } \right) \right) \otimes \left( \left( \llbracket B \rrbracket \boxplus \llbracket U \rrbracket \right) \otimes \bot \right) \right)$$
$$\begin{bmatrix} !T_1.T_2 \end{bmatrix} = \llbracket T_1 \rrbracket \otimes \overline{\llbracket T_2 \rrbracket} \qquad \begin{bmatrix} ?T_1.T_2 \rrbracket = \llbracket T_1 \rrbracket \otimes \llbracket T_2 \rrbracket$$
$$\begin{bmatrix} \oplus T \rrbracket = \oplus \llbracket T \rrbracket \qquad \llbracket \& T \rrbracket = \& \llbracket T \rrbracket \qquad \llbracket \operatorname{end}_! \rrbracket = \bot \qquad \llbracket \operatorname{end}_! \rrbracket = \mathbf{1}$$



deterministic choice between two pure expressions that do not return values; this allows us to define, e.g., non-deterministic sessions.

In general, the (two-sided) typing rules in Figure 4 encompass a notion of duality, in the sense that a connective appearing in the left-hand side of the turnstile in the Figure 4 corresponds to its dual in the right-hand side of the turnstile. This intuition will be captured in our embedding of functional expressions as processes, detailed next.

*Example 4.1.* We can now return to the code snippet in Fig. 2 and give some typings using the type structure just introduced. As mentioned in the Introduction, there is a precise stage of the protocol along (dual) names res and f after which abortion is safe. In our type structure we can precisely delineate such a place. We would have:

$l:?{ t string.? t string.end}_?$	$log: !string.!string.end_!$
$f: ? \texttt{string}. \texttt{!int}. \oplus (? \texttt{string.end}_?)$	$res: !string.?int. & (!string.end_!)$

These typings require minor modifications in the code of Fig. 2: we add prefix 'SOME? f; before RECV(f, bk); and prefix SOME! res; prefix before SEND(res, book);.

**Embedding**  $\lambda^{\text{exc}}$  **Into Session Typed Processes.** We now present a typeful encoding of  $\lambda^{\text{exc}}$  into the logically motivated typed process model of Section 2, and establish its correctness (Theorem 4.2). The encoding has two main components: the encoding of (functional) types into linear logic based session types, and the encoding of  $\lambda^{\text{exc}}$  expressions into (non-deterministic) concurrent processes.

Figure 5 gives the encoding of types. We use the abbreviations:  $\uparrow \llbracket T \rrbracket \triangleq \llbracket T \rrbracket \otimes 1$  and  $\llbracket U \rrbracket \boxplus \llbracket T \rrbracket \triangleq (\llbracket U \rrbracket \otimes 1) \oplus (\llbracket T \rrbracket \otimes 1)$  and  $\& (\llbracket T \rrbracket \div \llbracket U \rrbracket) \triangleq \& (\llbracket T \rrbracket \otimes ((\llbracket U \rrbracket \boxplus \oplus \llbracket T \rrbracket) \otimes 1))$ . Also, we assume the expected extension of the encoding of types to typing environments: given  $\mathsf{D} = x_1:T_1, \cdots, x_n:T_n$  then  $\llbracket \mathsf{D} \rrbracket = x_1: \llbracket T_1 \rrbracket, \cdots, x_n: \llbracket T_n \rrbracket$ .

The encoding of expressions is typeful: for each typing rule in Figure 4 we give a corresponding type derivation for session-typed processes. Figures 6 and 7 give a complete account; for the sake of readability, in the figures we show only the conclusion (final judgment) in the derivation.

As usual in encodings of (call-by-value) functional languages into the  $\pi$ -calculus (cf. [36]), our encoding of expressions is indexed by names, which are used by to interact with the environment; they can be seen as *continuations* or as *locations* where the value returned by an expression will be made available. In our case, these names are directly related to the effects of the source expression *e*:

- If e is pure (i.e., effectless) then its encoding [e] will be indexed by a single continuation name y. This will be denoted  $[e]_y$ .

