MATHEMATICS MASTERS EXAMINATION

OPTION 4
COMPUTER SCIENCE

NOTE: Any student whose answers require clarification may be required to submit to an oral examination.
Each of the twelve numbered questions is worth 20 points. All questions will be graded, but your score for the examination will be the sum of your scores on your eight best questions.

Please observe the following:

- **DO NOT** answer two or more questions on the same sheet
  (not even on both sides of the same sheet).

- **DO NOT** write your name on any of your answer sheets.

You will be given separate instructions on the use of these answer sheets. When you have completed a question, place it in the large envelope provided.
Computer Algorithms

1. (a) Sketch an algorithm which, when given a directed graph $G$ and two vertices $u$, $v$ of $G$, will determine whether or not there is a path from $u$ to $v$. Your answer must run in time $O(n)$, where $n$ is the size of $G$.

(b) A 2-CNF expression is the conjunction of terms, each term being a disjunction of two literals (a literal is a boolean variable or a negated boolean variable). For example, $(x \lor \overline{y}) \land (y \lor z) \land (\overline{x} \lor \overline{y})$ is a 2-CNF. 2SAT is the problem to decide whether a 2-CNF expression has an assignment of the variables which makes it true. Design an algorithm which runs in polynomial time to solve 2SAT. You might consider applying part (a) to the directed graph whose arcs correspond to the conjunctions. Why is your algorithm correct? What is the asymptotic running time of your algorithm?

2. Consider two algorithms to find the $k$ largest entries of an array $A$ with $n$ integer entries:

   1. put the entries of $A$ into a heap and then remove the the largest, $2^{\text{nd}}$ largest, $\ldots$, $k^{\text{th}}$ largest;

   2. use a modification of insertion sort to insert the $k$ largest entries into an auxiliary array of size $k$ (if an entry becomes too small, it is discarded)

Compute and compare the asymptotic running times of these two algorithms

Combinatorics

3. Let $n$ be a positive integer.

   (a) Show that for $0 \leq k < n$ the inequality \[ \binom{2n}{k} < \binom{2n}{n} \] holds.

   (b) Show that for all such $n$ the inequality \[ \binom{2n}{n} \leq 2^{2n} \leq 2n \binom{2n}{n} \] holds.

4. Let $n$ be a positive integer. Find, with proof, a closed-form expression for \[ \sum_{k=0}^{n} \binom{n}{k}^2 \].

Graph Theory

5. Let $0 \leq e \leq n$, and let $G$ be a simple undirected graph with $n$ vertices and $e$ edges. Show that $G$ has at least $n - e$ components.
6. Either draw a simple undirected plane graph that meets the given description or prove that no such graph exists.
(a) A graph with 6 vertices and 13 edges.
(b) A graph with 5 vertices and 9 edges.
(c) A bipartite graph with 6 vertices and 8 edges.

Error-Correcting Codes and Cryptography

7. A third-order LFSR sequence starts 100111. Find the next 4 elements of the sequence.

8. Let $C$ be the linear binary $[10,3]$-code with generator matrix

$$G = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

(a) Write down the eight codewords of $C$.
(b) What is the minimum weight of $C$? How many errors (per 10-bit block transmitted) can $C$ correct?
(c) How would you decode the received vector 1 1 1 1 0 1 0 0 1? (You don't need to use a systematic method; make use of your answer in (a).)
(d) If the code $C$ is used to transmit a 24-bit message (length before encoding) over a channel with white noise (error probability 0.04 per bit), what is the probability that the received message is decoded correctly?

Theory of Computation

9. Let $\Sigma = \{0,1\}$. Construct an NFA for the language

$$L = \{w \mid w \text{ has a substring which is two 0's separated by an even number of 1's}\}.$$ 

For example 101101, 11001 $\in L$ and 10101110 $\notin L$. Justify your answer.

10. Let $\Sigma = \{a, b, c\}$. Is the language

$$L = \{a^m b^n c^p \mid m \leq n \leq p\}$$

context-free? Justify your answer.

