

**UNIVERSITY OF CALIFORNIA  
COLLEGE OF ENGINEERING**

**E7: INTRODUCTION TO COMPUTER PROGRAMMING  
FOR SCIENTISTS AND ENGINEERS**

Professor Raja Sengupta

Spring 2015

Final Exam—May 12, 2015

8:00a – 11:00a

**[50 points ~ 3 hours]**

Question	Points
Part I	10
Part II	10
Part III	10
Part IV	20
<b>TOTAL</b>	<b>50</b>

Notes:

1. Your exam should have 20 pages. Check this before you begin.
2. Before you leave the exam hall, make sure your name has been cross-verified and marked on the class list by atleast one of the GSIs while you are submitting your exam. If not, your exam will not be graded.
3. You may use your notes, calculator or laptop on this examination provided that you do not impede those sitting next to you. No other electronic devices are permitted. Charging of devices is not allowed in the exam venue.
4. Use a #2 pencil and a green scantron sheet to record your answers. Mark your solution to each question on the corresponding space on your scantron. There is one correct answer for each question. Multiple markings, incomplete markings, or stray marks will cause your solution to be marked incorrect.
5. Please write your **name, subject, date, test number, student ID number, and discussion and lab section** on your scantron for identification purposes.

**Part I:****Questions 1-10: Basic Programming**

Given the following matrices:

$$A = \begin{bmatrix} 82 & 84 & 77 \\ 53 & 58 & 69 \\ 23 & 45 & 56 \end{bmatrix}, \quad B = \begin{bmatrix} 324 & 435 \\ 23 & 34 \\ 345 & 678 \end{bmatrix}$$

1. What would the variable x be after executing the following command:

```
>> x = A(size(B))
```

- A. [ 82 84 ]  
 B. [ 82; 84 ]  
 C. [ 53 23 ]  
 D. [ 23 53 ]
2. Zhi is given the following functions:

```
function y=f(x)
y = x.^2;
end
```

```
function y=g(x)
y = -f(-x);
end
```

```
function y=h(x)
y = x*x';
end
```

What is the value of y computed by Zhi using the following expressions:

```
>> x = [ 2 9 5 6 8 9 5 7 2 8 9 2 ];
```

```
>> y = h(g(f(x)));
```

- A. [13 87 54 78 897 79 45 212 48 15 657 153]  
 B. [13 87 54 78 897 79 45 212 48 15 657 153]'  
 C. 170921030  
 D. diag ([13 87 54 78 897 79 45 212 48 15 657 153]')

3. Consider the following function:

```
function [c, b, a] = myFunc(a,b,c)
```

```
...some code....
```

```
end
```

Suppose we evaluate >> [a, b, c] = myFunc(1,2,3) ? Then >> whos a,b,c must show us the values

- A. a = [], b = [], c = []  
 B. Values of a,b,c are unknown without the rest of the code  
 C. a = 3, b = 2, c = 1  
 D. a = 1, b = 2, c = 3

4. The Ackermann function is defined as:

$$A(m, n) = \begin{cases} n + 1, & \text{if } m = 0 \\ A(m - 1, 1), & \text{if } m > 0 \text{ and } n = 0 \\ A(m - 1, A(m, n - 1)), & \text{if } m > 0 \text{ and } n > 0 \end{cases}$$

The value of  $A(1,3)$  is:

- A. 5
  - B. 4
  - C. 3
  - D. 6
5. Ninh is given an incomplete function in Matlab that calculates  $A(m,n)$  for any given non-negative integers  $m$  and  $n$  as below:

```
function result = A(m,n)
    if m==0
        ... (1) ... = n+1;
    elseif (m>0) && (n==0)
        ... (1) ... = ... (2) ... ;
    else
        ... (1) ... = A(... (3) ..., ... (4) ...) ;
    end
end
```

In order to complete the code, Ninh must fill the blanks (1), (2) and (3), respectively, with:

- A. result,  $A(m-1,1)$ ,  $m-1$ ,  $A(m,n-1)$
- B. A,  $A(m-1,1)$ ,  $A(m-1)$ ,  $A(m,n-1)$
- C. A,  $result(m-1,1)$ ,  $A(m-1)$ ,  $result(m,n-1)$
- D. result,  $result(m-1,1)$ ,  $(m-1)$ ,  $result(m,n-1)$

