

Indice

1	Introduzione	3
1.1	Reversibilità	4
1.2	Debugger reversibile	5
2	Estensione con feature imperative	7
2.1	Sintassi del linguaggio	7
2.2	Semantica del linguaggio	8
2.3	Semantica reversibile	17
2.4	Semantica rollback	17
3	Conclusione	19

Capitolo 1

Introduzione

Erlang è un linguaggio di programmazione funzionale e concorrente basato sul paradigma ad attori (concorrenza basata sul *message-passing*).

BEAM è la macchina virtuale al centro di Erlang, fa parte dell'Erlang Run-Time System (ERTS), che compila il codice sorgente di Erlang in bytecode, che viene quindi eseguito su BEAM. Core Erlang [1] è uno dei linguaggi intermedi in cui Erlang viene compilato.

In questa tesi si mostra un'estensione della semantica descritta nell'articolo "A theory of reversibility for Erlang" [2] con delle feature imperative.

Nella Sezione 2.1 viene modificato il linguaggio presentato nell'articolo [2] aggiungendo le operazioni di registrazione di atomi e pid, deregistrazione di pid e di lookup. Inoltre viene aggiunto un non terminale *end* per poter identificare quando un processo è arrivato in uno stato finale.

Nella Sezione 2.2 viene modificata la nozione di sistema in cui viene introdotta la mappa formata dalle coppie di atomi, pid che sono stati registrati data nell'articolo [2]. Inoltre sono state introdotte le regole per la semantica del linguaggio, in questo caso sono state aggiunte le regole per la valutazione delle nuove espressioni e inoltre sono state aggiunte le regole per descrivere il fallimento di un'espressione, che nell'articolo [2] non erano presenti. Inoltre sono state aggiunte le regole di sistema per le nuove espressioni e per i casi di fallimento.

Nella Sezione 2.3 viene descritta la semantica reversibile divisa in semantica *forward* e *backward*. La prima è un'estensione delle regole introdotte precedentemente in cui nell'articolo [2] viene aggiunta una storia dei processi mentre in questo lavoro è stata l'aggiunta di una storia della mappa. La semantica *backward* invece descrive come viene fatto un passo all'indietro. In questo caso viene introdotta una nozione di operazione in lettura e scrittura per generalizzare le *side-condition* delle regole inserite e per dare una nozione di operazione concorrente più generale.

Nella Sezione 2.4 è stato esteso l'operatore di reversibilità che può essere utilizzato per annullare le azioni di un dato processo fino a raggiungere un determinato punto di controllo, introdotto dal programmatore. Per garantire la coerenza causale, l'azione di rollback potrebbe essere propagata ad altri processi dipendenti.

1.1 Reversibilità

Normalmente la computazione avviene in una direzione, in un programma le istruzioni vengono eseguite in un determinato ordine, per computazione reversibile ci si riferisce alla possibilità di eseguire un programma sia in avanti (o anche computazione *forward*) che indietro (o anche computazione *backward*).

Dato un programma non è ovvio che sia possibile eseguire una computazione backward, infatti tutti i linguaggi di programmazione tradizionali consentono operazioni con perdita di informazione.

Durante l'esecuzione di un programma, a meno che non sia reversibile o senza perdita di informazioni, i dati intermedi vengono persi mentre viene calcolato l'output finale.

Esempio 1. *L'assegnamento $x = 96$ elimina il vecchio valore di x che deve essere memorizzato se si vuole annullare questa assegnazione.*

Landauer [3] notò anche che ogni computazione irreversibile può essere trasformata in una computazione reversibile, includendola in una computazione più ampia nella quale nessuna informazione viene persa, salvando ogni volta gli stati intermedi della computazione così da non avere perdita di informazione.

L'idea alla base di questo lavoro [3] è che qualsiasi linguaggio di programmazione o formalismo può essere reso reversibile aggiungendo la cronologia del calcolo a ogni stato, questo metodo è di solito chiamato incorporamento di Landauer.

Questa idea fu ulteriormente migliorata da Bennett [4] al fine di evitare la generazione di dati "spazzatura", applicando una serie di analisi al fine di limitare il più possibile le informazioni richieste nella storia.

La reversibilità in un contesto sequenziale è facile da capire, per invertire l'esecuzione di un programma sequenziale è sufficiente annullare ricorsivamente l'ultima azione eseguita dal programma. La definizione di reversibilità in un contesto in cui vengono considerati anche i sistemi distribuiti è più complicata, poiché non esiste il concetto di "ultima azione". Poiché, molte azioni vengono eseguite contemporaneamente.

Una definizione adeguata di reversibilità in uno scenario concorrente è stata proposta da Danos e Krivine nel loro articolo [5]. Intuitivamente, la definizione afferma che qualsiasi azione può essere annullata a condizione che tutte le sue eventuali conseguenze siano annullate preventivamente.

Il calcolo reversibile potrebbe essere applicato per migliorare il modo in cui alcune attività vengono risolte, per esempio nel debug.

1.2 Debugger reversibile

I Debugger non sono altro che programmi, che consentono di analizzare se un programma è sintatticamente corretto, in modo tale da effettuare una ricerca del bug in modo più veloce e accurato possibile. Grazie ai Debugger si ha la possibilità di scovare errori o malfunzionamenti all'interno del programma sfruttando funzioni specifiche per il debugging: l'attività che consiste proprio nell'individuazione della porzione di software affetta da bug.

