CS320: Problems and Solutions for Day 4, Winter 2023

Problem 1 Let L be a language over alphabet $\{a, b\}$ with the following property:

For every word $w \in L$:
if $|w|$ is odd, then the middle symbol of w is a;
if $|w|$ is even, then w ends with bb.

Write a complete formal definition of a context-free grammar G that generates L. If such grammar G does not exist, explain why.

Answer: $G = \{V, \Sigma, P, S\}$, where:
$\Sigma = \{a, b\}$ is the set of terminals;
$V = \{S, E, O\}$ is the set of variables;
S is the start symbol;
and the set of productions P comprises:

$$\begin{aligned} S &\to O \mid E \\ O &\to aOa \mid aOb \mid bOa \mid bOb \mid a \\ E &\to aaE \mid abE \mid baE \mid bbE \mid bb \end{aligned}$$

Problem 2 Construct a context-free grammar G over alphabet $\{a, b, c\}$ that generates the language

$$L(G) = \{a^m b^n c^i \mid m + n < i\}$$

where m, n, i are non-negative integers.

Answer: Let $G = \{V, \Sigma, P, S\}$, where $\Sigma = \{a, b, c\}$ is the set of terminals; $V = \{S, A, B\}$ is the set of variables; S is the start symbol. The set of productions P comprises:

$$\begin{aligned} S &\to Sc \mid Ac \\ A &\to aAc \mid B \\ B &\to bBc \mid \lambda \end{aligned}$$

Problem 3 **(a)** Does there exist a pair of languages L_1 and L_2 such that all of the following three conditions hold?

- L_1 is regular;
- $L_2 \subseteq L_1$;
- L_2 is not regular, but is context-free.

If your answer to this part is "no" go to part **(d)**, else complete parts **(b)**–**(c)**.

Answer: Yes. In fact, for any alphabet Σ, the language Σ^* is regular, while every (non-regular) language over Σ is its subset. See parts **(b)**–**(c)** for a more interesting example where L_1 is infinite, with an infinite complement.

(b) Write a regular expression that defines L_1 (as in part **(a)**).

Answer:

$$\boldsymbol{a^* b^*}$$

(c) Write a complete formal definition of a context-free grammar that generates L_2 (as in part **(a)**). Describe L_2 briefly, using words and set-notation.

Answer: Let L_2 be the language of strings over $\{a, b\}$ where all a's precede all b's and the number of a's in the string equals the number of b's:

$$L_2 = \{a^n b^n \mid n \geq 0\}$$

L_2 is generated by the context-free grammar: $G = \{V, \Sigma, P, S\}$, where $\Sigma = \{a, b\}$ is the set of terminals; $V = \{S\}$ is the set of variables; S is the start symbol, and the production set P is:

$$S \to aSb \mid \lambda$$

The argument that shows that L_2 is not regular is almost identical to the one given in Problem **??**.

(d) Explain why such a pair of languages L_1 and L_2 (as in part **(a)**) does not exist.

Problem 4 Let L be the language generated by the context-free grammar $G = \{V, \Sigma, P, S\}$, where $\Sigma = \{a, b, c\}$, $V = \{S, A, B\}$, and the production set P is:

$$\begin{array}{l} S \to aSb \mid A \\ A \to cAb \mid B \\ B \to cb \end{array}$$

(a) Write 8 distinct strings that belong to L. If such strings do not exist, explain why.

(b) Write 8 distinct strings over alphabet $\{a, b, c\}$ that do not belong to L. If such strings do not exist, explain why.

Answer: Observe that:

$$L = \{a^m c^n b^{m+n} \mid m \geq 0, n \geq 1\}$$

whence:

$\in L$	$\notin L$
cb	b
$ccbb$	bc
$cccbbb$	ba
$acbb$	bca
$accbbb$	bcb
$acccbbbb$	aba
$aaacbbbb$	ab
$aaccbbbb$	abc

Problem 5 Write a complete formal definition of a context-free grammar G that generates language L, defined as follows:

$$L = \{a^m b^n ac \mid m \geq 0, n > m\}$$

If such a grammar does not exist, explain why.

Answer: $G = \{V, \Sigma, P, S\}$, where:
$\Sigma = \{a, b, c\}$, $V = \{S, A, B\}$, and P is:

$$\begin{array}{l} S \to Aac \\ A \to aAb \mid B \\ B \to bB \mid b \end{array}$$

Problem 6 Let:

$$L = \{a^m b^m \mid m \geq 0\}$$

Write a complete formal definition of a context-free grammar that generates the complement of L in $\{a, b\}^*$. If such grammar does not exist, prove it.

