APS - Notes de Cours

Jordi Bertran de Balanda

Cours 1

Introduction

Programme

Description statique (fichier) d'un comportement dynamique (exécution).

Statique: exécutable/bytecode/assembleur/code source

Code source: suite de caractères obéissant à des règles de syntaxe.

Syntaxe

- Composée de:
 - Lexique: ensemble d'unités lexicales (mots/symboles)
 - Grammaire: règles d'agencement des mots
- définit les suites de caractères qui sont des programmes
- donne un procédé de construction d'un mécanisme de reconnaissance des programmes (automates, langages formels)

Sémantique

- opérationnelle (grand pas/petit pas)
- dénotationnelle (lambda-calcul)
- axiomatique (cf. logique de Hoare en CPS)

Comportement dynamique: résulte du traitement de données statiques (via CPU/interp bytecode/interp code source)

Règles/mécanismes d'interprétation = fonction d'interprétation

Analyse

Dirigée par la syntaxe.

- Analyse syntaxique: éliminer les sources non évaluables
- Analyse de type: éliminer les erreurs prévisibles à l'exécution.

Noyau impératif

- Expressions: bool, int, void
- Instructions: affectation, alternative, boucle, séquence, déclaration = Blocs

Syntaxe

Lexique Mots clés (réservés):

Key	Bool	Int
VAR	true	add
CONST	false	sub
SET	and	mul
IF	or	div
WHILE		

Symboles réservés:

Description du lexique des noms de variables et des entiers: expressions régulières. En lex:

num : [0-9][0-9]* ident : [a-z][a-z]*

Grammaire Formalisation avec la Backus-Naur Form (BNF)

```
Prog ::= '[' Cmds ']'
Cmds ::= Cmd
      | Cmd ';' Cmds
Cmd ::= Dec
             #(declaration)
      | Stat #(statement)
Dec ::= 'CONST' ident Exp
      | 'VAR' ident TypeExp
TypeExp ::= 'bool' | 'int'
Stat ::= 'SET' ident Exp
      | 'IF' Exp Prog Prog
       | 'WHILE' Exp Prog
UnOp ::= '(' 'not' Exp ')'
BinOp ::= '(' 'and' Exp Exp ')'
        | '(' 'or' Exp Exp ')'
        | '(' 'add' Exp Exp ')'
        | '(' 'sub' Exp Exp ')'
        | '(' 'mul' Exp Exp ')'
        | '(' 'div' Exp Exp ')'
Exp ::= 'true'
       | 'false'
       | num
       | ident
       | UnOp
       | BinOp
```

Cours 2

Typage

Définition: assigner un type (de données)

- aux expressions
- aux instructions
- aux programmes

Contexte de typage

Un ensemble d'assignations de types. Noté Γ .

```
\Gamma \vdash x : t \text{ si } (x : t) \in \Gamma \ \forall \text{ type } t
```

Dans le cadre du langage réalisé en APS:

Expressions

Instructions Relation entre:

- Un contexte de typage
- Une expression
- Un type

```
\Gamma \vdash (\text{SET x e}) : \text{void si } \Gamma \vdash \text{x} : \text{t et } \Gamma \vdash \text{e} : \text{t} \ \forall \text{ type t}
\Gamma \vdash (\text{IF e } b_1 \ b_2) : \text{void si } \Gamma \vdash \text{e} : \text{bool et } \Gamma \vdash b_1 : \text{void et } \Gamma \vdash b_2 : \text{void}
\Gamma \vdash (\text{WHILE e b}) : \text{void si } \Gamma \vdash \text{e} : \text{bool et b} : \text{void}
```

Suite de commandes Relation entre:

- Un contexte de typage
- Un bloc (programme)
- void

Deux cas à considérer:

- [declaration; command sequence]
- [statement; command sequence]

```
\begin{split} \Gamma \vdash [\ ] : \ void \\ \Gamma \vdash (s::cs) : \ void \ si \ s: \ void \ et \ \Gamma \vdash cs: \ void \ \forall \ s \in statement \\ \Gamma \vdash ((CONST \ x \ e)::cs) : \ void \ si \ \Gamma; x:t \vdash cs: \ void \ et \ \Gamma \vdash e: \ t \\ \Gamma \vdash ((VAR \ x \ t)::cs) : \ void \ si \ \Gamma; x:t \vdash cs: \ void \\ Notation \ ci-dessus: \ \Gamma; x:t \ abrège \ \Gamma \cup \{x:t\} \end{split}
```

