LINGUAGENS FORMAIS E AUTÔMATOS

Prova Final – 18/04/2018 – Prof. Marcus Ramos

1ª Questão (1,5 ponto): Dada uma linguagem L qualquer, discorra sobre as estratégias que você poderia usar para: (i) provar que L é regular; (ii) provar que L não é regular; (iii) provar que L é livre de contexto; (iv) provar que L não é livre de contexto.

(i) apresentar uma gramática linear à direita, linear à esquerda, um autômato finito ou uma expressão regular, ou ainda usar as propriedades de fechamento das linguagens regulares;

(ii) derivar uma contradição a partir da aplicação do Pumping Lemma para as linguagens regulares;

(iii) apresentar uma gramática livre de contexto ou um autômato de pilha, ou ainda usar as propriedades de fechamento das linguagens livres de contexto;

(iv) derivar uma contradição a partir da aplicação do Pumping Lemma para as linguagens livres de contexto.

2ª Questão (1,5 ponto): Prove que a classe das linguagens regulares está contida na classe das linguagens livres de contexto, usando como argumento o fato de que autômatos de pilha podem simular autômatos finitos. Explique como esta simulação pode ser feita.

Seja L uma linguagem regular qualquer. Então, por definição, existe um autômato finito M que aceita L. No entanto, todo autômato finito pode ser simulado um algum autômato de pilha, bastando para isso que, em cada transição, o símbolo do topo da pilha (qualquer que seja ele), lido e removido no início da transição, seja reinserido no topo da pilha ao término da mesma. Logo, existe um autômato de pilha que aceita L, e L é também livre de contexto.

3ª Questão (1,5 ponto): Suponha que você não sabe que a linguagem $a^k b^k, k \geq 0$ é livre de contexto, e tente fazer uma prova de que ela não é livre de contexto usando o Pumping Lemma. Considere que n é a constante definida pelo Pumping Lemma e que a sentença escolhida é $a^n b^n$. Em que ponto, e por qual motivo, a sua prova falhará?

Supondo que a hipótese de que a linguagem seja livre de contexto seja verdadeira (o que de fato é), temos que $a^n b^n = uvwxy$, com $|vwx| \leq n$. Três são as possibilidades para a cadeia vwx:

- Contém apenas símbolos a (pelo menos um);
- Contém apenas símbolos b (pelo menos um);
- Contém símbolos a e símbolos b.

Logo, a cadeia vx tem as seguintes propriedades:

- Contém apenas símbolos a (pelo menos um);
- Contém apenas símbolos b (pelo menos um);
- Contém símbolos a e/ou símbolos b (pelo menos um).

Considere agora os casos acima e a cadeia uv^0wx^0y. Então:

- O primeiro caso gera um desbalanceamento e a cadeia resultante não pode pertencer à linguagem;
- Idem ao caso anterior;
- Se a cadeia contiver símbolos a e símbolos b simultaneamente, não será possível determinar as quantidades de cada um, e conseqüentemente não será possível afirmar que a cadeia uv^0wx^0y não pertence à linguagem. Portanto, este caso não pode ser contestado e não há como prosseguir com a prova. Logo, não é possível chegar numa contradição com a sentença proposta mas isto não nos permite provar nada. Nem que a linguagem é livre de contexto, nem que ela não é livre de contexto. A prova falha e deve-se buscar uma outra sentença ou então a prova de que a linguagem é livre de contexto (como de fato é), apresentando uma gramática livre de contexto que a gera ou então um autômato de pilha que a aceite.

4ª Questão (1,5 ponto): O Pumping Lemma para as linguagens livres de contexto pode ser usado para provar que uma linguagem é livre de contexto? Justifique a sua resposta.

Não pode, por dois motivos. O primeiro é que as linguagens livres de contexto de maior interesse são infinitas. Logo, a aplicação do quantificador universal para estas linguagens geraria a necessidade de provar que a propriedade é válida para todas as sentenças de cada linguagem, o que seria impossível do ponto de vista prático. Ainda que isto fosse possível, há um segundo impedimento: o Pumping Lemma para as linguagens livres de contexto é verificado não apenas para as linguagens desta classe, mas também para outras linguagens que não são livres de contexto. Logo, a tentativa de provar que uma linguagem é livre de contexto por meio da aplicação do PL poderia resultar num falso positivo.

5ª Questão (2 pontos): Prove que, para toda linguagem livre de contexto, existe um autômato de pilha que decide a mesma (isto é, é capaz de determinar se uma cadeia qualquer pertence ou não à linguagem, sem entrar em loop).

Conforme visto em sala de aula, é possível determinar o pertencimento de uma cadeia a uma linguagem livre de contexto usando uma gramática na Forma Norma de Chomsky (FNC) que gera a mesma. Supondo que m seja o maior número de regras de um mesmo símbolo não-terminal desta gramática, entre todos os não-terminais que ela possui, e que n seja o comprimento da cadeia a ser analisada, então é suficiente gerar m^{n} formas sentenciais para determinar o pertencimento. Por outro lado, sabemos que toda linguagem gerada por uma gramática livre de contexto pode ser aceita por algum autômato de pilha. Em particular, é possível construir um autômato de pilha não-determinístico com um único estado que simula as derivações mais à esquerda da gramática. Este autômato não entra em loop quando a gramática está na FNC, pois não existem regras unitárias. Logo, para toda linguagem livre de contexto existe um autômato de pilha que aceita a mesma sem entrar em loop.

6ª Questão (2 pontos): O uso da gramática na Forma Normal de Chomsky (FNC) é estritamente necessário para a prova do Pumping Lemma para as linguagens livres de contexto? Seria possível desenvolver uma prova similar sem no entanto fazer uso da FNC? Como isso poderia ser feito?

Sim, uso da FNC não é estritamente necessário para esta prova. Na verdade, basta apenas provar que uma sentença de comprimento suficientemente longo produz uma árvore (não necessariamente binária) com uma certa altura mínima, usando a gramática original. E que nesta altura mínima é possível identificar a repetição de um símbolo não-terminal. Especificamente, suponha que a maior quantidade de símbolos do lado direito de alguma regra da gramática seja m. Então, basta usar $n = m^k$ (ao invés de $n = 2^k$), onde k é o número de símbolos não-terminais da gramática em questão. O restante da prova procede da mesma forma.