Regular Languages and Expressions

Lecture 2

Thursday, August 27, 2020

LaTeXed: July 24, 2020 22:02

2.1
Regular Languages

Regular Languages

A class of simple but useful languages.

The set of regular languages over some alphabet Σ is defined inductively as:

1. $\emptyset$ is a regular language.
2. $\{\epsilon\}$ is a regular language.
3. $\{a\}$ is a regular language for each $a \in \Sigma$. Interpreting a as string of length 1.
4. If L_1, L_2 are regular then $L_1 \cup L_2$ is regular.
5. If L_1, L_2 are regular then $L_1 L_2$ is regular.
6. If L is regular, then $L^* = \cup_{n \geq 0} L^n$ is regular.
   The $\cdot^*$ operator name is **Kleene star**.
7. If L is regular, then so is $\overline{L} = \Sigma^* \setminus L$.

Regular languages are closed under operations of union, concatenation and Kleene star.

Regular Languages

A class of simple but useful languages.

The set of regular languages over some alphabet Σ is defined inductively as:

1. $\emptyset$ is a regular language.
2. $\{\epsilon\}$ is a regular language.
3. $\{a\}$ is a regular language for each $a \in \Sigma$. Interpreting a as string of length 1.
4. If L_1, L_2 are regular then $L_1 \cup L_2$ is regular.
5. If L_1, L_2 are regular then $L_1 L_2$ is regular.
6. If L is regular, then $L^* = \cup_{n \geq 0} L^n$ is regular.
   The $\cdot^*$ operator name is **Kleene star**.
7. If L is regular, then so is $\overline{L} = \Sigma^* \setminus L$.

Regular languages are closed under operations of union, concatenation and Kleene star.

Regular Languages

A class of simple but useful languages.

The set of regular languages over some alphabet Σ is defined inductively as:

1. $\emptyset$ is a regular language.
2. $\{\epsilon\}$ is a regular language.
3. $\{a\}$ is a regular language for each $a \in \Sigma$. Interpreting a as string of length 1.
4. If L_1, L_2 are regular then $L_1 \cup L_2$ is regular.
5. If L_1, L_2 are regular then $L_1 L_2$ is regular.
6. If L is regular, then $L^* = \cup_{n \geq 0} L^n$ is regular.
   The $\cdot^*$ operator name is **Kleene star**.
7. If L is regular, then so is $\overline{L} = \Sigma^* \setminus L$.

Regular languages are closed under operations of union, concatenation and Kleene star.

Regular Languages

A class of simple but useful languages.

The set of regular languages over some alphabet Σ is defined inductively as:

1. $\emptyset$ is a regular language.
2. $\{\epsilon\}$ is a regular language.
3. $\{a\}$ is a regular language for each $a \in \Sigma$. Interpreting a as string of length 1.
4. If L_1, L_2 are regular then $L_1 \cup L_2$ is regular.
5. If L_1, L_2 are regular then $L_1 L_2$ is regular.
6. If L is regular, then $L^* = \cup_{n \geq 0} L^n$ is regular.
   The $\cdot^*$ operator name is **Kleene star**.
7. If L is regular, then so is $\overline{L} = \Sigma^* \setminus L$.

Regular languages are closed under operations of union, concatenation and Kleene star.

Regular Languages

A class of simple but useful languages.

The set of regular languages over some alphabet Σ is defined inductively as:

1. $\emptyset$ is a regular language.
2. $\{\epsilon\}$ is a regular language.
3. $\{a\}$ is a regular language for each $a \in \Sigma$. Interpreting a as string of length 1.
4. If L_1, L_2 are regular then $L_1 \cup L_2$ is regular.
5. If L_1, L_2 are regular then $L_1 L_2$ is regular.
6. If L is regular, then $L^* = \cup_{n \geq 0} L^n$ is regular.
   The $\cdot^*$ operator name is **Kleene star**.
7. If L is regular, then so is $\overline{L} = \Sigma^* \setminus L$.

Regular languages are closed under operations of union, concatenation and Kleene star.

Regular Languages

A class of simple but useful languages.

The set of regular languages over some alphabet Σ is defined inductively as:

1. $\emptyset$ is a regular language.
2. $\{\epsilon\}$ is a regular language.
3. $\{a\}$ is a regular language for each $a \in \Sigma$. Interpreting a as string of length 1.
4. If L_1, L_2 are regular then $L_1 \cup L_2$ is regular.
5. If L_1, L_2 are regular then $L_1 L_2$ is regular.
6. If L is regular, then $L^* = \cup_{n \geq 0} L^n$ is regular.
   The $\cdot^*$ operator name is **Kleene star**.
7. If L is regular, then so is $\overline{L} = \Sigma^* \setminus L$.

Regular languages are closed under operations of union, concatenation and Kleene star.

Regular Languages

Have basic operations to build regular languages.
Important: Any language generated by a finite sequence of such operations is regular.

Lemma

Let $L_1, L_2, \ldots,$ *be regular languages over alphabet* Σ*. Then the language* $\cup_{i=1}^{\infty} L_i$ *is not necessarily regular.*

Some simple regular languages

Lemma

If w is a string then $L = \{w\}$ is regular.

Example: $\{aba\}$ or $\{abbabbab\}$. Why?

Lemma

Every finite language L is regular.

Examples: $L = \{a, abaab, aba\}$. $L = \{w \mid |w| \leq 100\}$. Why?

Some simple regular languages

Lemma

If w is a string then $L = \{w\}$ is regular.

Example: $\{aba\}$ or $\{abbabbab\}$. Why?

Lemma

Every finite language L is regular.

Examples: $L = \{a, abaab, aba\}$. $L = \{w \mid |w| \leq 100\}$. Why?

More Examples

- $\{w \mid w$ is a keyword in Python program$\}$
- $\{w \mid w$ is a valid date of the form mm/dd/yy$\}$
- $\{w \mid w$ describes a valid Roman numeral$\}$
  $\{I, II, III, IV, V, VI, VII, VIII, IX, X, XI, \ldots\}$.
- $\{w \mid w$ contains "CS374" as a substring$\}$.

Review questions

1. $L_1 \subseteq \{0,1\}^*$ be a finite language. L_1 is a set with finite number of strings. T/F?
2. $L_2 = \{0^i \mid i = 0, 1, \ldots, \infty\}$. The language L_2 is regular. T/F?
3. $L_3 = \{0^{2i} \mid i = 0, 1, \ldots, \infty\}$. The language L_3 is regular. T/F?
4. $L_4 = \{0^{17i} \mid i = 0, 1, \ldots, \infty\}$. The language L_4 is regular. T/F?
5. $L_5 = \{0^i \mid i$ is **not** divisible by $17\}$. L_5 is regular. T/F?
6. $L_6 = \{0^i \mid i$ is divisible by $2, 3$, or $5\}$. L_6 is regular. T/F?
7. $L_7 = \{0^i \mid i$ is divisible by $2, 3$, and $5\}$. L_7 is regular. T/F?
8. $L_8 = \{0^i \mid i$ is divisible by $2, 3$, but not $5\}$. L_8 is regular. T/F?
9. $L_9 = \{0^i 1^i \mid i$ is divisible by $2, 3$, but not $5\}$. L_9 is regular. T/F?
10. $L_{10} = \{w \in \{0,1\}^* \mid w$ has at most 374 1s$\}$. L_{10} is regular. T/F?

THE END

...

(for now)