

Compiladors (CL) GEI (2017-18)

Pràctica 1: Flex + Bison + Symtab

Objectiu

Fer la part frontal d'un compilador que interpreti programes senzills.

Aprenentatges

Utilització conjunta de Flex, Bison i Symtab.

Llenguatge font

Literals i comentaris:

- S'admeten literals enters (e.g. *123*, *-45*), reals (e.g. *3.1416*, *-0.5*) i cadenes (e.g. *"Hola"*).
- Les cadenes estan delimitades per dobles cometes ("), i no poden ocupar més d'una línia.
- Poden haver comentaris com en Java, i.e. comentaris que comencen amb *//* fins a final de línia, i comentaris que es poden estendre vàries línies, iniciats amb */** i finalitzats amb **/*.

Identificadors i expressions:

- Els identificadors i expressions tenen un tipus, que pot ser: enter, real o cadena.
- Les expressions aritmètiques poden estar formades per literals enters, literals reals, cadenes de caràcters, identificadors (de tipus no booleà), parèntesis, l'operador de concatenació de cadenes (+), operadors aritmètics unaris (+, -) i operadors aritmètics binaris (+, -, *, /, **mod**, **).
- L'ordre de precedència dels operadors aritmètics binaris és: potència (**) més precedència que producte, divisió i mòdul (**mod**), producte igual precedència que divisió i mòdul, aquests major precedència que suma i resta, i suma igual precedència que resta.
- Els operadors aritmètics unaris (manteniment i canvi de signe) tenen igual precedència que la suma i la resta binàries.

Sentències:

- Les sentències poden ser: expressions aritmètiques o assignacions.
- Les expressions aritmètiques i assignacions ocupen una línia.
- Les assignacions són del tipus
$$id := expressió$$
on l'expressió és una expressió aritmètica, i *id* és un identificador.
- En una assignació, el tipus de l'identificador *id* és igual al tipus del resultat de l'expressió corresponent.
- Si s'operen expressions aritmètiques del mateix tipus, el resultat és del mateix tipus. Operadors aritmètics disponibles: +, -, *, /, **, mod.

- Si s'operen aritmèticament (+, -, *, /) enters amb reals, el tipus resultant és real.
- Si es concatena (amb l'operador +) una expressió numèrica amb una cadena, el resultat és una cadena.
- Els identificadors que apareguin en una expressió qualsevol han d'haver estat prèviament inicialitzats, ja que cal conèixer els seus tipus.

Programa:

- La sintaxi d'un programa és:
llista_de_sentències

Opcions:

- Podeu afegir opcions diverses a la calculadora: taules/l·listes, registres, funcions predeterminades (e.g. sinus, longitud de cadena, etc.).

Sortides

- Les assignacions mostraran el nom, el seu tipus i el seu valor.
- Les expressions mostraran el seu tipus i el seu valor.
- Convé anar guardant en un arxiu de log cada producció de la gramàtica reconeguda, per així controlar el progrés i correcció de les anàlisis lèxica i sintàctica. També es pot guardar al mateix arxiu qualsevol altre missatge informatiu.
- Quan hi ha un error al codi font s'ha d'emetre un missatge d'error indicant la posició actual a l'arxiu font, i si l'error és lèxic, sintàctic o semàntic. Mentre més informació es doni dels errors, millor.

Exemple d'arxiu d'entrada

```
x := 2.3
i := 5
z := i * (x + i) - i / 7.2
s := "Hola"

// un comentari
x
i
z
s
17 + (3 * 1.0)

/*
* cometari multilínea
*/
i + i * 2
s + (s + i) + x
```

Lliurament

- Aquesta pràctica és individual.
- El lliurament es farà via Moodle, en les dates indicades al mateix.
- Cal lliurar un arxiu comprimit que contingui:
 - tot el codi font (sense arxius generats en el procés de compilació)
 - exemples i la seva sortida corresponent
 - scripts o Makefile per compilar-ho tot, executar el programa amb els exemples, i netejar-ho tot llevat dels arxius font i els d'exemple.
 - arxiu README (en format txt o pdf) amb instruccions per la compilació i execució, i descripció de tot allò que vulgueu destacar de la vostra pràctica (decisions de disseny, funcionalitat addicional, limitacions, etc.).
- El nom de l'arxiu lliurat ha de contenir el vostre nom i primer cognom.