# Estudio y desarrollo de técnicas para el testing de programas concurrentes

CABS: traducción de C a ABS

Marco Antonio Garrido Rojo<sup>1</sup>

17 de junio de 2018

## Introducción: historia y origen

- En un principio, la mayoría de sistemas solo podían ejecutar una única secuencia de instrucciones a la vez.
- Con el avance de los ordenadores durante las últimas décadas del siglo pasado, se permite la ejecución simultánea de varios hilos de ejecución.
- Hoy en día es común el uso de la programación concurrente en la mayoría de aplicaciones y programas de todo ámbito.

#### Introducción: inconvenientes

- La ejecución paralela conlleva riesgos adicionales no presentes en los programas secuenciales.
- El principal problema que presenta es la presencia de una memoria compartida sobre la que los distintos hilos realizan modificaciones en cualquier momento.
- Asociado a esto, pueden ocurrir deadlocks, carreras de datos o comportamientos impredecibles.

#### Introducción: soluciones

- Podemos evitar que ocurran estableciendo restricciones mediante el uso de semáforos y cerrojos o implementaciones de algoritmos de más bajo nivel como el tie-breaker.
- El uso de estos mecanismos puede ser bastante complejo.

#### Introducción: soluciones

```
bool in1 = false, in2 = false;
int last = 1:
process CS1 {
  while (true) {
    last = 1; in1 = true; /* entry protocol */
    while (in2 and last == 1) skip;
    critical section;
    in1 = false;
                             /* exit protocol */
    noncritical section:
process CS2 {
  while (true) {
    last = 2; in2 = true; /* entry protocol */
    while (in1 and last == 2) skip;
    critical section:
    in2 = false;
                              /* exit protocol */
    noncritical section;
```

Figura: Algoritmo tie-breaker para dos procesos. (Andrews')

- En este terreno se sigue investigando para encontrar un modo definitivo que permita solucionar los problemas o detectarlos.
- Uno de los primeros asuntos estudiados es determinar el estado o estados finales a los que se llega ejecutando un programa concurrente.

```
int var;
proc f1() {
  var := 1;
proc f2() {
  var := 2;
```

- Explosión exponencial en el número de interleavings en proporción al número de hilos y al número de instrucciones.
- No siempre hay una correlación con el número de estados finales.
- No todos los interleavings tienen la misma importancia y algunos de ellos son equivalentes entre sí.

```
int var1;
int var2;
proc f1() {
  var1 := 1; # 1
  var1 := 2; # 2
proc f2() {
  var2 := 2; # 3
  var2 := 1; # 4
```

## Introducción: grano

- ¿Cuál es la atomicidad?
  - Asignaciones atómicas: grano grueso.
  - ¿Lectura de valores?: grano fino.
  - ¡Secciones de código enteras!: ...
- ¿De qué depende el grano? ¿Qué ocurre si no hay memoria compartida?

#### Introducción: actores

- Cada objeto ejecuta sus tareas de forma concurrente con respecto a las del resto de objetos.
- Una única tarea a la vez por objeto. El resto de tareas esperan en una cola cuyo orden no es determinable.
- El paso de mensajes indica qué método desea ejecutar un objeto (pudiendo ser uno propio o perteneciente a otro objeto).
- Dependiendo del tipo de llamada, una tarea puede dejar paso a otra si aún no cuenta con los valores necesarios para proseguir.
- Se trata de una concurrencia donde el scheduler es non-preemptive.

## Introducción: ventajas

- Reducción notable del número de interleavings: importante cuando se aplican técnicas de testing sistemático.
- Presente en lenguajes como Erlang, Scala y ABS.
- Cuenta con una amplia cantidad de herramientas, entre ellas, SYCO, que implementa DPOR para obtener los posibles resultados finales de un programa, reduciendo los interleavings redundantes.

## Introducción: objetivo

- Acercar las herramientas de ABS a un lenguaje que no sea de modelado, como C.
- Un lenguaje que permita recrear una concurrencia entre procesos a nivel de grano fino.
- Un lenguaje sencillo pero completo: CABS.