Fall 05 CS 3/4
Numerical Analysis

11. Consider the matrix \( A = \begin{bmatrix} 2.436\times10^{-1} & 5.380\times10^{-1} & 2.600\times10^{-1} \\ 6.633\times10^{-1} & 1.391\times10^{-1} & 4.580\times10^{-1} \\ 5.862\times10^{-1} & 3.466\times10^{-1} & 6.548\times10^{-1} \end{bmatrix} \).

(a) Compute the LU decomposition of \( A \) with partial pivoting.

Calculate with four decimal places, using rounding: write the answer of every step rounded to four decimal places in scientific format, and use the rounded number in the calculations of the next step.

(b) What is the determinant of \( A \)?

12. The composite trapezoidal rule to approximate \( \int_a^b f(x)dx \) using \( n \) subintervals is

\[
T(n) = \frac{h}{2} (f(a) + f(b)) + h \sum_{k=1}^{n-1} f(a + kh), \quad h = \frac{b-a}{n}.
\]

(a) Derive the formula for \( T(n) \). For \( n = 1 \), \( T(1) \) is the trapezoidal rule.

(b) Derive the relationship between \( T(n) \) and \( T(2n) \), emphasizing how the \( n + 1 \) function values computed for \( T(n) \) can be reused in the computation of \( T(2n) \).
SOLUTIONS

MATHEMATICS MASTERS EXAMINATION

OPTION 4
COMPUTER SCIENCE
August 16, 2005
2–5 PM
NOTE: Any student whose answers require clarification may be required to submit to an oral examination.
Each of the twelve numbered questions is worth 20 points. All questions will be graded, but your score for the examination will be the sum of your scores on your eight best questions.

Please observe the following:

• DO NOT answer two or more questions on the same sheet (not even on both sides of the same sheet).

• DO NOT write your name on any of your answer sheets.

You will be given separate instructions on the use of these answer sheets. When you have completed a question, place it in the large envelope provided.
Computer Algorithms

1. Show how to find the maximum and the minimum integer in an array \( A \) containing \( n \) integers \( A[1], \ldots, A[n] \) by using at most \( 3 \left\lfloor \frac{n}{2} \right\rfloor \) comparisons. Prove your answer.

**Solution:** The strategy is to maintain the minimum and maximum elements seen thus far. Rather than processing each element of the input by comparing it to the current minimum and maximum (at a cost of 2 comparisons per element), we process elements in pairs. We compare pairs of elements from the input first with each other, and then we compare the smaller to the current minimum and the larger to the current maximum (at a cost of 3 comparisons for every 2 elements). Initialization depends on the parity of \( n \). If \( n \) is odd, we set both the minimum and the maximum to the first element, and then we process the elements in pairs. Thus we perform \( 3 \left\lfloor \frac{n}{2} \right\rfloor \) comparisons. If \( n \) is even, we perform 1 comparison on the first 2 elements to determine the initial values of the minimum and the maximum. Thus we perform \( 1 + \frac{3(n-2)}{2} = \frac{3n}{2} - 2 \) comparisons.

2. A Boolean expression is in 3-CNF form if it is written as

\[(t_{i1} \lor t_{i2} \lor t_{i3}) \land \cdots \land (t_{k1} \lor t_{k2} \lor t_{k3}),\]

where each literal \( t_{ij} \) is a variable or its negation. The 3-CNF-SAT problem is to decide whether there is an assignment that satisfies a given Boolean expression in 3-CNF form. The CLIQUE problem is to decide, given a graph \( G \) and a natural number \( k \), whether \( G \) has a complete subgraph with \( k \) vertices.

(a) Show that CLIQUE is in NP.

(b) Show that 3-CNF-SAT is polynomially reducible to CLIQUE.

Suggestion for (b): construct a graph whose vertices are the occurrences of literals in the given 3-CNF, where two vertices are adjacent when the corresponding literals belong to different clauses and are consistent (one is not the negation of the other). Show that this is a suitable reduction.