Generalmente il tipo di debugger più utilizzato è il debugger a runtime che consente la ricerca dei bug tramite funzionalità standard come *breakpoint*, *controlpoint* e *viste da watch*, che consentono al programmatore di analizzare più accuratamente determinate parti di codice per identificare gli errori con una scansione in avanti del codice.

Di contro però, la maggior parte dei debugger fornisce infatti, un'assistenza limitata all'esecuzione in avanti del codice per la navigazione temporale, per cui i programmatori devono spesso ricorrere alla simulazione dell'esecuzione del programma mentalmente per cercare di immaginare i flussi di istruzioni che vengono eseguiti. Infatti, con un debugger a *runtime*, il programmatore può cercare il bug inserendo dei breakpoint all'interno dei thread e sperare di scovare il bug facendo dei tentativi, magari entrando in watch per osservare come vengono modificate le variabili durante l'esecuzione del programma. A tal proposito sono nati anche debug più complessi i cosiddetti: debugger reversibili che a differenza dei debugger tradizionali, quest'ultimi consentono agli sviluppatori di registrare le attività del programma in esecuzione, per poi riavvolgere e riprodurre tali istruzioni, compresi eventuali errori, per ispezionare lo stato del programma, questi debugger sono molto utili negli scenari concorrenti [7].

Riavvolgere le azioni di un particolare processo significa anche annullare ogni conseguenza dell'operazione (cioè tutte le azioni "collegate" a quell'operazione) che si vuole annullare. Questa nozione è chiamata *consistenza causale* [5].

Capitolo 2

Estensione con feature imperative

Nel seguito di questo capitolo, verrà presentata la sintassi del linguaggio esteso supportato, seguita dalla semantica (reversibile). In contrasto con la semantica descritta nell'articolo "A theory of reversibility for Erlang" [2] quella che verrà presentata qui includerà funzioni imperative integrate. Le modifiche effettuate all'articolo [2] sono evidenziate in giallo.

2.1 Sintassi del linguaggio

In questa sezione, verrà mostrata la sintassi di un linguaggio funzionale di prim'ordine, concorrente e distribuito basato sul paradigma ad attori utilizzato. Questo linguaggio è un sotto insieme di Core Erlang [1], che è uno dei linguaggi intermedi in cui un programma Erlang viene compilato.

Nella Figura 2.1 viene presentata la sintassi del linguaggio, si può notare che vengono considerate solo espressioni di prim'ordine quindi il primo argomento delle applicazioni di funzioni e della spawn è un nome di funzione (invece che un'espressione o chiusura arbitraria) e il primo argomento nelle chiamate è un'operazione built-in *Op*. Rispetto alla sintassi precedente sono state aggiunte le funzioni built-in:

- **whereis** che dato in input un atomo restituisce il pid associato, se non è registrato l'atomo **undefined**,
- **registered** che restituisce una lista di tutti gli atomi nella mappa, se non sono presenti atomi registrati restituisce una lista vuota,

sono anche state aggiunte le funzioni:

$$\begin{aligned}
\text{module} &::= \text{module } Atom = fun_1 \dots fun_n \\
\text{fun} &::= \text{fname} = \text{fun } (Var_1, \dots, Var_n) \rightarrow \text{expr} \\
\text{fname} &::= Atom / Integer \\
\text{lit} &::= Atom \mid Integer \mid Float \mid Pid \mid [] \\
\text{expr} &::= Var \mid \text{fname} \mid [expr_1 \mid expr_2] \mid \{expr_1, \dots, expr_n\} \\
&\mid \text{call } Op \ (expr_1, \dots, expr_n) \mid \text{apply } \text{fname} \ (expr_1, \dots, expr_n) \\
&\mid \text{case } expr \text{ of } clause_1; \dots; clause_m \text{ end} \\
&\mid \text{let } Var = expr_1 \text{ in } expr_2 \mid \text{receive } clause_1; \dots; clause_n \text{ end} \\
&\mid \text{spawn}(\text{fname}, [expr_1, \dots, expr_n]) \mid expr ! expr \mid \text{self}() \\
&\mid \text{register}(expr, expr) \mid \text{unregister}(expr) \mid \text{end} \\
\text{clause} &::= pat \text{ when } expr_1 \rightarrow expr_2 \\
pat &::= Var \mid lit \mid [pat_1 \mid pat_2] \mid \{pat_1, \dots, pat_n\} \\
Op &::= \dots \mid \text{whereis} \mid \text{registered} \\
\text{end} &::= lit \mid [end_1 \mid end_2] \mid \{end_1, \dots, end_n\}
\end{aligned}$$

Figura 2.1: Regole di sintassi del linguaggio

- **register** che dati in input un atomo e un pid inserisce la in una mappa la coppia atomo,pid e restituisce l'atomo **true**. Altrimenti se l'atomo o il pid sono già presenti nella mappa il processo fallisce.
- **unregister** che dato in input un atomo toglie dalla mappa la coppia atomo,pid e restituisce l'atomo **true**. Altrimenti se l'atomo non è presente nella mappa il processo fallisce.

Inoltre è stato creato il non terminale *end* che per riuscire a determinare quando un processo è in uno stato finale, così è possibile determinare se un processo è fallito e in questo modo si è riusciti ad avere un comportamento il più fedele possibile a quello di Erlang nel caso delle funzioni aggiunte. Nelle regole che verranno descritte successivamente il simbolo ϵ rappresenterà un non terminale *end*.