- 20 Luís Caires and Jorge A. Pérez
  - If e is effectful (i.e., may raise an exception) then its encoding [e] will be indexed by names y and x. This will be denoted  $[e]_{y,x}$ : name y represents an *non-deterministic* continuation, along which the value to which e reduces may be produced; name x represents the continuation to the enclosing try-catch block exception handler.

These intuitive distinctions between the encoding of pure (effectless) and effectful expressions are made precise in our main technical result:

**Theorem 4.2** (Typability). Suppose  $D \vdash^U e : T$ . Then, for some names y, x, we have:

- $\llbracket e \rrbracket_y \vdash \overline{\llbracket \mathsf{D} \rrbracket}, y \colon \uparrow \llbracket T \rrbracket$ , if U = 0.
- $\ \overline{[e]}_{y,x} \vdash \overline{\llbracket \mathsf{D} \rrbracket}, y : \& \left( \llbracket T \rrbracket \div \overline{\llbracket U \rrbracket} \right), x : \llbracket U \rrbracket \boxplus \llbracket T \rrbracket, \text{ if } U \neq \mathbf{0}.$

Consequently, we obtain that our source language involving higher-order functions, concurrency, non-determinism—including abortion—, and exceptions will inherit basic safety guarantees from the target process language, namely preservation and global progress (deadlock absence and lock-freedom).

In the remainder of this section, we discuss a few notable cases of the typeful encoding of expressions. To reduce notational burden, we write  $y\langle z\rangle$ . P (where z is free in P) for the free output process (represented as  $\overline{y}(w)$ .( $[w \leftrightarrow z] | P$ ) - cf. Section 2). Also, y.0; P and  $y.\overline{0}$  will be shorthand notations for y.close; P and  $y.\overline{close}$ , respectively. Moreover, we define  $S_q$  as the process q(u).q.0; u.0; 0. Notice that  $S_q \vdash q : \uparrow [unit]$ .

As mentioned above, the typing of a let expression LET  $a = e_1 \text{ IN } e_2$  considers different possibilities for the interplay of effectless and effectful computations in  $e_1$  and  $e_2$ . If both expressions are pure (cf. Rule (LET2)) then the encoding is simple:

$$[\text{LET } a = e_1 \text{ IN } e_2]_y = (\nu q)([e_1]_q \mid q(a).q.0; [e_2]_y)$$

Since  $e_1$  is pure, we know  $[e_1]_q$  will surely produce a value, which will be made available to  $[e_2]_y$  along the private (linear) name q. The case in which both  $e_1$  and  $e_2$  may raise exceptions (cf. Rule (LET1)) is more interesting:

$$[\text{LET } a = e_1 \text{ IN } e_2]_{y,x} = (\nu q)([e_1]_{q,x} \mid q.\text{some}_D; q(a).q(s).q.0; [e_2]_{y,s})$$

In this case, since  $e_1$  may raise an exception, we account for this possibility via the prefix  $q.\texttt{some}_D$ , which requires all values (including sessions) in  $[e_2]_{y,s}$  (excepting a) to be in abortable state. The production of a value within  $[e_1]_{q,x}$  will be signaled by a prefix q.some, while throwing of an exception will be signaled by a prefix q.none (see next). Therefore, if  $[e_1]_{q,x}$  produces a value (q.some is executed) then this value will be passed to  $[e_2]_{q,s}$  using the private name q; subsequently, the reference to the enclosing try-block x will also be passed to  $[e_2]_{q,s}$  as parameter s, exploiting linearity of name-passing (delegation). Otherwise, if  $[e_1]_{q,x}$  ever raises an exception (q.none is executed) then all the values in D will be safely and hereditarily discarded.