In MATLAB, the function `sort(X)` allows the user to sort the elements of `X` in either ascending or descending order. When `X` is a cell array of strings, `sort(X)` sorts the strings in ASCII dictionary order. For example:

```
>> X = {'Watch', 'Out', 'Heisenberg', 'Winter', 'Is', 'Coming'};
>> sort(X)
ans =
    'Coming' 'Heisenberg' 'Is' 'Out' 'Watch' 'Winter'
```

In the case of `sort`, we can use iteration (for and while loops) to achieve the same results. The `iterationsort` function code below takes a cell array of strings `words` as an input, sorts them in ASCII dictionary order, and outputs the sorted `order` cell array. Study the code for the function, and answer the questions that follow:

```
function [order] = iterationsort(words)

m = length(words);
for i = 1:m
    j = 1;
    while j < m-i+1
        min_len_word = min(length(words{i}), length(words{i+j}));
        k = 1;
        while k <= min_len_word
            b = words{i}(k);
            c = words{i+j}(k);
            if b > c
                a = words{i};
                words{i} = words{i+j};
                words{i+j} = a;
                BLANK
            elseif words{i}(k) < words{i+j}(k)
                k = min_len_word + 1;
            else
                k = k+1;
            end
        end
        j = j+1;
    end
end
order = words;
end
```

6. What role does the inner while loop play in the overall function?
  - A. It iterates across the letters of a single word
  - B. It iterates across the different word lengths
  - C. It iterates across the elements of the input cell array
  - D. None of the above
  
7. Which line of code is the right fill for the blank?
  - A. `k = min_len_word;`
  - B. `k = min_len_word + i;`
  - C. `k = k;`
  - D. None of the above

Andrea randomly came across the following code for a random walk animation (i.e. the animation of the random path taken by the walker, say Andrea herself)

```
r=[0 0];
for t = 1:100
    if t==99
        t=1;
    end
    B1=rand(1,2);
    B=B1>0.5;
    if B==1
        rnew=r+[1 0];
    elseif B(1)==1
        rnew=r+[0 1];
    elseif B(2)==1
        rnew=r+[-1 0];
    else
        rnew=r+[0 -1];
    end

    hold on;
    plot([r(1) rnew(1)], [r(2) rnew(2)]);
    pause(0.1)

    %UPDATE the new position
    r=rnew;
    t = t+1;
end
```

8. Which of the following statements is true?
- A. This code plots 100 random paths.
  - B. This code plots a maximum of 100 random paths and a minimum of 1 random path.
  - C. This code plots the walker's path indefinitely.
  - D. None of the above.

9. Which of the Matlab programs below computes a function satisfying the following recursive definition:

$$f(x) = \begin{cases} 0 & \text{if } x \leq 0.5 \\ f(x/2) * f(x - 0.5) & \text{otherwise} \end{cases}$$

A. `function y = recur(x)`  
`y = 0;`  
`end`

B. `function [fx] = recur(x)`  
`if lt(x,0.5)`  
`fx = fx;`  
`else`  
`fx = recur(x/2)*recur(x-0.5);`  
`end`  
`end`

C. `function [y] = recur(y)`  
`if le(x,0.5)`  
`fx = recur(0);`  
`else`  
`fx = recur(x/2)*recur(x-0.5);`  
`end`  
`end`

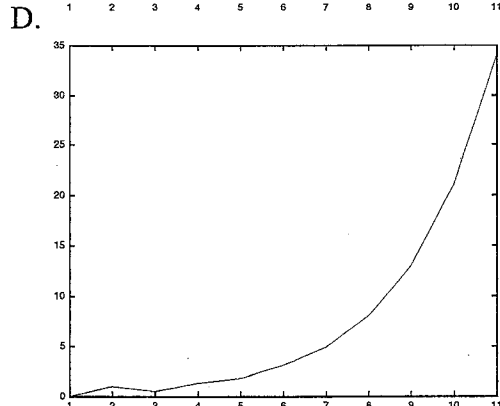
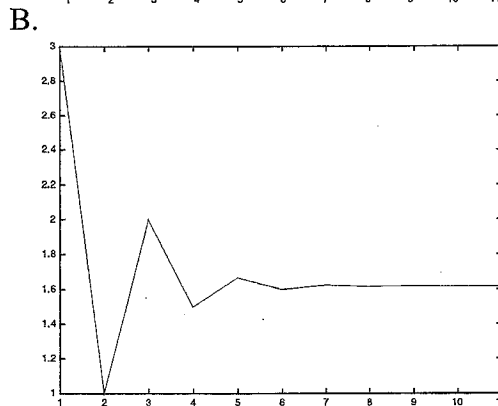
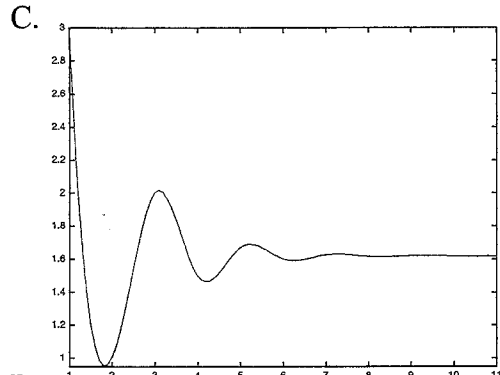
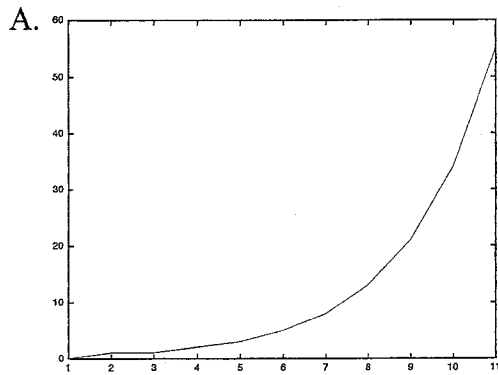
D. None of the above

