Computer Science 313

Nonregular Languages, Part Two

Instructor: Wayne Eberly

Department of Computer Science
University of Calgary

Lecture #11

Goal for Today

Goals for Today:

- Describe how the ***closure properties for regular languages*** can *also* be used to prove that a language $L \subseteq \Sigma^\star$ is ***not*** regular.

Closure Properties

If $L_1, L_2 \subseteq \Sigma^\star$ for some alphabet Σ, and L_1 and L_2 are both regular languages, then the following languages are regular languages as well.

(a) $L_1 \cup L_2$

(b) $L_1 \circ L_2$

(c) $L_1 \cap L_2$

(d) $L_1^\star$

(e) $L_1^C = \{\omega \in \Sigma^\star \mid \omega \notin L_1\}$

Parts (a), (b) and (d) were proved in class, while parts (c) and (e) were considered in the tutorial exercise on closure properties and regular operations.

Since $P \Rightarrow Q$ implies that $\neg Q \Rightarrow \neg P$, these imply the following.

Closure Properties

Suppose that $L_1, L_2 \subseteq \Sigma^\star$ for some alphabet Σ, and that any of the following languages are ***not*** regular languages:

(a) $L_1 \cup L_2$

(b) $L_1 \circ L_2$

(c) $L_1 \cap L_2$

Then either or both of L_1 and L_2 is not a regular language either. Furthermore, if any of

(d) $L_1^\star$

(e) $L_1^C = \{\omega \in \Sigma^\star \mid \omega \notin L_1\}$

is ***not*** a regular language then L_1 is ***not*** a regular language either.

Now that we know of *some* nonregular languages, this gives us ways to find others.

First Example

If $\Sigma = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$ then the language

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

is a subset of $\Sigma^\star$ that is not a regular language. This was proved during the last lecture using the Pumping Lemma for Regular Languages, when L was considered as a subset of $\widehat{\Sigma}^\star$ for the alphabet $\widehat{\Sigma} = \{\mathtt{a}, \mathtt{b}\}$.

However, exactly the same proof establishes that L is not a regular language, when considered as a subset of $\Sigma^\star$, as well.

First Example

Now consider another language $L_1 \subseteq \Sigma^\star$, namely

$$L_1 = \{\mathrm{a}^n \mathrm{b}^n \mathrm{c}^m \mid n, m \geq 0\} \subseteq \Sigma^\star.$$

Claim: L_1 is not a regular language.

Proof: Consider the language

$$L_2 = \{\mathrm{a}^k \mathrm{b}^\ell \mid k, \ell \geq 0\} \subseteq \Sigma^\star.$$

- Note that

  $$L_1 \cap L_2 = L,$$

  the language mentioned on the previous slide, and L is ***not*** a regular language.

First Example

- It now follows by the previous result that at least one of L_1 and L_2 is not a regular language, either.
- However, L_2 *is* a regular language, because it is the language of the regular expression $\mathtt{a}^\star \circ \mathtt{b}^\star$.
- It follows that L_1 is not a regular language, as claimed.

Second Example

Now let $\Sigma = \{\mathtt{a}, \mathtt{b}\}$. Once again, we already know that the language

$$L = \{\mathtt{a}^n\mathtt{b}^n \mid n \geq 0\} \subseteq \Sigma^\star$$

is ***not*** a regular language.

Claim: The language

$$\widehat{L} = \{\mathtt{a}^n\mathtt{b}^m \mid n, m \geq 0 \text{ and } n \neq m\}$$

is not a regular language.

Second Example

Proof: To begin, consider the language

$$L_1 = \{\mathtt{a}^n\mathtt{b}^m \mid n, m \geq 0 \text{ and } n \neq m\} \cup \{\omega \in \Sigma^* \mid \mathtt{ba} \text{ is a substring of } \omega\}.$$

With a bit of work one can establish that $L_1^C = L$, the language that we already know is not regular. It follows by the claim that L_1 is not a regular language, either.

Note next that $L_1 = \widehat{L} \cup L_2$ where

$$L_2 = \{\omega \in \Sigma \mid \mathtt{ba} \text{ is a substring of } \omega\}.$$

Second Example

It now follows by (a second application of) the claim that at least one of $\widehat{L}$ and L_2 is a nonregular language as well.

However L_2 *is* regular, since it is the language of the regular expression $\Sigma^\star \circ \mathtt{ba} \circ \Sigma^\star$.

Thus $\widehat{L}$ is a nonregular language (that is, a language that is not regular) as claimed.

Summary of Process

To prove that a language $L \subseteq \Sigma^\star$ is not regular...

- Conclude this immediately (after mentioning that this follows from the claim) if you already know that either or both of $L^\star$ or L^C is not regular.

- Otherwise, try to find (and give) a ***regular*** language $L_2 \in \Sigma^\star$ such that one or more of the languages
  - $L \cup L_2$
  - $L \circ L_2$ (or $L_2 \circ L$)
  - $L \cap L_2$

  is already known not to be regular. Since L_2 *is* regular, it follows from the claim that L cannot be regular, once again.

Why Do We Need the Pumping Lemma?

This process can only be used if we ***know about*** some languages that are not regular already!

It is necessary to use something like the "Pumping Lemma for Regular Languages" to prove that some languages are not regular, before that is the case!

With that noted: Every time we discover *another* language that is not regular, it gets easier to use the closure properties in this way (at least, in principle).