SCHOOL OF MATHEMATICAL AND COMPUTER SCIENCES

Computer Science

F29LP2

Language Processors (Mock)

Semester 2 201314

Sometime before 5 May 2014

Duration: As long as you like

ANSWER BOTH QUESTIONS (ACTUAL EXAM WILL BE THREE)

Answer each question in a separate script book.

Some words on using this mock paper

There is no concept in this paper that you have not seen already in the lecture notes and exercises. However, I have tried to pitch the difficulty level of this paper slightly above what you will face in the exam. Exam conditions are always harder, because of the stress.

I believe that if you can understand and do these questions, then you are certain to get a decent grade in the exam.

You *must* attempt this entire paper *before* looking at the answers. Have you attempted the paper yet?

Good luck.

1. (a) Explain in clear and precise English the precise meaning of the term *formal language*, in the context of this course. (2)

(b) Consider the following regular expressions:

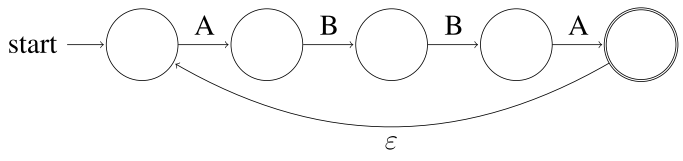
1. .?
2. .+
3. .*
4. .!
5. .$
6. .
7. .$$
8. .^

In English or otherwise, explain what languages (over ASCII characters) these regular expressions specify. (8)

(c) 1. Explain in English what a non-deterministic finite automaton (NFA) with ε-moves is.
2. Explain intuitively how an NFA with ε moves can be considered to specify a language.
3. Explain the connection with regular expressions.

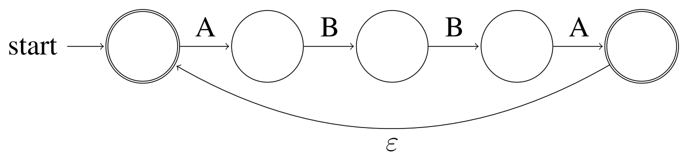
(3)

(d) Express as a regular expression the language accepted by the following automaton:



(2)

(e) Explain precisely, in English or otherwise, what the difference is between the previous regular expression and the one determined by this automaton:



(1)

(f) Draw a PDA that recognises the language $\{a^i b^j a^{i+j} \mid i \geq 1, j \geq 0\}$. Your answer must clearly state the acceptance mode used. (4)

2. (a) Give one example each of

- a left-recursive grammar, (1)
- a right-recursive grammar, (1)
- a grammar that is both left- and right-recursive. (1)

(b) Write a context-free grammar for the English language with nonterminals $\langle sentence \rangle$, $\langle noun \rangle$, $\langle verb \rangle$, $\langle definite\text{-}article \rangle$ (words like 'the' or 'that'), and $\langle adverb \rangle$ ('quickly', 'happily'). Your grammar should be sufficiently developed to produce the following sentences:

- The cat scratched the mat.
- Linux rocks.
- Jamie happily writes questions.

We do not care if your grammar also produces a incorrect sentences, such as "The the cat scratched the mat". You may ignore case. (6)

(c) Consider the following grammars:

$$\begin{aligned} T &::= T0 \mid T1 \mid \varepsilon \\ S &::= 0S \mid 1S \mid \varepsilon \\ U &::= UU \mid 0 \mid 1 \mid \varepsilon \end{aligned}$$

- All three grammars generate the same language. What is it? (1)
- Rank the grammars in order from best to worst from an implementational point of view, and explain your ranking. (2)

(d) Take a *natural number* to be an element of the language determined by the regex $0 \mid$ [1-9][0-9]*, and a *decimal number* to be an element of the language determined by the regex $(0 \mid$ [1-9][0-9]$^*)\backslash.$[0-9]$^+$ (so 00 is not a number but 10 is, and $1.$ is not a decimal number but 0.00 and 0.01 are decimal numbers).

Write a grammar (which need not be context-free) that will generate sentences over tokens $\{0, \ldots, 9, ., \approx\}$ of the form "$D \approx N$", where D is a decimal number and N is a natural number and N is equal to D rounded down to the nearest whole number.

So for instance, your grammar should recognise $10.9 \approx 10$ and $0.49 \approx 0$.

You may use dots notation to indicate evident repetition of a succession of rules, as in "$S ::= 0 \mid \cdots \mid 9$". Answers that are not evidently correct may score zero marks; if in doubt, provide a clear English explanation of how your answer works. Clearly state the start symbol. (4)

(e) Write a grammar to recognise sentences over $\{1, 2\}$ such that the sum of the 1s is equal to the sum of the 2s (in other words: there are twice as many 1s as 2s). Clearly state the start symbol. (2)

(f) Can your grammar be left-factored and so made deterministic to eliminate potential backtracking? Explain. (2)

END OF PAPER