Mailbox Types for Unordered Interactions¹

Presentación por Jonathan Bekenstein

Materia optativa sobre Tipos Comportamentales y Contratos

2024

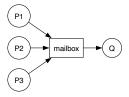


¹Ugo de'Liguoro, Luca Padovani (2018). https://arxiv.org/abs/1801.04167

Introducción

Buscamos modelar protocolos de comunicación en diferentes topologías de procesos concurrentes. La comunicación es **multi-party** y sucede mediante el uso de **mailboxes** no ordenados, en donde:

- procesos pueden escribir mensajes identificados por un tag con argumentos opcionales,
- procesos pueden consumir (leer) los mensajes en un orden arbitrario (out-of-order o selective processing).



Procesos bien tipados respetan el protocolo y no tienen deadlocks.

Sintaxis: gramática

```
Process P, Q := done (termination)
                            \mid \mathsf{X}[\overline{u}] \qquad \text{(invocation)} \\ \mid G \qquad \text{(guarded process)} \\ \mid u! \mathsf{m}[\overline{v}] \qquad \text{(stored message)} \\ \mid P \mid Q \qquad \text{(parallel composition)} \\ \mid (\nu a) P \qquad \text{(mailbox restriction)} 
  Guard G, H := fail u (runtime error)
                             free u.P (mailbox deletion)
                             |u?m(\overline{x}).P (selective receive)
                                    G + H (guard composition)
```

 \overline{u} denota la secuencia u_1, \ldots, u_n

Sintaxis: mensajes

Enviar mensajes

 $u!\mathbf{m}[\overline{v}]$

Guarda un mensaje identificado con el tag ${\bf m}$ y argumentos \overline{v} en el mailbox u.

Recibir mensajes

 $u?m(\overline{x}).P$

Consume selectivamente el mensaje con tag m del mailbox u y continúa con P reemplazando \overline{x} por los argumentos del mensaje.

Sintaxis: procesos

Invocación

 $X[\overline{u}]$ representa la invocación de un proceso llamado X con parámetros \overline{u} . Asumimos que existe una definición global de procesos de la forma $X(\overline{x}) \triangleq P$.

Paralelo y restricción

 $P \mid Q$ denota la composición paralela de procesos, y $(\nu a)P$ representa un mailbox a restringido al scope de P.

Terminación

Un proceso done representa un proceso terminado y que no realiza ninguna otra acción.

Sintaxis: guardas

Guardas

Las guardas G y la composición de guardas G+H nos permite modelar distintas "ramas" de ejecución en función del mensaje consumido del mailbox. Luego se usa exclusivamente la continuación de la guarda que consumió el mensaje.

Errores

 ${\tt fail}\ u$ permite modelar un runtime error al recibir un mensaje inesperado.

Eliminar mailbox

free u.P permite eliminar el mailbox u si ya no se va a utilizar y continuar la ejecución con P.

Semántica operacional

Reglas de reducción

```
 \begin{array}{lll} & \text{[R-READ]} & a ! \mathbf{m}[\overline{c}] \mid a ? \mathbf{m}(\overline{x}) . P + G \rightarrow P \{ \overline{c} / \overline{x} \} \\ & \text{[R-FREE]} & (\nu a) (\text{free } a . P + G) \rightarrow P \\ & \text{[R-DEF]} & \mathsf{X}[\overline{c}] \rightarrow P \{ \overline{c} / \overline{x} \} & \text{if } \mathsf{X}(\overline{x}) \triangleq P \\ & \text{[R-PAR]} & P \mid R \rightarrow Q \mid R & \text{if } P \rightarrow Q \\ & \text{[R-NEW]} & (\nu a) P \rightarrow (\nu a) Q & \text{if } P \rightarrow Q \\ & \text{[R-STRUCT]} & P \rightarrow Q & \text{if } P \equiv P' \rightarrow Q' \equiv Q \\ \end{array}
```

Relación de congruencia estructural

```
 \begin{split} \text{fail } a+G &\equiv G & G+H \equiv H+G & G+(H+H') \equiv (G+H)+H' \\ \text{done } \mid P \equiv P & P \mid Q \equiv Q \mid P & P \mid (Q \mid R) \equiv (P \mid Q) \mid R \\ & (\nu a)(\nu b)P \equiv (\nu b)(\nu a)P & (\nu a)P \mid Q \equiv (\nu a)(P \mid Q) & \text{if } a \not\in \text{fn}(Q) \end{split}
```

