

Instituto de Informática – Universidade Federal do Rio Grande do Sul

INF05005 Linguagens Formais e Autômatos - Trabalho Prático

Profa. Aline Villavicencio

1) Implemente os seguintes algoritmos de parsing:

- a) CYK (vale 33%);**
- b) Early (vale 33%);**
- c) Opcional: Chart-parser (2 pontos extra)**

para fazer o reconhecimento de uma gramática probabilística e

d) mostre a(s) árvore(s) de derivação gerada(s) (vale 33%).

Gramáticas probabilísticas estão descritas no capítulo 11 e (12) do livro:

- *Foundations of Statistical Natural Language Processing*. Chris Manning and Hinrich Schütze. MIT Press. Cambridge, MA: May 1999.**

E os algoritmos de reconhecimento são apresentados no livro *Linguagens Formais e Autômatos* de Paulo Blauth Menezes, e no livro *Natural Language Processing in PROLOG* do Gerald Gazdar, Chris Mellish (1989), disponíveis no xeróx na pasta da disciplina. As implementações devem seguir estes algoritmos.

A interface é definida por cada grupo, e a entrada será lida de um arquivo, a ser disponibilizado no Moodle (exemplo-gramatica.txt). As gramaticas estarão na FNC.

O trabalho é para ser feito em grupos de no máximo 3 alunos, que devem enviar esta informação por email para pschreiner@inf.ufrgs.br até o dia 20/05/2009. A linguagem de programação a ser usada é de escolha do próprio grupo. As alocações dos grupos será feita em ordem inversa ao envio do email informando os integrantes, e será divulgada no Moodle.

A avaliação será feita com base em apresentações dos grupos, tanto do código fonte quanto do executável, que será testado com exemplos a serem providos durante as apresentações.

O atraso na entrega do trabalho significa redução na nota de 0,5 pontos por dia.

Dúvidas podem ser esclarecidas comigo ou com o Paulo Schreiner durante horario de monitoria.

Boa sorte.