UNIVERSITY OF BRISTOL

Winter 2024 Examination Period

SCHOOL OF COMPUTER SCIENCE

Second Year PRACTICE Examination for the Degrees of Bachelor of Science Master of Engineering

COMS20007W
Programming Languages and Computation

TIME ALLOWED: 3 Hours

This paper contains *three* questions, worth 40, 30 and 30 marks respectively. Answer *all* questions. The maximum for this paper is 100 marks. Credit will be given for partial answers.

Other Instructions:

Candidates may bring to the exam room 1 double-sided A4 page of notes in any format. A reminder of key definitions is provided at the back of this paper.

TURN OVER ONLY WHEN TOLD TO START WRITING

- Q1. This question is about syntax.
 - *(a) Consider the following grammar over terminal symbols $\{a, b\}$:

$$S \longrightarrow aSa \mid bSb \mid \epsilon$$

- i. Give two examples of words over $\{a, b\}$ that are derivable in the grammar.
- ii. Give two examples of words over $\{a, b\}$ that are not derivable in the grammar.
- iii. Is the following statement true or false? Every word derivable in the grammar has even length.

[5 marks]

- *(b) Consider each of the following grammars over the alphabet $\{a,b,c\}$. In each case, the start symbol is S.
 - 1.

$$S \longrightarrow aSaS \mid bS \mid cS \mid \epsilon$$

2.

$$\begin{array}{ccc} S & \longrightarrow & TabbT \mid TbbaT \\ T & \longrightarrow & aT \mid bT \mid cT \mid \epsilon \end{array}$$

3.

$$\begin{array}{ccc} S & \longrightarrow & bTb \\ T & \longrightarrow & aT \mid bT \mid cT \mid \epsilon \end{array}$$

4.

$$\begin{array}{ccc} S & \longrightarrow & XSX \mid \epsilon \\ X & \longrightarrow & a \mid b \mid c \end{array}$$

5.

$$S \longrightarrow bS \mid cS \mid \epsilon$$

Match each of the following descriptions of languages to the regular expression above that denotes it:

- i. The language of all words that start and end with b.
- ii. The language of all words that do not contain a.
- iii. The language of all even length words.
- iv. The language of all words containing an even number of a.
- v. The language of all words that either contain abb or bba as a substring.

[5 marks]

*(c) Consider the following grammar for the syntax of Combinatory Logic:

$$M \longrightarrow \mathsf{var} \mid k \mid s \mid M \ M \mid (\ M\)$$

whose 5 terminal symbols are:

The nullable, first and follow maps for the non-terminals are:

- Nullable(M) = false
- $First(M) = \{var, k, s, (\}$
- Follow $(M) = \{ \mathsf{var}, k, s, (,) \}$
- i. Draw the parsing table for this grammar.
- ii. Is the grammar LL(1)?

[10 marks]

- ** (d) For each of the following sets of words over $\{a,b\}$, design a context-free grammar that expresses the set:
 - i. All words whose length is a multiple of 3, e.g. abb, ababba.
 - ii. All words that start and end with a different letter, e.g. abbaab.
 - iii. All words that contain a letter b exactly two places from the end, e.g. aabab, baa.
 - iv. All words that do not contain the substring aa.

[6 marks]

** (e) Give an LL(1) grammar equivalent to the following context-free grammar:

$$S \longrightarrow \emptyset \mid (S) \mid \mathsf{atom} \mid S \cup S \mid S \cap S \mid S^c$$

whose terminal symbols are:

$$\emptyset$$
 () atom \cup \cap c

[4 marks]

*** (f) Show that the following language over $\{0,1\}$ can be expressed by a context-free grammar and justify your construction.

$$\{1^k w \mid k \ge 1, w \in \Sigma^*, \#_1(w) \ge k\}$$

where $\#_1(v)$ counts the number of 1 characters in the word v, e.g. $\#_1(0010110) = 3$. [5 marks]

***(g) Define the following indexed family of words w_i by recursion on $i \in \mathbb{N}$:

$$w_0 = a$$
$$w_{k+1} = a + w_k$$

For example, $w_3 = a + a + a + a + a$ and $w_5 = a + a + a + a + a + a + a$.