Programmes/blocs Bloc: suite de déclarations et d'instructions

```
\vdash [ cs ] : void si \vdash cs : void
```

Notation d'une règle d'inférence:

```
\begin{array}{l} \Gamma\vdash e:bool \qquad \Gamma\vdash b:void \\ \Gamma\vdash (\text{WHILE e b}):void \\ \hline \Gamma\vdash (\text{VAR x t})::cs:void \\ \hline \Gamma\vdash (\text{VAR x t})::cs:void \\ \hline x:bool\vdash true:bool \qquad \frac{x:bool\vdash x:bool}{x:bool\vdash (notx):bool} \\ \hline \Gamma\vdash (\text{and true (not x)}):bool \end{array}
```

Simplification

On peut enlever des règles de typage en les ajoutant à la description du programme.

Par exemple:

```
true : bool; false : bool \vdash [ cs ] : void
```

On peut ainsi enlever toutes les opérations dont les paramètres sont identiques à la valeur de retour.

$$\frac{\Gamma \vdash op: t_1 * t_2 -> t \qquad \Gamma \vdash e_1: t_1 \qquad \Gamma \vdash e_2: t_2}{\Gamma \vdash (\text{op } e_1, e_2): t}$$

En généralisant:

$$\frac{\Gamma \vdash op:t_1*..*t_n -> t \quad \Gamma \vdash e_1:t_1 \ .. \ \Gamma \vdash e_n:t_n}{\Gamma \vdash (ope_1,e_2,..e_n):t}$$

Ajout des types dans la grammaire

Cours 2

Sémantique opérationnelle

- Expressions: fonctionnel \leadsto valeurs
- Instructions: impératif \leadsto état mémoire
- $D\'{e}claration: \leadsto \'{e}tat$ mémoire, environnement

Environnement

Association entre identificateurs et * valeurs * adresses mémoire

Notation:

#v - valeurs
@a - adresses

Soit r un environnement.

• Accès, avec x : Id

$$\begin{array}{l} - \ r(x) \\ - \ [\](x) = error \end{array}$$

• Ajout, avec w une valeur:

$$\begin{aligned} & - \ r[x = w] \\ & - \ r[x = w](x) = w \\ & - \ r[x = w](y) = r(y) \ avec \ x \not = y \end{aligned}$$

Mémoire

Associations entre adresses et

- valeurs
- autres adresses

Avec m une mémoire:

• Accès:

• Extension (allocation):

```
-m[a:=new]-m(a) est non défini
```

• Modification: m[a := w]

$$\begin{array}{l} -\ m[a:=w][a:=w'] = m[a:=w'] \\ -\ m[a':=w][a:=w'] = m[a:=w'][a':=w] \ avec \ a \ != a' \\ -\ [\][a:=w] = error \end{array}$$

- (m/r) (a) = m(a) si il existe x tq r(x) = @a
- error sinon

Jugement sémantique

- $Expression: r, m \vdash e \leadsto v$
- Instructions: $r, m \vdash s \leadsto m'$
- $D\'{e}clarations: r, m \vdash d \leadsto r', m'$
- Suite de commandes/Bloc: r, m \vdash [cs] \rightsquigarrow m'

Constantes

- r, m \vdash true $\leadsto \#t$
- r, m \vdash false $\rightsquigarrow \#f$
- $r, m \vdash n \rightsquigarrow \#n (n : num)$

Opérateurs booléens

- $r, m \vdash (not e) \leadsto \#f \text{ si } r, m \vdash e \leadsto \#t$
- $r, m \vdash (not e) \leadsto \#t \text{ si } r, m \vdash e \leadsto \#f$
- r, m \vdash (or e_1 e_2) \leadsto #t si r, m \vdash $e_1 \leadsto$ #t
- r, m \vdash (or e_1 e_2) \leadsto #t si r, m \vdash e_1 \leadsto #f et r, m \vdash e_2 \leadsto #t
- $\mathbf{r}, \mathbf{m} \vdash (\mathbf{or} \ e_1 \ e_2) \leadsto \# \mathbf{f} \ \mathbf{si} \ \mathbf{r}, \mathbf{m} \vdash e_1 \leadsto \# \mathbf{f} \ \mathbf{et} \ \mathbf{r}, \mathbf{m} \vdash e_2 \leadsto \# \mathbf{f}$
- r, m \vdash (and $e_1 \ e_2$) \rightsquigarrow #f si r, m \vdash $e_1 \rightsquigarrow$ #f
- r, m \vdash (or e_1 e_2) \leadsto #f si r, m \vdash e_1 \leadsto #t et r, m \vdash e_2 \leadsto #f
- $r, m \vdash (or e_1 e_2) \leadsto \#t \text{ si } r, m \vdash e_1 \leadsto \#t \text{ et } r, m \vdash e_2 \leadsto \#t$