**Solution:** (a) Given a graph \( G = (V, E) \) and a proposed clique \( K \) of size \( k \) in \( G \), we can verify this in polynomial time. The verification that \( |K| = k \) can be done in time \( O(|K|) = O(|V|) \) by counting the members of \( K \). The verification that \( K \) is a clique can be done in time \( O(|K|^2) = O(|V|^2) \) by checking that every two vertices of \( K \) are neighbors in \( G \).

(b) Let \( \phi = C_1 \land \cdots \land C_k \) be an instance of 3-CNF, that is, \( C_i = t_{i1} \lor t_{i2} \lor t_{i3} \), and let \( G \) be the graph constructed as in the suggestion. Assume that \( \phi \) has a satisfying assignment. Then for each \( i \) there is a vertex \( v_{ij} \) that corresponds to a true literal in \( \phi \). These vertices

Fall 05 CS 2/4
form a set $K$ of size $k$. Any two vertices of $K$ correspond to consistent literals, since both are true in $\phi$. Hence $K$ is a clique of size $k$ in $G$. Conversely, assume that $G$ has a clique $K$ of size $k$. Since literals of the same clause are not adjacent and there are $k$ clauses in all, each clause must have exactly one literal corresponding to a vertex in $K$. Assign the corresponding literals the value true; this is consistent since the corresponding vertices form a clique. This assignment of values (to some of the variables) satisfies $\phi$.

**Combinatorics**

3. 

(a) Consider the recurrence relation given by $h_n = 3h_{n-1} + 6h_{n-2}$ for all $n \geq 2$. Find the general solution of this recurrence relation.

(b) Let $s_n$ denote the number of strings $w$ of length $n$ formed using the digits $1,2,3,4,5$ in which no two even numbers that appear in $w$ are adjacent. For example, the string 3325214 is a string of length 7 of the desired form but 3324514 is not. Define $s_0 = 1$ and note that $s_1 = 5$. Prove that the recurrence relation for $s_n$ is the same as that for $h_n$.

(c) Use the initial values for $s_n$ to completely solve the recurrence relation that you obtained for $s_n$.

**Solution:** (a) We have $h_n = 3h_{n-1} + 6h_{n-2}$. If $h_n = q^n$ we get the polynomial equation $q^2 - 3q - 6 = 0$. The solutions to this are $\frac{3 \pm \sqrt{33}}{2}$. Hence the general solution is

$$h_n = c_1 \left( \frac{3 + \sqrt{33}}{2} \right)^n + c_2 \left( \frac{3 - \sqrt{33}}{2} \right)^n$$

for constants $c_1$ and $c_2$.

(b) Let $w$ be an arbitrary string of length $n \geq 2$ of the required form. Write $w = dx$ for a single digit $d$ and $x$ a string of length $n-1$. If $d$ is odd, then there are $s_{n-1}$ choices for $x$ and thus $3s_{n-1}$ possible $w$ with the left-most digit odd. If $d$ is even, then $w = dx$ where $d$ is an even digit, $c$ must be an odd digit, and $x$ is any string of the desired form of length $n-2$. Thus there are $2 \times 3 \times s_{n-2}$ possible $w$ with $d$ even. It follows that $s_n = 3s_{n-1} + 6s_{n-2}$ for all $n \geq 2$.

(c) Since the recurrence relation for $s_n$ is the same as for $h_n$ we need only solve the system of equations

$$c_1 \frac{3 + \sqrt{33}}{2} + c_2 \frac{3 - \sqrt{33}}{2} = 1$$

$$c_1 \frac{3 + \sqrt{33}}{2} + c_2 \frac{3 - \sqrt{33}}{2} = 5$$

Fall 05 CS 3/4
for $c_1$ and $c_2$. This gives

$$c_1 = \frac{7}{2\sqrt{33}} + \frac{1}{2} \quad \text{and} \quad c_2 = -\frac{7}{2\sqrt{33}} + \frac{1}{2}.$$  

These coefficients may be used in the general solution for part (a) to give a formula for $s_n$.

