Construction d'une sémantique concurrente par instrumentation d'une sémantique opérationnelle structurelle

L'exemple du langage Orc

SOS

Sémantiques opérationnelles structurelles [2]

Définition

Système de transition d'états $(F, L, \rightarrow)$:

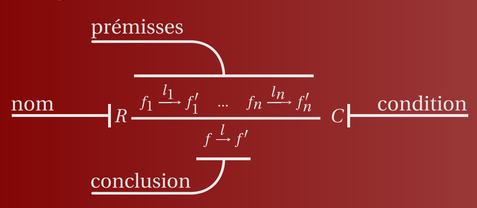
- F : ensemble d'états
- L : ensemble d'étiquettes
- $\rightarrow \in F \times L \times F$: fonction de transition définie par des règles d'inférence

Sémantique

Soit $f \in F$

- $[\![f]\!]$: l'ensemble des chemins finis partant de f
- $\sigma = l_1 ... l_n \in [\![f]\!]$ si il existe $f_1, ..., f_n$ tels que $f \xrightarrow{l_1} f_1 ... \xrightarrow{l_n} f_n$

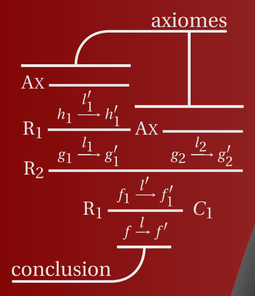
Règles d'inférence



$(f, l, f') \in \rightarrow$ si :

- pour tout i, $(f_i, l_i, f'_i) \in \rightarrow$
- C est vérifiée

Arbres de dérivation



Le langage Orc

Le langage Orc [1]

Philosophie :

- La modularité est essentielle
- Des sites et services existent
- Orc peut les *orchestrer*

Sites et publications

- Un site s'appelle comme une fonction dans un langage fonctionnel
- Une expression Orc publie zéro, une ou plusieurs valeurs

Le calcul Orc

- Sous-ensemble de Orc
- Modèle mathématique de la programmation concurrente
- Parfaitement défini mathématiquement

Syntaxe

$$
\begin{array}{lll}
f, g, h & ::= & p \| p(p) \| ?k \| f|g \| f > x > g \\
 & & \| f < x < g \| f; g \| D \# f \| \bot \\
D & ::= & \textbf{def}\ y(x) = f \\
v & ::= & V \| D \\
p & ::= & v \| \textbf{stop} \| x \\
w & ::= & NT(v) \| T(v) \| Neg \\
n & ::= & ?V_k(v) \| ?D \| h(\omega) \| h(!v) \\
l & ::= & !v \| n \| \omega
\end{array}
$$

Exécutions

Les étiquettes produites sont :

- $!v$ publication de v
- ω fin du calcul
- $?V_k(v)$ appel du site externe $V(v)$
- $?D$ appel du site interne D
- $h(!v)$ liaison de v à une variable
- $h(\omega)$ fin d'une sous-expression

Exemple

24 exécutions maximales possibles pour $y + z < y < ((2|3) > x > x) < z < 1$.
En particulier :

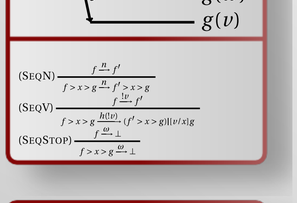
$h(!1).h(!3).h(!3).? + (3,1).!4.\omega$,
$h(!3).h(!1).h(!3).? + (3,1).!4.\omega$,
$h(!3).h(!3).h(!1).? + (3,1).!4.\omega$,
$h(!2).h(!1).h(!3).h(!3).? + (3,1).!4.\omega$,
$h(!1).h(!2).h(!3).h(!3).? + (3,1).!4.\omega$,
$h(!1).h(!3).h(!2).h(!3).? + (3,1).!4.\omega$,
$h(!2).h(!3).h(!1).h(!3).? + (3,1).!4.\omega$,
$h(!3).h(!2).h(!1).h(!3).? + (3,1).!4.\omega$,
$h(!3).h(!1).h(!2).h(!3).? + (3,1).!4.\omega$,
$h(!2).h(!3).h(!3).h(!1).? + (3,1).!4.\omega$,
$h(!3).h(!2).h(!3).h(!1).? + (3,1).!4.\omega$,
...

