Lecture L&M 1

Languages, Regular Expressions, Deterministic Finite Automata

1.1 Induction over words

A. Using induction on i, prove that $(w^R)^i = (w^i)^R$ for any word w and all $i \geq 0$, where ^{R} is the reversal operator, defined recursively by $\lambda^R = \lambda$ and $(wa)^R = aw^R$ for $a \in \Sigma$, $w \in \Sigma^+$.

   Hint: Use that $(uv)^R = v^R u^R$ for $u, v \in \Sigma^*$ (Theorem 2.1.6 in the book).

B. Use induction on the length of w to prove that $(w^R)^R = w$ for all words $w \in \Sigma^*$.

   Hint: Remember that every word w with $|w| \geq 1$ can be considered as a letter $a \in \Sigma$ followed by a word $v \in \Sigma^*$ with $|v| = |w| - 1$. You may also need Theorem 2.1.6 again.

C. *(extra)* Given an alphabet Σ consisting of k elements, let us define $f(n) := \#\{w \mid |w| \leq n\}$, i.e. $f(n)$ is the number of words over Σ with length smaller than or equal to n. Give a (non-recursive summation) formula for $f(n)$. Prove its correctness with (natural) induction.

1.2 Regular expressions

For each of the following languages L, give a regular expression E such that $\mathcal{L}(E) = L$:

D. L is the language of all words over $\{a, b, c\}$ in which all the occurrences of a precede all the occurrences of b and c, and in turn all the b precede all the c.

E. L is the same as in Exercise D, except it does not contain λ.

F. L is the language of all words over $\{a, b, c\}$ of length three.

G. L is the language of all words in $\{a, b\}^*$ that contain the substring ab and have length > 2.

H. L is the language over $\{a, b\}$ such that all words contain both aa and bb.

I. *(extra)* L is the language over $\{a, b\}$ such that all words contain both ab and ba.

1.3 Deterministic finite automata

J. Let M be the deterministic finite automaton (DFA) defined by $Q = \{q_0, q_1, q_2\}$ (with q_0 being the initial state), $\Sigma = \{a, b\}$, $F = \{q_2\}$ and transition function δ according to the following transition table:

δ	a	b
q_0	q_0	q_1
q_1	q_2	q_1
q_2	q_2	q_0

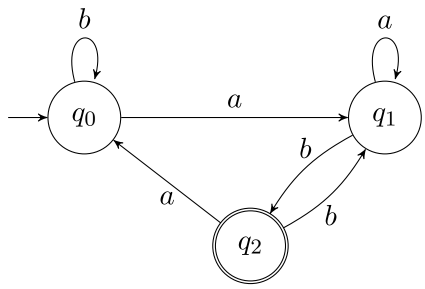
(a) Draw the state diagram for M.

(b) Trace the computations of M (i.e. the inference steps according to relation $\vdash$) when processing the words $abaa$, $bbbabb$, $bababa$ and $bbbaa$.

(c) Which of the above four words are accepted by M?

(d) Give a regular expression for $\mathcal{L}(M)$.

K. Let M be the DFA defined by the following state diagram:



(a) Construct the transition table of M.

(b) Which of the words $baba$, $baab$, $abab$ and $abaaab$ are accepted by M?

(c) Give a regular expression for $\mathcal{L}(M)$.

L. For each of the following languages L, draw the state diagram of a DFA that accepts the language. Can you also come up with a regular expression describing the language?

(a) L is the language of all words over $\{a, b\}$ that do not begin with aaa.

(b) L is the language of all words over $\{a, b\}$ that do not contain the substring aaa.

(c) *(extra)* L consists of all words of even length over $\{a, b, c\}$ that contain exactly one a.