Prove that every word in the language $\{w_i \mid i \in \mathbb{N}\}$ is derivable in the following grammar (whose start symbol is S):

$$\begin{array}{ccc} S & \longrightarrow & a \; U \\ U & \longrightarrow & + \; a \; U \mid \epsilon \end{array}$$

[5 marks]

- Q2. This question is about semantics.
 - *(a) For each of the following, indicate whether it represents a valid arithmetic expression, a valid Boolean expression, or neither. In each case, if the expression is valid, evaluate the appropriate denotation function in the state $[x \mapsto 1, y \mapsto 2, z \mapsto 3]$.

i.
$$x + 10 < 6 * (-42 - y)$$

ii.
$$x \leftarrow z - (42 + y)$$

iii. true && (false
$$|| 42 * x < 0$$
)

$$v. \ w * 2 = c + d$$

[5 marks]

** (b) Suppose we add a new form of arithmetic expressions — the *integer exponentiation* operator so that the grammar of arithmetic expressions is now defined as follows:

$$A \longrightarrow n \mid x \mid A + A \mid A - A \mid A * A \mid A ^ A$$

We extended the denotation function for arithmetic expressions with the equation:

$$[\![e_1 \hat{} e_2]\!]_{\mathcal{A}}(\sigma) = \begin{cases} 0 & \text{if } [\![e_2]\!]_{\mathcal{A}}(\sigma) < 0 \\ [\![e_1]\!]_{\mathcal{A}}(\sigma)^{[\![e_2]\!]_{\mathcal{A}}(\sigma)} & \text{otherwise} \end{cases}$$

- i. Find two arithmetic expressions $e_1 \in \mathcal{A}$ and $e_2 \in \mathcal{A}$ such that the arithmetic expression $x \hat{\ } (e_1 + e_2)$ is *not* semantically equivalent to the arithmetic expression $(x \hat{\ } e_1) \cdot (x \hat{\ } e_2)$.
- ii. Prove that the arithmetic expression e^2 is semantically equivalent to the arithmetic expression e*e for an any given arithmetic expression $e \in \mathcal{A}$.
- iii. Let $S_1 \in \mathcal{S}$ and $S_2 \in \mathcal{S}$ be arbitrary While statements. Prove that " $\langle \text{if } x = 1 \text{ then } x \leftarrow x \hat{\ } x; \ S_1 \text{ else } S_2, \ \sigma \rangle \rightarrow^* \sigma'$ " if, and only if, the statement " $\langle \text{if } x = 1 \text{ then } S_1 \text{ else } S_2, \ \sigma \rangle \rightarrow^* \sigma'$.

[10 marks]

***(c) Consider the While program shown in Figure 1.

while
$$b \le a do$$

 $a \leftarrow a - b;$
 $q \leftarrow q + 1$

Figure 1: A simple While program

i. For each of the following states, indicate whether the program terminates when executed in that initial state, and the values of q and a in the final state (if it exists). You do not need to state the corresponding trace.

1.
$$[a \mapsto 25, b \mapsto 3]$$

- 2. $[a \mapsto 25, b \mapsto -12]$
- 3. $[a \mapsto 25, b \mapsto 0]$
- 4. $[a \mapsto -25, b \mapsto 10]$
- 5. $[a \mapsto 10, b \mapsto 3]$
- ii. Find a loop invariant ${\cal I}$ from which you may conclude:

$${a = n \&\& q = 0} P {n < b * (q + 1)}$$

for some fixed integer $n \in \mathbb{Z}$.

