Examination
Formal Languages and Automata Theory
TDD14

(Formella Språk och Automatateori)

2013–10–30, 08.00 – 12.00

1. Allowed help materials
   - A sheet of notes - 2 sided A5 or 1 sided A4.
     The contents is up to you.
     The notes should be signed in the same way as the exam sheets and returned together with the exam.
   - English dictionary

   (Tillåtna hjälpmedel:
   - Ett papper med valfria anteckningar - 2 sidor A5 eller 1 sida A4.
     Anteckningarna ska signeras på samma sätt som tentamensarken och bifogas tentamen vid inlämnandet.
   - Engelsk ordbok)

2. You may answer in Swedish or English.
3. Total number of credits is 31: 3: 15 p, 4: 20 p, 5: 25 p.
4. *Jour* (person on duty): Włodek Drabent, tel. (013 28) 89 29.

GOOD LUCK !

Make sure that you justify your answers! Unexplained answers will be granted 0 points. (For example, assume that you are writing a grammar for a given language. Then you should also explain that the grammar indeed generates the language).

1. *(2p)* Consider the NFAϵ whose transition function is given by the table. (Its set of states is $Q = \{1, 2, 3, 4\}$, the input alphabet $\Sigma = \{a, b, c\}$, the start state and the final state is 4.) Using a standard method construct an equivalent DFA.

	ϵ	a	b	c
1	$\emptyset$	$\emptyset$	$\emptyset$	$\{4\}$
2	$\{3\}$	$\emptyset$	$\{1, 2\}$	$\emptyset$
3	$\{1\}$	$\emptyset$	$\emptyset$	$\emptyset$
$\rightarrow$ 4 **F**	$\{2\}$	$\{3, 4\}$	$\emptyset$	$\emptyset$

2. *(2p)*

	a	b
$\rightarrow$ 1	3	5
2 **F**	1	5
3 **F**	4	6
4	2	5
5	6	2
6	6	3

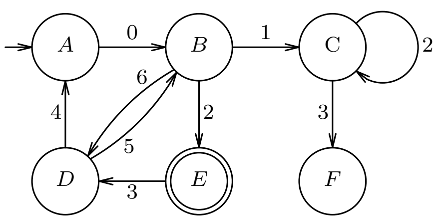
Using a standard method, construct the minimal DFA equivalent to the DFA given by the table. (Its set of states is $Q = \{1, 2, 3, 4, 5, 6\}$, the input alphabet $\Sigma = \{a, b\}$, the start state is 1, and $\{2, 3\}$ is the set of final states.)

3. *(2p)* Let L be the language accepted by the DFA from the previous problem. Consider the relation $\equiv_L$ on strings over $\Sigma = \{a, b\}$, defined by

$$x \equiv_L y \;\Leftrightarrow\; \forall z \in \Sigma^* \, (xz \in L \Leftrightarrow yz \in L).$$

How many equivalence classes does $\equiv_L$ have? Why? Choose two of the equivalence classes, and give two DFA's defining them.

4. *(2p)* Using a standard method, construct a regular expression defining the same language as the given NFA. (Its set of states is $Q = \{A, B, C, D, E, F\}$, the input alphabet $\Sigma = \{0, 1, 2, 3, 4, 5, 6\}$, the start state is A and the final state is E.)



5. *(2p)* Build a regular expression for the language

   (a) $\{\, a^i b^j c^k \mid i + k \text{ is even} \,\}$,

   (b) $\{\, a^i b^j c^k d^l \mid j + k \text{ is even} \,\} \cap \{\, a^i b^j c^k d^l \mid i + l \text{ is odd} \,\}$.

6. *(6p)* For each of the following languages answer whether it is regular, context-free but not regular, or not context-free. (Here a brief explanation is sufficient.)

(a) L_1 is the set of the strings over $\{a, b, c\}$ with even number of b's, not containing a substring abc and with each a immediately preceded by b.

(b) L_2 is the set of those strings from L_1 that are palindromes.

(c) L_3 is the set of those strings from L_2 that contain more c's than b's.

(d) $L_4 = \{\, x^R y x \mid x, y \in \{a, b\}^*,\ |y| = 3 \,\}$.

(e) $L_5 = \{\, x y x \mid x, y \in \{a, b\}^*,\ |x| = 3 \,\}$.

(Remember that x^R denotes the string x reversed.)

7. *(3p)* Prove that the language

$$L = \left\{ z \in \{a, b, c, d\}^* \;\middle|\; 0 \leq \#a(z) < \#b(z),\ 0 \leq \#c(z) < \#d(z) \right\}$$

is not regular <u>or</u> that it is not context-free. Use the appropriate pumping lemma or employ reasoning similar to the proof of the lemma.

Hint: Choose a simple string for pumping. Remember that $\#a(w)$ denotes the number of occurrences of symbol a in string w.