10. Jordan has a funky idea. He defines the following two functions in Matlab:

```
function j = funky(n)
j=0;
if n==1
    j = 0;
elseif n==2;
    j = 1;
elseif n>2
    j= funky(n-1)+funky(n-2);
end
```

```
function funkyplot(n)
i=2;
gr(1) = 3;
while i<n
    gr(i)= funky(i+1)/funky(i);
    i=i+1;
end
plot(1:n-1,gr(1:n-1))
```

What is the output of the following command: `>>funkyplot (12)`



## Part II:

### Questions 11-15: Computer Representation of Numbers

11. Which of the following numbers is the decimal representation of the IEEE 754 single precision floating point number `0 10000011 010000000000000000000000` ?
- A. 20
  - B. 10
  - C. 4
  - D. 2

Now suppose that Mogeng developed a new structure storing the decimal numbers. As a reminder, in IEEE754:

- 1) the first bit represents the sign of the number.
- 2) the next 8 bits represents the base 2 exponential.
- 3) the last 23 bits represents base 2 fractions.

Due to some reason, Mogeng doesn't have 32 bits to store this single floating number. Instead, he decides to use 18 bits with the following breakdown:

- 1) the first bit represents the sign of the number.
- 2) the next 6 bits represents the base 2 exponent. (The bias in this case is 31. We follow the convention in IEEE754, i.e. we preserve the largest exponent for special numbers NAN and zero)
- 3) the last 11 bits represents base 2 fractions.

12. What will be the greatest non-infinite number that can be represented by this structure?

- A.  $2^{32} * (2 - 2^{-12})$
- B.  $2^{31} * (2 - 2^{-11})$
- C.  $2^{63} * (2 - 2^{-12})$
- D.  $2^{62} * (2 - 2^{-11})$

13. How many numbers can be represented by this structure excluding special numbers?

- A.  $2 \times 61 \times (2^{11} + 1)$
- B.  $2 \times 62 \times 2^{11} + 1$
- C.  $2 \times 62 \times 2^{11}$
- D.  $2 \times 61 \times (2^{11} + 1)$

14. Let Nmax and Nmin be the largest and smallest positive numbers that can be represented by this structure. What should be the representation of Nmax - Nmin?

- A. 0 111111 1111111111
- B. 0 111110 1111111111
- C. 0 111110 1111111110
- D. 0 111111 1111111110

15. Which of the following number can be exactly represented by this structure?

- A.  $2^{13} + 4$
- B.  $2^{13} + 3$
- C.  $2^{13} + 2$
- D. All of the above

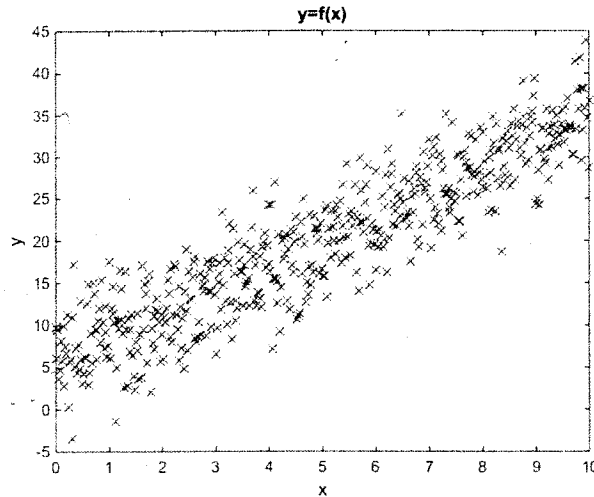


### Questions 16-20: Least Squares Regression and Interpolation

Consider the following code:

```
x=linspace(0,10,500)
y=3*x+5+4*randn(1,length(x))
plot(x,y,'x')
p=polyfit(x,y,1)
v=polyval(p,x)-y;
```