Opérateurs arithmétiques

- r, m \vdash (add $e_1 e_2$) $\rightsquigarrow v_1 + v_2$ si r, m $\vdash e_1 \rightsquigarrow v_1$ et r, m $\vdash e_2 \rightsquigarrow v_2$
- r, m \vdash (sub e_1 e_2) $\rightsquigarrow v_1$ v_2 si r, m $\vdash e_1 \rightsquigarrow v_1$ et r, m $\vdash e_2 \rightsquigarrow v_2$
- r, m \vdash (mul e_1 e_2) \leadsto v_1 * v_2 si r, m \vdash e_1 \leadsto v_1 et r, m \vdash e_2 \leadsto v_2
- \mathbf{r} , $\mathbf{m} \vdash (\operatorname{div} e_1 e_2) \leadsto v_1 / v_2 \operatorname{si} \mathbf{r}$, $\mathbf{m} \vdash e_1 \leadsto v_1 \operatorname{et} \mathbf{r}$, $\mathbf{m} \vdash e_2 \leadsto v_2 \operatorname{et} v_2 \stackrel{!}{=} 0$

Instructions

- $r, m \vdash (SET \times e) \leadsto m[a := v] \text{ si } r, m \vdash e \leadsto v \text{ et } r(x) = @a$
- r, m \vdash (IF b e_1 e_2) \rightsquigarrow m' si r, m \vdash b \rightsquigarrow #t et r, m \vdash $e_1 \rightsquigarrow$ m'
- r, m \vdash (IF b e_1 e_2) \rightsquigarrow m' si r, m \vdash b \rightsquigarrow #f et r, m \vdash e_2 \rightsquigarrow m'
- r, m \vdash (WHILE e b) \leadsto m si r, m \vdash e \leadsto #f
- r, m \vdash (WHILE e b) \leadsto m''si r, m \vdash e \leadsto #t et r, m \vdash b \leadsto m' et r, m' \vdash (WHILE e b) \leadsto m"

Suite de commandes

- Si s une instruction: $r, m \vdash s; cs \leadsto r, m' \vdash s \leadsto m'$
- Si d \equiv CONST x e: r, m \vdash d; cs \leadsto r[x := #v], m si r, m \vdash e \leadsto #v
- Si d \equiv VAR x: r, m \vdash d; cs \leadsto r[x := @a], m[a = new]

Bloc

•
$$r, m \vdash [cs] \leadsto r, (m'/r) si r, m \vdash cs \leadsto r', m'$$

Typage

Distinguer les identificateurs modifiables des immuables.

Introduction du type "pointeur vers": @t "pointeur vers t"

- Déclaration/suite de commandes
 - $G \vdash (VAR \ c \ t; cs) : void si G, x:@t \vdash cs : void$
- Expression

$$-G$$
, x:@t \vdash x:t $\mid\mid G \vdash$ x:t si $G(x) = t$ ou $G(@x) = t$

- Affectation
 - $-G \vdash (SET \times e) : void si G \vdash x:@t et G \vdash e:t$

Sémantique

•
$$r, m \vdash (SET \times e) \leadsto m[r(x) := v] \text{ si } r, m \vdash e \leadsto v$$

Cours 4

Retour sur la portée lexicale

À la fin de l'exécution du programme, x vaut 1. Si on retire /r de la règle de sémantique du bloc, on peut avoir le programme suivant où x vaut 4:

```
VAR x int;
IF (true)
     [ SET x 3 ]
     [ ... ]
SET x (x + 1)
]
```

Fonctions et procédures

- Fonctions: abstractions vis-à-vis d'une expression
- Procédures: abstractions vis-à-vis d'un bloc d'expressions

Fonction

- Domaine de départ, codomaine d'arrivée.
- Abstraction d'une expression avec ses variables associées
- En APS: [x:int] (x + 1): expression fonctionnelle.

