Lecture 1: Decision Problems and Regular Languages

Date: August 22, 2023.

Decision Problem is a computational problem that expects a Boolean answer on each input.

Given input \( w \), \( f \) : Inputs \( \rightarrow \{ T, \bot \} \)

\[
L = \{ w \in \Sigma^* | f(w) = T \}
\]

\[
L \subseteq \text{Inputs}
\]

\[
L = \{ \epsilon \} \text{ if } n = 0
\]

\[
L_\cdot(L^{n-1}) \text{ if } n > 0
\]

Inputs encoded as strings over some set of symbols.

Strings:

1. An alphabet is a finite set of symbols. For example \( \Sigma = \{0, 1\} \), \( \Sigma = \{a, b, c, \ldots, z\} \),
   \( \Sigma = \{\text{moveforward}, \text{moveback}\} \) are alphabets.

2. A string/word over \( \Sigma \) is a finite sequence of symbols over \( \Sigma \). For example, ‘0101001’, ‘string’,
   ‘(moveback)(rotate90)’

3. \( \epsilon \) is the empty string.

4. The length of a string \( w \) (denoted by \( |w| \)) is the number of symbols in \( w \). For example, \( |101| = 3 \),
   \( |\epsilon| = 0 \), \( |(\text{moveback})(\text{rotate90})| = 2 \).

5. \( \Sigma^* \) is the set of all strings over \( \Sigma \); \( \Sigma^n = \{ w \in \Sigma^* | |w| = n \} \)

6. Concatenation of two strings \( x \) and \( y \), denoted either \( x \cdot y \) or simply \( xy \), is the unique string containing
   the symbols of \( x \) in order, followed by the symbols in \( y \) in order.

   \[
   x = 10 \quad y = 01 \quad xy = 1001
   \]

7. \( y \) is a substring of \( w \) if there are strings \( x, z \) such that \( w = x \cdot y \cdot z \). If \( x = \epsilon \) then \( y \) is a prefix of \( w \).
   If \( z = \epsilon \) then \( y \) is a suffix of \( w \).

Language over \( \Sigma \) is a set \( L \subseteq \Sigma^* \). Examples include \( \{\epsilon\}, \{w | |w| > 5\} \).

- For languages \( A, B \subseteq \Sigma^* \), the concatenation of \( A \) and \( B \) is

   \[
   AB = A \cdot B = \{ u \cdot v | u \in A \text{ and } v \in B \}
   \]

- For languages \( A, B \subseteq \Sigma^* \), their union is \( A \cup B \), intersection is \( A \cap B \), and difference is \( A \setminus B \).

- For \( A \subseteq \Sigma^* \), the complement of \( A \) is \( \overline{A} = \Sigma^* \setminus A \).

Powers and Kleene Closure: For a language \( L \subseteq \Sigma^* \) and \( n \in \mathbb{N} \), define \( L^n \) inductively as follows.

\[
L^n = \begin{cases} 
\{ \epsilon \} & \text{if } n = 0 \\
L \cdot (L^{n-1}) & \text{if } n > 0
\end{cases}
\]

And define \( L^\ast = \cup_{n \geq 0} L^n \), and \( L^+ = \cup_{n \geq 1} L^n \).

Alternatively, \( L^\ast \) set of all strings formed by concatenating \( n \) strings from \( L \), \( L^\ast \) is the set of all strings formed by concatenating some (finite) number of strings from \( L \).
Problem 1. Answer the following questions taking $\Sigma = \{0, 1\}$.

1. What is $\Sigma^0$? $\epsilon \in \Sigma^0$

2. How many elements are there in $\Sigma^3$? $|\Sigma^3| = 8$.

3. How many elements are there in $\Sigma^n$? $2^n$

4. For what values of $n$, is $\Sigma^n \subseteq \Sigma^{n+1}$? Never. $\Sigma_2 = \{0, 1\}$.

5. For what values of $n$, is $\Sigma^n \subseteq \Sigma^*$? Always.

6. Let $u$ be an arbitrary string $\Sigma^*$. What is $\epsilon \cdot u$? What is $u \cdot \epsilon$? $\epsilon \cdot u = u \cdot \epsilon = u$.

Problem 2. Consider languages over $\Sigma = \{0, 1\}$.

1. What is $\emptyset^0$? $\emptyset^0 = \emptyset \subseteq \Sigma^n$.

2. Let $L \subseteq \Sigma^*$. What is $|L^*|$? Is it finite? Infinite? $|L^*| = 2^{2^n}$, $|\emptyset^*| = 1$; $\emptyset^* = \Sigma^*$ in all other cases. $|L^*|$ infinite.

3. What is $\emptyset^+, \{\epsilon\}^+$? $\emptyset^+ = \emptyset$, $\{\epsilon\}^+ = \{\epsilon\}$.

For set $A$, $|A| = \# \text{ elements in } A$.

Regular Languages over alphabet $\Sigma$ are inductively defined as follows.

- $\emptyset$ is a regular language.
- $\{\epsilon\}$ is a regular language.
- $\{a\}$ is a regular language for every $a \in \Sigma$.
- If $A, B$ are regular languages then $A \cup B$ is regular.
- If $A, B$ are regular then $AB$ is regular.
- If $A$ is regular then $A^*$ is regular.

Regular Expression Conventions: To avoid excessive use of parenthesis, the following notational convention will be adopted.

- Precedence order: $\cdot$, $\circ$, $\cdot$. For example $r + s^t$ denotes $(r + ((s)^t))$.
- Associativity: $r + s + t = ((r + s) + t) = (r + (s + t))$ and $rst = ((rs)t) = (r(st))$.

$\phi$. Given input $w$, answer $w$. 

2
Problem 3. Prove the following statements.

1. For any $w \in \Sigma^*$, $L = \{w\}$ is a regular language.
2. For any finite set $L \subseteq \Sigma^*$, $L$ is regular.
3. The set of all strings $\Sigma^*$ is regular.

Let $w = a_1a_2 \ldots a_n$ ($a_i \in \Sigma$).

$\epsilon w^j = \epsilon a_1a_2 \ldots a_j \epsilon a_{j+1} \ldots a_n \epsilon$.

$\rightarrow \text{Tw. } \epsilon w^j$ is regular language. $\iff \forall n \exists w. |w|=n, \epsilon w^j$ is regular language.

证明 by induction on $|w|$.

Base case: $n=0$. $\epsilon \in \Sigma^*$ is regular language. (by def)

Ind Hyp: Assume $\forall k < n, \exists w. |w|=k, \epsilon w^j$ is regular.

Ind Step: $w = a \cdot u$, $|u|=n-1$, and $a \in \Sigma$.

$\epsilon w^j = \epsilon a \cdot \epsilon u^j$ regular.

regular \iff regular (ind hyp)

$\Sigma^*$ is regular because $\Sigma$ is regular (by 2) and $\Sigma^*$ is closure of a regular languages.

Problem 4. Describe the following regular expressions in English.

1. $(0+1)^*$
2. $\emptyset$
3. $0^* + (0*10^*10^*)^*$
4. $(0+1)^*001(0+1)^*$
5. $(10)^* + (01)^* + 0(10)^* + 1(01)^*$
6. $(\epsilon + 1)(01)^*(\epsilon + 0)$
7. $(0 + \epsilon)(1 + 10)^*$

Problem 5. Describe the following languages as a regular expression.

1. All binary strings that have 00 as a substring
2. All binary strings such that the third character from the end is 1
3. All binary strings that have 00 as a substring but do not contain 011 as a substring