

UE INF404 - Projet Logiciel

Calculatrice : étape 3

Analyse d'une expression arithmétique générale

L2 Informatique

Année 2023 - 2024

Rappel des précédents épisodes (1)

Ecrire un interpréteur d'expressions arithmétiques

Version 1 = “Expressions Arithmétiques Simples” (EAS)

- les opérandes sont des entiers
- opérateurs arithmétiques usuels (+, -, *, /)
- pas de priorités (évaluation de gauche à droite)

Exemples :

$25 + 2$	\rightsquigarrow	27
$25 - 4 * 2$	\rightsquigarrow	42
25	\rightsquigarrow	25
$25 + *2$	\rightsquigarrow	erreur !
-25	\rightsquigarrow	erreur !
$25 \# 2$	\rightsquigarrow	erreur !
$25/0$	\rightsquigarrow	erreur !

Rappel des précédents épisodes (1)

Ecrire un interpréteur d'expressions arithmétiques

Version 1 = “Expressions Arithmétiques Simples” (EAS)

- les opérandes sont des entiers
- opérateurs arithmétiques usuels (+, -, *, /)
- pas de priorités (évaluation de gauche à droite)

Exemples :

$25 + 2$	\rightsquigarrow	27
$25 - 4 * 2$	\rightsquigarrow	42
25	\rightsquigarrow	25
$25 + *2$	\rightsquigarrow	erreur !
-25	\rightsquigarrow	erreur !
$25 \# 2$	\rightsquigarrow	erreur !
$25/0$	\rightsquigarrow	erreur !

Rappel des précédents épisodes (1)

Ecrire un interpréteur d'expressions arithmétiques

Version 1 = “Expressions Arithmétiques Simples” (EAS)

- les opérandes sont des entiers
- opérateurs arithmétiques usuels (+, -, *, /)
- pas de priorités (évaluation de gauche à droite)

Exemples :

$25 + 2$	\rightsquigarrow	27
$25 - 4 * 2$	\rightsquigarrow	42
25	\rightsquigarrow	25
$25 + *2$	\rightsquigarrow	erreur !
-25	\rightsquigarrow	erreur !
$25 \# 2$	\rightsquigarrow	erreur !
$25/0$	\rightsquigarrow	erreur !

Rappel des précédents épisodes (2)

Version 2 : EAEP

Expressions Arithmétiques Entièrement Parenthésées

↪ écrire **toute opération** entre parenthèses ...

Exemples :

25	↪	25
(25 + 2)	↪	27
((25 - 4) * 2)	↪	42
(25 - (4 * 2))	↪	17
(25 \$ 2)	↪	erreur! (<i>lexicale</i>)
25 + *2	↪	erreur! (<i>syntaxique</i>)
25 + 4 * 2)	↪	erreur! (<i>syntaxique</i>)
25 + (4 * 2)	↪	erreur! (<i>syntaxique</i>)
(25/(5 - 5))	↪	erreur! (<i>semantique</i>)

Rappel des précédents épisodes (2)

Version 2 : EAEP

Expressions Arithmétiques Entièrement Parenthésées

↪ écrire **toute opération** entre parenthèses ...

Exemples :

25	↪	25
(25 + 2)	↪	27
((25 - 4) * 2)	↪	42
(25 - (4 * 2))	↪	17
(25 \$ 2)	↪	erreur! (<i>lexicale</i>)
25 + *2	↪	erreur! (<i>syntaxique</i>)
25 + 4 * 2)	↪	erreur! (<i>syntaxique</i>)
25 + (4 * 2)	↪	erreur! (<i>syntaxique</i>)
(25/(5 - 5))	↪	erreur! (<i>semantique</i>)

Langage des EAEP

Lexique

- ENTIER = chiffre.(chiffre)*
- OPERATEUR = '+' + '-' + '*' + '/'
- PARO = '(' et PARF = ')'

Syntaxe

langage **non régulier**

Ne peut être décrite par un automate \Rightarrow **grammaire**

$exp \rightarrow eaep \text{ FIN_SEQUENCE}$
 $eaep \rightarrow \text{ENTIER}$
 $eaep \rightarrow \text{PARO } eaep \text{ op } eaep \text{ PARF}$
 $op \rightarrow \text{PLUS}$
 $op \rightarrow \text{MOINS}$
 $op \rightarrow \text{MUL}$

