Pumping Lemma Proofs

Alyssa Lytle

Fall 2025

Recall 1

Lemma 1: The Pumping Lemma for Regular Languages

If A is a regular language, then there is a number p (the pumping length) where if s is any string in A of length at least p, then s may be divided into three pieces, $s = xyz$, satisfying the following conditions:

1. for each $i \geq 0$, $xy^iz \in A$
2. $|y| > 0$
3. $|xy| \leq p$

Lemma 2: The Pumping Lemma for Context-Free Languages

If A is a context-free language, then there is a number p (the pumping length) where, if s is any string in A of length $\geq p$, then s may be divided into $s = uvxyz$ satisfying these conditions:

1. for each $i \geq 0$, $uv^ixy^iz \in A$
2. $|vy| > 0$ (at least one of v or y is not the empty string ϵ)
3. $|vxy| \leq p$

1 General Strategy

This is a good general strategy to follow when proving a language is nonregular/not context-free. This is very similar to the rubric that will be used to grade for correctness. We will also be looking for you to clearly mark each of these steps and add explanation, like in the example solutions.

(If you use a different strategy and it is a correct and complete proof, you will still get full credit.)

1. Assume a contradiction. (The language A *is* regular/context-free.)
2. Choose an input string s that would be accepted by this language. (Usually it's helpful to make it a string whose length is a multiple of pumping length p.)
3. Apply the pumping lemma to s.
4. Consider all different ways your can breakup s (for regular languages that would be $s = xyz$ and for context-free languages that would be $s = uvxyz$) and show there is no way to divide s such that when it is "pumped" you still get a string that is in A. (This is often demonstrated with a *proof by cases.*)
5. Point out the contradiction found in steps 3 and 4.

Example 1: HW3 Question 5

We want to prove that $C = \{a^i b^j c^k | 0 \leq i \leq j \leq k\}$ is not context-free. Let's start by assuming it is and finding a contradiction.

$$C = \{a^i b^j c^k | 0 \leq i \leq j \leq k\} \text{ is a context-free lanugage} \quad \text{(Assumption)} \tag{1}$$

$$\text{Choose } s = a^p b^p c^p \in C \quad \text{(Chose an } s \in C \text{ based on the definition of } C \text{ in line 1.)} \tag{2}$$

$$s \text{ can be split into } s = uvxyz\text{, where for any } i \geq 0 \text{ the string } uv^i xy^i z \in B, \tag{3}$$
$$|vy| > 0 \text{ and } |vxy| \leq p \quad \text{(Pumping lemma)}$$

$$\text{It is impossible to split } s \text{ into } s = uvxyz\text{, where for any } i \geq 0 \text{ the string } uv^i xy^i z \in B, \tag{4}$$
$$|vy| > 0 \text{ and } |vxy| \leq p \quad \text{(*Shown below)}$$

$$\text{Lines 3 and 4 contradict.} \rightarrow\leftarrow \tag{5}$$

***Mini Proof: Line 4 (Verbose Version)**

We can prove line 4 with a proof by cases.

There are two main "cases" we want to prove: the case where both v and y are all one letter and the case where v or y crosses some "boundary" of letters (e.g. is a string of as followed by bs) so either v and/or y contains more than one unique letter. Note that since $|vxy| \leq p$, we know vxy can not contain all three letters, therefore neither can v and y.

1. *v and y are all one letter.*
   This can be broken down further into these cases:

   (a) *No a's in vxy.* In other words, vxy is all bs and/or cs AND v and y each only contain one unique character.
   The problem with this is that uv^0xy^0z would then result in fewer bs or cs and would violate the fact that there are no more as than bs and cs. ($C = \{a^i b^j c^k | 0 \leq i \leq j \leq k\}$, so $uv^0xy^0z \notin C$)

   (b) *No b's in vxy.* This means either all as or cs (but not both) must appear in vxy. If all as appear, then uv^2xy^2z would result in more as than bs, violating the definition, so $uv^2xy^2z \notin C$
   If all cs appear, then we get the same issue as part (a) where $uv^0xy^0z \notin C$.

