Försättsblad till skriftlig tentamen vid Linköpings universitet

(fylls i av ansvarig)

Datum för tentamen	2010-10-21
Sal	TER3
Tid	14-18
Kurskod	TDDD14
Provkod	TEN1
Kursnamn/benämning	Formella språk och automatateori/ Formal Languages and Automata Theory
Institution	*IDA*
Antal uppgifter som ingår i tentamen	12
Antal sidor på tentamen (inkl. försättsbladet)	4 blad
Jour/Kursansvarig	Wlodek Drabent
Telefon under skrivtid	28 8929
Besöker salen ca kl.	2 ggr (första gången ung. 14.40)
Kursadministratör (namn + tfnnr + mailadress)	Madeleine Häger Dahlqvist 28 2360, madha@ida.liu.se
Tillåtna hjälpmedel	Se tentamens första sida
Övrigt (exempel när resultat kan ses på webben, betygsgränser, visning, övriga salar tentan går i m.m.)	
Vilken typ av papper ska användas, rutigt eller linjerat	rutat
Antal exemplar i påsen	52

Examination
Formal Languages and Automata Theory
TDDD14

(Formella Språk och Automatateori)

2010–10–21, 14.00 – 18.00

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. 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 $Q = \{0, 1, 2, 3\}$, the input alphabet $\Sigma = \{a, b\}$, the start state is 0, and 1 is the only final state.) Using a standard method construct an equivalent DFA.

	ϵ	a	b
$\rightarrow 0$	$\{1\}$	$\{1\}$	$\{2\}$
1 **F**	$\emptyset$	$\emptyset$	$\{1, 3\}$
2	$\{3\}$	$\emptyset$	$\emptyset$
3	$\emptyset$	$\{0, 3\}$	$\{2\}$

2. *(3p)* Construct a DFA for the language $L_1 \subseteq \{0, 1\}^*$ of the strings not containing a substring 101.

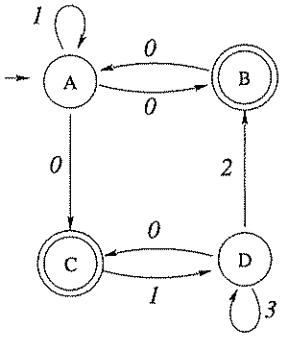
   Construct a DFA for the language $L_2 \subseteq \{0, 1\}^*$ of the strings with an even number of 1's and not containing a substring 101

   (The justification of your answer should contain explanation of the role of each state of your automata.)

3. *(2p)* Using a standard method, construct the minimal DFA equivalent to your DFA for language L_1 from the previous problem.

   Your DFA may already be minimal. In such case you may provide a proof that it is minimal (instead of constructing a minimal automaton in a standard way

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



5. *(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)
$$L_1 = \left\{ xcby \;\middle|\; \begin{array}{l} x, y \in \{a,b,c\}^*, \;\; \#c(x) \equiv 3 \bmod 4, \\ \text{each } b \text{ in } x \text{ is immediately preceded by } a, \\ y \text{ does not contain a substring } cba \end{array} \right\}$$

($\#a(w)$ denotes the number of occurrences of symbol a in string w),

(b) $L_2 = \{ a^j b^{m+l} c^m d^{j+l} \mid j, l, m \geq 0 \}$,

(c) $L_3 = \{ a^j b^{l+2} c^{j+3} d^l \mid j, l \geq 0 \}$,

(d) L_4 is the image of L_3 under the homomorphism $h\colon \{a,b,c,d\}^* \to \{0,1,2\}^*$ such that $h(a) = 10$, $h(b) = \epsilon$, $h(c) = 1010$, $h(d) = 101010$.

(e) $L_5 = \{ a^j b^{l+2} c^{j+3} d^l \mid 0 < l,\ 0 \leq j < 3 \}$,

6. *(3p)* Prove that the language
$$L_6 = \{ x \in \{a,b,c\}^* \mid 3 \cdot \#a(x) < |x| - 3 \}$$
is not regular, or that
$$L_7 = \{ a^j b^m c^{2j} \mid 0 < j < m \}$$
is not context-free. Use the appropriate pumping lemma or employ reasoning similar to the proof of the lemma.

7. *(1p)* What are practical consequences of a fact that a certain problem is undecidable?

8. *(3p)* Let us say that a state of a Turing machine M is useless if it is never reached in any computation of M.

Show that the problem "a state q of a Turing machine M is useless" is undecidable. Use the fact that it is undecidable whether a given Turing machine halts on some input. In other words, show that the language of the considered problem, this means
$$US = \left\{ \langle M, q \rangle \;\middle|\; \begin{array}{l} \text{Turing machine } M \\ \text{does not reach state } q \\ \text{on any input} \end{array} \right\},$$
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.