Syntaxe

Typage

- Si G; $x_1:t_1;x_2:t_2;...;x_n:t_n\vdash \text{e:t alors }G\vdash [x_1:t_1;...;x_n:t_n]\text{e}:(t_1*...*t_n)->\text{t}$
- Si G \vdash e:(t1*..*tn)->t et G \vdash e1:t1 .. G \vdash en:tn alors G \vdash (e e1 .. en):t

Sémantique

Ajouter les définitions des expressions fonctionnelles, des valeurs fonctionnelles (=> fermetures).

$$(([x_1:t_1;..;x_n:t_n]e') e_1 ... e_n)$$

Fermeture: (notation)
$$< e, r + [x_1 = \#]..[x_n:\#] >$$

Sémantique de la fermeture: r, m
$$\vdash [x_1:t_1;..;x_n:t_n]$$
e \leadsto < e, r + $[x_1=\#]..[x_n:\#]$ >

Application - Liaison statique

```
• Si r, m \vdash e \leadsto < e', r" + [x_1 = \#]..[x_n = \#] >
• et r, m \vdash e<sub>1</sub> \leadsto v<sub>1</sub> .. r, m \vdash e<sub>n</sub> \leadsto v<sub>n</sub>
• et r'[x_1 = v_1]..[x_n = v_n], m \vdash e' \leadsto v
• alors r, m \vdash (e e<sub>1</sub>..e<sub>n</sub>) \leadsto v

[ CONST f [x:int] (add x 1)
CONST g [x:int] (add (f x) 1)
CONST f [x:int] (add x 3)
VAR r int
SET r (g 42)
]
```

Liaison statique: la valeur de g est fabriquée au moment de sa déclaration, g capture la valeur de f à l'instant de la déclaration de g. La valeur de r est donc 44.

Procédures

```
Avec d = PROC p [x_1 : t_1; ..; x_n : t_n] blk
```

Syntaxe

Typage Si G; p: $t_1 * ... * t_n$ ->void \vdash cs : void et G \vdash blk : void alors G \vdash d;cs : void

Appel Si G \vdash p: $t_1*..*t_n$ -> void et G \vdash $e_1:t_1$.. G \vdash $e_n:t_n$ alors G \vdash CALL p $e_1...e_n:$ void

Sémantique

```
Fermeture "procédurale": < blk, \mathbf{r} + [x_1 = \#]..[x_n = \#] >
Si \mathbf{r} [\mathbf{p} = < blk, \mathbf{r} + [x_1 = \#]..[x_n = \#] >], \mathbf{m} \vdash \mathbf{cs} \leadsto (\mathbf{r}', \mathbf{m}') alors \mathbf{r}, \mathbf{m} \vdash \mathbf{d}; \mathbf{cs} \leadsto (\mathbf{r}', \mathbf{m}')
```

Application - Liaison statique

- Si $r(p) = \langle blk, r + [x_1 = \#]..[x_n = \#] \rangle$
- et r, m $\vdash e_1 \leadsto v_1$.. r, m $\vdash e_n \leadsto v_n$
- et r' $[x_1 = v_1..x_n = v_n]$, m \vdash blk \rightsquigarrow m'
- alors r, m \vdash (CALL $pe_1...e_n$) \rightsquigarrow m'

Cours 5

L'instruction RETURN

Typage

- Expression:
 - $-Si G \vdash e:t \ alors G \vdash (RETURN e):t$
- Séquences: soit s != (RETURN e)
 - Si G \vdash void et G \vdash cs:t alors G \vdash (s;cs):t
 - $-Si G \vdash s:t avec t != void et G \vdash cs:t alors G \vdash (s; cs):t$
- $Si \ G \vdash e:t \ et \ G \vdash x:t \ alors \ G \vdash (SET \ x \ e):void$
- $Si \ G \vdash$ e:bool $et \ G \vdash$ cs:t $alors \ G \vdash$ (WHILE e cs):t
- $Si \ G \vdash e:bool \ et \ G \vdash cs1:t \ et \ G \vdash cs2:t \ alors \ G \vdash (IF \ e \ cs1 \ cs2):t$

Sémantique

Instruction:

- une non valeur
- une valeur v

On note ω pour v ou \emptyset .

