

Processamento de Linguagens e Compiladores

LCC (2º ano)

2º Exame

Data: 15 de Julho de 2009
Hora: 09:30

Dispõe de 3 horas para realizar este exame

1 Filtros de Texto e Analisadores Léxicos (3v)

Supondo que vai usar o gerador Flex para construir automaticamente os programas abaixo descritos, escreva a respectiva especificação:

- a) Um verificador que receba um texto qualquer e faça uma validação simples (*em termos de número*) aos sintagmas nominais: assim sempre que encontrar um artigo definido ou indefinido no plural (terminado por 's') deve verificar se o substantivo respectivo também está no plural (também termina por 's') (admita que o substantivo é a palavra que se segue). Sempre que tal não se verifique, deve enviar para a saída o par de palavras que infringe essa regra básica. Todo o restante texto, deve ser retirado da saída.
Como poderia melhorar o seu filtro, de modo a assegurar com maior probabilidade que a palavra a seguir ao artigo é mesmo um substantivo?
- b) Um filtro que leia um texto formato numa versão leve de \LaTeX e o converta para \LaTeX correcto. Nessa versão leve de formatação, podem ser usados comandos \LaTeX normais, os quais serão preservados, Além disso pode-se usar uma notação simplificada de acordo com as seguintes regras:
- As secções começam com o caracter '=' no início da linha (as sub-secções com dois '==' e assim sucessivamente); o título da secção será o resto do texto até fim-de-linha.
 - Texto para destacar será colocado entre dois caracteres iguais: entre underscores '_' para itálico; entre mais '+' para bold; e entre ponto-de-exclamação '!' para teletype.
 - Listas de itens estão entre '[' e ']' após uma linha que termine em ':'; cada item começa por: '-' se é um *itemize*; e '+' se é um *enumerate*.

2 Expressões Regulares e Autómatos (7v)

Considere as seguintes ERs:

$$\begin{aligned}e1 &= a (b c^* + c b^*) d \\e2 &= a b d + (a c b^* + a b c^+) d \\e3 &= (a + b)^* c \\e4 &= (a c)^* + (b c)^*\end{aligned}$$

e os seguintes Autómatos:

```

AD1 :: < {'a', 'b', 'c', 'd'}, {1,2,3,4,5}, 1, {5}, delta >
  sendo delta 1 'a' = 2
         delta 2 'a' = 2
         delta 2 'b' = 3
         delta 2 'c' = 4
         delta 3 'c' = 3
         delta 3 'd' = 5
         delta 4 'b' = 4
         delta 4 'd' = 5

AND2 :: < {'a', 'b', 'c'}, {1,2,3,4,5,6,7,8,9,10}, 1, {10}, delta >
  sendo delta 1 & = 2      delta 3 'a' = 5      delta 7 & = 2
         delta 1 & = 8      delta 4 'b' = 6      delta 7 & = 8
         delta 2 & = 3      delta 5 & = 7      delta 8 & = 9
         delta 2 & = 4      delta 6 & = 7      delta 9 'c' = 10

```

Responda, então, às seguintes questões:

- mostre que e_1 e e_2 são equivalentes.
- usando os respectivos Autômatos Deterministas (que terá de construir informalmente) justifique a não-equivalência de e_3 e e_4 .
- usando a respectiva *cadeia de derivação*, mostre que a frase "acbbd" pertence à linguagem gerada por e_2 .
- construa, aplicando as regras sistemáticas, um autômato não-determinista equivalente a e_4 .
- escreva a expressão regular que define a linguagem aceite pelo autômato AD1.
- diga justificando se a frase "abba" pertence à linguagem gerada por AND2.
- considerando o autômato AND2, calcule o ϵ -fecho do estado 1.

3 Desenho/especificação de uma Linguagem (3v)

No contexto de uma disciplina de mestrado, pretende-se uma linguagem para criação da lista de projectos oferecidos e também da respectiva alocação aos alunos.

Por cada projecto terá de ser indicado, além do seu número de referência, o título, a área (descrita por um ou mais termos-chave), a data da proposta, o ou os supervisores, a descrição, a lista de capacidades a adquirir.

Quanto à alocação, deve ser indicado o número do projecto e o nome e número do aluno, ou alunos, que lhe foram associados.

Escreva então uma Gramática Independente de Contexto, *GIC*, que especifique a Linguagem pretendida (note que o estilo da linguagem (mais ou menos verbosa) e o seu desenho são da sua responsabilidade).

4 Gramáticas, Linguagens, Parsing e Tradução (7v)

A gramática independente de contexto, *GIC*, abaixo apresentada, define uma linguagem específica para descrever as análises feitas a jogos de diferentes plataformas para disponibilizar no site de jogos Eurogamer.

O Símbolo Inicial é Eurogamer, os Símbolos Terminais são escritos em maiúsculas (pseudo-terminais) ou em maiúscula entre apostrofes (palavras-reservadas e sinais de pontuação), e a string nula é denotada por &; os restantes (sempre em minúsculas) serão os Símbolos Não-Terminais.

```

p0: Eurogamer    --> Analises
p1: Analises     --> Analise
p2:              | Analises ',,' Analise

```

```

p3: Analise      --> ANALISE Descricao COMENTARIOS Comentarios ';'
p4: Descricao   --> '{' Referencia Tipo Titulo Texto Jogo Plataforma Autor '}'
p5: Referencia  --> id
p6: Tipo        --> ANALISE
p7:             | PREVISAO
p8:             | PRIMEIRAS_IMPRESSOES
p9: Titulo      --> string
p10: Texto      --> string
p11: Jogo       --> string
p12: Plataforma --> PS3
p13:           | XBOX360
p14:           | WII
p15:           | PC
p16: Autor      --> string
p17: Comentarios --> Comentario
p18:           Comentarios ';' Comentario
p19: Comentario --> USER Username COMENTARIO string DATA Data
p20: Data       --> Ano '-' Mes '-' Dia
p21: Ano        --> num
p22: Mes        --> num
p23: Dia        --> num

```

Neste contexto e após analisar a *GIC* dada, responda às alíneas seguintes.

- a) escreva uma Frase válida da linguagem gerada pela *GIC* dada, mostrando a respectiva Árvore de Derivação.
- b) altere a gramática de modo a agrupar os comentários por utilizador.
- c) o par de produções p1/p2 define uma lista com recursividade à esquerda. Altere esse par para usar recursividade à direita e mostre, através das árvores de derivação, a diferença entre ambos os esquemas iterativos.
- d) escreva as funções de um parser RD-puro (recursivo-descendente) para reconhecer o Símbolo Comentario e seus derivados.
- e) construa o estado inicial do autómato LR(0) pra gramática dada e os estados que dele saiem.
- f) transforme a *GIC* dada numa **gramática tradutora**, *GT*, reconhecível pelo yacc, para:
 - calcular e imprimir: o número de análises existentes; e o número de comentários por análise.
 - identificar e listar por ordem alfabética os jogos da PS3 com análise.
 - verificar se não existem análises com a mesma referência.