Langage des EAEP

Lexique

- ENTIER = chiffre.(chiffre)*
- OPERATEUR = '+' + '-' + '*' + '/'
- PARO = '(' et PARF = ')'

Syntaxe

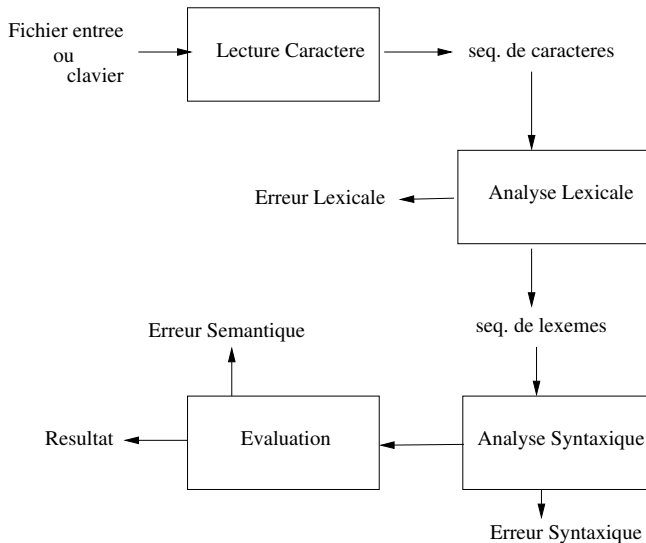
langage **non régulier**

Ne peut être décrite par un automate \Rightarrow **grammaire**

$$\begin{aligned} exp &\rightarrow eaep \text{ FIN_SEQUENCE} \\ eaep &\rightarrow \text{ENTIER} \\ eaep &\rightarrow \text{PARO } eaep \text{ op } eaep \text{ PARF} \\ op &\rightarrow \text{PLUS} \\ op &\rightarrow \text{MOINS} \\ op &\rightarrow \text{MUL} \end{aligned}$$

Structure de la calculatrice

Quatre composants/modules principaux ...



La suite ? ...Expressions Arithmétiques Générales (EAG)

Généralisation : priorités, donc parenthèses ...

ex : $5 + 3 * 4 = 17$ $(5 + 3) * 4 = 32$

Nouveau langage

- modifier la syntaxe :
règle de “bonne imbrication des parenthèses”
...mais parenthèses non systématiques ...!
⇒ définir une grammaire des EAG ?
- modifier la sémantique :
règles de calcul avec priorité des opérateurs

Conséquences en terme de programmation ?

- ① analyse syntaxique : à partir de la grammaire ?
- ② évaluation ?

La suite ? ...Expressions Arithmétiques Générales (EAG)

Généralisation : priorités, donc parenthèses ...

ex : $5 + 3 * 4 = 17$ $(5 + 3) * 4 = 32$

Nouveau langage

- modifier la syntaxe :
règle de “bonne imbrication des parenthèses”
...mais parenthèses non systématiques ...!
⇒ définir une grammaire des EAG ?
- modifier la sémantique :
règles de calcul avec priorité des opérateurs

Conséquences en terme de programmation ?

- 1 analyse syntaxique : à partir de la grammaire ?
- 2 évaluation ?

La suite ? ...Expressions Arithmétiques Générales (EAG)

Généralisation : priorités, donc parenthèses ...

ex : $5 + 3 * 4 = 17$ $(5 + 3) * 4 = 32$

Nouveau langage

- modifier la syntaxe :
règle de “bonne imbrication des parenthèses”
...mais parenthèses non systématiques ...!
⇒ définir une grammaire des EAG ?
- modifier la sémantique :
règles de calcul avec priorité des opérateurs

Conséquences en terme de programmation ?

- ① analyse syntaxique : à partir de la grammaire ?
- ② évaluation ?

Retour sur la grammaire des EAEP

$$\begin{aligned}exp &\rightarrow eaep \text{ FIN_SEQUENCE} \\eaep &\rightarrow \text{ENTIER} \\eaep &\rightarrow \text{PARO } eaep \text{ } op \text{ } eaep \text{ PARF} \\op &\rightarrow \text{PLUS} \\op &\rightarrow \text{MOINS} \\op &\rightarrow \text{MUL}\end{aligned}$$

- $\{\text{FIN_SEQUENCE}, \text{ENTIER}, \text{PARO}, \text{PARF}, \text{PLUS}, \text{etc}\} =$ lexèmes du langage = vocabulaire **terminal** de la grammaire
- $\{exp, op, eaep\} =$ vocabulaire **non-terminal** de la grammaire
- $exp =$ **axiome** de la grammaire
- Six **règles** dans cette grammaire ...