4. Let $n, p$ be integers, with $0 \leq p \leq n$. Define the Stirling number of the first kind $s(n, p)$. In other words, describe what is counted by $s(n, p)$. Find and prove a formula for $s(n, n - 1)$.

**Solution:** $s(n, p)$ is the number of ways to arrange $n$ (distinct) objects into $p$ nonempty circular permutations. When $p = n - 1$, we have $n - 1$ nonempty circular permutations, hence $n - 2$ of them must have one object, and one of them must have two objects. Given one or two objects, there is only one circular permutation. Hence it suffices to choose which two objects make up the circular permutation of size two, and there are $\binom{n}{2}$ such choices. Therefore $s(n, n - 1) = \binom{n}{2}$.

**Graph Theory**

5. Identify, with proof, all pairs $(p, q)$ of positive integers for which the complete bipartite graph $K_{p,q}$ has a Hamiltonian cycle.

**Solution:** $K_{p,q}$ is Hamiltonian if and only if $p = q > 1$. First suppose that $K_{p,q}$ is Hamiltonian. Then a Hamiltonian cycle alternates between the two parts, and starts and ends in the same part. Hence $p = q > 1$. On the other hand, if $p = q > 1$, then it is easy to construct a Hamiltonian cycle, for example, if the parts are $\{x_1, \ldots, x_p\}$ and $\{y_1, \ldots, y_p\}$, then one possibility is $x_1 y_1 x_2 y_2 \ldots x_p y_p x_1$.

6. Let $G$ be a forest with 73 vertices and 43 edges. Determine the number of connected components of $G$.

**Solution:** Suppose $G$ has $n$ vertices and $k$ connected components, with $n_i$ vertices in the $i$th component. Then, since each component is a tree, the $i$th component has $n_i - 1$ edges. Summing over $i$, we obtain that the number of edges $e$ satisfies $e = n - k$. This gives $k = n - e = 30$.

**Error-Correcting Codes and Cryptography**

7. The following cipher-text was encrypted by an affine cipher modulo 26:

CRWWZ

The plaintext starts ha. Decrypt the message.

**Solution:** happy. The affine cipher is

$$qx + 17$$

Fall 05 CS 4/4
p is 15: \( 9 \times 15 + 17 = 22 \pmod{26} \), which is W.
y is 24: \( 9 \times 24 + 17 = 25 \pmod{26} \), which is Z.
h is 7: \( 9 \times 7 + 17 = 2 \pmod{26} \), which is C.
a is 0: \( 0 \times 7 + 17 = 17 \pmod{26} \), which is R.

8. To be supplied

Theory of Computation

9. Let \( \Sigma = \{0, 1\} \). Construct a DFA for the following language:

\[
L = \{ w \in \Sigma^* \mid \text{successive 0's in } w \text{ are separated by exactly one or two 1's} \}.
\]

For example, 110101101, 1110111, 111 \in L, but 100,01110 \notin L. Justify your construction.

10. Is the language

\[
L = \{ a^n b^{2n} c^{3n} \mid n \geq 0 \}
\]

context-free? Justify your answer.

Solution: Not context-free. Using the pumping lemma, assume \( a^p b^{2p} c^{3p} \) can be written as \( uvxyz \) satisfying the usual three conditions. Then \( |uvy| \leq p \), so \( vxy \) contains at most two kinds of symbols, hence it does not contain at least one of the symbols. But, as \( |vy| > 0 \), in the string \( uv^2xy^2z \), the number of letters of at least one kind changes, and the number of letters of at least one kind does not change. Hence, the pumped string cannot be in \( L \), a contradiction.
Problem 8

Coding Theory

1. Let $C$ be the linear binary [7,3] code with parity check matrix

$$H = \begin{pmatrix}
1 & 0 & 1 & 1 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 1 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 1
\end{pmatrix}$$

a) What is the dimension of the $C$?