Answer: $G = \{V, \Sigma, P, S\}$, where:
$\Sigma = \{a, b\}$, $V = \{S_{ba}, S_<, S_>, A, B, D\}$,
and the production set P is:

$$\begin{array}{l} S \to S_{ba} \mid S_< \mid S_> \\ S_{ba} \to DbaD \\ D \to aD \mid bD \mid \lambda \\ S_< \to aS_<b \mid B \\ B \to bB \mid b \\ S_> \to aS_>b \mid A \\ A \to aA \mid a \end{array}$$

Problem 7 Let L_1 be the language defined by the regular expression:

$$(\boldsymbol{a} \cup \boldsymbol{b})^* \, \boldsymbol{c} \, (\boldsymbol{a} \cup \boldsymbol{b})^* \, \boldsymbol{c} \, (\boldsymbol{a} \cup \boldsymbol{b})^* \, \boldsymbol{c} \, (\boldsymbol{a} \cup \boldsymbol{b})^*$$

Let L_2 be the language generated by the context-free grammar $G = \{V, \Sigma, P, S\}$, where $\Sigma = \{a, b, c\}$, $V = \{S, A\}$, and the production set P is:

$$\begin{array}{l} S \to AAS \mid \lambda \\ A \to a \mid b \mid c \end{array}$$

(a) Write 5 distinct strings that belong to L_1 and do not belong to L_2 (belong to $L_1 \cap \overline{L_2}$). If such strings do not exist, explain why.

Answer: Observe that L_1 is the set of strings over alphabet $\{a, b, c\}$ that contain exactly 3 occurrences of letter c, while L_2 is the set of strings over alphabet $\{a, b, c\}$ that have even length.

$$ccc, cccaa, abccc, cacac, cbbcc$$

(b) Write 5 distinct strings that belong to L_2 and do not belong to L_1 (belong to $\overline{L_1} \cap L_2$). If such strings do not exist, explain why.

Answer:

$$\lambda, aa, ab, ac, abca$$

(c) Write 5 distinct strings that belong to L_1 and L_2 (belong to $L_1 \cap L_2$). If such strings do not exist, explain why.

Answer:

$$accc, cbccaa, abcacc, cacaca, cbbcca$$

(d) Write 5 distinct strings over alphabet $\{a, b, c\}$ that do not belong to L_1 and do not belong to L_2 (belong to $\overline{L_1} \cap \overline{L_2}$). If such strings do not exist, explain why.

Answer:

$$a, b, c, aaa, aac$$

Problem 8 **(a)** Let L_1 be the set of all strings over alphabet $\{a, b, d\}$ that do not contain the substring dba.

Write a complete formal definition of a context-free grammar G_1 that generates language L_1. If such a grammar does not exist, explain why.

Answer: $G_1 = \{V, \Sigma, P, S\}$, where:
$\Sigma = \{a, b, d\}$, $V = \{S, A, B\}$,
and the production set P is:

$$\begin{aligned}
&S \to aS \mid bS \mid dA \mid \lambda \\
&A \to aS \mid bB \mid dA \mid \lambda \\
&B \to bS \mid dA \mid \lambda
\end{aligned}$$

(b) Let L_2 be the set of all strings over alphabet $\{a, b\}$ that have even length or contain an even number of a's.

Write a complete formal definition of a context-free grammar G_2 that generates language L_2. If such a grammar does not exist, explain why.

Answer: $G_2 = \{V, \Sigma, P, S\}$, where:
$\Sigma = \{a, b\}$, $V = \{S, S_1, S_2, A, Z\}$,
and the production set P is:

$$\begin{aligned}
&S \to S_1 \mid S_2 \\
&S_1 \to S_1 S_1 \mid \lambda \mid ZZ \\
&Z \to a \mid b \\
&S_2 \to aA \mid bS_2 \mid \lambda \\
&A \to bA \mid aS_2
\end{aligned}$$

Problem 9 Let L_1 be the language defined by the regular expression:

$$(\boldsymbol{a} \cup \boldsymbol{ba} \cup \boldsymbol{ca})^* \, (\boldsymbol{b} \cup \boldsymbol{c})$$

Let L_2 be the language generated by the context-free grammar $G = \{V, \Sigma, P, S\}$, where $\Sigma = \{a, b, c\}$, $V = \{S, A, B\}$, and the production set P is:

$$\begin{aligned}
&S \to ABABA \\
&A \to AA \mid a \mid c \mid \lambda \\
&B \to b
\end{aligned}$$

(a) Write 5 distinct strings that belong to L_1 and do not belong to L_2 (belong to $L_1 \cap \overline{L_2}$). If such strings do not exist, explain why.