Ejemplo 1: Lock

```
\label{eq:free_lock} \begin{aligned} & \text{FreeLock}(self) \triangleq \texttt{free} \ self.\texttt{done} \\ & + self?\texttt{acquire}(owner). \texttt{BusyLock}[self, owner] \\ & + self?\texttt{release.fail} \ self \\ & \text{BusyLock}(self, owner) \triangleq owner!\texttt{reply}[self] \mid self?\texttt{release.FreeLock}[self] \\ & \text{User}(self, lock) \triangleq lock!\texttt{acquire}[self] \mid self?\texttt{reply}(l).(l!\texttt{release} \mid \texttt{free} \ self.\texttt{done}) \\ & (\nu lock)(\nu alice)(\nu carol)(\texttt{FreeLock}[lock] \mid \texttt{User}[alice, lock] \mid \texttt{User}[carol, lock]) \end{aligned}
```

Observaciones

- FreeLock consume de manera no determinística los mensajes acquire.
- User utiliza la referencia l para enviar el mensaje release ya que es ésta la referencia al mailbox que tiene la capabilidad de procesar este mensaje.

Caracterizaciones operacionales

Contextos de procesos

$$\mathscr{C} ::= [] \mid \mathscr{C} \mid P \mid P \mid \mathscr{C} \mid (\nu a)\mathscr{C}$$

Los contextos de procesos buscan identificar un "unguarded hole", es decir un agujero que no tiene prefijada una acción sobre un mailbox.

Def 4: Mailbox conformant

P es mailbox conformant si $P \not \to^* \mathscr{C}[\mathtt{fail}\ a]$ para todo \mathscr{C} y a. En el ejemplo del lock, ser mailbox conformant significa nunca liberar el lock antes de adquirirlo.

Caracterizaciones operacionales

Def 5: Deadlock free

P es deadlock free si $P \to^* Q \not\to \mathrm{implica}\ Q \equiv \mathtt{done}.$

Un proceso se considera *deadlock free* si al terminar tenemos que (1) no hay subprocesos esperando un mensaje que nunca se va a producir y (2) todos los mailbox están vacíos.

Def 7: Fairly terminating

P es fairly terminating si $P\to^* Q$ implica que $Q\to^* \mathrm{done}.$

Es una propiedad más fuerte que deadlock freedom. Si un proceso es fairly terminating entonces se garantiza junk freedom (no quedan mensajes sin consumir en ningún mailbox).

Los patrones son *expresiones regulares conmutativas* que describen las configuraciones válidas de los mensajes dentro de un mailbox.

- 0: *unreliable mailbox* que recibió un mensaje inesperado.
- ▶ A + B: contiene un mensaje A o un mensaje B pero no ambos.
- ▶ A + 1: contiene un mensaje A o está vacío.
- ► A · B: contiene un mensaje A y un mensaje B.
- A*: contiene un cantidad arbitraria (incluso 0) de mensajes A.

Mailbox type system Capabilities

Un mailbox type consiste en un capability (? o !) junto a un patrón. Un proceso debe cumplir ciertas obligaciones y tiene ciertas garantías descriptas por el mailbox type asociado al mailbox que usa.

- ▶ !A: el proceso **debe** escribir un mensaje A en el mailbox.
- ?A: el proceso tiene garantizado recibir un mensaje A.

Capabilities: más ejemplos

- ► !(A + 1): el proceso puede escribir un mensaje A en el mailbox, pero no está obligado a hacerlo.
- ► !(A + B): el proceso debe escribir un mensaje A o B, pero puede elegir cuál.
- ?(A + B): el proceso debe estar preparado para recibir tanto un mensaje A como B.
- ?(A · B): el proceso tiene garantizado recibir un mensaje A y otro B, y puede elegir en qué orden recibirlos.
- ▶ !(A · B): el proceso **debe** escribir un mensaje A y otro B.
- ▶ !A*: el proceso **elige** cuántos mensajes A escribir.
- ➤ ?A*: el proceso **debe** estar preparado para recibir una cantidad arbitraria de mensajes A.

Semántica de patrones

La semántica de los patrones se define como conjuntos de multiconjuntos de átomos: $m[\overline{\tau}]$.

Dada una relación preorder $\mathscr R$ sobre los tipos básicos, escribimos $E \sqsubseteq_{\mathscr R} F$ para decir que E es un subpatrón de F si $\langle \mathtt{m}_i[\overline{\tau}_i] \rangle_{i \in I} \in \llbracket E \rrbracket$ implica $\langle \mathtt{m}_i[\overline{\sigma}_i] \rangle_{i \in I} \in \llbracket F \rrbracket$ y además $\overline{\tau}_i \mathscr R \ \overline{\sigma}_i$ para todo $i \in I$.