Problème

Comment différencier concurrence et non-déterminisme ?

$f > x > g$

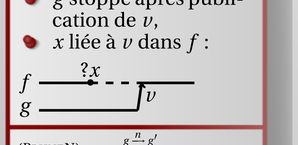
- Nouvelle instance de g lancée à chaque publication par f :



$$(\text{SeqN})\ \frac{f \xrightarrow{n} f'}{f > x > g \xrightarrow{n} f' > x > g} \qquad (\text{SeqV})\ \frac{f \xrightarrow{!v} f'}{f > x > g \xrightarrow{h(!v)} (f' > x > g)[v/x]g} \qquad (\text{SeqStop})\ \frac{f \xrightarrow{\omega} \bot}{f > x > g \xrightarrow{\omega} \bot}$$

$f < x < g$

- f et g lancés en parallèle,
- g stoppé après publication de v,
- x liée à v dans f :

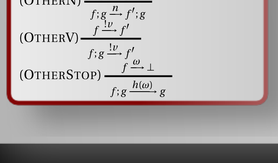


$$(\text{PruneN})\ \frac{g \xrightarrow{n} g'}{f < x < g \xrightarrow{n} f < x < g'} \qquad (\text{PruneLeft})\ \frac{f \xrightarrow{l} f'}{f < x < g \xrightarrow{l} f' < x < g}\ l \neq \omega$$

$$(\text{PruneV})\ \frac{g \xrightarrow{!v} g'}{f < x < g \xrightarrow{h(!v)} [v/x]f} \qquad (\text{PruneStop})\ \frac{g \xrightarrow{\omega} \bot}{f < x < g \xrightarrow{h(\omega)} [\textbf{stop}/x]f}$$

$f; g$

- g lancée ssi f s'arrête sans publier :

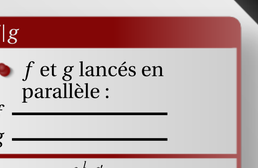


$$(\text{OtherN})\ \frac{f \xrightarrow{n} f'}{f;g \xrightarrow{n} f';g} \qquad (\text{OtherV})\ \frac{f \xrightarrow{!v} f'}{f;g \xrightarrow{!v} f'} \qquad (\text{OtherStop})\ \frac{f \xrightarrow{\omega} \bot}{f;g \xrightarrow{h(\omega)} g}$$

Sémantique

$f|g$

- f et g lancés en parallèle :



$$(\text{ParLeft})\ \frac{f \xrightarrow{l} f'}{f|g \xrightarrow{l} f'|g}\ l \neq \omega \qquad (\text{ParRight})\ \frac{g \xrightarrow{l} g'}{f|g \xrightarrow{l} f|g'}\ l \neq \omega$$

$$(\text{ParStop})\ \frac{f \xrightarrow{\omega} \bot \quad g \xrightarrow{\omega} \bot}{f|g \xrightarrow{\omega} \bot}$$

Valeurs et sites

- v : publication
- **stop** : termine le programme
- V : paramètres liés avant les appels de sites externes
- $D \equiv \textbf{def}\ y(x) = g$: appel indulgent pour les sites internes

$$(\text{Publish})\ \frac{}{v \xrightarrow{!v} \textbf{stop}} \qquad (\text{Stop})\ \frac{}{\textbf{stop} \xrightarrow{\omega} \bot} \qquad (\text{DefDeclare})\ \frac{[D/y]f \xrightarrow{l} f'}{D \# f \xrightarrow{l} f'}$$

$$(\text{IntCall})\ \frac{}{D(p) \xrightarrow{?D} [D/y][p/x]g} \qquad (\text{StopCall})\ \frac{}{\textbf{stop}(p) \xrightarrow{\omega} \bot} \qquad (\text{ExtCall})\ \frac{}{V(v) \xrightarrow{?V_k(v)} ?k}\ k \text{ fraîche}$$

$$(\text{ExtStop})\ \frac{}{V(\textbf{stop}) \xrightarrow{\omega} \bot} \qquad (\text{NtRes})\ \frac{}{?k \xrightarrow{!v} ?k}\ ?k \text{ reçoit } NT(v) \qquad (\text{Tres})\ \frac{}{?k \xrightarrow{!v} \textbf{stop}}\ ?k \text{ reçoit } T(v)$$

$$(\text{NegRes})\ \frac{}{?k \xrightarrow{\omega} \bot}\ ?k \text{ reçoit } Neg$$

