

PRINCE OF SONGKLA UNIVERSITY
FACULTY OF ENGINEERING
Department of Computer Engineering

Midterm Examination: Semester 1

Academic Year: 2007-2008

Date: 5th August, 2007

Time: 13:30 – 15:30 (2 hours)

Subject Number: 240-340

Room: A201

Subject Title: Compiler Structures

Lecturer: Aj. Andrew Davison

Exam Duration: 2 hours

This paper has 3 pages.

Authorized Materials:

- Writing instruments (e.g. pens, pencils).
- Books (e.g. dictionaries) and calculators are **not** permitted.

Instructions to Students:

- *Answer questions in English.* Perfect English is **not** required.
- Attempt all questions.
- Write your answers in an answer book.
- Start your answer to each question on a new page
- Clearly number your answers.
- Any unreadable parts will be considered wrong.
- When writing programs, use good layout, and short comments; marks will not be deducted for minor syntax errors.
- The marks for each part of a question are given in brackets (...).

Question 1

(20 minutes; 20 marks)

- Write a regular expression for the time of day (e.g. 9:17am, 12:30pm). (5)
- Write a regular expression for a dollar amount, with optional cents (e.g. \$23, \$5.23). (5)
- Write a regular expression that matches a lowercase word containing the five vowels in the order a, e, i, o, u. Each vowel appears only once in the word. (5)
- Write a lex program which counts the number of lines in an input file. (5)

Question 2

(20 minutes; 20 marks)

- Briefly explain the four parts of a grammar. Do **not** include any program code. (5)
- Write a context-free grammar that generates all palindromes of a 's and b 's. A palindrome is a string that reads the same backwards and forwards (e.g. abba, babab).

Draw a parse tree for the string abaaba using your grammar. (5)

- Write a context-free grammar for all strings of a 's and b 's in any order, containing the same number of a 's and b 's (e.g. abab, abbaab).

Draw a parse tree for the string bbbaaa using your grammar. (10)

Question 3

(25 minutes; 25 marks)

- What are FIRST sets? Define FIRST() using FIRST_SEQ(). (8)
- What are FOLLOW sets? Define FOLLOW() using FIRST_SEQ(). (8)
- Calculate the FIRST and FOLLOW sets for the grammar:

$$S \rightarrow a S e \mid B$$

$$B \rightarrow b B e \mid C$$

$$C \rightarrow c C e \mid d$$

Show all your working. (9)

Question 4 is on the Next Page.

Question 4

(55 minutes; 55 marks)

Consider the following Cmds grammar:

```
Cmds → Cmds Cmd | Cmd
Cmd → 'move' [ INT ] 'fwd' |
      'move' [ INT ] 'back' |
      'turn' [ INT ] 'left' |
      'turn' [ INT ] 'right'
```

INT is a token representing a positive integer.

- Explain the problems with left recursion for LL(1) grammars, and remove the left recursion from the Cmds grammar. (5)
- Explain left factoring for LL(1) grammars, and carry out left factoring on the Cmds grammar. (5)
- Calculate the FIRST and FOLLOW sets for the rewritten Cmds grammar. (10)
- What are PREDICT sets? Define PREDICT() using FIRST() and FOLLOW(). Do not define FIRST() and FOLLOW(). (10)
- Use PREDICT() to check that the rewritten Cmds grammar is LL(1). (10)
- Explain how a recursive descent parser is generated from a LL(1) grammar by translating the rewritten Cmds grammar into C functions. Include code for main(), but **not** for nextToken(), match(), error(), or scanner().

Show all your working, including how you obtain the if-tests in the functions. (15)

--- End of Examination ---