Answer:

$$ab, ac, aab, aac, cab$$

(b) Write 5 distinct strings that belong to L_2 and do not belong to L_1 (belong to $\overline{L_1} \cap L_2$). If such strings do not exist, explain why.

Answer:

$$bb, abb, cbb, aabb, ccbb$$

(c) Write 5 distinct strings that belong to L_1 and L_2 (belong to $L_1 \cap L_2$). If such strings do not exist, explain why.

Answer:

$$bab, babac, abab, ababac, cabab$$

(d) Write 5 distinct strings over alphabet $\{a, b, c\}$ that do not belong to L_1 and do not belong to L_2 (belong to $\overline{L_1} \cap \overline{L_2}$). If such strings do not exist, explain why.

Answer:

$$a, aa, aaa, aaaa, aaaaa$$

Note that L_2 is the set of strings over $\{a, b, c\}$ that contain exactly 2 occurrences of letter b. It is given by the regular expression:

$$\boldsymbol{(a \cup c)^* \, b \, (a \cup c)^* \, b \, (a \cup c)^*}$$

Problem 10 Let L be the language generated by the context-free grammar $G = \{V, \Sigma, P, S\}$, where $\Sigma = \{a, b, c\}$, $V = \{S, A, B, D\}$, and P is:

$$\begin{aligned} S &\to AB \mid D \\ A &\to AA \mid \lambda \mid abc \\ B &\to bB \mid \lambda \\ D &\to aaD \mid bbbD \mid cD \mid \lambda \end{aligned}$$

Write a regular expression that defines L. If such regular expression does not exist, prove it.

Answer:

$$\boldsymbol{(abc)^* b^* \cup (aa \cup bbb \cup c)^*}$$

Problem 11 Write a complete formal definition of a context-free grammar $G = \{V, \Sigma, P, S\}$ over alphabet $\{a, b, c\}$ such that G generates the language of all strings whose length is even or gives remainder 1 if divided by 3. If such grammar does not exist, prove it.

Answer: $G = \{V, \Sigma, P, S\}$, where $\Sigma = \{a, b, c\}$,
$V = \{S_1, S_2, D\}$, and P is:

$$\begin{aligned} S &\to S_1 \mid S_2 \\ S_1 &\to DDS_1 \mid \lambda \\ S_2 &\to DDDS_2 \mid D \\ D &\to a \mid b \mid c \end{aligned}$$

Problem 12 Let L_1 be the language defined by the regular expression:

$$\boldsymbol{(ab)^*}$$

Let L_2 be the language generated by the context-free grammar $G = \{V, \Sigma, P, S\}$, where $\Sigma = \{a, b\}$, $V = \{S\}$, and the production set P is:

$$S \to aSa \mid bSb \mid a \mid b \mid \lambda$$

(a) Write a regular expression that defines $L_1 \cap L_2$.
If such regular expression does not exist, explain why.

Answer:

$$\boldsymbol{\lambda}$$

To verify the answer, observe first that L_2 is the language of all palindromes over $\{a, b\}$. In contrast, every non-empty string from L_1 begins with a and ends with b, and cannot be a palindrome. Hence, $L_1 \cap L_2$ contains no non-empty strings. However, the empty string is derivable in G by the last rule; it also belongs to L_1, because L_1 is a Kleene star of a non-empty language.

(b) What is the cardinality of $L_1 \cap L_2$? (If possible, state the exact number. If the set is infinite, specify if it is countable or not.)

Answer:

$$|L_1 \cap L_2| = |\{\lambda\}| = 1$$

(c) Compare the cardinalities of L_1 and L_2, and explain which one (if any) is greater.

Answer:

$$|L_1| = |L_2| = \aleph_0$$

Every language is countable, and so are L_1 and L_2. To see that L_1 is infinite, observe that it is a Kleene star of a non-empty language. To see that L_2 is infinite, observe that any of the first two rules can be applied an unbounded number of times, before a non-empty terminal string is produced by any of the next two rules.