### CABS: sintaxis

- Subconjunto de C con sintaxis para hebras.
- Tipos estáticos y arrays.

```
type global_var1;
type global_var2;

type nombre_de_funcion(args...) {
    ...
codigo
    ...
return exp
}
```

## CABS: sintaxis

```
int array[10];
1 int global_var;
                       10 int main() {
                           array[0] = 10;
                           array[1] = 5;
4 int f(int res, int
   arr[10]) {
                       13
                           int res;
                       res = array[0] + 1;
int var;
                       global_var = f(res,
var = 2 * res +
   arr[1];
                           array);
                           return 0;
 return var;
                       17 }
```

## CABS: sintaxis

```
1 int var;
3 int main() {
 thread f(1);
 thread f(2);
   return 0;
void f(int value) {
10
 var = value;
11 }
```

## CABS: semántica

- Estado: (G, F, RP).
- Ámbito global de varibles (G): asocia nombres de variables a valores de V.
- Definición de funciones:  $\mathbf{F}(\mathbf{func}) = (t, S_F, args_F, a_{ret})$ .
- Lista de marcos de ejecución:  $RP \rightsquigarrow (local : s, S)$

#### CABS: semántica

$$\begin{aligned} & \mathsf{init}(\varepsilon) = (\mathsf{nil}, \mathsf{nil}, []) \\ & \mathsf{init}(\mathit{int} \; \mathsf{var}; \mathsf{P}) = (\mathsf{G} \left[ \; \mathsf{var} \mapsto 0 \right], \mathsf{F}, \mathsf{RP}) \\ & \mathsf{donde} \; \mathsf{init}(\mathsf{P}) = (\mathsf{G}, \mathsf{F}, \mathsf{RP}) \\ & \mathsf{init}(\mathit{bool} \; \mathsf{var}; \mathsf{P}) = (\mathsf{G} \left[ \; \mathsf{var} \mapsto \mathsf{FALSE} \right], \mathsf{F}, \mathsf{RP}) \\ & \mathsf{donde} \; \mathsf{init}(\mathsf{P}) = (\mathsf{G}, \mathsf{F}, \mathsf{RP}) \\ & \mathsf{init}(\mathit{int} \; \mathsf{func}(\mathsf{arg}) \{S; \mathsf{return} \; a\} \mathsf{P}) = (\mathsf{G}, \mathsf{F} \left[ \mathsf{func} \mapsto (\mathit{int}, S, \mathsf{arg}, a) \right], \mathsf{RP}) \\ & \mathsf{donde} \; \mathsf{init}(\mathsf{P}) = (\mathsf{G}, \mathsf{F}, \mathsf{RP}) \\ & \mathsf{init}(\mathit{bool} \; \mathsf{func}(\mathsf{arg}) \{S; \mathsf{return} \; b\} \mathsf{P}) = (\mathsf{G}, \mathsf{F} \left[ \mathsf{func} \mapsto (\mathit{bool}, S, \mathsf{arg}, b) \right], \mathsf{RP}) \\ & \mathsf{donde} \; \mathsf{init}(\mathsf{P}) = (\mathsf{G}, \mathsf{F}, \mathsf{RP}) \\ & \mathsf{init}(\mathit{void} \; \mathsf{func}(\mathsf{arg}) \{S\} \mathsf{P}) = (\mathsf{G}, \mathsf{F} \left[ \mathsf{func} \mapsto (\mathit{void}, S, \mathsf{arg}, \varepsilon) \right], \mathsf{RP}) \\ & \mathsf{donde} \; \mathsf{init}(\mathsf{P}) = (\mathsf{G}, \mathsf{F}, \mathsf{RP}) \end{aligned}$$

◆□▶ ◆□▶ ◆臺▶ ◆臺▶ · 臺 · 夕久②

## CABS: semántica de las Aexp

$$[\mathsf{num}_{\mathcal{A}}] \ \overline{\langle n, (\mathbf{G}, \mathbf{F}, \mathsf{RP} \leadsto (s, S)) \rangle \to_{\mathsf{Aexp}} \langle \mathcal{N} \llbracket n \rrbracket, (\mathbf{G}, \mathbf{F}, \mathsf{RP} \leadsto (s, S)) \rangle}$$