The encoding of values takes in the account that a value may occur in an abortable context. The encoding of a variable z of type T is as follows:

$$[z]_y = y \langle z \rangle. y.\overline{0} \qquad [z]_{y,x} = y.\overline{\texttt{some}}; y \langle z \rangle. y \langle x \rangle. y.\overline{0}$$

If the variable z occurs in an effectless context/expression, then its encoding, given on the left, is standard; name y will have type  $\uparrow [[T]]$ . Otherwise, if z occurs in an

effectful (abortable) context, then its encoding, given on the right, first announces the production of a value using prefix  $y.\overline{\texttt{some}}$ ; after z is sent along y, name x (representing the continuation of the enclosing try-catch block exception handler) will be sent along y. The type of x will be  $\llbracket U \rrbracket \boxplus \llbracket T \rrbracket$ , where U is the type of the enclosing exception. Thus, intuitively, x encompasses the potential for a normal execution ( $\llbracket T \rrbracket$ ) but also contains information on the (exceptional) behavior to be triggered upon abortion ( $\llbracket U \rrbracket$ ). A more concrete justification for the typing  $x: \llbracket U \rrbracket \boxplus \llbracket T \rrbracket$  will become apparent when discussing the deterministic choice that underlies the encoding of try-catch and throw expressions.

To encode an abstraction  $\lambda z.e$ , we distinguish several cases, depending on whether e and  $\lambda z.e$  are effectful or not. The simplest case is when both e and  $\lambda z.e$  are pure:

$$[\lambda z.e]_{y} = \overline{y}(f).(y.\overline{0} \mid f(z).f(k).f.0; [e]_{k})$$

We follow closely known encodings of  $\lambda$ -calculus in the  $\pi$ -calculus, here adapted to a linear setting in which the continuation y and the reference to the function body f are session-typed [39]. When both  $\lambda z.e$  and e are effectful we follow a similar principle:

$$[\lambda z.e]_{y,x} = y.\overline{\mathtt{some}}; \overline{y}(f).(y\langle x \rangle.y.\overline{\mathtt{0}} \mid f(z).f(k).f(j).f.\mathtt{0}; [e]_{k,j})$$

The prefix  $y.\overline{\text{some}}$  declares the production of a value, namely the reference to the function body f. An invocation to f must supply the parameter of the function (z) but also the two continuations k and j, to be linearly used by the encoding  $[e]_{k,j}$ .

The encoding of applications goes hand in hand with the encoding of let expressions. Given the let-expanded semantics (which forces an expression's context to deal with potentially abortable expressions), the encoding of applications (f a) is simple:

$$[(f a)]_{y,x} = f\langle a \rangle . f\langle y \rangle . f\langle x \rangle . f.\overline{0}$$

We may now discuss the encodings of try-catch and throw expressions:

$$[\operatorname{TRY} e_1 \operatorname{CATCH} z. e_2]_y = (\nu j)((\nu k)([e_1]_{k,j} \mid \underbrace{k.\operatorname{some}_{\emptyset}; k(u).k(z).k.0; z.\operatorname{inl}; z\langle u \rangle. z.\overline{0})}_{(I)})$$

$$\underbrace{j.\operatorname{case}(j(u).j.0; y\langle u \rangle. y.\overline{0}, j(z).j.0; [e_2]_y)}_{(II)})$$

 $[\texttt{THROW}\, z]_{y,x} = y.\overline{\texttt{none}} \mid x.\texttt{inr}; x\langle z\rangle. x.\overline{\texttt{0}}$ 

Intuitively, the encoding of TRY  $e_1$  CATCH z.  $e_2$  is in two parts, denoted (I) and (II) above. Part (I) concerns normal behaviors only; Part (II) accounts for both normal and exceptional behaviors, and specifies what to do in either case:

- Suppose that  $e_1$  does not raise an exception: this means that  $[e_1]_{k,j}$  will execute a process  $k.\overline{\texttt{some}}$  which will synchronize with Part (I). Subsequently, the obtained value and the reference to the enclosing exception block will be passed around; in this case, z will be substituted by j, and the prefix j.inl will synchronize with the choice on j (Part (II)) to send the resulting value along y. This (deterministic) choice discards the right branch, which contains  $[e_2]_y$ .

- Suppose now that  $e_1$  raises an exception: because of the encoding of throw, this means that  $[e_1]_{k,j}$  will execute a process k.<u>none</u> which will synchronize with Part (I). As a result, the remaining behavior on k and z will be discarded. However, the choice on j (Part (II)) will continue to be available: this is used by the encoding of throw, which by executing j.inr will select the right branch of Part (II). The value raised by the exception will be then passed to  $[e_2]_u$ , which can now be executed.

