CS 154
Formal Languages and Computability
Assignment #3 Solutions

Department of Computer Science
San Jose State University

Spring 2016

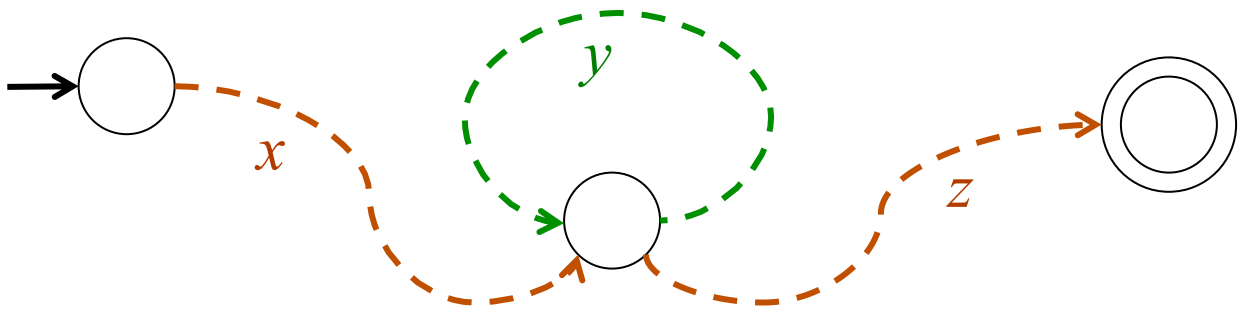
Instructor: Ron Mak

www.cs.sjsu.edu/~mak

The Pumping Lemma for Regular Languages

- Let L be an infinite regular language.
- Then for <u>any</u> string $w \in L$, there exists a positive integer m such that we can decompose w into xyz, where
  - $|w| \geq m$
  - $|xy| \leq m$
  - $|y| \geq 1$
  - And $w_i = xy^iz$ is also in L for <u>all</u> $i = 0, 1, 2, \ldots$

> We might not know <u>what value</u> m has, only that it <u>exists</u>.



> In particular, when $i = 0$, the string xz is in L.

The Pumping Lemma for RLs, *cont'd*

- We use the pumping lemma to prove that a given language L is <u>not</u> regular.
- We do so using a proof by contradiction.
- <u>Assume</u> that L is regular.

A language is either regular or not.

- And so the pumping lemma must hold for all strings w in L.
- Show that this leads to a contradiction.
- Therefore, the original assumption that L is regular must not be true and L is <u>not</u> regular.

Assignment #3: Question 1

- Use the pumping lemma to show that the language $L = \{a^n : n \text{ is a perfect cube: } 0, 1, 8, 27, \ldots\}$ is not regular.

All string lengths are perfect cubes.

- Assume that L is regular and so the pumping lemma must hold for any string w in L.
- Choose $w = xyz = a^{m^3}$ with $|xy| \le m$ and $|y| \ge 1$.

We don't know what m equals, only that it exists.

- Since $w_i = xy^iz$ is in L for all $i = 0, 1, 2, \ldots$, choose $i = 3$: $w_3 = xy^3z$
- $|xy^3z| = |xyz| + 2|y| = m^3 + 2|y|$
- But $|y| \le m$ since $|xy| \le m$, and so $|xy^3z| = m^3 + 2|y| \le m^3 + 2m < (m+1)^3$

$(m+1)^3 = m^3 + 3m^2 + 3m + 1$

- Also since $|y| \ge 1$, $|xy^3z| = m^3 + 2|y| > m^3$
- So $m^3 < |xy^3z| < (m+1)^3$ and so $|xy^3z|$ cannot be a perfect cube since it's between two consecutive perfect cubes.
- Therefore, xy^3z is not in L, a contradiction, and so L is not regular.

Assignment #3: Question 1 (Alternate)

- Use the pumping lemma to show that the language $L = \{a^n : n \text{ is a perfect cube: } 0, 1, 8, 27, \ldots\}$ is not regular.
- Assume that L is regular and so the pumping lemma must hold for any string w in L.
- Choose $w = xyz = a^{m^3}$ with $|xy| \leq m$ and $|y| \geq 1$.
- Then $y = a^k$ for some $1 \leq k \leq m$.
- The pumped strings will be $w_i = a^{m^3+(i-1)k}$ for $i = 0, 1, 2, \ldots$
  - w_0 is the string that does not contain y.
- But $w_2 = a^{m^3+k} \notin L$ because $m^3 < m^3 + k < m^3 + m < (m+1)^3$ *i.e.*, $|w_2|$ is between two consecutive perfect cubes.
- This is a contradiction, so L is not regular.

Assignment #3: Question 2

- Use the pumping lemma to show that the language $L = \{a^n : n \text{ is a power of 2: } 1, 2, 4, 8, \ldots\}$ is not regular.

- Assume that L is regular and so the pumping lemma must hold for any string w in L.
- Let p be the smallest integer such that $2^p > m$.
- Choose $w = xyz = a^{2^p}$ and then $y = a^k$ for some $1 \leq k \leq m$.
- The pumped strings will be $w_i = a^{2^p+(i-1)k}$ for $i = 0, 1, 2, \ldots$
  - w_0 is the string that does not contain y.
- But $w_2 = a^{2^p+k} \notin L$ because $2^p < 2^p + k \leq 2^p + m < 2^p + 2^p = 2^{p+1}$
  i.e., $|w_2|$ is between two consecutive powers of 2.
- This is a contradiction, so L is not regular.