The plot of y as a function of y vs x looks like the following:



16. What will be the approximate value of the variable  $p$ , after running the above code?
- A. [4 5 3]
  - B. [3 5]
  - C. [3 5 4]
  - D. [4 3]
17. What will be the mean and variance of the array  $v$  ?
- A. Mean=0 Variance=16
  - B. Mean=0 Variance=9
  - C. Mean=4 Variance=0
  - D. Mean=1 Variance=0

18. Christiaan is given mean rainfall data for each of the 12 months for California (from a year when we had good rains):

88mm, 96mm, 36mm, 54mm, 22mm, 18mm, 39mm, 45mm, 68mm, 79mm, 48mm, 29mm.

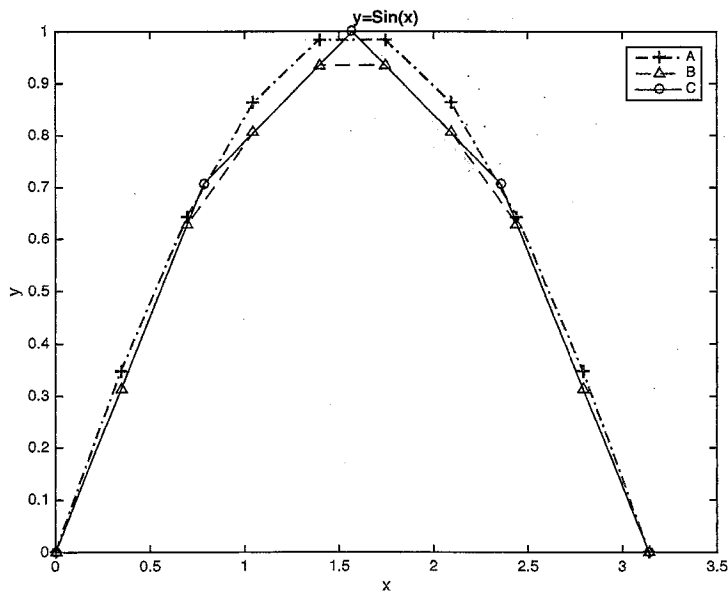
He wants to understand the trend of rainfall variation throughout the year but the data may contain noise. So, he fits a third degree polynomial to this data and then uses it to estimate the mean rainfall for each of the 12 months. Now, he is asked to report the resulting mean monthly rainfall for that year. What should he do to get the best estimate of it?

- A. Compute it based on either the estimated data or the provided data.
- B. Compute it based on the estimated data but not the provided data.
- C. Compute it based on the provided data but not the estimated data.
- D. None of the above.

19. Consider the following code:

```
x1=linspace(0,pi,5);
x2=linspace(0,pi,10);
y1=sin(x1);
y2=interp1(x1,y1,x2);
y3=spline(x1,y1,x2);
```

Which line represents the vector y3 on the graph below?



- A. A
- B. B
- C. C
- D. None of the Above

20. We want to write our own function to fit a polynomial to given data. Consider the function given below:

```
function coefficients = myPolynomialFit(xVec, yVec, n)
% xVec: a vector of the independent variable values
% yVec: a vector of the dependent variable values
% n: the desired order of the polynomial
% coefficients: a row vector of the coefficients of the calculated
% polynomial fit.

%first, we build the A and B matrices, row by row
A=[];
B=[];
for i=1:length(xVec)
    newRow=[];
    for j=n:-1:0
        _____missing code_____; % create all elements of each row
    end
    A=[A; newRow];
    B=[B; yVec(i)];
end
coefficients = A\B;
```

If the function should return the coefficients of the polynomial of degree  $n$ , what code should be added at the indicated line?