2.2 Semantica del linguaggio

In questa sezione descriveremo formalmente la semantica del linguaggio presentato nella sezione 2.1.

Definizione 2 (Processo). Un processo è indicato da una tupla $\langle p, (\theta, e), q \rangle$ dove p è il pid del processo (ed è unico), (θ, e) è il controllo, che consiste di un ambiente (una sostituzione) e di un'espressione da valutare, e q è la casella

di posta del processo, una coda FIFO con la sequenza di messaggi che sono stati inviati al processo.

Consideriamo le seguenti operazioni sulle cassette postali locali. Dato un messaggio v e una casella di posta locale q , $v : q$ denota una nuova casella di posta con il messaggio v sopra (cioè v è il messaggio più recente). Indichiamo anche con $q \backslash v$ una nuova coda che risulta da q rimuovendo l'occorrenza più vecchia del messaggio v (che non è necessariamente il messaggio più vecchio nel coda). \square

Un *sistema* in esecuzione può quindi essere visto come un insieme di processi, che definiamo formalmente come segue:

Definizione 3 (Sistema). Un sistema è indicato da $\Gamma; \Pi; \mathbf{M}$, dove Γ , la *casella di posta globale*, è un insieme di coppie nella forma (*destinazione, messaggio*) e Π è un insieme di processi, indicato da un'espressione della forma

$$\langle p_1, (\theta_1, e_1), q_1 \rangle \mid \cdots \mid \langle p_n, (\theta_n, e_n), q_n \rangle$$

dove “ \mid ” denota un operatore associativo e commutativo. Data una casella di posta globale Γ , $\Gamma \cup \{(p, v)\}$ denota una nuova casella di posta che include anche la coppia (p, v) , usiamo “ \cup ” come unione multiset. Spesso denotiamo un sistema con un'espressione nella forma $\Gamma; \langle p, (\theta, e), q \rangle \mid \Pi; \mathbf{M}$ per far notare che $\langle p, (\theta, e), q \rangle$ è un processo arbitrario del pool (grazie al fatto che “ \mid ” è associativo e commutativo). Infine \mathbf{M} rappresenta l'insieme delle coppie (*atomo, pid*) registrate, indicheremo con “ \cup ” l'aggiunta di una coppia alla mappa e con “ \backslash ” la rimozione di una coppia dalla mappa. \square

Intuitivamente, Γ memorizza i messaggi dopo che sono stati inviati e prima che vengano inseriti nella casella di destinazione, quindi rappresenta i messaggi che si trovano nella rete.

La semantica è definita tramite due relazione di transizione: \longrightarrow per le espressioni e \hookrightarrow per il sistema. Mostreremo prima la relazione di transizione etichettata

$$\longrightarrow : (Env, Exp) \times Label \times (Env, Exp)$$

dove *Env* e *Exp* rappresentano rispettivamente l'ambiente (cioè le sostituzioni) e le espressioni mentre *Label* denota un elemento dell'insieme

$$\{\tau, \text{send}(v_1, v_2), \text{rec}(\kappa, \overline{cl_n}), \text{spawn}(\kappa, a/n, [\overline{v_n}]), \text{self}(\kappa), \text{register}(\kappa, a, p), \text{unregister}(\kappa, a), \tau\text{MM}', \perp\}$$

Verrà utilizzato ℓ per indicare un'etichetta fra quelle appena indicate.

Per chiarezza le regole di transizione della semantica verranno divise in quattro insiemi: quelle per le espressioni sequenziali sono raffigurate nella Figura 2.2, quelle per i fallimenti delle espressioni sequenziali che si trovano nella

Figura 2.3, quelle per le espressioni concorrenti sono nella Figura 2.4 e infine quelle per i fallimenti delle espressioni concorrenti che sono nella Figura 2.5.