b) Find a systematic generator matrix $G$ for the $C$. (The leftmost $k$ columns of $G$ should form an identity matrix, where $k$ is your answer to (a) above.)

c) What is the minimum weight of $C$? How many errors (in a seven-bit block) can $C$ always correct? In addition to correcting these errors, how many errors (in a seven-bit block) can $C$ always detect?

d) Write down part of the table of syndromes and coset leaders for $C$. Include only the cosets containing a vector of weight 0 or 1.

e) Say the sender transmits a codeword $c$ to us, and we receive the vector $v = (0 0 1 1 1 0 0)$. Compute the syndrome of $v$, and use the syndrome and your table of part (d) to determine which codeword the sender transmitted, assuming nearest neighbor decoding.
Solution to Coding Theory Question, Fall 2005 Masters Exam

a) \( H \) is an \( (n-k) \times n \) matrix, where \( n \) is the length and \( k \) is the dimension. So \( n-k = 4 \) and \( n = 7 \), implying that \( k = 3 \).

b) Note \( H = (P | I_4) \), so the generator matrix \( G = (I_3 | P^T) \), or

\[
\begin{pmatrix}
1 & 0 & 0 & 1 & 1 & 1 & 0 \\
0 & 1 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 1 & 1 & 0 & 1 & 1 \\
\end{pmatrix}
\]

c) The minimum weight of \( C \) is 4. To see this, note that no sum of three or fewer columns of \( C \) is the zero vector.

\( C \) can correct \( \lfloor (4-1)/2 \rfloor = 1 \) error per 7-bit encoded block. Since 4 is even, \( C \) can in addition detect two errors in a 7-bit block.

d) \[
\begin{array}{ll}
\text{syndrome} & \text{coset leader} \\
0000 & 0000000 \\
0001 & 0000001 \\
0010 & 0000010 \\
0100 & 0001000 \\
0111 & 0100000 \\
1000 & 0010000 \\
1011 & 0010000 \\
1110 & 1000000 \\
\end{array}
\]

e) The syndrome is \( vH^T = (0111) \). The coset leader for syndrome \( (0111) \) is 0100000. Under nearest neighbor decoding, we assume the error vector \( e \) is 0100000, so the codeword \( c \) that was transmitted was \( v+e = 0111000 \).
Solution to Problem 9 (Fall '05, CS)
Two new questions for MCS 471 Fall 05 Master’s Exam


(a) Compute the LU decomposition of \( A \) with partial pivoting.

Calculate with four decimal places, using rounding: write the answer of every step rounded to four decimal places in scientific format, and use the rounded number in the calculations of the next step.

(b) How many arithmetical operations have you used for this LU decomposition?

What is the order of the number of arithmetical operations for a LU decomposition on a general \( n \)-by-\( n \) matrix?

2. To solve the initial value problem \( \frac{dy}{dx} = f(x, y) \), \( y(x_0) = y_0 \), we could choose a step size \( h > 0 \), \( x_{n+1} = x_n + h \), and approximate \( y(x_n) \) by \( y_n \), using the third-order Adams-Bashforth method:

\[
y_{n+1} = y_n + \frac{h}{12} (23f_n - 16f_{n-1} + 5f_{n-2}) .
\]

(a) Derive this formula.

(b) What is the local error of this formula? Justify.

At step \( n \), estimate the global error \( |y_n - y(x_n)| \).
Answers to the two new questions for MCS 471 Fall 05 Master’s Exam

1. 