- A. `newRow = [newRow, xVec(i)^j];`
- B. `newRow = [xVec(i)^j];`
- C. `newRow = [newRow, sum(xVec.^(i+j-2))];`
- D. `newRow = [sum(xVec.^(i+j-2))];`

**Part III:****Questions 21-30: Numerical Methods – Differentiation, Integration & ODEs**

21. The Boole's rule for numerical integration including the error term is expressed by:

$$\int_{x_1}^{x_5} f(x) dx = \frac{2h}{45} (7f(x_1) + 32f(x_2) + 12f(x_3) + 32f(x_4) + 7f(x_5)) - \frac{8}{945} h^7 f^{(6)}(c)$$

where  $h$  is defined as  $h = (x_5 - x_1)/4$ , and  $c$  is some number in the interval  $[x_1, x_5]$ .

Cyril wants to implement this in Matlab. Given a function handle `fun` and an integration interval `x`, which of the following codes must he use to implement the Boole's rule? ( $f^{(6)}()$  means sixth order differential of  $f()$ )

A.

```

myBoole1.m
function int = myBoole1(fun, x)
int=0;
for i=1:length(x)-1
    h=(x(i+1)-x(i))/4;
    int=int+2*h/45*(7*fun(x(i))+32*fun(x(i)+h)+12*fun(x(i)+2*h)...
        +32*fun(x(i)+3*h)+7*fun(x(i)+4*h));
end
end

```

B.

```

myBoole2.m
function int = myBoole2(fun, x)
int=0;
for i=1:length(x)-1
    h=(x(i+1)-x(i))/4;
    int=int+2*h/45*(7*fun(x(i))+32*fun(x(i)+h)+12*fun(x(i)+2*h)...
        +32*fun(x(i)+3*h)+7*fun(x(i)+4*h))-8/945*h^7*fun(x(i))^6;
end
end

```

C.

```

myBoole3.m
function int = myBoole3(fun, x)
int=0;
for i=1:length(x)
    h=(x(i+1)-x(i))/5;
    int=int+2*h/45*(7*fun(x(i))+32*fun(x(i)+h)+12*fun(x(i)+2*h)...
        +32*fun(x(i)+3*h)+7*fun(x(i)+4*h));
end
end

```

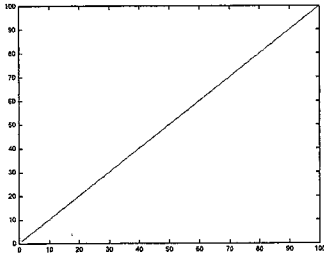
D. None of the above

22. Which of the following is true when you compare Boole's rule with Simpson's?
- A. The Boole's rule is a better estimate than the Simpson's rule.
  - B. The Simpson's rule is a better estimate than the Boole's rule.
  - C. There is not enough information on the Boole's rule to answer this question.
  - D. It depends on the input.
23. Which of the following statement is true about the Backward Difference method?
- A. It has the same order of accuracy as the Central difference method.
  - B. It has the same order of accuracy as the Forward difference method.
  - C. Both A and B are true.
  - D. None of the above
24. You are given the following differential equation and initial condition:  
 $dx/dt = t, x(0) = 0$   
Which of the following statements is true?
- A. It will be solved with zero error by the Euler method but not by using the ode45 solver.
  - B. It will be solved with zero error by the ode45 solver but not by using the Euler method.
  - C. It will be solved with zero error by using either the Euler method or the ode45 solver
  - D. None of the above.
25. You solve the following ode using the ode45 solver:  $dx/dt = t^6$ .  
Now, assume  $e$  is the error with respect to the analytical solution. What will be the output of the following command?  
`>> max(abs(e))`
- A. None of the below
  - B. Always 0
  - C. Sometimes a negative number
  - D. Always a positive number

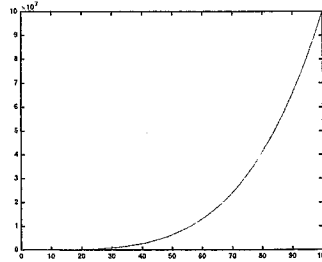
For 26 and 27, assume you solved the same ode ( $dx/dt = t^6$ ) using the RK4 method.

26. Which of the following plots represents the error growth with respect to the interval length? (Answer based on the shape. x axis is interval length and y axis is the error. You may ignore the y axes values)

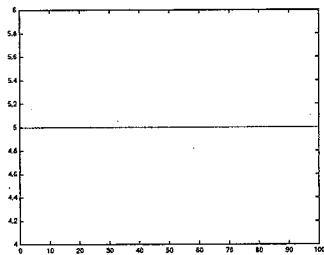
A.