$$\begin{array}{c}
\text{(Var)} \frac{}{\theta, X \xrightarrow{\tau} \theta, \theta(X)} \quad \text{(Tuple)} \frac{\theta, e_i \xrightarrow{\ell} \theta', e'_i}{\theta, \{\overline{v_{1,i-1}}, e_i, \overline{e_{i+1,n}}\} \xrightarrow{\ell} \theta', \{\overline{v_{1,i-1}}, e'_i, \overline{e_{i+1,n}}\}} \\
\text{(List1)} \frac{\theta, e_1 \xrightarrow{\ell} \theta', e'_1}{\theta, [e_1|e_2] \xrightarrow{\ell} \theta', [e'_1|e_2]} \quad \text{(List2)} \frac{\theta, e_2 \xrightarrow{\ell} \theta', e'_2}{\theta, [v_1|e_2] \xrightarrow{\ell} \theta', [v_1|e'_2]} \\
\text{(Let1)} \frac{\theta, e_1 \xrightarrow{\ell} \theta', e'_1}{\theta, \text{let } X = e_1 \text{ in } e_2 \xrightarrow{\ell} \theta', \text{let } X = e'_1 \text{ in } e_2} \quad \text{(Let2)} \frac{}{\theta, \text{let } X = v \text{ in } e \xrightarrow{\tau} \theta[X \mapsto v], e} \\
\text{(Case1)} \frac{\theta, e \xrightarrow{\ell} \theta', e'}{\theta, \text{case } e \text{ of } cl_1; \dots; cl_n \text{ end} \xrightarrow{\ell} \theta', \text{case } e' \text{ of } cl_1; \dots; cl_n \text{ end}} \\
\text{(Case2)} \frac{\text{match}(\theta, v, cl_1, \dots, cl_n) = \langle \theta_i, e_i \rangle}{\theta, \text{case } v \text{ of } cl_1; \dots; cl_n \text{ end} \xrightarrow{\tau} \theta \theta_i, e_i} \\
\text{(Call1)} \frac{\theta, e_i \xrightarrow{\ell} \theta', e'_i \quad i \in \{1, \dots, n\}}{\theta, \text{call } op \ (\overline{v_{1,i-1}}, e_i, \overline{e_{i+1,n}}) \xrightarrow{\ell} \theta', \text{call } op \ (\overline{v_{1,i-1}}, e'_i, \overline{e_{i+1,n}})} \\
\text{(Call2)} \frac{\text{eval}(op, v_1, \dots, v_n) = v}{\theta, \text{call } op \ (v_1, \dots, v_n) \xrightarrow{\tau} \theta, v} \\
\text{(Call3)} \frac{\text{evalM}(M, op, v_1, \dots, v_n) = (v, M')}{\theta, \text{call } op \ (v_1, \dots, v_n) \xrightarrow{\tau MM'} \theta, v} \\
\text{(Apply1)} \frac{\theta, e_i \xrightarrow{\ell} \theta', e'_i \quad i \in \{1, \dots, n\}}{\theta, \text{apply } a/n \ (\overline{v_{1,i-1}}, e_i, \overline{e_{i+1,n}}) \xrightarrow{\ell} \theta', \text{apply } a/n \ (\overline{v_{1,i-1}}, e'_i, \overline{e_{i+1,n}})} \\
\text{(Apply2)} \frac{\mu(a/n) = \text{fun } (X_1, \dots, X_n) \rightarrow e}{\theta, \text{apply } a/n \ (v_1, \dots, v_n) \xrightarrow{\tau} \theta \cup \{X_1 \mapsto v_1, \dots, X_n \mapsto v_n\}, e}
\end{array}$$

Figura 2.2: Semantica standard: valutazione espressioni sequenziali

Le transizioni sono etichettate con τ (una riduzione sequenziale senza side-effects) o con $\tau MM'$ che indica una riduzione che accede alla mappa senza side-effects o con l'etichetta \perp che indica la propagazione di un fallimento o con un'etichetta che identifica la riduzione di un'azione con alcuni side-effects. Le etichette sono usate nelle regole di sistema (Figure 2.6, 2.7) per determinare gli effetti collaterali associati e/o le informazioni da recuperare. Nella Figura 2.2 sono presenti le regole di transizione per la valutazione delle espressioni sequenziali.

Come in Erlang, consideriamo che l'ordine di valutazione degli argomenti in una tupla, lista, ecc. È fisso da sinistra a destra.

Per la valutazione dei case, si assume una funzione ausiliaria **match** che seleziona la prima clausola, $cl_i = (pat_i \text{ when } e'_i \rightarrow e_i)$, in modo tale che v corrisponda a pat_i , ovvero $v = \theta_i(pat_i)$, e che la guardia sia soddisfatta, cioè, $\theta\theta_i, e'_i \rightarrow^* \theta', true$. Come in Core Erlang, assumiamo che i patterns possano solo contenere nuove variabili (ma le guardie potrebbero avere variabili legate, quindi passiamo l'ambiente corrente θ alla funzione **match**).

Le funzioni possono essere definite nel programma (in questo caso sono invocate da **apply**) o essere un built-in (invocate da **call**). In quest'ultimo caso vengono valutati utilizzando la funzione ausiliaria **eval**.

Nella regola *Apply2*, si considera che la mappatura μ memorizza tutte le definizioni di funzione nel programma, cioè, mappa ogni nome di funzione a/n in una copia della sua definizione **fun** $(X_1, \dots, X_n) \rightarrow e$, dove X_1, \dots, X_n sono nuove variabili (distinte) e sono le uniche variabili che possono essere libere in e . Per quanto riguarda le applicazioni, si noti che consideriamo solo il primo ordine funzioni. Per estendere la nostra semantica a considerare anche funzioni di ordine superiore, si dovrebbe ridurre il nome della funzione a *chiusura* della forma $(\theta', \text{fun } (X_1, \dots, X_n) \rightarrow e)$.

Alle regole della semantica standard delle espressioni sequenziali è stata aggiunta la regola *Call3* utilizzata per la valutazione delle funzioni built-in che devono accedere alla mappa. Le funzioni built-in come si può vedere dalla regola *Call2* vengono valutate tramite la funzione ausiliaria **eval**, per questo motivo si è scelto di utilizzare la funzione ausiliaria **evalM** per valutare le funzioni built-in che utilizzano una mappa. Infatti per la valutazione **evalM** oltre a prendere in input i parametri prende in input pure la mappa, indicata con **M**. La funzione **evalM** e la funzione **eval** sono quindi due funzioni parziali visto che non è definito l'output per ogni input. **M'** corrisponde alla parte della mappa acceduta dalla funzione **evalM**.