$$\left[\mathsf{var}_{\mathcal{A}}^{L}\right] \frac{\mathsf{local}(x) = \mathsf{v}}{\langle x, (\mathbf{G}, \mathbf{F}, \mathsf{RP} \leadsto (\mathsf{local} : s, S)) \rangle \to_{\mathsf{Aexp}} \langle v, (\mathbf{G}, \mathbf{F}, \mathsf{RP} \leadsto (\mathsf{local} : s, S)) \rangle}$$

$$\begin{bmatrix} \mathsf{var}_{\mathcal{A}}^{\mathcal{G}} \end{bmatrix} \ \frac{\mathcal{G}(x) = v \quad \mathit{local}(x) = \mathit{undef}}{\langle x, (\mathbf{G}, \mathbf{F}, \mathsf{RP} \leadsto (\mathit{local} : s, S)) \rangle \rightarrow_{\mathit{Aexp}} \langle v, (\mathbf{G}, \mathbf{F}, \mathsf{RP} \leadsto (\mathit{local} : s, S)) \rangle}$$

# CABS: semántica de las **Aexp**

$$\left[\bigcirc_{\mathcal{A}}^{1}\right] \ \frac{\langle a_{1}, (\mathbf{G}, \mathbf{F}, \mathbf{RP} \leadsto (s, S)) \rangle \rightarrow_{Aexp} \langle a'_{1}, (\mathbf{G}, \mathbf{F}, \mathbf{RP} \leadsto (s', S')) \rangle}{\langle a_{1} \bigodot a_{2}, (\mathbf{G}, \mathbf{F}, \mathbf{RP} \leadsto (s, S)) \rangle \rightarrow_{Aexp} \langle a'_{1} \bigodot a_{2}, (\mathbf{G}, \mathbf{F}, \mathbf{RP} \leadsto (s', S')) \rangle}$$

$$\left[\bigodot_{\mathcal{A}}^{2}\right] \ \frac{\langle a_{2}, (\mathbf{G}, \mathbf{F}, \mathsf{RP} \leadsto (s, S))) \rangle \rightarrow_{\mathsf{Aexp}} \langle a'_{2}, (\mathbf{G}, \mathbf{F}, \mathsf{RP} \leadsto (s', S'))) \rangle}{\langle v \bigodot a_{2}, (\mathbf{G}, \mathbf{F}, \mathsf{RP} \leadsto (s, S))) \rangle \rightarrow_{\mathsf{Aexp}} \langle v \bigodot a'_{2}, (\mathbf{G}, \mathbf{F}, \mathsf{RP} \leadsto (s', S'))) \rangle}$$

$$\left[\bigcirc_{\mathcal{A}}^{3}\right] \frac{}{\langle v_{1} \bigcirc v_{2}, (\mathbf{G}, \mathbf{F}, \mathbf{RP} \leadsto (s, S))\rangle \rightarrow_{Aexp} \langle v_{1} \bigcirc_{\mathcal{N}} v_{2}, (\mathbf{G}, \mathbf{F}, \mathbf{RP} \leadsto (s, S))\rangle}$$

## CABS: semántica de las Aexp

$$\begin{bmatrix} \mathsf{call}_{\mathcal{A}}^2 \end{bmatrix} \ \frac{\mathsf{F}(\mathsf{func}) = (\mathsf{int}, \mathsf{S}_{\mathsf{F}}, \mathsf{args}_{\mathsf{F}}, \mathsf{a}) \qquad \mathsf{check\_args}(\mathsf{args}_{\mathsf{F}}, \mathsf{args}) \qquad \mathsf{eval}(\mathsf{args}) = \mathsf{v}_1 : \cdots : \mathsf{v}_n \\ \\ \langle \mathsf{func}(\mathsf{args}), (\mathsf{G}, \mathsf{F}, \mathsf{RP} \leadsto (\mathsf{local} : \mathsf{s}, \mathsf{S})) \rangle \rightarrow_{\mathsf{Aexp}} \langle \mathsf{UNSTACK}(\mathsf{a}), (\mathsf{G}, \mathsf{F}, \mathsf{RP} \leadsto (\mathsf{nil} \left[\mathsf{args} \mapsto \mathsf{v}_1 : \cdots : \mathsf{v}_n\right] : \mathsf{local} : \mathsf{s}, \mathsf{S}_{\mathsf{F}}; \mathsf{S})) \rangle \\ \\ \rangle \\ \end{pmatrix}$$

