

The Questions

1. [COMPULSORY]

- (a) Provide three example functions that the Context Handling phase of compilation performs. Subsequently, provide three reasons for incorporating an intermediate code generation stage during compilation. [6]
- (b) Provide the formal definition of a Linearly Bounded Automaton (LBA) and describe the differences between a Push Down Automaton (PDA) and a Linearly-Bounded Automaton (LBA). [6]
- (c) Within Chomsky's hierarchy of grammars, describe the main difference between a type-2 and a type-3 grammar, and provide two example production rules (one for each type of grammar) that illustrate this difference. [6]
- (d) Describe the data structures involved, and the steps performed by the LR parsing algorithm. [6]
- (e) Within the context of converting a Non-Deterministic Finite Automaton (NFA) to a Deterministic Finite Automaton (DFA), describe the subset construction algorithm by providing the two functions required by the algorithm, as well as the steps involved in the operation of the algorithm. [8]
- (f) Provide the definition of LL(1) grammars; explain why the grammar below is not LL(1), and use left-factoring to transform it to its LL(1) equivalent. [8]
- $X \rightarrow yXa$
 $X \rightarrow yXb$
 $X \rightarrow c$

2. You are required to construct the minimal deterministic finite state automaton (DFA) for the regular expression $a^*(b|c|d)^*e$ following the steps below.

- (a) Construct a non-deterministic finite automaton (NFA) using Thompson's algorithm. [12]
- (b) Construct the equivalent DFA using the subset construction algorithm. *Explain the intermediate steps you have taken.* [12]
- (c) Describe the DFA minimization algorithm, and subsequently apply it to the DFA you have constructed in (b). Show whether your DFA was already minimal or not. *Explain the intermediate steps of the application of the DFA minimization algorithm* [6]

3. (a) In the context of shift-reduce parsing, and in order to construct the parsing table, the functions $Closure(I)$ and $goto(I, X)$ need to be defined (where I is a set of items, and X is a grammar symbol). Provide the definitions of these functions. [6]

(b) Provide the algorithm for computing the canonical LR(0) collection of sets of items. [6]

(c) For the grammar below, construct the canonical set of LR(0) items and provide the DFA that can recognize viable prefixes for the grammar:

$$\begin{aligned} G' &\rightarrow G \\ G &\rightarrow G - X \mid X \\ X &\rightarrow X * F \mid F \\ F &\rightarrow (G) \mid a \end{aligned}$$

[18]

4. (a) Calculate the FIRST and FOLLOW sets for all non-terminal symbols for the grammar below where $\{a, -, *, (,)\}$ are terminals, and $\{G', G, T, T', F\}$ are non-terminals:

$$\begin{aligned} (1) \quad &G \rightarrow T G' \\ (2) \quad &G' \rightarrow - T G' \mid \epsilon \\ (3) \quad &T \rightarrow F T' \\ (4) \quad &T' \rightarrow * F T' \mid \epsilon \\ (5) \quad &F \rightarrow (G) \mid a \end{aligned}$$

[10]

(b) A partially constructed parsing table for the grammar above is given below, with $\$$ denoting the end of input marker. Complete ALL remaining table entries, clearly marking any error entries as shown.

(NB: Make sure you copy this table to your exam booklet)

[20]

Non-terminal	Input Symbol					
	a	-	*	()	\$
G			Error			Error
G'					$G' \rightarrow \epsilon$	
T				$T \rightarrow FT'$		
T'			$T' \rightarrow *FT'$			
F	$F \rightarrow a$	Error				