   (c) *No c's in vxy.* In other words, vxy is all bs and/or cs AND v and y each only contain one unique character. This gives us a similar issue as part (b) where uv^2xy^2z would result in either more a's than c's and/or more b's than c's, violating the definition. So $uv^2xy^2z \notin C$

2. *v and/or y contain more than one unique letter*
   This case is pretty simple. You just need to consider that any string that contains more than one unique letter will cause out-of-order repeats when pumped. (E.g. if $v = aa \ldots abb \ldots b$ then $v^2 = aa \ldots abb \ldots baa \ldots abb \ldots b$ so $uv^2xy^2z \notin C$.

I included the verbose version of the proof of line 4 so you could fully understand the reasoning, but for the purposes of an exam, a shorter version like the following would be acceptable.

***Mini Proof: Line 4 (Acceptable Shortened Version)**

We can prove line 4 with a proof by cases.

Note that since $|vxy| \leq p$, we know vxy can not contain all three letters, therefore neither can v and y, so we get these cases:

1. *v and y are all one letter.*
   This can be broken down further into these cases:
   (a) *No a's in vxy.* uv^0xy^0z would result in fewer bs or cs than as, so $uv^0xy^0z \notin C$
   (b) *No b's in vxy.* This means either all as or cs (but not both) must appear in vxy. If all as appear, uv^2xy^2z would result in more as than bs, so $uv^2xy^2z \notin C$
   If all cs appear, then we get the same issue as part (a)
   (c) *No c's in vxy.* uv^2xy^2z would result in either more a's than c's and/or more b's than c's, so $uv^2xy^2z \notin C$
2. *v and/or y contain more than one unique letter*
   Any string that contains more than one unique letter will cause out-of-order repeats when pumped.

Example 2: HW3 Question 7

We want to prove that Let the language of all palindromes over $\{0, 1\}$ containing equal numbers of 0s and 1s is not context-free. Let's start by assuming it is and finding a contradiction.

The language B of all palindromes over $\{0, 1\}$
containing equal numbers of 0s and 1s is a context-free lanugage (Assumption) (1)
Choose $s = 0^p1^{2p}0^p \in B$ (Chose an $s \in B$ based on the definition of B in line 1.) (2)
s can be split into $s = uvxyz$, where for any $i \geq 0$ the string $uv^ixy^iz \in B$, (3)
$|vy| > 0$ and $|vxy| \leq p$ (Pumping lemma)
It is impossible to split s into $s = uvxyz$, where for any $i \geq 0$ the string $uv^ixy^iz \in B$, (4)
$|vy| > 0$ and $|vxy| \leq p$ (*Shown below)
Lines 3 and 4 contradict. $\rightarrow\leftarrow$ (5)

***Mini Proof: Line 4 (Verbose Version)**

Because $|vxy| \leq p$ and the three blocks $0^p, 1^{2p}, 0^p$ have lengths $p, 2p, p$ respectively, the substring vxy can either lie entirely inside one block of 0s or 1s or else cross exactly one "boundary".

1. *vxy lies entirely in the left block 0^p.*

   Then v and y consist only of 0's, so pumping changes the total number of 0's but does not change the number of 1's. Take $i = 0$ (or $i = 2$). The pumped string has a different number of 0's than 1's, so it cannot be in B. This contradicts the pumping lemma's requirement that $uv^ixy^iz \in B$ for all i.

2. *vxy lies entirely in the middle block 1^{2p}.*

   Then v and y consist only of 1's, so pumping changes the total number of 1's but not the number of 0's. Again choose $i = 0$ (or $i = 2$). The pumped string has a different number of 0's than 1's, so it cannot be in B. This contradicts the pumping lemma's requirement that $uv^ixy^iz \in B$ for all i.

3. *vxy lies entirely in the right block 0^p.*

   This is symmetric to Case 1: pumping changes only the number of 0's, so for some i the numbers of 0's and 1's differ and the pumped string is not in B.

4. *vxy crosses the boundary between the left 0^p and the middle 1^{2p}.*

   It is possible that this can be divided in a way where we maintain the property of the same *amounts* of 0s and 1s, but it wouldn't keep the palindrome structure. For example, uv^0xy^0z would make the front half of the string shorter, so you wouldn't have the balanced property (you'd have more 0's in the back half of the string), so $uv^0xy^0z \notin B$.

5. *vxy crosses the boundary between the middle 1^{2p} and the right 0^p.*

   This follows the same logic as case 4. The balance of the palindrome would be off, so $uv^ixy^iz \notin B$.