La sémantique instrumentée

Exécution

$\sigma \in [\![f_0]\!]_i, \sigma_i = (k_i, l_i, c_i, a_i)$
$\overline{\overline{\sigma}} = (E, L, \leq, \nearrow, \Lambda)$

- k_i : identifiant unique de l'événement
  $E = \{k_0, ..., k_n\}$
- l_i : étiquette dans la sémantique standard
  $L = \{l_0, ..., l_n\} \quad \Lambda(k_i) = l_i$
- c_i : causes de k
  $k_i \leq k_j \Leftrightarrow k_i \in c_j$
- a_i : causes faibles de k
  $k_i \nearrow k_j \Leftrightarrow k_i \in a_j$

Quelques règles

$$(\text{Publish})\ \frac{}{v \xrightarrow{k,!v,\emptyset,\emptyset}_i \langle \textbf{stop}, \{k\}, \emptyset \rangle_l}\ k \text{ fraîche}$$

$$(\text{ParStop})\ \frac{f \xrightarrow{k,\omega,c,a}_i f' \quad g \xrightarrow{k',\omega,c',a'}_i g'}{f|g \xrightarrow{k,\omega,c \cup c',a \cup a'}_i \bot}$$

$$(\text{SeqV})\ \frac{f \xrightarrow{k,!v,c,a}_i f'}{f > x > g \xrightarrow{k,h(!v),c,a}_i (f' > x > g)|\langle [v/x]g, c \cup \{k\}, \emptyset \rangle_l}$$

$$(\text{PruneV})\ \frac{g \xrightarrow{k,!v,c,a}_i g'}{f < x < g \xrightarrow{k,h(!v),c,a}_i (|\langle v, c \cup \{k\}, \emptyset \rangle_l / x]f, c \cup \{k\}, \emptyset \rangle_\omega}$$

$$(\text{PruneN})\ \frac{g \xrightarrow{k,n,c,a}_i g'}{f < x < g \xrightarrow{k,n,c,a}_i f < x < \langle g', \emptyset, \{k\} \rangle_{!v}}$$

Le problème des sites

- *Substring* : appels indépendants
- *Read*/*Write* : appels liés causalement
- Solution : laisser les sites gérer la causalité

$$(\text{NtRes})\ \frac{}{?k \xrightarrow{!j,!v,c,a \cup c}_i \langle ?k, c \cup \{j\}, \emptyset \rangle_\omega}\ j \text{ fraîche}$$

$\langle f, c, a \rangle_L$

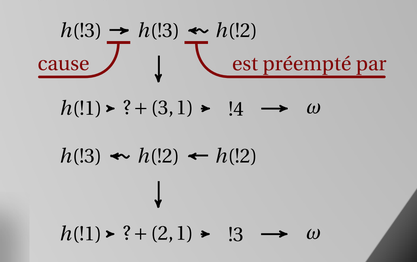
- f : l'expression évaluée
- L : désigne un type d'événement
- c : causes des prochains événements de type L
- a : causes faibles des prochains événements de type L

$$(\text{CauseYes})\ \frac{f \xrightarrow{k,l,c,a}_i f'}{\langle f, c', a' \rangle_L \xrightarrow{k,l,c \cup c', a \cup a'}_i \langle f', c', a' \rangle_L}\ l \in L$$

$$(\text{CauseNo})\ \frac{f \xrightarrow{k,l,c,a}_i f'}{\langle f, c', a' \rangle_L \xrightarrow{k,l,c,a}_i \langle f', c', a' \rangle_L}\ l \notin L$$

Exemple

Deux exécutions maximales possibles pour $y + z < y < ((2|3) > x > x) < z < 1$



Concurrence +

Théorèmes

Soit f_0 un programme Orc.