Grammaire (Hors-Contexte) : définition

Quadruplet $G = (V_t, V_n, z, P)$ tel que :

- V_t : vocabulaire terminal (les lexèmes)
ensemble fini de symboles, notés en majuscules
- V_n : vocabulaire non-terminal
ensemble fini de symboles, notés en minuscules ($V_t \cap V_n = \emptyset$)
- $Z \in V_n$: axiome
- P : ensemble des règles (ou productions)
 $P \subseteq V_n \times (V_t \cup V_n)^*$
Les éléments (X, α) de P seront notés $X \rightarrow \alpha$.

Reconnaitre une EAEP : exemples de dérivations

$exp \rightarrow eaep \text{ FIN_SEQUENCE}$
 $eaep \rightarrow \text{ENTIER}$
 $eaep \rightarrow \text{PARO } eaep \text{ op } eaep \text{ PARF}$
 $op \rightarrow \text{PLUS}$
 $op \rightarrow \text{MOINS}$
 $op \rightarrow \text{MUL}$

$exp \Rightarrow_G eaep \text{ FIN_SEQUENCE}$
 $\Rightarrow_G \text{ENTIER FIN_SEQUENCE}$

$exp \Rightarrow_G eaep \text{ FIN_SEQUENCE}$
 $\Rightarrow_G \text{PARO } eaep \text{ op } eaep \text{ PARF FIN_SEQUENCE}$
 $\Rightarrow_G \text{PARO ENTIER op } eaep \text{ PARF FIN_SEQUENCE}$
 $\Rightarrow_G \text{PARO ENTIER PLUS } eaep \text{ PARF FIN_SEQUENCE}$
 $\Rightarrow_G \text{PARO ENTIER PLUS ENTIER PARF FIN_SEQUENCE}$

Reconnaitre une EAEP : exemples de dérivations

$exp \rightarrow eaep \text{ FIN_SEQUENCE}$
 $eaep \rightarrow \text{ENTIER}$
 $eaep \rightarrow \text{PARO } eaep \text{ op } eaep \text{ PARF}$
 $op \rightarrow \text{PLUS}$
 $op \rightarrow \text{MOINS}$
 $op \rightarrow \text{MUL}$

$exp \Rightarrow_G eaep \text{ FIN_SEQUENCE}$
 $\Rightarrow_G \text{ENTIER FIN_SEQUENCE}$

$exp \Rightarrow_G eaep \text{ FIN_SEQUENCE}$
 $\Rightarrow_G \text{PARO } eaep \text{ op } eaep \text{ PARF FIN_SEQUENCE}$
 $\Rightarrow_G \text{PARO ENTIER op } eaep \text{ PARF FIN_SEQUENCE}$
 $\Rightarrow_G \text{PARO ENTIER PLUS } eaep \text{ PARF FIN_SEQUENCE}$
 $\Rightarrow_G \text{PARO ENTIER PLUS ENTIER PARF FIN_SEQUENCE}$

Reconnaitre une EAEP : exemples de dérivations

$exp \rightarrow eaep \text{ FIN_SEQUENCE}$
 $eaep \rightarrow \text{ENTIER}$
 $eaep \rightarrow \text{PARO } eaep \text{ op } eaep \text{ PARF}$
 $op \rightarrow \text{PLUS}$
 $op \rightarrow \text{MOINS}$
 $op \rightarrow \text{MUL}$

$exp \Rightarrow_G eaep \text{ FIN_SEQUENCE}$
 $\Rightarrow_G \text{ENTIER FIN_SEQUENCE}$

$exp \Rightarrow_G eaep \text{ FIN_SEQUENCE}$
 $\Rightarrow_G \text{PARO } eaep \text{ op } eaep \text{ PARF FIN_SEQUENCE}$
 $\Rightarrow_G \text{PARO ENTIER op } eaep \text{ PARF FIN_SEQUENCE}$
 $\Rightarrow_G \text{PARO ENTIER PLUS } eaep \text{ PARF FIN_SEQUENCE}$
 $\Rightarrow_G \text{PARO ENTIER PLUS ENTIER PARF FIN_SEQUENCE}$

Dérivations

$G = (V_t, V_n, Z, P)$ une grammaire hors-contexte, $w, w' \in (V_t \cup V_n)^*$.