Escribimos $\simeq_{\mathscr{R}}$ para denotar $\sqsubseteq_{\mathscr{R}} \cap \sqsupseteq_{\mathscr{R}}$.

Notemos que $\sqsubseteq_{\mathscr{R}}$ es covariante respecto a \mathscr{R} , pues $\overline{\tau} \mathscr{R} \overline{\sigma}$ implica $\mathtt{m}[\overline{\tau}] \sqsubseteq_{\mathscr{R}} \mathtt{m}[\overline{\sigma}]$.

Mailbox type system Subtipado

Decimos que \mathscr{R} es una relación de subtipado si $\tau \, \mathscr{R} \, \sigma$ implica

- 1. $\tau = ?E$ y $\sigma = ?F$ y $E \sqsubseteq_{\mathscr{R}} F$, o bien
- 2. $\tau = !E \text{ y } \sigma = !F \text{ y } F \sqsubseteq_{\mathscr{R}} E$

Escribimos \leqslant para denotar la mayor relación de subtipado y decimos que τ es un subtipo de σ si $\tau \leqslant \sigma$.

Escribimos \lessgtr para $\leqslant \cap \geqslant$, \sqsubseteq para \sqsubseteq_{\leqslant} y \simeq para \simeq_{\leqslant} .

Las 2 reglas se corresponden directamente con las reglas covariantes y contravariantes usuales para canales con capacidades de entrada y salida.

Subtipado: ejemplos

- !(A + B) ≤ !A: un mailbox de tipo !(A + B) es más permisivo que otro de tipo !A. Si un proceso necesita un mailbox para escribir un mensaje A, también le sirve un mailbox donde se puede escribir mensajes A o B.

Relaciones de subtipado especiales

Algunos mailbox types tienen patrones que están en cierta relación particular con las constantes $\mathbb O$ (unreliable mailbox) y $\mathbb 1$ (empty mailbox). Estos tipos tienen la siguiente clasificación:

- ▶ relevant: si $\tau \not \le !1$ (caso contrario *irrelevant*) Un mailbox *relevant* debe usarse, mientras que uno *irrelevant* puede descartarse. Todos los mailbox con input capability son *relevantes*.
- reliable: si τ ≤ ?0 (caso contrario unreliable)
 Un mailbox reliable no recibió mensajes inesperados. Todos los mailbox con output capability son reliable.
- ▶ **usable**: si $!0 \nleq \tau$ (caso contrario *unusable*) Un mailbox *usable* puede ser usado. Todos los mailbox con input capability son *usable*.

Ejemplo 11: lock type

```
\begin{aligned} \mathsf{FreeLock}(self) &\triangleq \mathtt{free} \ self.\mathtt{done} \\ &+ self?\mathtt{acquire}(owner).\mathtt{BusyLock}[self,owner] \\ &+ self?\mathtt{release.fail} \ self \\ \mathsf{BusyLock}(self,owner) &\triangleq owner!\mathtt{reply}[self] \mid self?\mathtt{release}.\mathsf{FreeLock}[self] \end{aligned}
```

El mailbox usado por FreeLock tendrá diferentes tipos dependiendo del estado interno del lock y desde qué óptica lo miramos.

- Desde FreeLock: ?acquire[!reply[!release]]*
- ► Desde BusyLock: ?(release · acquire[!reply[!release]]*)
- Desde User hacia FreeLock: !acquire[!reply[!release]]*
- Desde Owner hacia BusyLock: !release

Grafo de dependencias: sintaxis

Dependency Graph
$$\varphi, \psi ::= \emptyset \mid \{u, v\} \mid \varphi \sqcup \psi \mid (\nu a) \varphi$$

El grafo de dependencias es un multigrafo no dirigido donde los vértices del grafo son los nombres de los mailbox y el **objetivo es trackear las dependencias entre mailboxes**. Intuitivamente hay una dependencia entre u y v si:

- ightharpoonup v es un argumento de algún mensaje del mailbox u,
- $lackbox{ o bien }v$ aparece en la continuación de un proceso esperando un mensaje en u.