$$\begin{array}{c}
\text{(TupleF)} \frac{\theta, e_i \xrightarrow{\perp} \text{fail}}{\theta, \{\overline{v_{1,i-1}}, e_i, \overline{e_{i+1,n}}\} \xrightarrow{\perp} \text{fail}} \\
\\
\text{(List1F)} \frac{\theta, e_1 \xrightarrow{\perp} \text{fail}}{\theta, [e_1|e_2] \xrightarrow{\perp} \text{fail}} \quad \text{(List2F)} \frac{\theta, e_2 \xrightarrow{\perp} \text{fail}}{\theta, [v_1|e_2] \xrightarrow{\perp} \text{fail}} \\
\\
\text{(Let1F)} \frac{\theta, e_1 \xrightarrow{\perp} \text{fail}}{\theta, \text{let } X = e_1 \text{ in } e_2 \xrightarrow{\perp} \text{fail}} \\
\\
\text{(Case1F)} \frac{\theta, e \xrightarrow{\perp} \text{fail}}{\theta, \text{case } e \text{ of } cl_1; \dots; cl_n \text{ end} \xrightarrow{\perp} \text{fail}} \\
\\
\text{(Case2F)} \frac{\text{match}(\theta, v, cl_1, \dots, cl_n) = \emptyset}{\theta, \text{case } v \text{ of } cl_1; \dots; cl_n \text{ end} \xrightarrow{\perp} \text{caseFail}} \\
\\
\text{(Call1F)} \frac{\theta, e_i \xrightarrow{\perp} \text{fail}}{\theta, \text{call } op \ (\overline{v_{1,i-1}}, e_i, \overline{e_{i+1,n}}) \xrightarrow{\perp} \text{fail}} \\
\\
\text{(Call2F)} \frac{\text{eval}(op, v_1, \dots, v_n) = \emptyset \ \wedge \ \text{evalM}(M, op, v_1, \dots, v_n) = \emptyset}{\theta, \text{call } op(v_1, \dots, v_n) \xrightarrow{\perp} \text{callFail}} \\
\\
\text{(Apply1F)} \frac{\theta, e_i \xrightarrow{\perp} \text{fail}}{\theta, \text{apply } a/n \ (\overline{v_{1,i-1}}, e_i, \overline{e_{i+1,n}}) \xrightarrow{\perp} \text{fail}} \\
\\
\text{(Apply2F)} \frac{\mu(a/n) = \emptyset}{\theta, \text{apply } a/n \ (v_1, \dots, v_n) \xrightarrow{\perp} \text{applyFail}}
\end{array}$$

Figura 2.3: Semantica standard: valutazione del fallimento delle espressioni sequenziali

Nella Figura 2.3 sono presenti le regole di transizione che gestiscono i fallimenti. La propagazione dell'errore è indicata dall'etichetta \perp , quindi se una valutazione di un'espressione fallisce allora il fallimento viene propagato. Con il simbolo \emptyset si intende:

- nel caso dei match che l'argomento v non corrisponde a nessuna clausola¹

¹Più precisamente un programma Erlang quando è tradotto nella rappresentazione intermedia Core Erlang, viene aggiunta una clausola catch-all quindi il simbolo \emptyset corrisponde al match della clausola aggiunta.

- Nei casi di eval, evalM e μ significa che le funzioni non sono definite sull'input dato.

$$\begin{array}{c}
\text{(Send1)} \frac{\theta, e_1 \xrightarrow{\ell} \theta', e'_1}{\theta, e_1 ! e_2 \xrightarrow{\ell} \theta', e'_1 ! e_2} \quad \text{(Send2)} \frac{\theta, e_2 \xrightarrow{\ell} \theta', e'_2}{\theta, v_1 ! e_2 \xrightarrow{\ell} \theta', v_1 ! e'_2} \\
\text{(Send3)} \frac{}{\theta, v_1 ! v_2 \xrightarrow{\text{send}(v_1, v_2)} \theta, v_2} \\
\text{(Receive)} \frac{}{\theta, \text{receive } cl_1; \dots; cl_n \text{ end} \xrightarrow{\text{rec}(\kappa, \overline{cl_n})} \theta, \kappa} \\
\text{(Spawn1)} \frac{\theta, e_i \xrightarrow{\ell} \theta', e'_i \quad i \in \{1, \dots, n\}}{\theta, \text{spawn}(a/n, [\overline{v_{1,i-1}}, e_i, \overline{e_{i+1,n}}]) \xrightarrow{\ell} \theta', \text{spawn}(a/n, [\overline{v_{1,i-1}}, e'_i, \overline{e_{i+1,n}}])} \\
\text{(Spawn2)} \frac{}{\theta, \text{spawn}(a/n, [\overline{v_n}]) \xrightarrow{\text{spawn}(\kappa, a/n, [\overline{v_n}])} \theta, \kappa} \\
\text{(Self)} \frac{}{\theta, \text{self}() \xrightarrow{\text{self}(\kappa)} \theta, \kappa} \\
\text{(Register1)} \frac{\theta, e_1 \xrightarrow{\ell} \theta', e'_1}{\theta, \text{register}(e_1, e_2) \xrightarrow{\ell} \theta', \text{register}(e'_1, e_2)} \\
\text{(Register2)} \frac{\theta, e_2 \xrightarrow{\ell} \theta', e'_2}{\theta, \text{register}(v_1, e_2) \xrightarrow{\ell} \theta', \text{register}(v_1, e'_2)} \\
\text{(Register3)} \frac{}{\theta, \text{register}(v_1, v_2) \xrightarrow{\text{register}(\kappa, v_1, v_2)} \theta, \kappa} \\
\text{(Unregister1)} \frac{\theta, e \xrightarrow{\ell} \theta', e'}{\theta, \text{unregister}(e) \xrightarrow{\ell} \theta', \text{unregister}(e')} \\
\text{(Unregister2)} \frac{}{\theta, \text{unregister}(v) \xrightarrow{\text{unregister}(\kappa, v)} \theta, \kappa}
\end{array}$$

