

Instituto de Informática – Universidade Federal do Rio Grande do Sul

INF05005 Linguagens Formais e Autômatos - *Trabalho Prático*

Profa. Aline Villavicencio

Desenvolver um sistema *Gerador Randômico de Sentenças*, de acordo com a especificação abaixo:

1) (1 ponto) Usando o formato do arquivo gramatica.txt disponível no Moodle da disciplina, construa uma gramática probabilística para gerar diálogos em inglês de um tópico de interesse (seriado/novela/futebol/conferência,...). Uma descrição de gramática probabilística pode ser encontrada nos capítulos 11 e 12 do livro do Foundations of Statistical Language Processing de Manning e Schutze.

2) (5 pontos) Implementar um parser probabilístico seguindo o Algoritmo de CYK descrito no livro *Linguagens Formais e Autômatos* de Paulo Blauth Menezes, que aceite como entrada uma gramática no formato de gramatica.txt e uma sentença para reconhecer como parte da linguagem ou não, retornando as árvores de derivação correspondentes e suas probabilidades.

3) (4 pontos) Modifique o parser para que, recebendo como entrada uma gramática no formato de gramatica.txt, gere como saída textos usando a mesma de acordo com as probabilidades das regras (a cada “Enter” ou click do usuário, um novo texto diferente é gerado).

A interface para o sistema é definida por cada grupo, conforme especificação acima. Para a implementação qualquer linguagem de programação pode ser utilizada.

Os grupos deverão ter exatamente 3 integrantes, e os nomes dos integrantes deverão ser informados por email até o dia 15/05/2010 (ver abaixo).

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.

A entrega de cada um dos módulos deverá ser feita pelo Moodle da disciplina, através das salas de entrega apropriadas. Atraso nas data de entrega (de qualquer um dos módulos) resultará em desconto de 0.5 pontos por dia da semana até a entrega ser realizada.

Datas Importantes:

-Definição dos grupos: até o dia 15/05/2010. Enviar por email para: arssantos@inf.ufrgs.br com cc para avillavicencio@inf.ufrgs.br.

-Entrega da gramática especificada em 1): até o dia 21/05/2010 pelo Moodle

-Entrega do parser especificado em 2): até o dia 04/06/2010 pelo Moodle

- Data de entrega do sistema completo especificado em 3): até o dia 14/06/2010**
- Data de apresentações: nos dias 15, 17, 22 e 24 de junho de 2010 conforme escala a ser divulgada.**

Boa sorte.