## CABS: semántica de las asignaciones

$$\begin{bmatrix} \mathsf{ass}^1_\mathsf{C} \end{bmatrix} \ \frac{\langle \mathsf{a}, (\mathsf{G}, \mathsf{F}, \mathsf{RP} \leadsto (\mathit{local} : s, \varepsilon)) \rangle \to_{\mathsf{Aexp}} \langle \mathsf{a}', (\mathsf{G}, \mathsf{F}, \mathsf{RP} \leadsto (s', S_\mathcal{A})) \rangle}{(\mathsf{G}, \mathsf{F}, \mathsf{RP} \leadsto (\mathit{local} : s, \ \mathsf{var} = a; S)) \to (\mathsf{G}, \mathsf{F}, \mathsf{RP} \leadsto (s', S_\mathcal{A}; \ \mathsf{var} = a'; S))} \\ \begin{bmatrix} \mathsf{is\_local}(\mathit{var}) \\ \hline (\mathsf{G}, \mathsf{F}, \mathsf{RP} \leadsto (\mathit{local} : s, \ \mathsf{var} = v; S)) \to (\mathsf{G}, \mathsf{F}, \mathsf{RP} \leadsto (\mathit{local} [ \ \mathsf{var} \mapsto v] : s, S)) \\ \hline (\mathsf{ass}^3_\mathsf{C}] \ \hline \\ (\mathsf{G}, \mathsf{F}, \mathsf{RP} \leadsto (\mathit{local} : s, \ \mathsf{var} = v; S)) \to (\mathsf{G} [ \ \mathsf{var} \mapsto v], \mathsf{F}, \mathsf{RP} \leadsto (\mathit{local} : s, S)) \\ \hline \end{cases}$$

#### CABS: semántica del if

$$[\mathsf{if}_\mathsf{C}] \ \frac{\langle b, (\mathsf{G}, \mathsf{F}, \mathsf{RP} \leadsto (\mathit{local} : s, \varepsilon)) \rangle \rightarrow_{\mathit{Bexp}} \langle b', (\mathsf{G}, \mathsf{F}, \mathsf{RP} \leadsto (s', S_{\mathbb{B}})) \rangle}{(\mathsf{G}, \mathsf{F}, \mathsf{RP} \leadsto (\mathit{local} : s, \mathsf{if}(b) \{S_1\} \mathsf{else} \{S_2\} S)) \rightarrow (\mathsf{G}, \mathsf{F}, \mathsf{RP} \leadsto (s', S_{\mathbb{B}}; \mathsf{if}(b') \{S_1\} \mathsf{else} \{S_2\} S))}$$

$$\left[\mathsf{if}^\mathsf{TRUE}_\mathsf{C}\right] \ \frac{}{\left(\mathbf{G},\mathbf{F},\mathsf{RP} \leadsto (\mathit{local}:s,\mathsf{if}(\mathsf{TRUE})\{S_1\}\mathsf{else}\{S_2\}S)\right) \to \left(\mathbf{G},\mathbf{F},\mathsf{RP} \leadsto (\mathit{local}:s,S_1;S)\right)}$$

$$\left[\mathsf{if}^{\mathsf{FALSE}}_{\mathsf{C}}\right] \ \frac{}{\left(\mathsf{G},\mathsf{F},\mathsf{RP} \leadsto (\mathit{local}:s,\mathsf{if}(\mathsf{FALSE})\{S_1\}\mathsf{else}\{S_2\}S)\right) \to \left(\mathsf{G},\mathsf{F},\mathsf{RP} \leadsto (\mathit{local}:s,S_2;S)\right)}$$

### CABS: semántica del thread

$$[\mathsf{end}_\mathsf{C}] \ \overline{(\mathsf{G},\mathsf{F},\mathsf{RP} \leadsto (\mathsf{s},\varepsilon))) \to (\mathsf{G},\mathsf{F},\mathsf{RP}))}$$

$$\begin{bmatrix} \mathsf{thread}^1_\mathsf{C} \end{bmatrix} \ \frac{\mathbf{F}(\mathsf{func}) = (t, S_F, \mathit{args}_F, e) \quad \mathit{check\_args}(\mathit{args}_F, \mathit{args}) \quad \mathit{eval}(\mathit{args}) = \mathit{args'}}{(\mathbf{G}, \mathbf{F}, \mathsf{RP} \leadsto (s, \mathsf{thread}\, \mathsf{func}(\mathit{args}); S)) \to (\mathbf{G}, \mathbf{F}, (\mathsf{RP} \leadsto (s, \mathsf{thread}\, \mathsf{func}(\mathit{args'}); S))}$$