(d) Compare the cardinalities of $L_1 \cup L_2$ and $L_1 \cap L_2$, and explain which one (if any) is greater.

Answer: By the answer to part (b):

$$|L_1 \cap L_2| = 1$$

By the answer to part (c), both L_1 and L_2 are infinite and countable—hence, their union is also infinite and countable:

$$|L_1 \cup L_2| = \aleph_0$$

Consequently:

$$|L_1 \cap L_2| < |L_1 \cup L_2|$$

Problem 13 **(a)** Let L_1 be a language over alphabet $\{a, b, c, d, e\}$, defined as follows:

$$L_1 = \{a^n b^m c^k d^{m+2} e^{n+3} \mid k, n, m \geq 0\}$$

Write a complete formal definition of a context-free grammar G_1 that generates language L_1. If such grammar does not exist, explain why.

Answer: $G_1 = \{V, \Sigma, P, S\}$, where $\Sigma = \{a, b, c, d, e\}$,
$V = \{S, A, B\}$, and the production set P is:

$$\begin{array}{l} S \to aSe \mid Aeee \\ A \to bAd \mid Bdd \\ B \to cB \mid \lambda \end{array}$$

(b) Let L_2 be a language over alphabet $\{a, b, c, d, e\}$, defined as follows:

$$L_2 = \{a^{2n} b^{3n} c^{k+2} d^{3k+1} e^{m+3} a^{2m+5} \mid k, n, m \geq 0\}$$

Write a complete formal definition of a context-free grammar G_2 that generates language L_2. If such grammar does not exist, explain why.

Answer: $G_2 = \{V, \Sigma, P, S\}$, where $\Sigma = \{a, b, c, d, e\}$,
$V = \{S, A, B, D\}$, and the production set P is:

$$\begin{array}{l} S \to ABD \\ A \to aaAbbb \mid \lambda \\ B \to cBddd \mid ccd \\ D \to eDaa \mid eeeaaaaa \end{array}$$

Problem 14 Let:

$$L = \{a^\ell b^{2j} c^k d^{2m} \mid \ell, j, k, m \geq 0\}$$

(a) Write a complete formal definition of a context-free grammar that generates L. If such grammar does not exist, prove it.

Answer: $G = \{V, \Sigma, P, S\}$, where
$\Sigma = \{a, b, c, d\}$, $V = \{S, A, B, D, E\}$, and P is:

$$\begin{array}{l} S \to ABED \\ A \to aA \mid \lambda \\ B \to bbB \mid \lambda \\ E \to cE \mid \lambda \\ D \to ddD \mid \lambda \end{array}$$

(b) Write a regular expression that defines L. If such regular expression does not exist, prove it.

Answer:

$$\boldsymbol{a^*(bb)^*c^*(dd)^*}$$

Problem 15 **(a)** Let:

$$L = \{a^i b^j c^k d^m \mid i = j + k \text{ and } m = 2\ell,\ i, j, k, m, \ell \geq 0\}$$

Write a complete formal definition of a context-free grammar that generates L. If such grammar does not exist, prove it.

Answer: $G = \{V, \Sigma, P, S\}$, where $\Sigma = \{a, b, c, d\}$,
$V = \{S, A, B, D\}$, and the production set P is:

$$\begin{aligned}
&S \to AD \\
&A \to aAc \mid B \\
&B \to aBb \mid \lambda \\
&D \to DD \mid \lambda \mid dd
\end{aligned}$$

(b) What is the cardinality of the set of context-free grammars? Answer by giving the exact number (if this set is finite) or by specifying if it is countable or uncountable.

Answer: The set of context-free grammars is infinite and countable (cardinality $\aleph_0$.)

Problem 16 **(a)** Let L be the language defined by the regular expression:

$$(\boldsymbol{a}\,(\boldsymbol{ca} \cup \boldsymbol{da})^*\,\boldsymbol{b}) \;\cup\; (\boldsymbol{b}\,(\boldsymbol{ca} \cup \boldsymbol{da})^*\,\boldsymbol{a})$$

Write a complete formal definition of a context-free grammar G_1 that generates language L. If such grammar does not exist, explain why.

