CS301: Theory of Computation
Section 1.4 - Nonregular Languages

Daniel White

Gettysburg College
Computer Science

Spring 2024

All material in this course so far has dealt with regular languages and how to describe them. But are all languages regular?

Regular languages are those recognized by finite automata, which have finite memory. Some languages may appear to require arbitrarily large memory to recognize though.

exercise: Convince yourself that the language $B = \{0^n 1^n \mid n \geq 0\}$ is not regular.

exercise: What about the language
$D = \{w \mid w \text{ has equal number of } 01 \text{ and } 10 \text{ as substrings}\}$?

Theorem 1.70 (pg. 78)

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:

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

In other words, if a string from any given regular language is long enough, it can be split into three parts, with the first two parts being sufficiently short, and the middle part can be "pumped" with the resulting string(s) staying within the language.

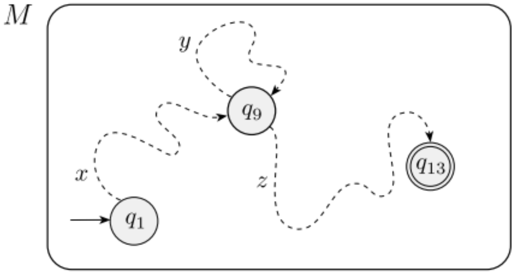
Proof. Let M be a DFA recognizing A and assign p to be the number of states of M.

If A contains no strings of length p, we are done. Therefore, consider some $s \in A$ with length at least p, and consider the sequence of states that M goes through when processing s, which has length at least $p+1$.

It follows by the pigeonhole principle that some state in the sequence above is repeated ... (for example, below)

$$\begin{array}{cccccccccccccccccc} s = & & s_1 & & s_2 & & s_3 & & s_4 & & s_5 & & s_6 & & \cdots & & s_n & \\ & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow \\ & q_1 & & q_3 & & q_{20} & & q_9 & & q_{17} & & q_9 & & q_6 & & q_{35} & & q_{13} \end{array}$$

exercise: Finish this proof. *(pgs. 78-79)*



homework: Formalize the proof of the pumping lemma *(pg. 79)*

exercise: Let B be the language $\{0^n 1^n \mid n \geq 0\}$. Use the pumping lemma to prove that B is not regular.

By way of contradiction, assume B is regular. Let p be the pumping length given in the pumping lemma and consider the string $s = 0^p 1^p \in B \ldots$

homework: 1.29 & 1.55ab

exercise: Let $C = \{w \mid w \text{ has an equal number of 0s and 1s}\}$. Use the pumping lemma to prove that B is not regular.

By way of contradiction, assume C is regular. Let p be the pumping length given in the pumping lemma and consider the string $s = 0^p1^p \in C$ ...

exercise: Why does the choice $s = (01)^p$ not work?

exercise: Repeat the first exercise above using the fact that B (previous slide) is not regular.

By way of contradiction, assume C is regular. It follows that $C \cap 0^*1^*$ is regular since ...

exercise: Show that $E = \{0^i 1^j \mid i > j\}$ is not regular.

By way of contradiction, assume E is regular. Let p be the pumping length for E and consider the string ...

homework: example 1.75