Our encoding of try-catch therefore elegantly amalgamates the key features of our process model: most notably, the presence of abortable behaviors in a pleasant coexistence with non-abortable behaviors, and the interplay between non-deterministic and deterministic choices—indeed, it is the deterministic choice that underlies the exception mechanism what ultimately justifies the type  $\llbracket U \rrbracket \boxplus \llbracket T \rrbracket$  for x, mentioned above.

In the typed model presented here (and its encoding into processes), we consider try-catch constructs TRY  $e_1$  CATCH z.  $e_2$  in which  $e_2$  is pure (cf. Fig. 4). However, there is no fundamental obstacle to address the general case in which both  $e_1$  and  $e_2$  may raise exceptions; the encoding given in Fig. 7 can be extended following expected lines.

Constructs for non-deterministic behaviors have fairly straightforward encodings:

$$\begin{split} & [\texttt{SOME}\,!\,z;e]_{y,x} = z.\overline{\texttt{some}};[e]_{y,x} & [\texttt{NONE}\,!\,z;e]_{y,x} = z.\overline{\texttt{none}} \mid [e]_{y,x} \\ & [\texttt{SOME}\,?\,z;e]_{y,x} = z.\mathtt{some}_D; (\nu q)([e]_q \mid \mathsf{S}_q) \mid \overline{y}(v).(v.\overline{\mathsf{O}} \mid y.\overline{\mathsf{O}}) \end{split}$$

In [SOME ?  $z; e]_{y,x}$ , notice that typing ensures that e does not return a value; also, set D enables to safely discard behaviors in e in the event of an exception. Given the conditions ensured by typing, the encoding of non-deterministic choices is unsurprising:

$$[e_1 \oplus e_2]_y = ((\nu z)([e_1]_z \mid \mathsf{S}_z) \oplus (\nu z)([e_2]_z \mid \mathsf{S}_z)) \mid [*]_y$$

In essence, processes  $S_z$  consume the (unit) value produced by  $[e_1]_y$  and  $[e_2]_y$  through z. The resulting processes can then be composed first in a non-deterministic choice, and then in an independent parallel composition with  $[*]_y$ . It would not be hard to extend this encoding to handle the general case in which  $e_1$  and  $e_2$  may raise exceptions and return values different from **unit**. To that end, typing should ensure that  $e_1$  and  $e_2$  are each typable in an abortable context (cf. Rule (T&)), but also that the name representing the continuation to the enclosing exception handlers (i.e., x) is given an abortable type.

The encoding of constructs for channel-based communication and forking expressions is also as expected, given our target language. We have, for instance:

$$\begin{split} [\texttt{SEND}(c,a);e]_{y,x} &= c\langle a\rangle.[e]_{y,x} & [\texttt{RECV}(c,z);e]_{y,x} = c(z).[e]_{y,x} \\ [\texttt{FORK}\,x.e]_y &= \overline{y}(x).((\nu q)([e]_q \mid \mathsf{S}_q) \mid y.\overline{0}) \end{split}$$

### **5** Further Related Work

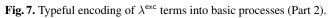
In the purely functional (and sequential) programming setting, control operators have been given Curry-Howard interpretations in the context of classical logic [22,32,2]. To our best knowledge, this paper presents the first attempt at tackling state-aware concurrent programming features, involving linearity (our main focus herein), while building