Figura 2.4: Semantica standard: valutazione delle espressioni concorrenti

Consideriamo ora la valutazione di espressioni concorrenti che producono side-effects (Figura 2.4). Abbiamo le regole *Send1*, *Send2* e *Send3* per “!”. In questo caso, sappiamo *localmente* a cosa dovrebbe essere ridotta l'espressione (cioè v_2 nella regola *Send3*). Per le restanti regole, questo non è noto localmente e, quindi, restituiamo un nuovo simbolo distinto, κ , viene trattato come una variabile, in modo che nelle regole di sistema nella figu-

re 2.6, 2.7 κ leggerà al suo valore corretto:² l'espressione selezionata nella regola *Ricevive*, un pid nelle regole *Spawn* e *Self* e **true** o **false** nelle regole *Register* e *Unregister*. In questi casi, l'etichetta della transizione contiene tutte le informazioni necessarie alle regole di sistema per eseguire la valutazione a livello di sistema, incluso il simbolo κ . Questo *trucco* ci permette di mantenere separate le regole per espressioni e sistemi (cioè, la semantica mostrata nelle Figura 2.2 e 2.4 è per lo più indipendente dalle regole nelle Figure 2.6, 2.7), in contrasto con altre semantiche di Erlang, ad esempio [6], dove sono combinate in una singola relazione di transizione. Nelle regole di valutazione delle espressioni concorrenti della semantica standard sono state aggiunte le regole per la valutazione della *register* e dell'*unregister*.

$$\begin{array}{c}
\text{(Send1F)} \frac{\theta, e_1 \xrightarrow{\perp} \text{fail}}{\theta, e_1 ! e_2 \xrightarrow{\perp} \text{fail}} \quad \text{(Send2F)} \frac{\theta, e_2 \xrightarrow{\perp} \text{fail}}{\theta, v_1 ! e_2 \xrightarrow{\perp} \text{fail}} \\
\\
\text{(Spawn1F)} \frac{\theta, e_i \xrightarrow{\perp} \text{fail}}{\theta, \text{spawn}(a/n, [\overline{v_{1,i-1}}, e_i, \overline{e_{i+1,n}}]) \xrightarrow{\perp} \text{fail}} \\
\\
\text{(Register1F)} \frac{\theta, e_1 \xrightarrow{\perp} \text{fail}}{\theta, \text{register}(e_1, e_2) \xrightarrow{\perp} \text{fail}} \quad \text{(Register2F)} \frac{\theta, e_2 \xrightarrow{\perp} \text{fail}}{\theta, \text{register}(v_1, e_2) \xrightarrow{\perp} \text{fail}} \\
\\
\text{(Unregister1F)} \frac{\theta, e \xrightarrow{\perp} \text{fail}}{\theta, \text{unregister}(e) \xrightarrow{\perp} \text{fail}}
\end{array}$$

Figura 2.5: Semantica standard: valutazione del fallimento delle espressioni concorrenti

Anche nelle regole della valutazione delle espressioni concorrenti sono state aggiunte le regole che gestiscono il fallimento di un processo e che propagano il fallimento della valutazione di un'espressione.

Ora si possono presentare le regole di sistema che sono rappresentate nelle figure 2.6, 2.7 e , in tutte le regole vengono considerato un sistema generale nella forma $\Gamma; \langle p, (\theta, e), q \rangle \mid \Pi; \mathbf{M}$ dove Γ è la casella di posta globale, $\langle p, (\theta, e), q \rangle \mid \Pi$ è l'insieme dei processi che contiene almeno un processo e \mathbf{M} è la mappa dei pid registrati. Rispetto all'articolo [2] in tutte le regole è stata aggiunta \mathbf{M} visto che la definizione di sistema (Definizione 3) è cambiata.

²Nota che κ assume valori nel dominio $\text{expr} \cup \text{Pid}$, al contrario delle variabili ordinarie che possono essere associate solo a valori.

Ora verranno descritte brevemente tutte le regole di transizione del sistema.

