

Assignatura 21742 Compiladors I

Curs 2017-2018

Enunciat de la pràctica

Aspectes generals

La pràctica podrà ser realitzada en grups de com a molt tres persones. La conformació de cada grup s'haurà de fer mitjançant l'apartat corresponent dins de la plana de l'assignatura a l'eina Aula Virtual.

La pràctica consisteix en el disseny i implementació d'un processador per un llenguatge de programació. Les tasques que haurà de realitzar el processador són les d'anàlisi lèxic, sintàctic i semàntic. S'haurà de contemplar necessàriament la implementació d'una taula de símbols i un gestor d'errors.

Funcionalitat del Processador

El processador que s'ha de desenvolupar ha de contemplar les següents funcionalitats:

- Haurà de ser capaç de processar el codi font subministrat en un arxiu de text.
- Haurà de generar una sèrie de fitxers com a resultat de la seva execució:
 - Fitxer de tokens: tots els tokens generats segons la seqüència d'entrada.
 - Taula de símbols: tota la informació de totes les dades introduïdes a la taula de símbols una vegada que s'hagin processat completament.
 - Parseig: relació de totes les regles utilitzades. Es tracta de l'arbre de derivació per al document d'entrada. El format del document serà `.dot`
 - Errors: si es detecten errors es generarà un document amb els errors detectats. Indicant per a cada error, la línia on s'ha detectat l'error, el tipus d'error (lèxic, sintàctic, semàntic) i un missatge explicatiu.

Característiques del llenguatge

La següent és la llista amb totes les característiques que ha de tenir el llenguatge. Hi ha algunes, les marcades amb un asterisc (*) que no són obligatòries. Altres, les marcades amb un nombre (2, 4), tenen una certa quantitat d'opcions, d'entre les que s'han de cobrir almenys el nombre indicat.

- Un cos general de programa on hi hagi d'haver els subprogrames, les declaracions i les instruccions del programa (de l'estil del `main` de *java* o *c++* o un apartat d'instruccions tipus *python*)
- Definició de subprogrames: funcions o procediments
- Tipus:
 - Enter
 - Cadena de caràcters
 - Lògic
- Valors
 - Declaració i ús de variables
 - Constants de qualsevulla dels tipus
- Operacions:
 - Assignació

- Condicional
- Selecció* múltiple (tipus switch)
- Bucles*
 - while
 - repeat until
 - for
- Crida a funcions
- Retorn de funcions
- Expressions aritmètiques i lògiques:
 - Fent ús de literals del tipus adient
 - Fent ús de constants i variables
- Operacions d'entrada i sortida
- Operadors⁶:
 - Aritmètics²: suma, resta, producte, divisió, mòdul
 - Relacionals²: igual, diferent, major, menor, major o igual, menor o igual
 - Lògics²: i, o, no
 - Especials*:
 - Pre increment/decrement
 - Post increment/decrement
 - Assignació i operació alhora és a dir la combinació de cada operador amb l'operació d'assignació (+=, ...)
 - Operació condicional (? :)

Presentació de la pràctica

S'haurà de lliurar una documentació descrivint la feina realitzada. Una vegada lliurada la pràctica s'haurà de realitzar una entrevista per discutir els diferents aspectes.

Elements d'avaluació:

- Documentació: correctament escrita, descrivint les tècniques utilitzades, el disseny i qualsevulla aspecte que es desitgi remarcar. No constarà dels llistats de codi font. Si per algun motiu especial es considera d'especial interès alguna part del codi, sí que es podrà adjuntar.
 - Per a la part de l'analitzador lèxic: tokens, gramàtica, autòmat, rutines semàntiques i errors
 - Per a la part de l'analitzador sintàctic: gramàtica, taules i mètodes de l'analitzador. Justificació de l'elecció del mètode d'anàlisi escollit.
 - Per a la part de l'analitzador semàntic: traducció dirigida per la sintaxi, amb les rutines semàntiques.
 - Per a la taula de símbols: descripció de l'estructura i organització.
- Codi font complet. Instruccions per a la correcta execució. La compilació del codi font o la interpretació no ha de tenir cap missatge d'error ni d'avís. L'execució del processador no ha de generar errors ni excepcions no controlades.
- Casos de prova, tant correctes³ com erronis³. Tots els casos s'han de poder reproduir
 - Pels casos correctes: el codi, els tokens, l'arbre d'anàlisi sintàctica, bolcat de la taula de símbols
 - Pels casos erronis: el codi, missatges d'error obtinguts