- w' est un dérivé direct par G de w , noté $w \Rightarrow_G w'$ ssi :
il existe $X \in V_n$, il existe $u, t, v \in (V_t \cup V_n)^*$ tels que

$$w = uXv \text{ et } w' = utv \text{ et } X \rightarrow t \in P.$$

- w' est un dérivé par G de w , noté $w \Rightarrow_G^* w'$, ssi :
il existe $w_0, w_1, \dots, w_n \in (V_t \cup V_n)^*$ tels que

$$w_0 = w, w_n = w' \text{ et pour tout } i \text{ } w_{i+1} \text{ est un dérivé direct de } w_i.$$

Remarque : \Rightarrow_G^* est la fermeture transitive de \Rightarrow_G .

Langages hors-contextes

$G = (V_t, V_n, Z, P)$ une grammaire hors-contexte.

Le langage défini par G (ou reconnu par G) est l'ensemble $L(G)$ tel que :

$$L(G) = \{w \in V_t^* \mid Z \Rightarrow_G^* w\}$$

Un langage L est dit **hors-contexte** s'il existe une grammaire hors-contexte qui le reconnaît.

Remarque : Tout langage régulier est un langage hors-contexte, il existe des langages hors-contexte qui ne sont pas réguliers.

Langages hors-contextes

$G = (V_t, V_n, Z, P)$ une grammaire hors-contexte.

Le langage défini par G (ou reconnu par G) est l'ensemble $L(G)$ tel que :

$$L(G) = \{w \in V_t^* \mid Z \Rightarrow_G^* w\}$$

Un langage L est dit **hors-contexte** s'il existe une grammaire hors-contexte qui le reconnaît.

Remarque : Tout langage régulier est un langage hors-contexte, il existe des langages hors-contexte qui ne sont pas réguliers.

Langages hors-contextes

$G = (V_t, V_n, Z, P)$ une grammaire hors-contexte.

Le langage défini par G (ou reconnu par G) est l'ensemble $L(G)$ tel que :

$$L(G) = \{w \in V_t^* \mid Z \Rightarrow_G^* w\}$$

Un langage L est dit **hors-contexte** s'il existe une grammaire hors-contexte qui le reconnaît.

Remarque : Tout langage régulier est un langage hors-contexte, il existe des langages hors-contexte qui ne sont pas réguliers.

Analyse Syntaxique

Soit $G = (V_t, V_n, Z, P)$ une grammaire hors-contexte,
soit ω un séquence de V_t^* .

Analyse syntaxique = déterminer si $\omega \in L(G)$?

Contraintes pratiques :

- algorithme **linéaire** en fonction de $|\omega|$
- accès **séquentiel** à ω , de gauche à droite

⇒ possible uniquement pour un **sous-ensemble strict** des langages hors-contexte

Problèmes posés :

- Est-ce que le **langage** considéré appartient à cette classe ?
- Si oui, comment construire (efficacement) un analyseur (efficace) ?

⇒ réponse en fonction de la **structure** de la **grammaire** considérée !

Analyse Syntaxique

Soit $G = (V_t, V_n, Z, P)$ une grammaire hors-contexte,
soit ω un séquence de V_t^* .

Analyse syntaxique = déterminer si $\omega \in L(G)$?

Contraintes pratiques :

- algorithme **linéaire** en fonction de $|\omega|$
- accès **séquentiel** à ω , de gauche à droite

⇒ possible uniquement pour un **sous-ensemble strict** des langages hors-contexte

Problèmes posés :

- Est-ce que le **langage** considéré appartient à cette classe ?
- Si oui, comment construire (efficacement) un analyseur (efficace) ?

⇒ réponse en fonction de la **structure** de la **grammaire** considérée !

Analyse Syntaxique

Soit $G = (V_t, V_n, Z, P)$ une grammaire hors-contexte,
soit ω une séquence de V_t^* .

Analyse syntaxique = déterminer si $\omega \in L(G)$?

Contraintes pratiques :

- algorithme **linéaire** en fonction de $|\omega|$
- accès **séquentiel** à ω , de gauche à droite

⇒ possible uniquement pour un **sous-ensemble strict** des langages hors-contexte

Problèmes posés :

- Est-ce que le **langage** considéré appartient à cette classe ?
- Si oui, comment construire (efficacement) un analyseur (efficace) ?

⇒ réponse en fonction de la **structure** de la **grammaire** considérée !