Answer: $G_1 = \{V_1, \Sigma, P_1, S\}$, where $\Sigma = \{a, b, c, d\}$,
$V_1 = \{S, A, B, D\}$, and the production set P_1 is:

$$\begin{aligned}
&S \to A \mid B \\
&A \to aDb \\
&B \to bDa \\
&D \to DD \mid \lambda \mid ca \mid da
\end{aligned}$$

(b) Let L be the language defined in part (a).

Write a complete formal definition of a context-free grammar G_2 that generates language L^*. If such grammar does not exist, explain why.

Answer: $G_2 = \{V_2, \Sigma, P_2, S_2\}$, where $\Sigma = \{a, b, c, d\}$,
$V_2 = \{S_2, S, A, B, D\}$, and the production set P_2 is:

$$\begin{aligned}
&S_2 \to S_2 S_2 \mid \lambda \mid S \\
&S \to A \mid B \\
&A \to aDb \\
&B \to bDa \\
&D \to DD \mid \lambda \mid ca \mid da
\end{aligned}$$

Problem 17 Let L be the set of strings over alphabet $\Sigma = \{a, b, c\}$ defined as follows:

$$L = \{a^m b^k c^\ell \mid m, k, \ell \geq 0 \wedge m = k + \ell\}$$

1. Write a complete formal definition of a context-free grammar G that generates L. If such grammar does not exist, explain why.

   Answer: The general template for strings in L is:

   $$a^\ell a^k b^k c^\ell \text{ for } k, \ell \geq 0$$

   whence the grammar: $G = (V, \Sigma, P, S)$, where
   $\Sigma = \{a, b, c\}$, $V = \{S, B\}$, and the production set P is:

   $$\begin{aligned}
   &S \to aSc \mid B \\
   &B \to aBb \mid \lambda
   \end{aligned}$$

2. List six different strings that belong to $\overline{L}$ (where $\overline{L} = \Sigma \setminus L$). If this is impossible, explain why.

   Answer:

$$cba, ca, cb, ba, abab, abca$$

3. List six different strings that belong to

$$\overline{L} \cap \boldsymbol{a^* b^* c^*}$$

   (where $\overline{L} = \Sigma \setminus L$). If this is impossible, explain why.

   Answer:

$$a, b, c, aaabc, abc, acc$$

4. State the cardinalities of L and $\overline{L}$, and determine which one is greater (if any.) If possible, give exact numbers, otherwise state if sets are countable or not. Explain your answer briefly.

   Answer: L and $\overline{L}$ are infinite and countable:

$$|L| = |\overline{L}| = \aleph_0$$

   To see that L is infinite, observe that it contains, for example, a word $(aa)^n b^n c^n$ for every $n \geq 0$. To see that $\overline{L}$ is infinite, note that it contains, for example, $\boldsymbol{aa^*}$. L and $\overline{L}$ are countable, since every language is countable.

Problem 18 Let L_1 be the language defined by the regular expression:

$$\boldsymbol{a^* b^*}$$

Let L_2 be the language generated by the context-free grammar $G = (V, \Sigma, P, S)$, where $\Sigma = \{a, b\}$, $V = \{S\}$, and the production set P is:

$$S \rightarrow aSb \mid \lambda$$

(a) Write 5 distinct strings that belong to $L_1 \setminus L_2$. If such strings do not exist, explain why.

Answer:

$$a, b, aab, abbb, abbbbbb$$

(b) Write 5 distinct strings that belong to $L_2 \setminus L_1$. If such strings do not exist, explain why.

Answer: It is impossible to list even one such string, since:

$$L_2 \setminus L_1 = \emptyset$$

Precisely:

$$L_2 \subset L_1$$

To see this, observe:

$$L_1 = \{a^m b^k \mid m, k \geq 0\}$$

$$L_2 = \{a^m b^k \mid m = k \text{ and } m, k \geq 0\}$$

(c) Write 5 distinct strings that belong to $L_1 \cap L_2$. If such strings do not exist, explain why.

Answer:

$$\lambda, ab, aabb, aaabbb, aaaabbbb$$

(Recall that $L_1 \cap L_2 = L_2$.)

(d) Write 5 distinct strings over alphabet $\{a, b\}$ that belong to $\overline{L_1 \cup L_2}$ (the complement of $L_1 \cup L_2$.) If such strings do not exist, explain why.

Answer:

$$ba, aabba, bab, baabb, aba$$

(Recall that $L_1 \cup L_2 = L_1$. Hence, $\overline{L_1 \cup L_2} = \overline{L_1}$, which is exactly the set of strings that contain ba as a substring.)