Examination
Formal Languages and Automata Theory
TDDD14 & TDDD85

(Formella Språk och Automatateori)

2015–10–21, 14.00 – 18.00

1. **NOT ALL PROBLEMS ARE FOR BOTH COURSES.**
   Pay attention to “only” comments.
2. 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)
3. You may answer in Swedish or English.
4. Total number of credits is 33. Limits:

   3: 16 p, 4: 22 p, 5: 28 p.
5. *Jour* (person on duty): Johannes Schmidt, tel. 07 25 72 18 03

GOOD LUCK !

Make sure that you justify your answers! Unexplained answers will be granted 0 points. (For instance, if you are writing a grammar for a given language then you should explain that the grammar indeed generates the language. If you apply some known method then you should explain each step. And so on.)

1. *(4p)* The DFA $M = (Q, \Sigma, \delta, s, F)$ is defined as follows:

$$Q = \{1, 2, 3, 4, 5, 6\} \qquad \Sigma = \{a, b\} \qquad s = 1 \qquad F = \{6\}$$

with the transition function δ given by

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

(a) Using the standard method, construct an equivalent, minimal DFA M_{min}.

(b) Let L be the language defined by M_{min} (and thus by M), and consider the relation $R_L \subseteq \Sigma^* \times \Sigma^*$ defined by

$$xR_Ly \Leftrightarrow (\forall z \in \Sigma^*(xz \in L \Leftrightarrow yz \in L))$$

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

2. *(2p)* Order the following formalisms according to their expressive power: placing A before B means that any language definable by A is definable by B. Also state which, if any, of them are equivalent.
(In this problem you are not required to provide a justification).

- Context-free Grammars (CFG)
- Deterministic Finite Automata (DFA)
- Deterministic Pushdown Automata (DPDA)
- LR(1) grammars
- Nondeterministic Finite Automata (NFA)
- Nondeterministic Finite Automata with ϵ-transitions (NFAϵ)
- Nondeterministic Turing Machines (NTM)
- Pushdown Automata (PDA)

- Regular expressions (reg.exp.)
- Turing Machines (TM)

Compare the expressive power of LR(0) grammars with the expressive power of the formalisms listed above.

3. *(2p)* Given the two context-free grammars $G_1 = (N_1, \Sigma, P_1, S_1)$ and $G_2 = (N_2, \Sigma, P_2, S_2)$ here is an *attempt* to construct a grammar $G = (N, \Sigma, P, S)$ defining the intersection $L(G_1) \cap L(G_2)$.

$$\begin{array}{rcl} N &=& N_1 \cup N_2 \\ S &=& S_1 \\ P &=& P_1 \cap P_2 \end{array}$$

Show that the construction is wrong. That is, find grammars G_1, G_2 such that $L(G) \neq L(G_1) \cap L(G_2)$.

4. *(3p)* Using a standard method, construct a regular expression defining the same language as the DFA whose transition function δ is given by

	a	b
$\rightarrow A$	A	C
B F	A	B
C	B	C

5. *(5p)* 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) $\{\, a^n b^n c^n \mid 0 \leq n,\ n = n + 2n \}$

   (b) $\{\, xy \mid x \in \{a^n b^n c^n \mid 0 \leq n\},\ |y| \leq 6,\ |x| < |y| \,\}$

   (c) $\{\, ww^R \mid w \in \{a^n b^n c^n \mid 0 \leq n\}$, w contains not more a's than b's$\}$

   (d) $\{\, ww^R \mid w \in \{a, b, c\}^*,\ w$ contains more a's than b's$\,\}$

   (e) the intersection of the languages from c) and d), unified with the language from a)

6. *(5p)* Which of the following statements are true, which are false? Why?

   (a) Every context free language whose complement is recursively enumerable is recursive.

   (b) The complement of a context free language which contains a recursive language is context free.

   (c) There exists a DFA for any recursive language.

(d) The language defined by a DFA is recursive.

(e) Any infinite language is regular, or context free, or recursive.