$$\begin{bmatrix} \frac{U}{z:T \vdash^{0} z:T} \end{bmatrix} y = \frac{1}{y(z).y.\overline{0} \vdash z:\overline{[T]}, y: \uparrow [T]} \\ \begin{bmatrix} \frac{U \neq 0}{z:T \vdash^{U} z:T} \end{bmatrix} y, x = \frac{1}{y.\overline{some}; y(z).y(x).y.\overline{0} \vdash z:\overline{[T]}, y: \& ([T]] \div \overline{[U]}), x: [U] \boxplus [T]} \\ \begin{bmatrix} D, z: A \vdash^{0} e: B \\ \overline{D} \vdash^{0} \lambda z. e: A \xrightarrow{0} B \end{bmatrix} y = \frac{1}{\overline{y}(f).(y.\overline{0} \mid f(z).f(k).f.0; [e]_k) \vdash \overline{[D]}, y: \uparrow [A \xrightarrow{0} B]} \\ \begin{bmatrix} D, z: A \vdash^{U} e: B \\ \overline{D} \vdash^{0} \lambda z. e: A \xrightarrow{U} B \end{bmatrix} y = \overline{y(f).(y.\overline{0} \mid f(z).f(k).f.0; [e]_{k,j}) \vdash \overline{[D]}, y: \uparrow [A \xrightarrow{U} B]} \\ \begin{bmatrix} D, z: A \vdash^{U} e: B \\ \overline{D} \vdash^{V} \lambda z. e: A \xrightarrow{U} B \end{bmatrix} y, x = \overline{y.\overline{some}; \overline{y}(f).(y(x).y.\overline{0} \mid f(z).f(k).f(j).f.0; [e]_{k,j})} \\ \vdash \overline{[D]}, y: \& ([A \xrightarrow{U} B] \Rightarrow \overline{[V]}), x: [V] \boxplus [A \xrightarrow{U} B] \end{bmatrix} \\ \begin{bmatrix} D, z: A \vdash^{U} e: B \\ \overline{D} \vdash^{V} \lambda z. e: A \xrightarrow{U} B \end{bmatrix} y, x = \overline{y.\overline{some}; \overline{y}(f).(y(x).y.\overline{0} \mid f(z).f(k).f(j).f.0; [e]_{k,j})} \\ \vdash \overline{[D]}, y: \& ([A \xrightarrow{U} B] \Rightarrow \overline{[V]}), x: [V] \boxplus [A \xrightarrow{U} B] \end{bmatrix} \\ \begin{bmatrix} D \downarrow^{U} f: A \xrightarrow{T} B \quad D' \vdash^{V} a: A \\ \overline{D}, D' \vdash^{T} (fa) : B \end{bmatrix} y, x = \overline{f(a).f(y).f(x).f.\overline{0}} + \overline{f:[A \xrightarrow{T} B]}, y: \uparrow [B], a: \overline{[A]} \end{bmatrix} \\ \begin{bmatrix} D \vdash^{U} f: A \stackrel{O}{\to} B \quad D' \vdash^{V} a: A \\ \overline{D}, D' \vdash^{T} (fa) : B \end{bmatrix} \end{bmatrix} y = \overline{f(a).f(y).f(x).f.\overline{0} \vdash f: \overline{[A \xrightarrow{O} B]}, y: \uparrow [B], a: \overline{[A]}} \\ \\ \begin{bmatrix} D \vdash^{U} f: A \quad D_{2}, a: A \vdash^{0} e_{2} : B \\ \overline{D_{1}, D_{2}} \vdash^{0} \text{LET} a = e_{1} \text{IN} e_{2} : B \end{bmatrix} y = \overline{(\nu q)([e_{1}]_{q} \mid q(a).q.0; [e_{2}]_{y,z})} + \overline{[D_{1}], \overline{[D_{2}]}, y: \uparrow [B]} \\ \\ \begin{bmatrix} D_{1} \vdash^{0} e_{1}: A \quad D_{2}, a: A \vdash^{T} e_{2}: B \\ \overline{D_{1}, D_{2}} +^{T} \text{LET} a = e_{1} \text{IN} e_{2} : B \end{bmatrix} y, x = \overline{(\nu q)([e_{1}]_{q} \mid q(a).q.0; [e_{2}]_{y,z})} + \overline{[D_{1}], x: [T]} \boxplus [B] \\ \\ \\ \begin{bmatrix} D_{1} \vdash^{T} e_{1}: A \quad D_{2}, a: A \vdash^{T} e_{2}: B \\ D_{1}, \oplus D_{2} \vdash^{T} \text{LET} a = e_{1} \text{IN} e_{2} : B \end{bmatrix} y, x = \overline{(\nu q)([e_{1}]_{q}, u], a: (D_{2}], y: \& ([B] \div [T]), x: [T] \boxplus B [B] \\ \\ \\ \begin{bmatrix} D \vdash^{1} e: B \\ D_{1} \oplus D_{2} \vdash^{T} \text{LET} a = e_{1} \text{IN} e_{2} : B \end{bmatrix} y, x = \overline{(\nu q)([e_{1}]_{q}, u], a: (D_{2}], y: \& ([B] \div [T]), x: [T] \boxplus B [B] \\ \\ \\ \begin{bmatrix} D \vdash^{1} u \text{IF} e: B \\ D \vdash^{1} u \text{IF} e: B \end{bmatrix} y, x = \overline{(\nu q)([e]_{q} \mid q(v).q.0; y.ssome}; y(v).y(x), y, \overline{0})} \\ \vdash \overline{[D], y: \& ([T] \div [B]), x: [T] \boxplus B [B] \\ \\ \\ \begin{bmatrix} \oplus D \vdash^{0} (v): T \\$$