$$\begin{array}{c}
(Seq) \frac{\theta, e \xrightarrow{\tau} \theta', e'}{\Gamma; \langle p, (\theta, e), q \rangle \mid \Pi; \mathbf{M} \hookrightarrow \Gamma; \langle p, (\theta', e'), q \rangle \mid \Pi; \mathbf{M}} \\
(Call3) \frac{\theta, e \xrightarrow{\tau MM'} \theta', e'}{\Gamma; \langle p, (\theta, e), q \rangle \mid \Pi; \mathbf{M} \hookrightarrow \Gamma; \langle p, (\theta', e'), q \rangle \mid \Pi; \mathbf{M}} \\
(Send) \frac{\theta, e \xrightarrow{\text{send}(p'', v)} \theta', e' \quad \text{isAtom}(p'') = false}{\Gamma; \langle p, (\theta, e), q \rangle \mid \Pi; \mathbf{M} \hookrightarrow \Gamma \cup (p'', v); \langle p, (\theta', e'), q \rangle \mid \Pi; \mathbf{M}} \\
(SendA) \frac{\theta, e \xrightarrow{\text{send}(a, v)} \theta', e' \quad \text{isAtom}(a) = true \quad \text{matchPid}(\mathbf{M}, a) = p''}{\Gamma; \langle p, (\theta, e), q \rangle \mid \Pi; \mathbf{M} \hookrightarrow \Gamma \cup (p'', v); \langle p, (\theta', e'), q \rangle \mid \Pi; \mathbf{M}} \\
(SendF) \frac{\theta, e \xrightarrow{\text{send}(a, v)} \theta', e' \quad \text{isAtom}(a) = true \quad \text{matchPid}(\mathbf{M}, a) = false}{\Gamma; \langle p, (\theta, e), q \rangle \mid \Pi; \mathbf{M} \hookrightarrow \Gamma; \langle p, (\theta, \text{sendFail}), q \rangle \mid \Pi; \mathbf{M}} \\
(Receive) \frac{\theta, e \xrightarrow{\text{rec}(\kappa, \overline{cl_n})} \theta', e' \quad \text{matchrec}(\theta, \overline{cl_n}, q) = (\theta_i, e_i, v)}{\Gamma; \langle p, (\theta, e), q \rangle \mid \Pi; \mathbf{M} \hookrightarrow \Gamma; \langle p, (\theta' \theta_i, e' \{ \kappa \mapsto e_i \}), q \setminus v \rangle \mid \Pi; \mathbf{M}} \\
(Spawn) \frac{\theta, e \xrightarrow{\text{spawn}(\kappa, a/n, [\overline{v_n}])} \theta', e' \quad p' \text{ is a fresh pid}}{\Gamma; \langle p, (\theta, e), q \rangle \mid \Pi; \mathbf{M} \hookrightarrow \Gamma; \langle p, (\theta', e' \{ \kappa \mapsto p' \}), q \rangle \mid \langle p', (id, \text{apply } a/n (\overline{v_n})), [] \rangle \mid \Pi; \mathbf{M}} \\
(Self) \frac{\theta, e \xrightarrow{\text{self}(\kappa)} \theta', e'}{\Gamma; \langle p, (\theta, e), q \rangle \mid \Pi; \mathbf{M} \hookrightarrow \Gamma; \langle p, (\theta', e' \{ \kappa \mapsto p \}), q \rangle \mid \Pi; \mathbf{M}} \\
(Sched) \frac{}{\Gamma \cup \{(p, v)\}; \langle p, (\theta, e), q \rangle \mid \Pi; \mathbf{M} \hookrightarrow \Gamma; \langle p, (\theta, e), v : q \rangle \mid \Pi; \mathbf{M}}
\end{array}$$

Figura 2.6: Regole standard della semantica del sistema

La regola *Seq* aggiorna semplicemente il controllo (θ, e) del processo considerato quando un'espressione sequenziale viene ridotta utilizzando le regole di espressione.

La regola *Call3* come *Seq* aggiorna semplicemente il controllo (θ, e) del processo considerato quando un'espressione sequenziale viene ridotta utilizzando le regole di espressione, però accede alla mappa, quindi utilizza l'etichetta $\tau MM'$ invece di τ .

Le regole *Send* e *SendA* aggiungono la coppia (p'', v) alla casella di posta globale Γ invece di aggiungerla alla coda del processo p'' . Ciò è necessario

per garantire che tutti i possibili intrecci di messaggi siano modellati correttamente. È stata introdotta la funzione ausiliaria **isAtom** che viene utilizzata per capire se la regola debba accedere alla mappa, l'unica differenza fra le due regole è che la *SendA* utilizza un atomo come destinatario. Si noti che e' è solitamente diverso da v poiché e può avere operatori annidati diversi. Ad esempio, se e ha la forma “**case** $p!V$ **of** $\{ \dots \}$,” allora e' sarà “**case** v **of** $\{ \dots \}$ ” con etichetta **send**(p, v).

La regola *SendF* introduce un fallimento, infatti l'espressione da valutare successivamente nel processo sarà *sendFail* un atomo che corrisponde ad un fallimento perché all'atomo che si utilizza per inviare un messaggio non è associato nessun pid che si capisce grazie alla funzione **matchPid**.

Nella regola *Receive*, usiamo la funzione ausiliaria **matchrec** per valutare un'espressione di ricezione. La differenza principale con **match** è che **matchrec** prende anche una coda q e restituisce il messaggio selezionato v . Più precisamente, la funzione **matchrec** esegue la scansione della coda q cercando il *primo* messaggio v che corrisponde a un modello dell'istruzione di ricezione. Quindi, κ viene associato all'espressione nella clausola selezionata, e_i , e l'ambiente viene esteso con la sostituzione corrispondente.

Se nessun messaggio nella coda q corrisponde ad alcuna clausola, la regola non è applicabile e il processo selezionato non può essere ridotto (cioè viene sospeso). Come nelle espressioni case, assumiamo che i pattern possano contenere solo nuove variabili.

Le regole presentate finora consentono di memorizzare i messaggi nella casella di posta globale, ma non di rimuoverne i messaggi. Questo è esattamente il compito dello scheduler, modellato dalla regola *Sched*. Questa regola sceglie in modo non deterministico una coppia (p, v) nella cassetta postale globale Γ e consegna il messaggio v al processo di destinazione p . Qui, ignoriamo deliberatamente la restrizione: “ i messaggi inviati, direttamente, tra due processi dati arrivano nello stesso ordine in cui sono stati inviati ”, poiché le attuali implementazioni lo garantiscono solo all'interno dello stesso nodo. In pratica, ignorare questa restrizione equivale a considerare che ogni processo è potenzialmente eseguito in un nodo diverso. Una definizione alternativa che garantisce questa restrizione può essere trovata in [8].