\[
A = \begin{bmatrix}
8.105E-2 & 2.114E-1 & 1.956E-2 \\
4.215E-1 & 7.097E-1 & 7.472E-1 \\
4.337E-2 & 2.140E-1 & 8.820E-1 \\
\end{bmatrix}
\]

\[
\rightarrow \begin{bmatrix}
4.215E-1 & 7.097E-1 & 7.472E-1 \\
8.105E-2 & 2.114E-1 & 1.956E-2 \\
4.337E-2 & 2.140E-1 & 8.820E-1 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
0.8105 \\
R2 := R2 - \begin{bmatrix}
4.215E-1 & 7.097E-1 & 7.472E-1 \\
R1 & 4.215 \\
\end{bmatrix} \\
\begin{bmatrix}
1.923E-1 & 7.492E-2 & -1.241E-1 \\
.4337 & 1 \\
\end{bmatrix} \\
R3 := R3 - \begin{bmatrix}
1.029E-1 & 1.410E-1 & 8.051E-1 \\
R1 & 4.215 \\
\end{bmatrix} \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
4.215E-1 & 7.097E-1 & 7.472E-1 \\
1.029E-1 & 1.410E-1 & 8.051E-1 \\
1.923E-1 & 7.492E-2 & -1.241E-1 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
.7492 \\
R3 := R3 - \begin{bmatrix}
4.215E-1 & 7.097E-1 & 7.472E-1 \\
R2 & 1.410 \\
\end{bmatrix} \\
\begin{bmatrix}
1.029E-1 & 1.410E-1 & 8.051E-1 \\
1.923E-1 & 5.313E-1 & -5.518E-1 \\
\end{bmatrix} \\
\end{bmatrix}
\]

The LU decomposition consists of \( L \), \( U \), and \( P \):

\[
L = \begin{bmatrix}
1.000E+0 & 0.000E+0 & 0.000E+0 \\
1.029E-1 & 1.000E+0 & 0.000E+0 \\
1.923E-1 & 5.313E-1 & 1.000E+0 \\
\end{bmatrix}
\]

\[
U = \begin{bmatrix}
4.215E-1 & 7.097E-1 & 7.472E-1 \\
0.000E+0 & 1.410E-1 & 8.051E-1 \\
0.000E+0 & 0.000E+0 & -5.518E-1 \\
\end{bmatrix}
\]

\[
P = \begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
1 & 0 & 0 \\
\end{bmatrix}
\]

2
(b) The number of arithmetical operations on this example is 13.

We need 3 divisions for the multipliers. The first reduction takes 2 multiplications and 2 subtractions for two rows (4 operations per row), so 8 operations in total. The second reduction takes 1 multiplication and 1 subtraction. In general, the order of arithmetical operations is $O(n^3)$, for the LU factorization on an $n$-by-$n$ matrix.

2. (a) We apply the method of undetermined coefficients for a quadrature rule with three function evaluations for

$$\int_{x_n}^{x_{n+1}} f(z) \approx c_0 f_n + c_1 f_{n-1} + c_2 f_{n-2}.$$ 

Without loss of generality, we may take $x_n = 0$, so that $x_{n+1} = h$. Asking that the quadrature rule integrates 1, $x$, and $x^2$ correctly leads to the three linear equations in the unknowns coefficients $c_0$, $c_1$, and $c_2$:

$$\int_{0}^{h} 1 \, dx = h = c_0 + c_1 + c_2$$
$$\int_{0}^{h} x \, dx = \frac{h^2}{2} = c_0 0 + c_1 (-h) + c_2 (-2h)$$
$$\int_{0}^{h} x^2 \, dx = \frac{h^3}{3} = c_0 0^2 + c_1 (-h)^2 + c_2 (-2h)^2$$

Solving this linear system gives $c_0 = \frac{23}{12} h$, $c_1 = -\frac{16}{12} h$, and $c_2 = \frac{5}{12} h$.

(b) The local error is $O(h^4)$. This can be computed by integrating the error of the polynomial, interpolating through the three points $(x_n, f_n)$, $(x_{n-1}, f_{n-1})$, and $(x_{n-2}, f_{n-2})$. This integral is

$$\int_{x_n}^{x_{n}+h} (z-x_n)(z-x_n+h)(z-x_n+2h) \, dz.$$ 

The global error at step $n$ is estimated to be of order $O(h^3)$, as we take $n$ steps, the local errors of order $O(h^4)$ add up, and $n = (x_n - x_0)/h$. 

3