- $\{\sigma_{|l} : \sigma \in [\![f_0]\!]_i\} = [\![f_0]\!]$
- $\forall \sigma_i \in [\![f_0]\!]_i, Lin\left(\overline{\overline{\sigma_i}}\right) \subset [\![f_0]\!]$
- $\forall \sigma_1, \sigma_2 \in [\![f_0]\!]^\omega_{ic}, \overline{\overline{\sigma_1}} \equiv \overline{\overline{\sigma_2}}$
- $\forall \sigma_{ic} \in [\![f_0]\!]^\omega_{ic}, Lin\left(\overline{\overline{\sigma_{ic}}}\right) = [\![f_0]\!]$

La sémantique concurrente

Exécution

$\sigma \in [\![f_0]\!]_{ic}, \sigma_i = (k_i, l_i, c_i, a_i, b_i)$
$\overline{\overline{\sigma}} = (E, L, \leq, \nearrow, \Lambda)$

- k_i : identifiant unique de l'événement
  $E = \{k_i | l_i \neq \varepsilon\}$
- l_i : étiquette dans la sémantique standard
  $L = \{l_0, ..., l_n\} \quad \Lambda(k_i) = l_i$
- c_i : causes de k
  $k_i \leq k_j \Leftrightarrow k_i \in c_j$
- a_i : causes faibles de k
- b_i : conséquences faibles de k
  $k_i \nearrow k_j \Leftrightarrow k_i \in a_j \vee (c_j \cup \{k_j\}) \cap b_i \neq \emptyset$

Exécutions équitables

$\sigma \in [\![f_0]\!]^\omega_{ic}$ si :

- σ est une exécution selon $\rightarrow_{ic}$
  $\forall \gamma \in \text{pref}(\sigma), \gamma \in [\![f_0]\!]_{ic}$
- tout événement possible sera tiré un jour
  $\forall \gamma \in \text{pref}(\sigma), \forall e, \gamma.e \in [\![f_0]\!]_{ic} \Rightarrow \overline{\overline{\gamma.e}} < \overline{\overline{\sigma}}$

Existence des exécutions équitables :
$\forall \gamma \in [\![f_0]\!]_{ic}, \exists \sigma, \gamma\sigma \in [\![f_0]\!]^\omega_{ic}$

$f \setminus k$

- f : l'expression évaluée
- k : ses conséquences sont ignorées

$$(\text{Hide})\ \frac{f \xrightarrow{k,l,c,a,b}_{ic} f'}{f \setminus k' \xrightarrow{k,l,c,a,b}_{ic} f' \setminus k'}\ k' \notin c$$

Quelques règles

$$(\text{ParStart})\ \frac{f|_{(\emptyset,\emptyset)} g \xrightarrow{k,l,c,a,b}_{ic} f'}{f|g \xrightarrow{k,l,c,a,b}_{ic} f'} \qquad (\text{Publish})\ \frac{}{v \xrightarrow{k,!v,\emptyset,\emptyset,\emptyset}_{ic} \langle \textbf{stop}, \{k\}, \emptyset, \emptyset \rangle_l}\ k \text{ fraîche}$$

$$(\text{ParLeftStop})\ \frac{f \xrightarrow{k,\omega,c,a,b}_{ic} f'}{f|_K g \xrightarrow{k,\varepsilon,\emptyset,\emptyset,\emptyset}_{ic} f'|_{(c,a,b) \triangleright K} g} \qquad (\text{SeqV})\ \frac{f \xrightarrow{k,!v,c,a,b}_{ic} f'}{f > x > g \xrightarrow{k,h(!v),c,a,b}_{ic} (f' > x > g)|\langle [v/x]g, c \cup \{k\}, a, b \rangle_l}$$

$$(\text{ParLeft})\ \frac{f \xrightarrow{k,l,c,a,b}_{ic} f'}{f|_K g \xrightarrow{k,l,c,a,b}_{ic} f'|_K g}\ l \neq \omega \qquad (\text{PruneV})\ \frac{g \xrightarrow{k,!v,c,a,b}_{ic} g'}{f < x, K < g \xrightarrow{k,h(!v),c,a,b}_{ic} [x|\langle v, c \cup \{k\}, a, b \rangle_l / x]f < x, K \lhd (c, a, b) < \langle g', \emptyset, \{k\}, \{k\} \rangle_{!v}}$$

