ST. CLOUD STATE UNIVERSITY

CSCI 402/502: Introduction to Theory of Computation
Instructor: Pranava K. Jha

Applying Pumping Lemma for Context-free Languages: Some Examples

Example 1: $L = \{a^m b^m c^m : m \geq 0\}$.

L is not context-free. Assume otherwise. Then L must satisfy the pumping lemma for context-free languages. Let n be the constant of the pumping lemma for L, and let $z = a^n b^n c^n$. Clearly, $z \in L$ and $|z| \geq n$. By the pumping lemma, z may be written as $uvwxy$ in such a way that

(i) $|vwx| \leq n$
(ii) $v \neq \varepsilon$ or $x \neq \varepsilon$, and
(iii) for all $k \geq 0$, $uv^k wx^k y \in L$.

Consider such a representation of z. First suppose that $v \neq \varepsilon$, and v contains a's as well as b's (in which case v must consist of some a's followed by some b's). Now examine uv^2wx^2y that would contain a's then b's then a's then b's, i.e., letters in the "wrong" order with respect to valid strings in L, hence $uv^2wx^2y \notin L$. The same argument applies for other combinations of two distinct letters. This shows that if $v \neq \varepsilon$, then v must be made up of one kind of letter only. An analogous statement holds for x.

Condition (i) of the pumping lemma asserts that v and x cannot be "too far" apart. In particular, if v contains a's, then x cannot contain c's.

Suppose that each of v and x consists of a's only, i.e., $v = a^i$ and $x = a^j$. Then $1 \leq i + j \leq n$. Now consider $uv^2wx^2y = a^{n+i+j} b^n c^n$ that cannot be in L, since $i + j \geq 1$. A similar argument holds if v and x are assumed to consist of b's only, or c's only.

Next suppose that $v = a^i$ and $x = b^j$. Then $1 \leq i + j \leq n$. Consider $uv^2wx^2y = a^{n+i} b^{n+j} c^n$, which is such that the number of c's is unaffected while the number of a's or the number of b's has increased by at least one, hence uv^2wx^2y cannot be in L. An identical argument holds if $v = b^i$ and $x = c^j$. Contradiction!

Conclusion: L is not context-free.

Example 2. $L = \{a^k b^m c^k d^m : k, m \geq 0\}$.

L is not context-free. Assume otherwise. Then L must satisfy the pumping lemma for context-free languages. Let n be the constant of the pumping lemma for L, and let $z = a^n b^n c^n d^n$. Clearly, $z \in L$ and $|z| \geq n$. By the pumping lemma, z may be written as $uvwxy$ in such a way that

(i) $|vwx| \leq n$
(ii) $v \neq \varepsilon$ or $x \neq \varepsilon$, and
(iii) for all $k \geq 0$, $uv^kwx^ky \in L$.

Consider such a representation of z. First suppose that $v \neq \varepsilon$, and v contains a's as well as b's (in which case v must consist of some a's followed by some b's). Now examine uv^2wx^2y that would contain a's then b's then a's then b's, i.e., letters in the "wrong" order with respect to valid strings in L, hence $uv^2wx^2y \notin L$. The same argument applies for other combinations of two distinct letters. This shows that if $v \neq \varepsilon$, then v must be made up of one kind of letter only. An analogous statement holds for x.

Condition (i) of the pumping lemma asserts that v and x cannot be "too far" apart. In particular, if v contains a's, then x cannot contain c's or d's. Similarly, if v contains b's, then x cannot contain d's. Accordingly, the following four cases need to be examined:

(A) v and x are made up of one kind of letter only.
(B) $v \in \boldsymbol{a}^*$ and $x \in \boldsymbol{b}^*$.
(C) $v \in \boldsymbol{b}^*$ and $x \in \boldsymbol{c}^*$.
(D) $v \in \boldsymbol{c}^*$ and $x \in \boldsymbol{d}^*$.

In case (A), uv^2wx^2y is of the form $a^{n+i+j}b^n c^n d^n$ or $a^n b^{n+i+j} c^n d^n$ or $a^n b^n c^{n+i+j} d^n$ or $a^n c^n d^{n+i+j}$ where $i = |v|$ and $j = |x|$. It is easy to see that none of these strings is in L, since $i + j \geq 1$.

In case (B), uv^2wx^2y is of the form $a^{n+i}b^{n+j}c^n d^n$ where $i = |v|$ and $j = |x|$. If $i \geq 1$, then the number of a's increases without a corresponding increase in the number of c's, and if $j \geq 1$, then the number of b's increases without a corresponding increase in the number of d's. Accordingly, $uv^2wx^2y \notin L$. Cases (C) and (D) are similar. Contradiction!

Conclusion: L is not context-free.

Remark: The sets $\{a^k b^m c^m d^k : k, m \geq 0\}$ and $\{a^k b^k c^m d^m : k, m \geq 0\}$ are easily seen to be (deterministic) context-free.

<u>Example 3</u>. $L = \{ww\colon w \in \{a, b\}^*\}$.

L is <u>not</u> context-free. Assume otherwise. Then L must satisfy the pumping lemma for context-free languages. Let n be the constant of the pumping lemma for L, and let $z = a^n b^n a^n b^n$. Clearly, $z \in L$ and $|z| \geq n$. By the pumping lemma, z may be written as $uvwxy$ in such a way that

(i) $|vwx| \leq n$
(ii) $v \neq \varepsilon$ or $x \neq \varepsilon$, and
(iii) for all $k \geq 0$, $uv^k wx^k y \in L$.

Note that z itself has the following form:

$$\underbrace{\overbrace{a\ \cdots\ a}^{n}}_{\substack{\text{first}\\ \text{block}}}\underbrace{\overbrace{b\ \cdots\ b}^{n}}_{\substack{\text{second}\\ \text{block}}}\underbrace{\overbrace{a\ \cdots\ a}^{n}}_{\substack{\text{third}\\ \text{block}}}\underbrace{\overbrace{b\ \cdots\ b}^{n}}_{\substack{\text{fourth}\\ \text{block}}}.$$

Consider an alleged representation of z as $uvwxy$. Condition (i) of the pumping lemma asserts that v and x cannot be "too far" apart. In particular, if v contains letters from the first block, then x cannot contain letters from the third block or from the fourth block. Alternatively, if v contains letters from the second block, then x cannot contain letters from the fourth block. Let $|v| = i$ and $|x| = j$, whence $1 \leq i + j \leq n$.

- If both v and x contain letters from the first block only, then uv^0wx^0y, i.e., uwy is of the form $a^{n-(i+j)}b^n a^n b^n$ that cannot be written as ww. An analogous statement holds if v and x contain letters from the second block (resp. third block or the fourth block) only.

- If v contains letters from the first block or a mix of letters from the first block and the second block, and x contains letters from the second block, then uv^0wx^0y, i.e., uwy is of the form $a^{n-i}b^{n-j}a^n b^n$ where $i \geq 1$ or $j \geq 1$. It is easy to see that $a^{n-i}b^{n-j}a^n b^n$ cannot be written as ww. An analogous statement holds if v contains letters from the second/third block and x contains letters from the third block or if v contains letters from the third/fourth block and x contains letters from the fourth block.

It follows that the given language L does not satisfy the pumping lemma. Therefore, L is not context-free.

—o—