8. *(1p)* Explain briefly the notion of the language of a decision problem.

9. *(3p)* Show that the problem "a Turing machine M reaches state q on some input" is undecidable. Use the fact that it is undecidable whether a given Turing halts on some input. In other words, show that the language of the considered problem, this means

$$RS2 = \left\{ \langle M, q \rangle \;\middle|\; \begin{array}{l} \text{Turing machine } M \\ \text{reaches state } q \\ \text{on some input} \end{array} \right\},$$

is not recursive, using the fact that the language

$$HP2 = \left\{ \langle M \rangle \;\middle|\; \begin{array}{l} \text{Turing machine } M \\ \text{halts on some input} \end{array} \right\}$$

is not recursive.

10. *(4p)* Which of the following statements are true, which are false?

(a) $LL = \{\, xx \mid x \in L \,\}$ for some language L.

(b) If L_1 and L_2 are context-free languages then $L_1 \cap L_2$ is not context-free.

(c) There exists a non-total Turing machine (i.e. one looping on some inputs) which accepts a context-free language.

(d) There exists a language L over an alphabet Σ such that for each homomorphism $h \colon \Sigma^* \to \Sigma^*$ there exists a string $x \in L$ for which $|h(x)| \leq |x|$.

11. *(3p)* In an attempt to construct LR parsers for certain grammars, we applied the standard method of constructing a DFA for the viable prefixes of a grammar. Some fragments of the obtained DFA's are given below.

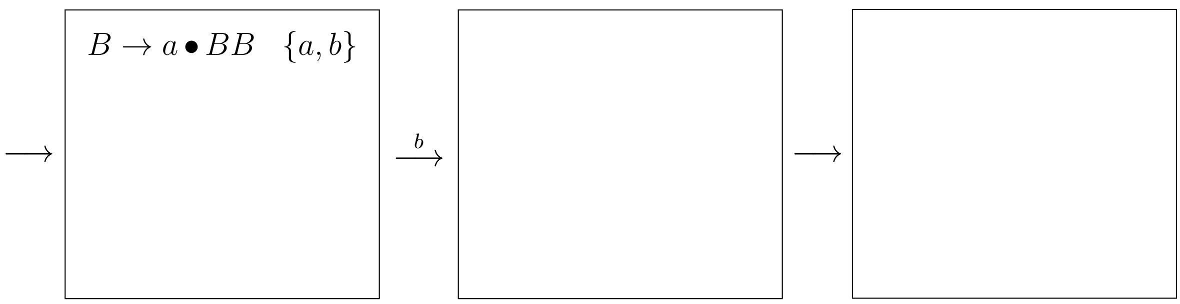
Complete the missing items in the given states, the missing lookahead sets and the missing symbols labelling the arrows. In each case answer the following questions. Justify your answers.

- Does the fragment of a DFA satisfy the conditions for the grammar to be LR(0)?
- The same question about the conditions for LR(1).

You may skip adding missing items or lookahead sets if they are not needed to answer the questions. For instance if you find the items in some state to violate the LR(1) conditions then you do not need to complete the other states.

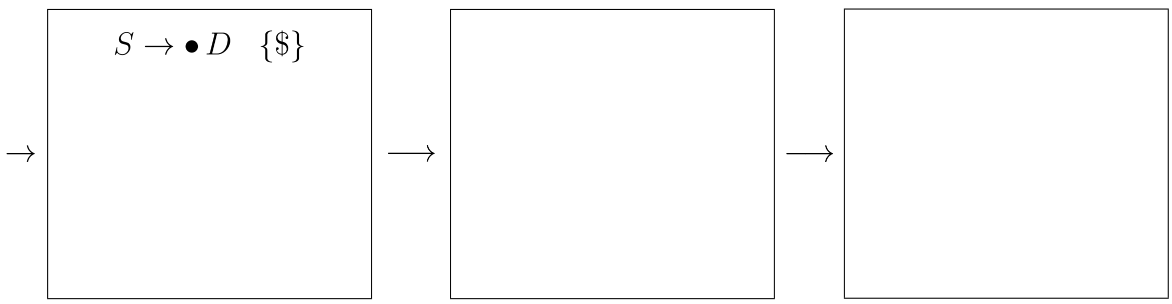
S, A, B, C, D are nonterminal symbols and a, b, c are terminal symbols of the grammars; S is the start symbol.

(a)



The productions of the grammar are $A \to aC \mid bAA$, $B \to bC \mid aBB$, $C \to aB \mid bA \mid \epsilon$, $S \to C$.

(b)



The productions of the grammar are $A \to a \mid bAA$, $B \to b \mid aBB$, $C \to aB \mid bA$, $D \to DC \mid \epsilon$, $S \to D$.

12. *(1p)* Consider the grammar from problem 11b with D removed from the alphabet and the rules for D and S replaced by $S \to \epsilon \mid CS$. Is the grammar LL(1)?