Justify why I is a loop invariant in relation to the strongest post-condition of the loops body, and how it can be used to conclude the above triple.

[15 marks]

- Q3. This question is about computability.
 - *(a) Show that the function $f: \mathbb{N} \longrightarrow \mathbb{N}$ defined by

$$f(x) \begin{cases} \simeq 2^x - 1 & \text{if } x \text{ is even} \\ \uparrow & \text{otherwise} \end{cases}$$

is computable. [5 marks]

- *(b) State whether each of the following statements is true or false.
 - The set of prime numbers is decidable.
 - If a function has an inverse, it must be an injection.
 - Every surjection has an inverse.
 - WHILE programs compute partial functions.
 - If a function is computable then it must be an injection.

[5 marks]

- ** (c) Let $f:A\to B$ and $g:B\to C$. Show that if $g\circ f:A\to C$ is injective, then so is f. [3 marks]
- ** (d) Show that the predicate

$$U = \left\{ \ulcorner S \urcorner \mid \text{for all } k \leq 2023 \text{ it is true that } \left[\!\left[S\right]\!\right]_{\mathbf{x}}(k) = \left[\!\left[S\right]\!\right]_{\mathbf{x}}(k+1) \right\}$$

is semi-decidable. (The use of "=" here means that both sides of the equality must be defined and equal.) [5 $\it marks$]

***(e) Show that the predicate

$$V = \left\{ \ulcorner S \urcorner \mid \text{there exists } k \in \mathbb{N} \text{ such that } \llbracket S \rrbracket_{\mathbf{x}} \left(k \right) = \llbracket S \rrbracket_{\mathbf{x}} \left(k + 1 \right) \right\}$$

is undecidable (The use of "=" here means that both sides of the equality must be defined and equal.) [5 marks]

*** (f) Show that the following predicate is undecidable:

$$P = \{ \langle \ulcorner S_1 \urcorner, \, \ulcorner S_2 \urcorner \rangle \mid \text{for all } n \in \mathbb{N} \colon [\![S_1]\!]_x(n) \simeq 1 \text{ iff } [\![S_2]\!]_x(n) \simeq k \text{ where } k \neq 1 \ \}$$

[7 marks]

Reminder of Important Definitions

Grammars

A Context Free Grammar (CFG) consists of four components:

- An alphabet of terminal symbols.
- A finite, non-empty set of *non-terminal* symbols, disjoint from the terminals.
- A finite set of *production rules*.
- A designated non-terminal called the start symbol.

A sentential form, usually α , β , γ and so on, is just a finite sequence of terminals and nonterminals.

The sentential form α can make a *derivation step* to β , written $\alpha \to \beta$, just if:

- α has shape $\gamma_1\,X\,\gamma_2$ and β has shape $\gamma_1\,\delta\,\gamma_2$
- and there is a production rule $X := \delta$ in the grammar

A *derivation sequence* is a non-empty sequence of sentential forms α_1 , α_2 , ... α_{k-1} , α_k in which consecutive elements of the sequence are derivation steps:

$$\alpha_1 \to \alpha_2 \to \cdots \to \alpha_{k-1} \to \alpha_k$$

A sentential form β is *derivable* from α , written $\alpha \to^* \beta$ just if there is a derivation sequence starting with α and ending with β .

We say that a word w is in the *language of a grammar* G with start symbol S, and write $w \in L(G)$ just if $S \to^* w$.