**Fig. 6.** Typeful encoding of  $\lambda^{\text{exc}}$  terms into basic processes (Part 1).

on a Curry-Howard interpretation of classical linear logic as session types. A very tentative sketch of some ideas behind this work was presented at Cardelli's Fest [6]; here we provide a complete account of non-determinism and abortion, introduce new computational primitives, present associated results, and provide non-trivial examples, including the typeful embedding of a realistic functional, concurrent language with exceptions.

The tensions between affinity, linearity and control effects have been widely investigated in different settings, and already referred in the Introduction. The work [41] con-



siders a form of affinity in stateful settings (including session types) and explores how to safely interface an affine language with a conventional one. We share several high level aims with [41], although following a fundamentally different approach, and obtaining results of different relevance; in particular, we consider a unified (concurrent) language that admits a fundamental Curry-Howard correspondence with linear logic, and offers

25

strong guarantees by static typing such as deadlock-freedom. Within the session types literature, the interplay of session types and functional languages (including encodings of functional calculi) has received much attention (see, e.g., [21,44,31,28]) but non-determinism/abortion do not seem to have been addressed. The paper [31] relates effect and session type systems, but effects such as exceptions are not addressed. A work exploring affinity in session calculi is [30]. Existing works on exception mechanisms for session types impose severe syntactic restrictions to typable programs and/or do not ensure progress: this observation applies to models of interactional exceptions and interruptible sessions based on both binary sessions (cf. [12,17]). Further work is required to connect our process model (based on binary session types) with multiparty structured interactions with exceptions/interruptions, following logic-based relationships between binary and multiparty session types [7].

As explained in the Introduction, a main aim of this work is not just to propose yet another point in the design space solution for exceptions, affinity, or linearity. Instead, we show how a minimal set of logically motivated primitives is expressive enough to model fairly general notions of (controlled) affinity and non-determinism in higherorder concurrent programs (including exception handling mechanisms) while preserving all the fundamental properties of a Curry-Howard interpretation for linear logic. We leave for future work a deeper study of the expressiveness of our model, as exceptions and compensations are key programming abstractions in models of service-oriented computing (see, e.g., [19] and references therein).

### 6 Concluding Remarks

We have presented the first type system that accommodates non-deterministic and abortable behaviors within session-based concurrent programs while building on a Curry-Howard correspondence with linear logic. Conceptually simple, our approach conservatively extends classical linear logic with two dual modal connectives, related to linear logic exponentials, but that express non-determinism and abortion rather than sharing.

We have shown that our type system enforces progress and session fidelity; its underlying operational semantics, based on Curry-Howard principles, is actually compatible with standard non-confluent formulations of internal non-determinism for process algebra, in the sense of our postponing result (Theorem 3.5). Our system is very expressive, as illustrated by several examples, including a typed embedding of a higher-order linear functional language with threads, sessions, non-determinism, and exceptions.