Assignment #3: Question 3

- Use the pumping lemma to show that the language $L = \{a^{pq} : p \text{ and } q \text{ are both prime numbers}\}$ is not regular.

- Assume that L is regular and so the pumping lemma must hold for any string w in L.
- Let p and q be the smallest primes such that $pq \geq m$.
- Choose $w = xyz = a^{pq}$ and then $y = a^k$ for some $1 \leq k \leq m$.
- The pumped strings will be $w_i = a^{pq+(i-1)k}$ for $i = 0, 1, 2, \ldots$
  - w_0 is the string that does not contain y.
- But $w_{pq+1} = a^{pq+pqk} \notin L$ because $pq + pqk = pq(1 + k)$ which is not a product of two primes.
- This is a contradiction, so L is not regular.

Assignment #3: Question 4

- Use the pumping lemma to show that the language $L = \{a^p b^q : p \text{ divided by } q \text{ is an integer quotient}\}$ is not regular.

- Assume that L is regular and so the pumping lemma must hold for any string w in L.
- Choose $w = a^m b^m$ and then $y = a^k$ for some $1 \leq k \leq m$.
- In the pumped strings, choose $i = 0$ to remove y from the first half of w and so $(m\text{-}k)/m$ is not integer.
- This is a contradiction, so L is not regular.

Assignment #3: Question 5

- Use the pumping lemma to show that the language $L = \{a^p b^q : p + q \text{ is a prime number}\}$ is not regular.
- Assume that L is regular and so the pumping lemma must hold for any string w in L.
- Choose $w = a^m b^{p-m}$ where $p > m$ is a prime number.
- Choose $y = a^k$ for some $1 \le k \le m$.
- The pumped strings will be $w_i = a^{m+(i-1)k} b^{p-m}$ for $i = 0, 1, 2, \ldots$
  - w_0 is the string that does not contain y.
- But $w_{p+1} = a^{m+pk} b^{p-m} \notin L$ because $(m + pk) + (p - m) = p(1 + k)$ and so the sum of the two exponents is not prime.
- This is a contradiction, so L is not regular.

Assignment #3: Question 6

- Let $\Sigma = \{0, 1, +, =\}$. Use the pumping lemma to show that the language $L = \{b_1{=}b_2{+}b_3 : b_1, b_2, b_3$ are binary integers, and b_1 is the sum of b_2 and $b_3\}$ is not regular. For example, the string $1001{=}10{+}111$ is in L.

- Assume that L is regular and so the pumping lemma must hold for any string w in L.
- Choose $w = xyz$ be the string $1^m{=}0^m{+}1^m$.
  - Example: $11111{=}00000{+}11111$
- And so $y = 1^k$ for some $1 \le k \le m$.
- Then xy^2z is the string $1^{m+k}{=}0^m{+}1^m$ which is not in L.
- This is a contradiction, so L is not regular.

Assignment #3: Question 7

- Let language L be denoted by the regular expression a^*b^*. What is wrong with the following "proof" that L is not regular?

Of course, L is regular.

> Assume that L is regular. Then it must be defined by a DFA with k states, for some integer $k > 0$. Take the string $w = a^k b^k$ and split it $w = xyz$, with $y = ab$. Then wy^2z is not in L, which contradicts the pumping lemma. Therefore, L cannot be regular.

- Since $|xy| \leq m$, setting $y = ab$ says that $m = |a^*| + 1$.
- But the pumping lemma states that there there exists a positive integer m, so even if $m = |a^*| + 1$ doesn't work for the lemma, there could be another value for m that does.
- For example, if $m \leq |a^*|$ and y is all a's, the lemma holds.

Assignment #3: Question 8

- Prove whether or not language
  $L = \{a^{p+qi} : p \text{ and } q \text{ are fixed integer values, and } i \geq 0\}$
  is regular.
- The language is regular because its strings are denoted by the regular expression $a^p(a^q)^*$.

Assignment #3: Question 9

- Prove whether or not language
  $L = \{a^p b^q : p \geq 100 \text{ and } q \geq 100 \text{ are fixed integer values}\}$
  is regular.
- The language is regular because its strings are denoted by the regular expression $a^{100}a^*b^{100}b^*$.

Assignment #3: Question 10

- Assume that `<stmt>`, `<if_stmt>`, `<boolexpr>`, and `<assign_stmt>` are nonterminal symbols, and `if`, `else`, `(`, and `)` are terminal symbols.

  Here's a grammar written in BNF for Java-style IF statements:

```
<stmt>    ::= <assign_stmt> | <if_stmt>
<if_stmt> ::= if ( <boolexpr> ) <stmt>
            | if ( <boolexpr> ) <stmt> else <stmt>
```

  How is this grammar ambiguous?
  Give an example of an ambiguity.

Assignment #3: Question 10, *cont'd*

```
<stmt>     ::= <assign_stmt> | <if_stmt>
<if_stmt> ::= if ( <boolexpr> ) <stmt>
             | if ( <boolexpr> ) <stmt> else <stmt>
```

- An if statement has an optional else part.
- An if statement contains one or two statements, each of which can in turn be an if statement.
- The grammar is <u>ambiguous</u>.
- In the statement

```
if (a == b) if (c == d) x = 1 else y = 1
```

- To which if statement does the else part belong?
  - Is it `if (a == b) if (c == d) x = 1 else y = 1`
  - Or `if (a == b) if (c == d) x = 1 else y = 1`

Most languages take the first choice.