$$\begin{bmatrix} \mathsf{thread}_{\mathsf{C}}^2 \end{bmatrix} \ \frac{\mathsf{F}(\mathsf{func}) = (\mathsf{t}, S_{\mathsf{F}}, \mathsf{args}_{\mathsf{F}}, \mathsf{e}) \qquad \mathsf{check\_args}(\mathsf{args}_{\mathsf{F}}, \mathsf{args}) \qquad \mathsf{eval}(\mathsf{args}) = \mathsf{v}_1 : \dots : \mathsf{v}_n \\ \frac{\mathsf{G}(\mathsf{F}, \mathsf{RP} \leadsto (\mathsf{s}, \mathsf{thread}\,\mathsf{func}(\mathsf{args}); S)) \to (\mathsf{G}, \mathsf{F}, (\mathsf{RP} \cup (\mathsf{nil}\,[\mathsf{args} \mapsto \mathsf{v}_1 : \dots : \mathsf{v}_n] : [], S_{\mathsf{F}})) \leadsto (\mathsf{s}, S))}{\mathsf{grade}}$$

# CABS: ejemplo

```
1 int var;
3 int main() {
 thread f(1);
 thread f(2);
   return 0;
void f(int value) {
 var = value;
10
11 }
```

## CABS: ejemplo

Excluyendo algunos pasos:

```
\begin{aligned} (\mathsf{G},\mathsf{F},(\mathsf{nil},\mathsf{thread}\,f(1);\mathsf{thread}\,f(2);)) &\to \\ &\quad (\mathsf{G},\mathsf{F},(\mathsf{nil}\,[\mathit{value}\mapsto 1]\,,\mathit{var}=\mathit{value};):(\mathsf{nil},\mathsf{thread}\,f(2);)) \to \\ &\quad (\mathsf{G},\mathsf{F},(\mathsf{nil}\,[\mathit{value}\mapsto 1]\,,\mathit{var}=\mathit{value};):(\mathsf{nil}\,[\mathit{value}\mapsto 2]\,,\mathit{var}=\mathit{value};):(\mathsf{nil},\varepsilon)) \to \\ &\quad (\mathsf{G}\,[\mathit{var}\mapsto 1]\,,\mathsf{F},(\mathsf{nil}\,[\mathit{value}\mapsto 1]\,,\varepsilon):(\mathsf{nil}\,[\mathit{value}\mapsto 2]\,,\mathit{var}=\mathit{value};):(\mathsf{nil},\varepsilon)) \to \\ &\quad (\mathsf{G}\,[\mathit{var}\mapsto 2]\,,\mathsf{F},(\mathsf{nil}\,[\mathit{value}\mapsto 1]\,,\varepsilon):(\mathsf{nil}\,[\mathit{value}\mapsto 2]\,,\varepsilon):(\mathsf{nil},\varepsilon)) \to \\ &\quad (\mathsf{G}\,[\mathit{var}\mapsto 2]\,,\mathsf{F},[]) \end{aligned}
```

#### **ABS**: sintaxis

- Sintaxis parecida a la de Java (aproximadamente).
- Clases que implementan interfaces para exponer los métodos públicos.
- Tipos futuro para esperar resultados del paso de mensajes (await).

```
interface nombre_de_interfaz {
    ...
    signaturas
    ...
}

Tipo nombre_metodo(args);

class nombre_de_clase(args) implements nombre_de_interfaz {
    ...
    declaraciones de atributos privados
    ...
    implementaciones
    ...
}
```

### ABS: semántica

- Estado: (**O**, **C**).
- Definición de las clases:  $\mathbf{C} \llbracket c \rrbracket = (\text{Inter}, attr_c, met, args_c).$
- Lista de objetos instanciados (actores):
  - $\mathbf{O} \sim (id, c, \mathsf{RT} \sim (loc, S, t), t, \mathsf{attr})$