Grafo de dependencias: LTS

$$\left\{ u,v \right\} \xrightarrow{u-v} \emptyset \text{ [G-AXIOM]} \qquad \frac{\varphi \xrightarrow{u-v} \varphi'}{\varphi \sqcup \psi \xrightarrow{u-v} \varphi' \sqcup \psi} \text{ [G-LEFT]} \qquad \frac{\psi \xrightarrow{u-v} \psi'}{\varphi \sqcup \psi \xrightarrow{u-v} \varphi \sqcup \psi'} \text{ [G-RIGHT]}$$

$$\frac{\varphi \xrightarrow{u-v} \psi \qquad a \neq u,v}{(\nu a)\varphi \xrightarrow{u-v} (\nu a)\psi} \text{ [G-NEW]} \qquad \frac{\varphi \xrightarrow{u-w} \psi \qquad \psi \xrightarrow{w-v} \varphi'}{\varphi \xrightarrow{u-v} \varphi'} \text{ [G-TRANS]}$$

La semántica del grafo de dependencias está dada por un LTS donde el label u-v representa un camino que conecta u con v. La relación $\varphi \xrightarrow{u-v} \varphi'$ describe que u está conectado con v en φ , y φ' describe el grafo residual luego de eliminar las aristas usadas para conectar u con v.

Grafo de dependencias: propiedades

Def 12: graph acyclicity and entailment

Sea $\operatorname{dep}(\varphi) \stackrel{\operatorname{def}}{=} \{(u,v) \mid \exists \varphi' : \varphi \xrightarrow{u-v} \varphi' \}$ la relación de dependencias generada por φ .

- lacktriangle Decimos que φ es *acíclico* si dep (φ) es irreflexiva.
- ▶ Decimos que φ entails (implica) ψ , escrito $\varphi \Rightarrow \psi$, si $dep(\psi) \subseteq dep(\varphi)$.

Ejemplo 2: Future variable

```
\begin{aligned} \mathsf{Future}(self) &\triangleq self ? \mathtt{put}(x) . \mathsf{Present}[self, x] \\ \mathsf{Present}(self, x) &\triangleq \mathtt{free} \ self . \mathtt{done} \\ &+ self ? \mathtt{get}(sender) . (sender ! \mathtt{reply}[x] \mid \mathsf{Present}[self, x]) \\ &+ self ? \mathtt{put} . \mathtt{fail} \ self \end{aligned}
```

- Una variable futura admite un único mensaje put para setear el valor de la variable.
- Puede recibir una cantidad arbitraria de mensajes get, antes y después de ser seteada. Si se reciben antes del put, estos mensajes quedan pendientes en el mailbox ya que el proceso definido por Future no tiene una guarda que selecciona los mensajes get.

Ejemplo 2: Future variable

Detección de deadlock con el grafo de dependencias

```
(\nu f)(\nu c)(\mathsf{Future}[f] \mid f! \mathtt{get}[c] \mid c?\mathtt{reply}(x).\mathtt{free}\ c.f! \mathtt{put}[x])
```

- ▶ El mailbox c es el argumento del mensaje get guardado en el mailbox f.
- ► El mailbox f aparece en la continuación luego de leer del mailbox c.



Claramente el grafo de dependencias **tiene un ciclo**. Vamos a utilizar estos grafos en las reglas de tipado para evitar deadlocks.

Reglas de tipado

Type environments

 Γ y Δ son funciones parciales de nombres a tipos: $\overline{u}:\overline{\tau}.$

Procesos

$$\Gamma \vdash P :: \varphi$$

P está bien tipado bajo Γ y genera el grafo de dependencias $\varphi.$ Decimos que el juicio de tipado está bien formado si

 $\operatorname{fn}(\varphi)\subseteq\operatorname{dom}(\Gamma)$ y φ es acíclico.

Guardas

$$\Gamma \vdash G$$

Reglas de tipado

Typing rules for processes

$$\Gamma \vdash P :: \varphi$$

Typing rules for guards

 $\Gamma \vdash G$

$$\frac{\Gamma \vdash P :: \varphi}{u : ?0, \Gamma \vdash \mathtt{fail} \ u} \quad \overset{[\mathtt{T} \vdash P :: \varphi}{=} \quad \frac{u : ?1, \Gamma \vdash \mathtt{free} \ u.P}{=} \quad \overset{[\mathtt{T} \vdash P :: \varphi}{=} \quad \frac{u : ?E, \Gamma, \overline{x} : \overline{\tau} \vdash P :: \varphi}{u : ?(\mathtt{m}[\overline{\tau}] \cdot E), \Gamma \vdash u ? \mathtt{m}(\overline{x}) \cdot P} \quad \frac{u : ?E_i, \Gamma \vdash G_i \quad \overset{(i=1,2)}{=}}{u : ?(E_1 + E_2), \Gamma \vdash G_1 + G_2} \quad \overset{[\mathtt{T} \vdash \mathtt{BRANCH}]}{=} \quad \frac{u : ?E_i, \Gamma \vdash G_i \quad \overset{(i=1,2)}{=}}{=} \quad \overset{[\mathtt{T} \vdash \mathtt{BRANCH}]}{=} \quad \overset{[\mathtt{T} \vdash \mathtt{BRANCH}]}{=}$$

$$\overline{u: \operatorname{!m}[\overline{\tau}], \overline{v}: \overline{\tau} \vdash u \operatorname{!m}[\overline{v}] :: \{u, \{\overline{v}\}\}} \quad \text{\tiny [T-MSG]}$$

Un mensaje está bien tipado si el mailbox u admite recibir un mensaje con tag m y argumentos de tipo $\overline{\tau}$, y además los valores de los argumentos \overline{v} efectivamente tienen tipo $\overline{\tau}$.