Nullable

On nonterminals:

$$\mathsf{Nullable}(X) \text{ iff } X \to^* \epsilon$$

On sentential forms:

$$\mathsf{Nullable}_s(\alpha) = \begin{cases} \mathsf{true} & \text{if } \alpha = \epsilon \\ \mathsf{false} & \text{if } \alpha \text{ is of shape } a\beta \\ \mathsf{Nullable}(X) \land \mathsf{Nullable}_s(\beta) & \text{if } \alpha \text{ is of shape } X\beta \end{cases}$$

First

On nonterminals:

$$\mathsf{First}(X) = \{ a \mid \exists \beta. \, X \to^* a\beta \}$$

On sentential forms:

$$\mathsf{First}_s(\alpha) = \begin{cases} \emptyset & \text{if } \alpha = \epsilon \\ \{a\} & \text{if } \alpha \text{ is of shape } a\beta \\ \mathsf{First}(X) & \text{if } \alpha \text{ is of shape } X\beta \text{ and } \neg \mathsf{Nullable}(X) \\ \mathsf{First}(X) \cup \mathsf{First}_s(\beta) & \text{if } \alpha \text{ is of shape } X\beta \text{ and } \mathsf{Nullable}(X) \end{cases}$$

Follow

On nonterminals:

$$\mathsf{Follow}(X) = \{ a \mid \exists \alpha \beta. \ S \to^* \alpha X a \beta \}$$

Parse Tables and LL(1)

We define the *parsing table*, usually T, for a given grammar as a 2d array indexed by pairs of a nonterminal and a terminal. Each entry T[X,a] is a set of production rules from the grammar, such that some rule $X \longrightarrow \beta$ is in the set T[X,a] just if, either:

- 1. $a \in \mathsf{First}_s(\beta)$
- 2. or, Nullable_s(β) and $a \in Follow(X)$

A grammar whose parsing table contains at most one rule in each cell is called LL(1).

Abstract Syntax of Arithmetic Expressions

An arithmetic expression is a tree described by the following grammar:

$$A := n \mid x \mid A + A \mid A - A \mid A * A$$

where n ranges over integer literals, and x ranges over variables. Parentheses are used to resolve ambiguity and to indicate the structure of the tree. We write \mathcal{A} for the set of arithmetic expressions.

Abstract Syntax of Boolean Expressions

A Boolean expression is a tree described by the following grammar.

$$B ::= \mathsf{false} \mid \mathsf{true} \mid !B \mid B \& \& B \mid B \parallel B \mid A = A \mid A \leq A$$

Parentheses are used to resolve ambiguity and to indicate the structure of the tree. We write \mathcal{B} for the set of Boolean expressions.

Abstract Syntax of Statements

A statement is a tree described by the following grammar:

$$S \coloneqq \mathsf{skip} \mid x \leftarrow A \mid S; S \mid \mathsf{if} \ B \ \mathsf{then} \ S \ \mathsf{else} \ S \mid \mathsf{while} \ B \ \mathsf{do} \ S$$

Braces " $\{\cdots\}$ " are used to resolve ambiguity and to indicate the structure of the tree. We write S for the set of statements.

States

A state is a total function from the set State = Var $\to \mathbb{Z}$, where Var is the set of variables. We write $[x_1 \mapsto v_1, \, x_2 \mapsto v_2, \, \dots, \, x_n \mapsto v_n]$ to indicate the state that maps the variable $x_i \in \mathsf{Var}$ to the value $v_i \in \mathbb{Z}$ for all $i \le n$. By convention, any variable not explicitly mentioned by a given state σ is assigned the value 0.

For a given state $\sigma \in \mathsf{State}$, we write $\sigma[x \mapsto v]$ for some variable $x \in \mathsf{Var}$ and $v \in \mathbb{Z}$ to denote the state that maps the variable x to v and any other variable y to the value $\sigma(y)$.

Semantics of Arithmetic Expressions

The denotation function for arithmetic expressions $[\![\cdot]\!]_{\mathcal{A}} \in \mathcal{A} \to (\mathsf{State} \to \mathbb{Z})$, which is defined by recursion in Figure 2. We say that two arithmetic expressions $e_1, e_2 \in \mathcal{A}$ are semantically equivalent if, and only if, $[\![e_1]\!]_{\mathcal{A}}(\sigma) = [\![e_2]\!]_{\mathcal{A}}(\sigma)$ for all states $\sigma \in \mathsf{State}$.