$$(\text{ParMid})\ \frac{K \xrightarrow{k,l,c,a,b}_{ic} K'}{f|_K g \xrightarrow{k,l,c,a,b}_{ic} f|_{K'} g} \qquad (\text{PruneN})\ \frac{g \xrightarrow{k,n,c,a,b}_{ic} g'}{f < x, K < g \xrightarrow{k,n,c,a,b}_{ic} f < x, K < \langle g', \emptyset, \{k\}, \emptyset \rangle_{!v}}$$

$$(\text{SeqN})\ \frac{f \xrightarrow{k,n,c,a,b}_{ic} f'}{f > x > g \xrightarrow{k,n,c,a,b}_{ic} f' > x > g}$$

Le problème de l'arrêt

- Exemple :
  $\textbf{stop} < x < (1|2|3)|\textbf{stop} < x < (4|5|6)$
- 9 causes possibles pour s'arrêter ⇒ distribution

$$e \triangleright (K_L, K_R) \lhd e' \xrightarrow{k,\omega,e_c \cup e'_c, e_a \cup e'_a, e_b \cup e'_b}_{ic} (K'_L, K'_R)$$

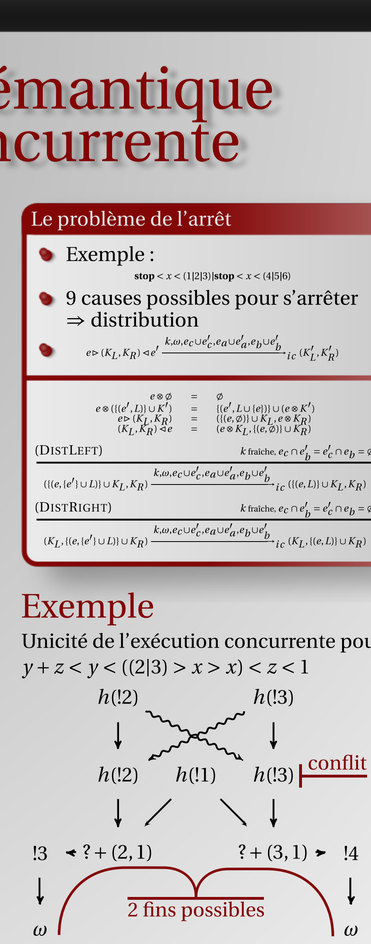
$$\begin{array}{rcl}
e \otimes \emptyset & = & \emptyset \\
e \otimes (\{(e', L)\} \cup K') & = & \{(e', L \cup \{e\})\} \cup (e \otimes K') \\
e \triangleright (K_L, K_R) & = & \{(\{e, \emptyset\}) \cup K_L, e \otimes K_R\} \\
(K_L, K_R) \lhd e & = & (e \otimes K_L, \{(e, \emptyset)\} \cup K_R)
\end{array}$$

$$(\text{DistLeft})\ \frac{k \text{ fraîche}, e_c \cap e'_b = e'_c \cap e_b = \emptyset}{(\{(e, \{e'\} \cup L)\} \cup K_L, K_R) \xrightarrow{k,\omega,e_c \cup e'_c, e_a \cup e'_a, e_b \cup e'_b}_{ic} (\{(e, L)\} \cup K_L, K_R)}$$

$$(\text{DistRight})\ \frac{k \text{ fraîche}, e_c \cap e'_b = e'_c \cap e_b = \emptyset}{(K_L, \{(e, \{e'\} \cup L)\} \cup K_R) \xrightarrow{k,\omega,e_c \cup e'_c, e_a \cup e'_a, e_b \cup e'_b}_{ic} (K_L, \{(e, L)\} \cup K_R)}$$

Exemple

Unicité de l'exécution concurrente pour $y + z < y < ((2|3) > x > x) < z < 1$



+ Conflit

LAES

Structures étiquetées asymétriques d'événements

Définition

$\mathcal{E} = (E, L, \leq, \nearrow, \Lambda)$

- E : événements
- L : étiquettes
- $\leq \in E^2$:
  causalité (ordre partiel)
- $\nearrow \in E^2$:
  causalité faible
- $\Lambda : E \mapsto L$:
  fonction d'étiquetage