We have not discussed the presence of intuitionistic (unrestricted) types in the functional language of Section 4, as the main focus in the paper is on linearity and its challenging combination with non-determinism and abortion. The combination of these ingredients with general (non-linear) functional values and shared sessions would not bring any surprise, resulting from the type discipline of the interpretation of the exponentials in the basic model. Expected properties of our type system that we do not formally establish in this paper, such as strong normalization and confluence, can be established along predictable lines [33]. An additional advantage of our approach is its natural compatibility with other extensions to the basic framework, for example behavioral polymorphism [8]. Another interesting direction for future work is to better understand the behavioral equivalences induced by our interpretation.

# References

- 1. J.-M. Andreoli. Logic Programming with Focusing Proofs in Linear Logic. J. Log. Comput., 2(3):297–347, 1992.
- Z. M. Ariola and H. Herbelin. Minimal Classical Logic and Control Operators. In *ICALP* 2003, pages 871–885, 2003.
- A. Barber. Dual Intuitionistic Linear Logic. Technical Report LFCS-96-347, Univ. of Edinburgh, 1996.
- P. N. Benton, G. M. Bierman, and V. de Paiva. Computational Types from a Logical Perspective. J. Funct. Program., 8(2):177–193, 1998.
- M. Boreale. On the Expressiveness of Internal Mobility in Name-Passing Calculi. *Theor. Comput. Sci.*, 195(2):205–226, 1998.
- L. Caires. Types and Logic, Concurrency and Non-Determinism. In M. Abadi, P. Gardner, A. D. Gordon, and R. Mardare, editors, *Essays for the Luca Cardelli Fest*, pages 69–83. Microsoft Research TR MSR-TR-2014-104, 2014.
- L. Caires and J. A. Pérez. Multiparty session types within a canonical binary theory, and beyond. In *FORTE 2016*, pages 74–95, 2016.
- L. Caires, J. A. Pérez, F. Pfenning, and B. Toninho. Behavioral Polymorphism and Parametricity in Session-Based Communication. In *ESOP'13*, number 7792 in LNCS, 2013.
- L. Caires and F. Pfenning. Session Types as Intuitionistic Linear Propositions. In CON-CUR'10, number 6269 in LNCS, pages 222–236, 2010.
- L. Caires, F. Pfenning, and B. Toninho. Towards concurrent type theory. In *Types in Language Design and Implementation*, pages 1–12, 2012.
- 11. L. Caires, F. Pfenning, and B. Toninho. Linear Logic Propositions as Session Types. *Math. Struct. in Comp. Sci.*, 2016.
- S. Capecchi, E. Giachino, and N. Yoshida. Global Escape in Multiparty Sessions. *Mathematical Structures in Computer Science*, 26(2):156–205, 2016.
- M. Carbone, K. Honda, and N. Yoshida. Structured Interactional Exceptions in Session Types. In *CONCUR 2008*, volume 5201 of *LNCS*, pages 402–417. Springer, 2008.
- M. Carbone, S. Lindley, F. Montesi, C. Schürmann, and P. Wadler. Coherence generalises duality: A logical explanation of multiparty session types. In *CONCUR 2016*, pages 33:1– 33:15, 2016.
- L. Cardelli. Typeful Programming. IFIP State-of-the-Art Reports: Formal Description of Programming Concepts, pages 431–507, 1991.
- R. DeLine and M. Fähndrich. Typestates for Objects. In ECOOP 2004 Object-Oriented Programming, 18th European Conference, pages 465–490, 2004.
- R. Demangeon, K. Honda, R. Hu, R. Neykova, and N. Yoshida. Practical Interruptible Conversations: Distributed Dynamic Verification with Multiparty Session Types and Python. *Formal Methods in System Design*, 46(3):197–225, 2015.
- T. Ehrhard and L. Regnier. Differential Interaction Nets. *Theor. Comput. Sci.*, 364(2):166– 195, 2006.
- C. Ferreira, I. Lanese, A. Ravara, H. T. Vieira, and G. Zavattaro. Advanced Mechanisms for Service Combination and Transactions. In *Results of the SENSORIA Project*, volume 6582 of *LNCS*, pages 302–325. Springer, 2011.
- P. Gardner, C. Laneve, and L. Wischik. Linear Forwarders. *Information and Compututation*, 205(10):1526–1550, 2007.
- S. Gay and V. T. Vasconcelos. Linear type theory for asynchronous session types. J. Funct. Programming, 20(1):19–50, 2010.
- 22. T. Griffin. A Formulae-as-Types Notion of Control. In POPL'90, pages 47-58, 1990.