Nella figura 2.7 si può notare che sono state aggiunte anche le regole per la **register** e per l'**unregister**. Queste regole sono divise in 2 casi:

- il caso di successo,
- il caso di fallimento (che introduce come prossima espressione da valutare un atomo che descrive un fallimento).

$$\begin{array}{c}
(RegisterT) \frac{\theta, e \xrightarrow{\text{register}(\text{true}, a, p')} \theta', e' \text{ matchMapReg}(M, a, p') = \text{true}}{\Gamma; \langle p, (\theta, e), q \rangle \mid \Pi; M \hookrightarrow \Gamma; \langle p, (\theta', e'), q \rangle \mid \Pi; M \cup (a, p')} \\
\\
(RegisterF) \frac{\theta, e \xrightarrow{\text{register}(\text{false}, a, p')} \theta', e' \text{ matchMapReg}(M, a, p') = \text{false}}{\Gamma; \langle p, (\theta, e), q \rangle \mid \Pi; M \hookrightarrow \Gamma; \langle p, (\theta, \text{regFail}), q \rangle \mid \Pi; M} \\
\\
(UnregisterT) \frac{\theta, e \xrightarrow{\text{unregister}(\text{true}, a)} \theta', e' \text{ matchMapUnreg}(M, a) = p'}{\Gamma; \langle p, (\theta, e), q \rangle \mid \Pi; M \hookrightarrow \Gamma; \langle p, (\theta', e'), q \rangle \mid \Pi; M \setminus (a, p')} \\
\\
(UnregisterF) \frac{\theta, e \xrightarrow{\text{unregister}(\text{false}, a)} \theta', e' \text{ matchMapUnreg}(M, a) = \text{false}}{\Gamma; \langle p, (\theta, e), q \rangle \mid \Pi; M \hookrightarrow \Gamma; \langle p, (\theta, \text{unregFail}), q \rangle \mid \Pi; M} \\
\\
(Catch) \frac{\theta, e \xrightarrow{\perp} \text{fail}}{\Gamma; \langle p, (\theta, e), q \rangle \mid \Pi; M \hookrightarrow \Gamma; \langle p, (\theta, \text{fail}), q \rangle \mid \Pi; M} \\
\\
(End) \frac{\text{matchEnd}(M, p) = a'}{\Gamma; \langle p, (\theta, \epsilon), q \rangle \mid \Pi; M \hookrightarrow \Gamma; \langle p, (\theta, \epsilon), q \rangle \mid \Pi; M \setminus (a', p)}
\end{array}$$

Figura 2.7: Regole standard della semantica del sistema che utilizzano la mappa

Il caso di successo corrisponde alle regole *RegisterT* e *UnregisterT* che rispettivamente aggiungo la coppia (a, p') alla mappa e tolgono la coppia (a, p') dalla mappa. Mentre il caso di fallimento corrisponde alle regole *RegisterF* e *UnregisterF*.

La regola *Catch* serve per gestire il caso di fallimento di un processo. In questo caso il fallimento viene catturato al top-level.

La regola *End* serve per gestire il caso di fine di un processo. In caso di fallimento o terminazione di un processo se il processo è registrato viene deregistrato.

Questa scelta è stata fatta per essere il più possibili fedeli al comportamento di Erlang, visto che nel linguaggio utilizzato non sono presenti le eccezioni.

2.3 Semantica reversibile

2.4 Semantica rollback

Capitolo 3

Conclusione

Bibliografía

- [1] Richard Carlsson. An introduction to core Erlang. In *In Proceedings of the PLI'01 Erlang Workshop*, 2001.
- [2] Ivan Lanese, Naoki Nishida, Adrián Palacios, and Germán Vidal. A theory of reversibility for Erlang. *Journal of Logical and Algebraic Methods in Programming*, 100:71–97, November 2018.
- [3] R. Landauer. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5(3):183–191, July 1961.
- [4] C. H. Bennett. Logical reversibility of computation. *IBM Journal of Research and Development*, 17(6):525–532, November 1973.
- [5] Vincent Danos and Jean Krivine. Reversible communicating systems. In *CONCUR 2004 - Concurrency Theory*, pages 292–307. Springer Berlin Heidelberg, 2004.
- [6] R. Caballero, E. Martín-Martín, A. Riesco, and S. Tamarit. A declarative debugger for concurrent erlang programs (extended version). Technical Report SIC-15/13, Dpto. Sistemas Informáticos y Computación, Universidad Complutense de Madrid, 2013.
- [7] Engblom, J. (2012, September). A review of reverse debugging. In *System, Software, SoC and Silicon Debug Conference (S4D)*, 2012 (pp. 1-6). IEEE.
- [8] N. Nishida, A. Palacios, and G. Vidal. A reversible semantics for Erlang. In M. Hermenegildo and P. López-García, editors, *Proc. of the 26th International Symposium on Logic-Based Program Synthesis and Transformation, LOPSTR 2016*, volume 10184 of *LNCS*, pages 259–274. Springer, 2017. Preliminary version available from <https://arxiv.org/abs/1608.05521>.