7. *(3p)* Prove using the corresponding Pumping lemma that the language

$$\{ a^n b^m c^r \mid 0 \leq n,\ 0 \leq m,\ 0 \leq r \text{ and } r = n + m \}$$

is not regular.

8. *(3p)* **Only TDDD14**
Conceive an algorithm that determines in finite time whether two given regular expressions define the same language.

9. *(3p)* **Only TDDD85**
Prove that a recursive enumerable language that contains infinitely many regular languages is not necessarily regular itself.

10. *(6p)* 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.

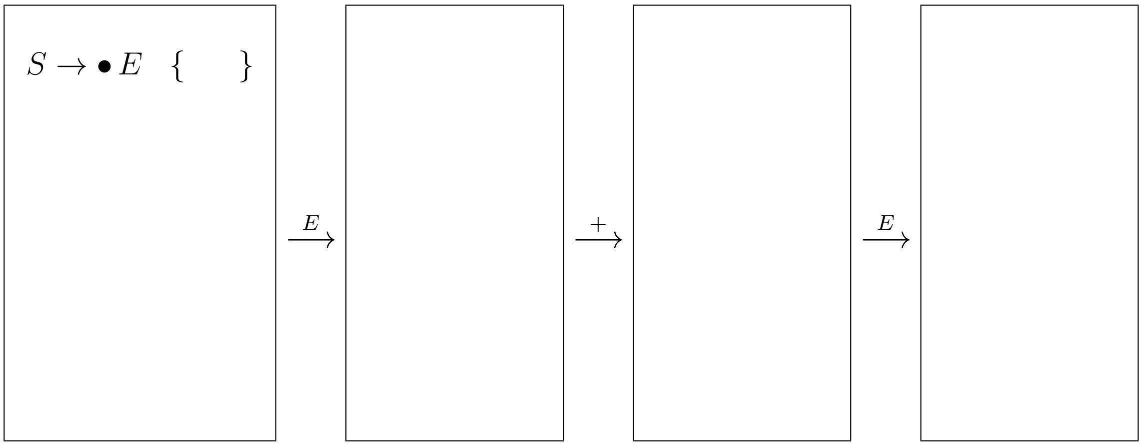
Complete the missing items in the given states, the missing lookahead sets, and the missing symbols labelling the arrows. In each case, answer the following questions. Justify your answers.

- 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. In such a case just make an appropriate comment.

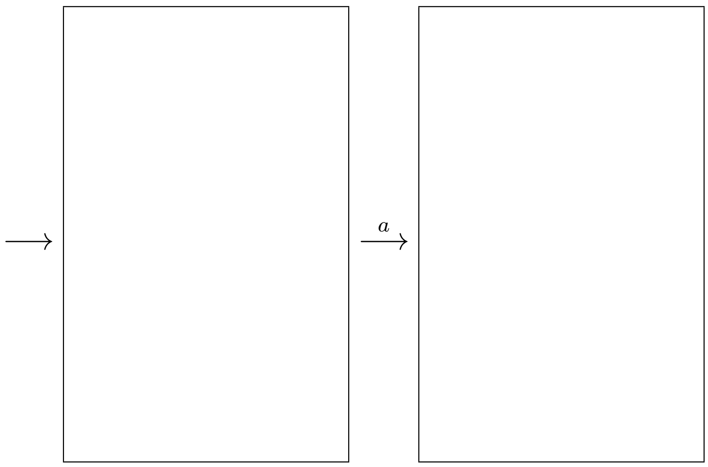
(a) x, y are terminal symbols and S, E are nonterminal symbols of this grammar; S is the start symbol.

↓



The productions of the grammar are $S \to E,\ E \to E + E \mid x \mid y$.

(b) a, b, c are terminal symbols and S, X, Y are nonterminal symbols of this grammar; S is the start symbol.



The first state is the start state.
The productions of the grammar are $S \to aXbY \mid XaYb, X \to a, Y \to c \mid Yc$.

THE END. Remember the warning from the beginning of page 2. Check that your answers are justified.