## ABS: algunas reglas semánticas

$$\begin{bmatrix} ass_{\mathsf{ABS}}^1 \end{bmatrix} \quad \frac{is \lrcorner local(x) \qquad is \lrcorner lnt(x)}{(\mathbf{0} \leadsto (id, c, \mathsf{RT} \leadsto (loc, x = a; S, t), t, \mathsf{attr}), \mathsf{C}) \to (\mathbf{0} \leadsto (id, c, \mathsf{RT} \leadsto (loc \left[x \mapsto \mathcal{A} \ \llbracket \mathsf{a} \rrbracket \rfloor_{loc, \mathsf{attr}}\right], S, t), t, \mathsf{attr}), \mathsf{C})}$$

$$\begin{array}{c} C \ \left[ c' \right] = (\text{Inter, attre, met, args_c}) & \textit{check\_args(args_c, args)} \\ \\ [\text{Obj}_{ABS}] \ \hline \\ (O \leadsto (\textit{id, c, RT} \leadsto (\textit{loc, Inter inter} = \textit{new c'(args)}; \textit{S, t)}, \textit{t, attr}), C) \rightarrow (O : \textit{o} \leadsto (\textit{id, c, RT} \leadsto (\textit{loc [inter} \mapsto \textit{id'}], \textit{S, t)}, \textit{t, attr}), C) \\ \end{array}$$

donde  $o=(id',c',[],\bot,\mathsf{attr'})$  con id' un nuevo identificador de objeto no utilizado y  $\mathsf{attr'} = \mathsf{attr}_c \left[ \mathsf{args}_c \mapsto \mathcal{E} \left[ \mathsf{args}_{||_{loc,\mathsf{attr}}} \right] \text{ los atributos del nuevo objeto creado.} \right]$ 



## ABS: algunas reglas semánticas

$$\begin{bmatrix} loc \cup attr \ \llbracket int \rrbracket = id' & C \ \llbracket c' \rrbracket = (lnter, attr_c, met, args_c) & contains(met, m) & check.args(args_m, args) \\ \hline \\ (O \leadsto (id', c', RT', t', attr')(id, c, RT \leadsto (loc, Int \times = await int!m(args); S, t), t, attr), C) \rightarrow st \\ \end{bmatrix}$$

donde o' = (id', c', RT' : tsk, t', attr'), tsk = (loc', S', t'') con t'' un identificador de tarea nuevo y  $loc' = nil [args_m \mapsto args]$ ,  $met [\![m]\!] = (S', args_m)$  y  $st = (\mathbf{O} \leadsto o'(id, c, RT \leadsto (loc, Int x = await t'; S, t), t, attr), \mathbf{C})$ 

$$[\text{ret}_{\mathsf{ABS}}^1] \frac{\mathit{tsk} = (\mathit{loc}', \varepsilon(\nu), t'')}{(\mathsf{O} \leadsto (\mathit{id}', c', \mathsf{RT}' \leadsto \mathit{tsk}, t', \mathsf{attr}')(\mathit{id}, c, \mathsf{RT} \leadsto (\mathit{loc}, \mathit{Int} \, x = \mathit{await} \, t'; S, t), t, \mathsf{attr}), \mathsf{C}) \to \mathit{st}}$$

 $\mathsf{donde}\ o' = (\mathit{id'}, c', \mathsf{RT'}, t', \mathsf{attr'})\ \mathsf{y}\ \mathsf{st} = (\mathbf{0} \leadsto o'(\mathit{id}, c, \mathsf{RT} \leadsto (\mathit{loc}\ [\mathsf{x} \mapsto \nu]\ , \mathsf{S}, t), t, \mathsf{attr}), \mathbf{C})$ 

$$\overline{ (\textbf{0} \leadsto (\textit{id},\textit{c}, \mathsf{RT} \leadsto (\textit{loc}, \mathsf{return}\;\textit{a}; S, \textit{t}), \textit{t}, \mathsf{attr}), \mathsf{C}) \rightarrow (\textbf{0} \leadsto (\textit{id},\textit{c}, \mathsf{RT} \leadsto (\textit{loc}, \varepsilon(\mathcal{A}\; \llbracket \textit{a} \rrbracket_{\textit{loc}, \mathsf{attr}}), \textit{t}), \bot, \mathsf{attr}), \mathsf{C}) }$$