Dans la suite ...

- On considère une technique d'analyse dite **descendante** :
 - ▶ produire ω à partir de Z
 - ▶ produire une dérivation gauche $Z \Rightarrow_G^* \omega$
- Cette technique se programme **directement à partir de la grammaire** par un ensemble de **procédures mutuellement récursives** (c.f “eaep”) :
A chaque non-terminal $X \rightsquigarrow$ une procédure $\text{Rec_}X$
- On doit donc :
 - ▶ Ecrire la grammaire de notre langage sous une forme qui facilite l'analyse descendante (**grammaire LL(1)**)
 - ▶ programmer l'**analyseur syntaxique**, à partir de la grammaire

Dans la suite ...

- On considère une technique d'analyse dite **descendante** :
 - ▶ produire ω à partir de Z
 - ▶ produire une dérivation gauche $Z \Rightarrow_G^* \omega$
- Cette technique se programme **directement à partir de la grammaire** par un ensemble de **procédures mutuellement récursives** (c.f “eaep”) :
A chaque non-terminal $X \rightsquigarrow$ une procédure $\text{Rec_}X$
- On doit donc :
 - ▶ Ecrire la grammaire de notre langage sous une forme qui facilite l'analyse descendante (**grammaire LL(1)**)
 - ▶ programmer l'**analyseur syntaxique**, à partir de la grammaire

Ecrire un analyseur à partir de la grammaire ?

Exemple des EAEP

eaep → ENTIER

eaep → PARO *eaep op eaep* PARF

```
Rec_eaep =  
  selon LC.nature  
    cas ENTIER : Avancer  
    cas PARO : Avancer ; Rec_eaep ; Rec_op ; Rec_eap ;  
      si LC.nature = PARF alors Avancer sinon Erreur  
    autre : Erreur  
fin
```

op → PLUS

op → MOINS

op → MUL

```
Rec_op =  
  selon LC.nature  
    cas PLUS, MUL, MOINS : Avancer  
    autre : Erreur  
fin
```

Ecrire un analyseur à partir de la grammaire ?

Exemple des EAEP

eaep → ENTIER

eaep → PARO *eaep op eaep* PARF

```
Rec_eaep =  
  selon LC.nature  
    cas ENTIER : Avancer  
    cas PARO : Avancer ; Rec_eaep ; Rec_op ; Rec_eap ;  
      si LC.nature = PARF alors Avancer sinon Erreur  
    autre : Erreur  
fin
```

op → PLUS

op → MOINS

op → MUL

```
Rec_op =  
  selon LC.nature  
    cas PLUS, MUL, MOINS : Avancer  
    autre : Erreur  
fin
```

Expressions Arithmétiques Générales (EAG)

→ parenthèses non obligatoires, priorités entre les opérateurs

Exemples : 5 ; 5 + 3 ; 5 + 3 * 2 ; (5+3) * 2 ; etc.

Grammaire 1

eag → ENTIER

eag → *eag op eag*

eag → PARO *eag* PARF

op → PLUS

op → MOINS

op → MUL

Problèmes

- priorités des opérateurs ?
- écriture de l'analyseur ? récursion infinie ?

Modifier la grammaire des EAG ? (1)

→ prendre en compte les priorités \Rightarrow structurer la grammaire¹

Grammaire 2

```
eag  → seq_terme
seq_terme  → seq_terme op1 terme
seq_terme  → terme
terme  → seq_facteur
seq_facteur  → seq_facteur op2 facteur
seq_facteur  → facteur
facteur  → ENTIER
facteur  → PARO eag PARF
op1  → PLUS
op1  → MOINS
op2  → MUL
```

1. cf. prochain cours !

Modifier la grammaire des EAG ? (2)

→ éliminer la “récursivité à gauche”

Grammaire 3

<i>eag</i>	→	<i>seq_terme</i>
<i>seq_terme</i>	→	<i>terme suite_seq_terme</i>
<i>suite_seq_terme</i>	→	<i>op1 terme suite_seq_terme</i>
<i>suite_seq_terme</i>	→	ϵ
<i>terme</i>	→	<i>seq_facteur</i>
<i>seq_facteur</i>	→	<i>facteur suite_seq_facteur</i>
<i>suite_seq_facteur</i>	→	<i>op2 facteur suite_seq_facteur</i>
<i>suite_seq_facteur</i>	→	ϵ
<i>facteur</i>	→	<i>ENTIER</i>
<i>facteur</i>	→	<i>PARO eag PARF</i>
<i>op1</i>	→	<i>PLUS</i>
<i>op1</i>	→	<i>MOINS</i>
<i>op2</i>	→	<i>MUL</i>