On exige :

- finitude des histoires causales
  $\forall e \in E, [e] = \{e' \in E | e' \leq e\}$ fini
- $\forall e, e', e < e' \Rightarrow e \not\nearrow e'$
- $\nearrow \cap [e]^2$ acyclique pour tout $e \in E$

Linéarisations

$\Lambda(e_0) ... \Lambda(e_n) \in Lin(\mathcal{E})$ si :

- les $e_i \in E$ sont distincts,
- $\forall e \in E, \forall e' \in \{e_0, ..., e_n\}, e \leq e' \Rightarrow e \in \{e_0, ..., e_n\}$,
- $\forall e_i, e_j \in \{e_0, ..., e_n\}, e_i \nearrow e_j \Rightarrow i < j$

Notions couvertes

La causalité

- $e_1 \leq e_2$ si e_1 est nécessaire pour que e_2 se produise

La causalité faible

- $e_1 \nearrow e_2$ si e_1 ne peut pas se produire après e_2

La préemption

- $e_1 \rightsquigarrow e_2$ si $e_1 \nearrow e_2$ et $e_1 \not\leq e_2$
- e_1 est préempté par e_2 si l'exécution de e_2 empêche celle de e_1

La concurrence

- $e_1 || e_2$ si $\neg(e_1 \nearrow e_2 \vee e_2 \nearrow e_1)$
- e_1 et e_2 sont concurrents s'ils peuvent se produire dans n'importe quel ordre

Le conflit

- $\#\{e_1, ..., e_n\}$ s'il existe $e'_1 \nearrow ... \nearrow e'_n \nearrow e'_1$ avec $e'_i < e_i$ pour tout i
- $e_1, ... e_n$ sont en conflit s'ils ne peuvent pas tous se produire

Correspondance

$\mathcal{E}_1 = (E_1, L_1, \leq_1, \nearrow_1, \Lambda_1)$
$\mathcal{E}_2 = (E_2, L_2, \leq_2, \nearrow_2, \Lambda_2)$
$\mathcal{E}_1 < \mathcal{E}_2$ si $\exists f : E_1 \rightarrow E_2$ tel que

- $\forall e \in E_1, \Lambda_1(e) = \Lambda_2(f(e))$
- $\forall e_1 \leq_1 e_2 \in E_1, f(e_1) \leq_2 f(e_2)$
- $\forall e_1 \nearrow_1 e_2 \in E_1, f(e_1) \nearrow_2 f(e_2)$
- $\forall e_2 \leq_2 f(e) \in E_2, \exists e_1 \in E_1, e_2 = f(e_1)$

$\mathcal{E}_\infty \equiv \mathcal{E}_\epsilon$ si $\mathcal{E}_\infty < \mathcal{E}_\epsilon$ et $\mathcal{E}_\epsilon < \mathcal{E}_\infty$

Travaux à venir

Étude des causes entre appels de sites

- formalisation des structures de données réparties

Création d'un débogueur

La sémantique instrumentée permet :

- d'analyser des causes premières
- de rejouer les traces

La sémantique concurrente permet :

- d'explorer complètement l'espace des exécutions
- d'étudier les situations de compétition

Bibliographie

[1] Kitchin D., Quark A., Cook W., Misra J., The Orc Programming Language, *Formal Techniques for Distributed Systems ; Proceedings de FMOODS/FORTE*, vol. 5522, Springer, 2009, p. 1-25.

[2] Plotkin G. D., A structural approach to operational semantics, *rapport no DAIMI FN-19, 1981, Université de Aarhus, Aarhus, Danemark, réimprimé dans J. Log. Algebr. Program.* 60-61 : 17-139 (2004).

MSR'13

Modélisation des Systèmes Réactifs

Rennes, 13-15 Novembre 2013

Matthieu Perrin
Claude Jard
Achour Mostefaoui

Laboratoire d'Informatique de Nantes-Atlantique, Université de Nantes