9. *(2p)* Explain briefly the notion of a universal Turing machine (UTM). What is the language accepted by it? To which class of languages does it belong?

10. *(3p)* Which of the following statements are true, which are false? Justify your answers.

    (a) The language $\{\epsilon\}$ is not regular.

    (b) Symbol A in grammar $S \to ab \mid aAb;\ \ A \to aAb \mid bAa$ is useless.

    (c) If h is a homomorphism and L a language then $h^{-1}(h(L)) = L$.

11. *(1p)* Explain why the grammar with productions

$$S \to Ab \qquad\qquad A \to aSc \mid bSac \mid \epsilon$$

    is not LL(1). (S, A are the nonterminal symbols, and a, b, c are the terminal symbols of the grammar.)

12. *(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.

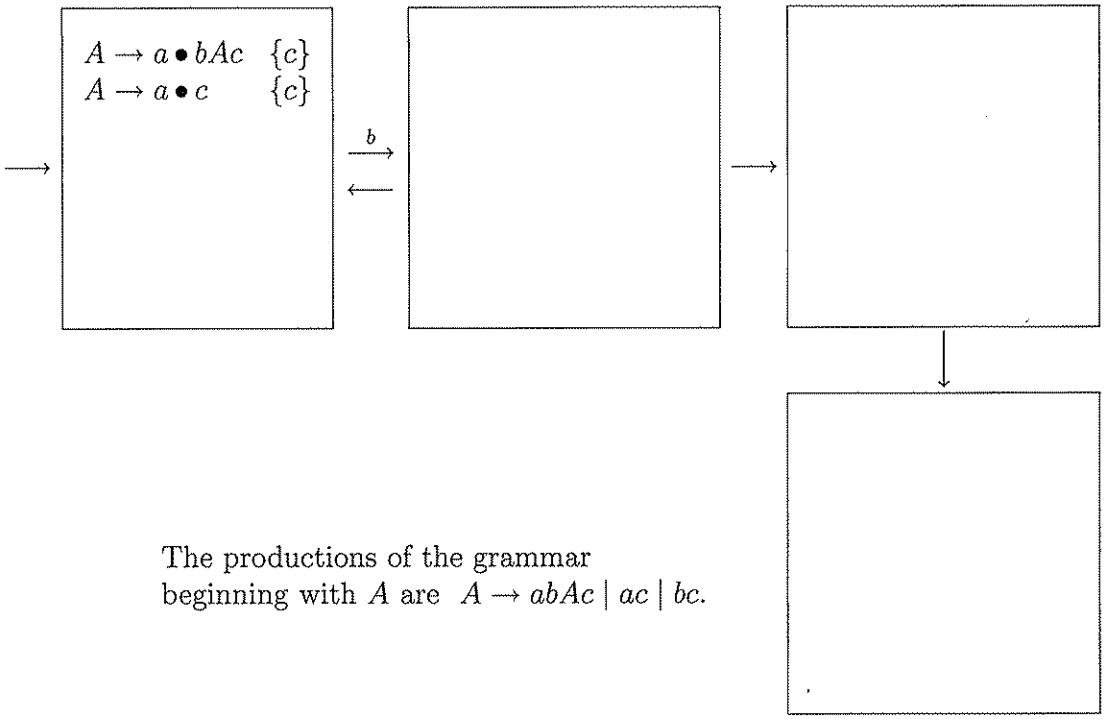
    Complete the missing items in the given states, the missing lookahead sets and the missing symbols labelling the arrows. In each case answer the following questions. Justify your answers.

    - 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.

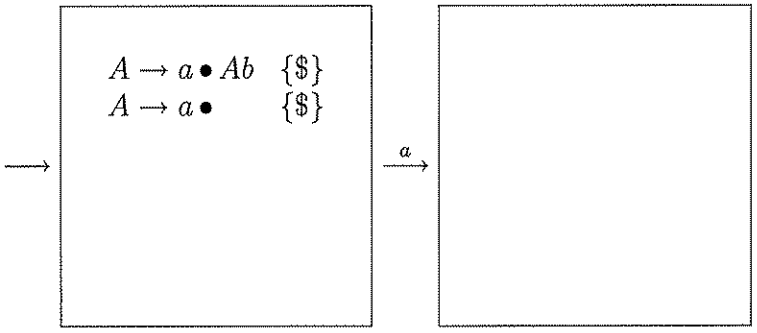
    a, b, c are terminal symbols and S, A, B are nonterminal symbols of the grammars; S is the start symbol.

(a)



The productions of the grammar beginning with A are $A \to abAc \mid ac \mid bc$.

(b)



The productions of the grammar beginning with A are $A \to aAb \mid bAc \mid a$.