Vérifier la syntaxe d'une EAG (1)

→ un algo récursif, **basé sur la grammaire**

eag → *seq_terme*

```
Rec_eag() =  
    Rec_seq_terme() ;
```

seq_terme → *terme suite_seq_terme*

```
Rec_seq_terme() =  
    Rec_terme() ; Rec_suite_seq_terme() ;
```


Vérifier la syntaxe d'une EAG (1)

→ un algo récursif, **basé sur la grammaire**

$$eag \rightarrow seq_terme$$

```
Rec_eag() =  
    Rec_seq_terme() ;
```

$$seq_terme \rightarrow terme \mid suite_seq_terme$$

```
Rec_seq_terme() =  
    Rec_terme() ; Rec_suite_seq_terme() ;
```

Vérifier la syntaxe d'une EAG (2)

lexèmes **directeurs** : indiquent la règle à dériver !

suite_seq_terme \rightarrow *op1 terme suite_seq_terme* [PLUS, MOINS]

suite_seq_terme \rightarrow ε

```
Rec_suite_seq_terme() =  
  selon LC().nature // LC est le lexeme_courant()  
    cas PLUS, MOINS :  
      Avancer() ; Rec_terme() ; Rec_suite_seq_terme()  
    autre : // Rien, epsilon !  
fin
```

Idem pour *Rec_seq_facteur()* et *Rec_suite_seq_facteur()* ...

Vérifier la syntaxe d'une EAG (2)

lexèmes **directeurs** : indiquent la règle à dériver !

suite_seq_terme \rightarrow *op1 terme suite_seq_terme* [PLUS, MOINS]

suite_seq_terme \rightarrow ε

```
Rec_suite_seq_terme() =  
  selon LC().nature // LC est le lexeme_courant()  
    cas PLUS, MOINS :  
      Avancer() ; Rec_terme() ; Rec_suite_seq_terme()  
    autre : // Rien, epsilon !  
fin
```

Idem pour *Rec_seq_facteur()* et *Rec_suite_seq_facteur()* ...

Vérifier la syntaxe d'une EAG (2)

lexèmes **directeurs** : indiquent la règle à dériver !

suite_seq_terme \rightarrow *op1 terme suite_seq_terme* [PLUS, MOINS]

suite_seq_terme \rightarrow ε

```
Rec_suite_seq_terme() =  
  selon LC().nature // LC est le lexeme_courant()  
    cas PLUS, MOINS :  
      Avancer() ; Rec_terme() ; Rec_suite_seq_terme()  
    autre : // Rien, epsilon !  
fin
```

Idem pour *Rec_seq_facteur()* et *Rec_suite_seq_facteur()* ...

Vérifier la syntaxe d'une EAG (3)

facteur → *ENTIER*

facteur → PARO *eag* PARF

```
Rec_facteur() =  
  selon LC().nature // LC est le lexeme_courant()  
    cas ENTIER : Avancer()  
    cas PARO : Avancer() ; Rec_eag() ;  
              si LC.nature = PARF alors Avancer sinon Erreur  
    autre : Erreur  
fin
```

op1 → PLUS

op1 → MOINS

```
Rec_op1 =  
  selon LC.nature  
    cas PLUS, MOINS : Avancer  
    autre : Erreur  
fin
```

Idem pour Rec_op2 ...

Vérifier la syntaxe d'une EAG (3)

facteur → *ENTIER*

facteur → PARO *eag* PARF

```
Rec_facteur() =  
  selon LC().nature // LC est le lexeme_courant()  
    cas ENTIER : Avancer()  
    cas PARO : Avancer() ; Rec_eag() ;  
              si LC.nature = PARF alors Avancer sinon Erreur  
    autre : Erreur  
fin
```

op1 → PLUS

op1 → MOINS

```
Rec_op1 =  
  selon LC.nature  
    cas PLUS, MOINS : Avancer  
    autre : Erreur  
fin
```

Idem pour Rec_op2 ...

Dans la suite (les prochains TPs) ...

- ❶ Ecrire l'analyseur syntaxique pour le langage *eag* à partir de la **Grammaire 3**.
- ❷ Etendre cet analyseur pour calculer la valeur de l'expression (après le prochain cours!)