Enviar un mensaje genera una dependencia entre el mailbox u y todos los argumentos \overline{v} .

$$\{u, \{\overline{v}\}\} \equiv \{u, v_1\} \sqcup \cdots \sqcup \{u, v_n\}$$

Mailbox type system [T-FAIL] y [T-FREE]

$$\frac{}{u: ?\mathbb{O}, \Gamma \vdash \mathtt{fail} \; u} \quad \overset{[\mathsf{T-FAIL}]}{\underbrace{u: ?\mathbb{I}, \Gamma \vdash \mathtt{free} \; u.P}} \quad \overset{[\mathsf{T-FREE}]}{\underbrace{u: ?}} \quad \overset$$

La regla [T-FAIL] permite tipar un *runtime error* si el mailbox u tiene tipo ?0 indicando que el mailbox es *unreliable* (recibió un mensaje inesperado).

La regla [T-FREE] dice que podemos liberar el mailbox u si tiene tipo ${\it ?1}$ indicando que está vacío, siempre y cuando la continuación P está bien tipada en el ambiente residual.

Mailbox type system [T-IN]

$$\frac{u: \mathbf{?}E, \Gamma, \overline{x}: \overline{\tau} \vdash P :: \varphi}{u: \mathbf{?}(\mathbf{m}[\overline{\tau}] \cdot E), \Gamma \vdash u \mathbf{?}\mathbf{m}(\overline{x}) \cdot P} \quad {}^{\scriptscriptstyle{\text{[T-IN]}}}$$

Consumir un mensaje m de un mailbox u requiere que el mailbox tenga tipo $?(\mathbf{m}[\overline{\tau}] \cdot E)$ que garantiza que hay al menos 1 mensaje m, posiblemente con otros mensajes acorde al patrón E.

La continuación P debe estar bien tipada en un ambiente donde el mailbox tiene tipo ?E que describe el contenido del mailbox luego de haber consumido el mensaje \mathtt{m} , y además incluye el tipo de los argumentos \overline{x} del mensaje.

$$\frac{u:?E_i,\Gamma \vdash G_i \stackrel{(i=1,2)}{-}}{u:?(E_1+E_2),\Gamma \vdash G_1+G_2} \quad \text{\tiny [T-BRANCH]}$$

La composición de guardas G_1+G_2 ofrece las acciones de G_1 y G_2 donde cada una se corresponde con el subpatrón E_1 y E_2 del mailbox type $?(E_1+E_2)$.

$$\frac{u:?E,\Gamma \vdash G \qquad \vDash E}{u:?E,\Gamma \vdash G :: \{u, \mathsf{dom}(\Gamma)\}} \quad \text{\tiny [T-GUARD]}$$

Esta regla tipa un guarded process G que posiblemente consume mensajes del mailbox u. El grafo de dependencias se define entre u y todas las variables libres que aparecen en la continuación.

La condición $\models E$ pide que el patrón E esté en **forma normal**.

Patrones en forma normal

$$u:?(\mathtt{A}\cdot\mathtt{C}+\mathtt{B}\cdot\mathtt{A})\vdash u?\mathtt{A}.P+u?\mathtt{B}.Q$$

Si usamos la regla <code>[T-IN]</code> para tipar ambas guardas, necesitamos tipar P y Q en un ambiente donde el mailbox u fue actualizado para reflejar que se consumió un mensaje. Podríamos inferir que para P tenemos u:?C y para Q tenemos u:?A.

Esto está mal porque el tipo de u especifica que un mensaje $\mathbf A$ puede estar acompañado de otro mensaje $\mathbf C$ o $\mathbf B$. La solución es llevar los patrones a una **forma normal** para evitar estos problemas. En este ejemplo, una forma normal podría ser: $u: \mathbf A \cdot (\mathbf B + \mathbf C)$.

Mailbox type system Propiedades

Teorema 23

Teorema 24

Teorema 25

Encoding binary sessions



Conclusiones