Figure 2: Definition of the denotational semantics of arithmetic expressions.

Semantics of Boolean Expressions

The denotation function for Boolean expressions $[\![\cdot]\!]_{\mathcal{B}} \in \mathcal{B} \to (\mathsf{State} \to \mathbb{B})$ is defined by recursion in Figure 3. We say that two Boolean expressions $e_1, e_2 \in \mathcal{B}$ are semantically equivalent if, and only if, $[\![e_1]\!]_{\mathcal{B}}(\sigma) = [\![e_2]\!]_{\mathcal{B}}(\sigma)$ for all states $\sigma \in \mathsf{State}$.

$$\begin{split} & \llbracket \mathsf{false} \rrbracket_{\mathcal{B}}(\sigma) &= \bot \\ & \llbracket \mathsf{true} \rrbracket_{\mathcal{B}}(\sigma) &= \top \\ & \llbracket !e \rrbracket_{\mathcal{B}}(\sigma) &= \neg \llbracket e \rrbracket_{\mathcal{B}}(\sigma) \\ & \llbracket e_1 \ \&\& \ e_2 \rrbracket_{\mathcal{B}}(\sigma) &= \llbracket e_1 \rrbracket_{\mathcal{B}}(\sigma) \wedge \llbracket e_2 \rrbracket_{\mathcal{B}}(\sigma) \\ & \llbracket e_1 \ \| \ e_2 \rrbracket_{\mathcal{B}}(\sigma) &= \llbracket e_1 \rrbracket_{\mathcal{B}}(\sigma) \vee \llbracket e_2 \rrbracket_{\mathcal{B}}(\sigma) \\ & \llbracket e_1 \ = e_2 \rrbracket_{\mathcal{B}}(\sigma) &= \llbracket e_1 \rrbracket_{\mathcal{A}}(\sigma) = \llbracket e_2 \rrbracket_{\mathcal{A}}(\sigma) \\ & \llbracket e_1 \le e_2 \rrbracket_{\mathcal{B}}(\sigma) &= \llbracket e_1 \rrbracket_{\mathcal{A}}(\sigma) \le \llbracket e_2 \rrbracket_{\mathcal{A}}(\sigma) \\ \end{split}$$

Figure 3: Definition of the denotational semantics of Boolean expressions.

Semantics of Statements

The small-step operational semantics relation $\to \subseteq \mathcal{C} \times \mathcal{C}$ is defined by the rules in Figure 4 where the set of configurations \mathcal{C} is $(\mathcal{S} \times \mathsf{State}) \cup \mathsf{State}$.

Figure 4: Definition of the operational semantics of statements.

Hoare Triples

A Hoare triple $\{P\}$ S $\{Q\}$ for P, $Q\subseteq$ State asserts that, for any state $\sigma\in$ State, if $\sigma\in P$ and $\langle S,\sigma\rangle\to^*\sigma'$, then $\sigma'\in Q$. The sets P and Q can be represented as Boolean expressions extended with quantifiers.

The rules for constructing Hoare triples are given in Figure 5.

Figure 5: Rules of Hoare logic.

Computable Functions

We write $[x \mapsto n]$ for the state that maps the variable x to the number $n \in \mathbb{N}$, and every other variable to 0.

A 'while' program S computes a partial function $f: \mathbb{N} \to \mathbb{N}$ (with respect to x) just if $f(m) \simeq n$ exactly when $\langle S, [\mathbf{x} \mapsto m] \rangle \Downarrow [\mathbf{x} \mapsto n]$.

A function $f: \mathbb{N} \to \mathbb{N}$ is *computable* just if there is a program S that computes f with respect to the variable x.