Relation sémantique des instructions:

- Env, Mem, Stat -> (Val, Mem)
- r, m \vdash s \leadsto (ω , m')

```
RETURN Si r, m \vdash e \leadsto v alors r, m \vdash (RETURN e) \leadsto (v, m)
```

Affectation Si r, m
$$\vdash$$
 (SET x e) \rightsquigarrow (\emptyset , m[r(x):=v])

Boucle

- Si r, m \vdash e \leadsto #f alors r, m \vdash (WHILE e blk) \leadsto (\emptyset , m)
- Si r, m \vdash e \leadsto #t alors
 - − Si r, m ⊢ blk \leadsto (∅, m') et r, m' ⊢ (WHILE e blk) \leadsto (ω, m") alors r, m ⊢ (WHILE e blk) \leadsto (ω, m")
 - Si r, m \vdash blk \rightsquigarrow (v, m') alors r, m \rightsquigarrow (WHILE e blk) \rightsquigarrow (v, m')

Alternative

- Si r, m \vdash e \leadsto #t et r, m \vdash blk1 \leadsto (ω , m') alors r, m \vdash (IF e blk1 blk2) \leadsto (ω , m')
- Si r, m \vdash e \leadsto #f et r, m \vdash blk2 \leadsto (ω , m') alors r, m \vdash (IF e blk1 blk2) \leadsto (ω , m')

Suites de commandes Pour toute instruction s et toute continuation cs: * Si r, m \vdash s \leadsto (\emptyset , m') et r, m' \vdash cs \leadsto (ω , m") alors r, m \vdash (s;cs) \leadsto (ω , m") * Si r, m \vdash s \leadsto (v, m') alors r, m \vdash (s;cs) \leadsto (v, m')

Déclarations

- Si d \equiv CONST x e
 - Si r, m \vdash e \leadsto v et r[x=v], m \vdash cs \leadsto (ω , m') alors r, m \vdash (d;cs) \leadsto (ω , m')
- Si $d \equiv VAR \times t$
 - Si r[x=a], m[x=new] \vdash cs \rightsquigarrow (ω , m') alors r, m \vdash (d;cs) \rightsquigarrow (ω , m')

Fonctions procédurales

Syntaxe

Typage

• Expression/application:

```
- Si G \vdash f:void->t alors G \vdash (f):t
```

• Déclaration/suite de commandes:

```
- d ≡ FUN f t [ ] blk

* Si G[f:void->t] ⊢ cs:t' et G ⊢ blk:t alors G ⊢ (d;cs):t'

- d ≡ FUN f t [x_1:t_1;...x_n:t_n] blk

* Si G[f:t_1*...*t_n] ⊢ cs:t' et G[x_1:t_1;...;x_n:t_n] ⊢ blk:t alors G ⊢ (d;cs):t'
```

Sémantique

Relation sémantique des expressions: r, m \vdash e \leadsto (v, m')

Valeur d'une fonction procédurale: < bd, r+[x1=#] >

Application

• Si

```
\begin{array}{l} - \ r, \ m \vdash e \leadsto (< r' + [x1 = \#..xn = \#] >, \ m') \\ - \ r, \ m' \vdash e1 \leadsto (v1, \ m1) \ et \ .. \ et \ r, \ mn-1 \vdash en \leadsto (vn, \ mn) \\ - \ r'[x1 = v1]..[xn = vn], \ mn \vdash bd \leadsto m" \end{array}
```

• alors r, m \vdash (e e1 .. en) \rightsquigarrow (v, m")

Fonctions récursives

Syntaxe

Sémantique

Valeur d'une fonction récursive: < bd, $r+[f=!][x_1=\#]..[x_n=\#] >$

Application

• Si

$$\begin{array}{l} -\text{ r, m} \vdash \text{e} \leadsto (<\text{r'}+[\text{f=!}][x_1=\#]..[x_n=\#]>,\text{ m'}) \\ -\text{ r, m'} \vdash e_1 \leadsto (v_1,\,m_1) \text{ et .. et r, } m_n-1 \vdash \text{en} \leadsto (v_n,\,m_n) \\ -\text{ r'}[\text{f=< bd, r+[f=!]}[\text{x1=\#}]..[\text{xn=\#}]>][x_1=v_1]..[x_n=v_n],\,m_n \vdash \text{bd} \\ \leadsto \text{m''} \end{array}$$

• alors r, m \vdash (e $e_1 \dots e_n$) \leadsto (v, m")