- K. Honda. Types for Dyadic Interaction. In CONCUR'93, number 715 in LNCS, pages 509–523, 1993.
- K. Honda, V. T. Vasconcelos, and M. Kubo. Language Primitives and Type Discipline for Structured Communication-Based Programming. In ESOP'98, number 1381 in LNCS, 1998.
- H. Hüttel, I. Lanese, V. T. Vasconcelos, and L. C. et. al. Foundations of Session Types and Behavioural Contracts. ACM Comput. Surv., 49(1):3, 2016.
- N. Kobayashi, B. C. Pierce, and D. N. Turner. Linearity and the pi-calculus. In 23rd Symposium on Principles of Programming Languages, POPL'96, pages 358–371. ACM, 1996.
- N. R. Krishnaswami, A. Turon, D. Dreyer, and D. Garg. Superficially Substructural Types. In *ICFP'12*, pages 41–54, 2012.
- 28. S. Lindley and J. G. Morris. Embedding session types in Haskell. In 9th International Symposium on Haskell, Haskell 2016, pages 133–145, 2016.
- F. Militão, J. Aldrich, and L. Caires. Rely-Guarantee Protocols. In ECOOP 2014 Object-Oriented Programming - 28th European Conference, pages 334–359, 2014.
- D. Mostrous and V. T. Vasconcelos. Affine Sessions. In *Proc. of COORDINATION 2014*, volume 8459 of *LNCS*, pages 115–130. Springer, 2014.
- D. A. Orchard and N. Yoshida. Effects as sessions, sessions as effects. In Proc. of POPL 2016, pages 568–581. ACM, 2016.
- M. Parigot. Lambda-Mu-Calculus: An Algorithmic Interpretation of Classical Natural Deduction. In LPAR'92, pages 190–201, 1992.
- J. A. Pérez, L. Caires, F. Pfenning, and B. Toninho. Linear Logical Relations for Session-Based Concurrency. In *ESOP'12*, number 7211 in LNCS, 2012.
- F. Pfenning. Structural Cut Elimination. In 10th Annual IEEE Symposium on Logic in Computer Science, LICS'95, pages 156–166. IEEE Computer Society, 1995.
- 35. G. D. Plotkin. A Powerdomain Construction. SIAM J. Comput., 5(3):452-487, 1976.
- 36. D. Sangiorgi and D. Walker. *The*  $\pi$ -*calculus: A Theory of Mobile Processes.* Cambridge University Press, 2001.
- A. Scalas and N. Yoshida. Lightweight Session Programming in Scala. In 30th European Conference on Object-Oriented Programming, ECOOP 2016, pages 21:1–21:28, 2016.
- B. Toninho, L. Caires, and F. Pfenning. Dependent session types via intuitionistic linear type theory. In *PPDP'11*, pages 161–172, 2011.
- B. Toninho, L. Caires, and F. Pfenning. Functions as Session-Typed Processes. In *FoS-SaCS'12*, number 7213 in LNCS, 2012.
- B. Toninho, L. Caires, and F. Pfenning. Higher-order processes, functions, and sessions: A monadic integration. In *ESOP'13*, pages 350–369, 2013.
- J. A. Tov and R. Pucella. Stateful Contracts for Affine Types. In ESOP 2010, pages 550–569, 2010.
- 42. J. A. Tov and R. Pucella. A theory of Substructural Types and Control. In *OOPSLA 2011*, pages 625–642, 2011.
- 43. J. A. Tov and R. Pucella. Practical Affine Types. In POPL 2011, pages 447-458, 2011.
- 44. P. Wadler. Propositions as Sessions. In ICFP'12, ACM, pages 273-286, 2012.