

Instituto de Informática – Universidade Federal do Rio Grande do Sul
INF05005 Linguagens Formais e Autômatos - Trabalho Prático

Simplificação de GLCs, Forma Normal de Chomsky e Parsing de Linguagem Natural com CYK

Implemente os algoritmos de simplificação de GLCs, da Forma Normal de Chomsky e o algoritmo de parsing de Cocke-Younger-Kasami, seguindo exatamente as definições dos algoritmos do livro *Linguagens Formais e Autômatos* de Paulo Blauth Menezes. A implementação deve aceitar como entradas (1) um arquivo texto contendo uma gramática qualquer (2) e uma sentença. A gramática de entrada deve ser simplificada e transformada na Forma Normal de Chomsky, mostrando os resultados intermediários de cada uma das etapas de simplificação e de transformação na FNC.

Com a gramática resultante, use o algoritmo de CYK para fazer o reconhecimento da sentença de entrada e retornar as árvores de derivação caso seja reconhecida. Para cada sentença de entrada devem ser mostradas, em caso de aceitação, todas as árvores de derivação produzidas para a sentença.

A interface é definida por cada grupo e a linguagem de programação é livre. A entrada será lida de um arquivo que segue o formato do exemplo disponível no Moodle.

Os grupos deverão ter exatamente 3 integrantes, e os nomes dos integrantes informados por Moodle (**até 04/05/18**).

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

1) (1 ponto) Implemente um leitor para o arquivo de gramática:

- Entrada: arquivo de gramática;
- Saída: definição formal da gramática.

2) (2.5 pontos) Implemente os 3 algoritmos de simplificação de GLCs e o de transformação na Forma Normal de Chomsky descritos no livro *Linguagens Formais e Autômatos* de Paulo Blauth Menezes:

- Entrada: arquivo de gramática;
- Saída: gramática simplificada e na forma normal.

3) (4.5 pontos) Implemente o parser CYK:

- Entrada: uma gramática e uma sentença
- Saída: tabela triangular do algoritmo CYK, aceita/rejeita e as árvores de derivação da palavra, em caso de aceite.

4) (2 pontos) Apresentação do trabalho. A presença de todos os integrantes de cada grupo é obrigatória para a apresentação do trabalho. A ausência de algum integrante resultará em desconto na nota.

Apresentações terão cerca de 15 minutos e devem executar o programa, discutir as estruturas de dados usadas e os detalhes do algoritmo implementado.

A entrega do trabalho deverá ser feita pelo Moodle da disciplina, através da sala de entrega apropriada, e deverá ser feita através de um arquivo compactado contendo todos os arquivos necessários para a execução do programa, além de um arquivo “README.txt” contendo nome completo e número do cartão dos integrantes. Atraso nas datas de entrega resultará em desconto de 0.5 pontos por dia da semana até a entrega ser realizada. A nota é composta pela implementação (8 pontos máximo) e pela apresentação (2 pontos máximo), de modo que todos os integrantes do grupo devem estar presentes na apresentação.

Datas Importantes:

- 04/05/2018 - Confirmação dos grupos pelo Moodle.
- 11/05/2018 - Entrega do módulo de leitura da gramática pelo Moodle.
- 18/05/2018 - Entrega do módulo de simplificação da gramática e transformação na Forma Normal de Chomsky pelo Moodle.
- 08/06/2018 - Entrega do módulo do parser e da implementação completa (integração dos 3 módulos) pelo Moodle.
- 12/06/2018 – Início das apresentações conforme escala a ser divulgada.