Predicates

The *characteristic function* of U is the function

$$\chi_U : \mathbb{N} \to \mathbb{N}$$

$$\chi_U(n) = \begin{cases} 1 & \text{if } n \in U \\ 0 & \text{if } n \notin U \end{cases}$$

The semi-characteristic function of U is the partial function

$$\xi_U: \mathbb{N} \to \mathbb{N}$$

$$\xi_U(n) \begin{cases} \simeq 1 & \text{if } n \in U \\ \uparrow & \text{otherwise} \end{cases}$$

A predicate $U \subseteq \mathbb{N}$ is *decidable* just if its characteristic function $\chi_U : \mathbb{N} \to \mathbb{N}$ is computable.

The 'while' program that computes the characteristic function χ_U of a predicate $U \subseteq \mathbb{N}$ is called a *decision procedure*. Any predicate for which there is no decision procedure is called *undecidable*.

A predicate $U \subseteq \mathbb{N}$ is *semi-decidable* just if its semi-characteristic function ξ_U is computable.

The *Halting Problem* is the following predicate:

$$\mathsf{HALT} = \{ \langle \lceil S \rceil, n \rangle \mid [\![S]\!]_{\mathbf{x}}(n) \downarrow \}$$

Bijections

A function $f: A \to B$ is *injective* (or 1-1) just if for any $a_1, a_2 \in \mathcal{A}$ we have that $f(a_1) = f(a_2)$ implies $a_1 = a_2$. We sometimes write $f: A \to B$ whenever f is an injection.

A function $f:A\to B$ is surjective just if for any $b\in\mathcal{B}$ there exists $a\in\mathcal{A}$ such that f(a)=b. We sometimes write $f:A\twoheadrightarrow B$ whenever f is a surjection.

A function $f: A \to B$ is a *bijection* just if it is both injective and surjective.

Let $f:A\to B$ be a function. f is an isomorphism just if it has an inverse. That is, if there exists a function $f^{-1}:B\to A$ such that:

- for all $a \in \mathcal{A}$ we have $f^{-1}(f(a)) = a$
- for all $b \in \mathcal{B}$ we have $f(f^{-1}(b)) = b$

Encoding Data

A pairing function is a bijection $\mathbb{N} \times \mathbb{N} \xrightarrow{\cong} \mathbb{N}$. We assume that we have a fixed pairing function

$$\langle -, - \rangle : \mathbb{N} \times \mathbb{N} \xrightarrow{\cong} \mathbb{N}$$

with the following inverse:

$$\mathsf{split}: \mathbb{N} \xrightarrow{\cong} \mathbb{N} \times \mathbb{N}$$

Reflections

Suppose we have two bijections:

$$\phi: A \xrightarrow{\cong} \mathbb{N} \quad \psi: B \xrightarrow{\cong} \mathbb{N}$$

The *reflection* of $f:A \rightarrow B$ under (ϕ, ψ) is the function

$$\tilde{f}: \mathbb{N} \to \mathbb{N}$$

$$\tilde{f}(n) = \psi(f(\phi^{-1}(n)))$$

Gödel Numbering

Let **Stmt** be the set of Abstract Syntax Trees of While. We assume that we have a Gödel numbering

$$\lceil - \rceil : \mathbf{Stmt} \xrightarrow{\cong} \mathbb{N}$$

which encodes While programs as natural numbers.

A *code transformation* is a function $f : \mathbf{Stmt} \to \mathbf{Stmt}$.

Universal Function

The universal function, U, is defined as follows:

$$U: \mathbf{Stmt} \times \mathbb{N} \to \mathbb{N}$$
$$U(P, n) = \llbracket P \rrbracket_{\mathbf{x}}(n)$$

Reductions

Let $U,W\subseteq\mathbb{N}$ be predicates, and let $f:\mathbb{N}\to\mathbb{N}$. The function f is a many-one reduction from U to W just if it is computable, and it is also the case that

$$n \in U \Leftrightarrow f(n) \in W$$

We may write $f: U \lesssim V$ (read "f is a reduction from U to V ").