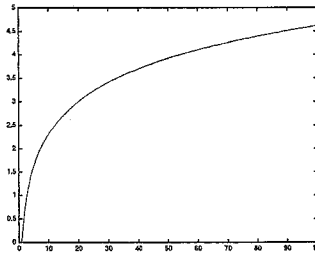
B.



C.

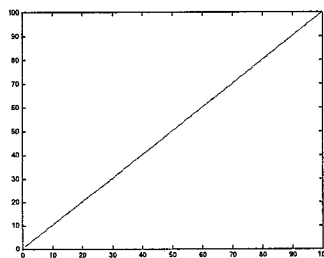


D.

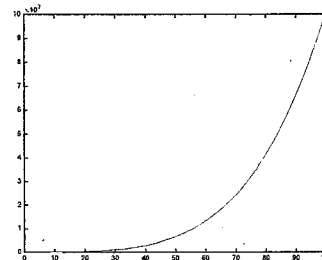


27. Which of the following plots represents the error with respect to the step size? (Answer based on the shape. x axis is interval length and y axis is the error. You may ignore the y axes values)

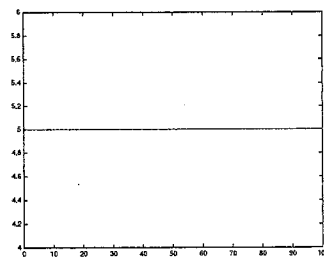
A.



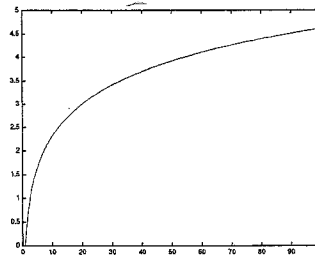
B.



C.



D.



28. Let a second order differential equation be described in the general form:

$$a\ddot{x} + b\dot{x} + cx = 0$$

Suppose we would like to solve the above ODE with the matlab command:

```
[...] = ode45(@(...) mySystemOfEquation(...),[0 12],[1;2])
```

The “...” is a placeholder for arguments that you may pass to the function or retrieve from it. Which of the following defines the mySystemOfEquation function?

- A. 

```
function [dx] = mySystemOfEquation(t,x,a,b,c)
    dx(1,1) = x(2)/t(2)
    dx(2,1) = (1/a) (-b*x(2) - c*x(1))
end
```
- B. 

```
function [dx] = mySystemOfEquation(t,x,a,b,c)
    dx(1,1) = diff(x(1))/diff(t)
    dx(2,1) = diff(x(2))/diff(t)
end
```
- C. 

```
function [dx] = mySystemOfEquation(t,x,a,b,c)
    dx(1,1) = x(2)
    dx(2,1) = (1/a) (-b*x(2) - c*x(1))
end
```
- D. None of the above

29. Using a standard ODE solver such as ode23 or ode45, which statement is true?

- A. The equations must all be first order.
- B. The equations must be independant, one from the other.
- C. There must be at least as many equations as variables.
- D. All of the above.

30. Consider the ODE  $\ddot{x}(t) + \ddot{x}(t) - \dot{x}(t) + x(t) = 0$  with initial conditions  $x(0) = 0$ ,  $\dot{x}(0)=2$  and  $\ddot{x}(0)=4$ .

Which of the following plots the solution  $x(t)$ ?

- A. 

```
[t,x]=ode45(@(t,x) [x(2);x(3); -x(3)+x(2)-x(1)],[0 10],[0 2 4]); plot(t,y(:,3))
```
- B. 

```
[t,x]=ode23(@(t,x) [x(2);x(3); -x(3)+x(2)-x(1)],[0 10],[4 2 0]); plot(t,y(:,1))
```
- C. 

```
[t,x]=ode45(@(t,x) [x(2);x(1); -x(1)+x(2)-x(3)],[0 10],[4 2 0]); plot(t,y(:,3))
```
- D. 

```
[t,x]=ode23(@(t,x) [x(2);x(3); -x(3)+x(2)-x(1)],[0 10],[0 2 4]); plot(t,y(:,1))
```

## Part IV:

### Questions 31-40: Data Structures, Pointers and Complexity

31. Which of the following consumes the least amount of memory without losing accuracy?

- A. `x = int8(5106642554)`
- B. `x = int32(5106642554)`
- C. `x = '5106642554'`
- D. `x = 5106642554`

32. Raja and Vishwa together wrote a single program to produce the E7 exam with  $n$  questions. Raja's code generated the questions and runs in time  $3n^2$  in the worst case. Vishwa's had to take the generated questions, arrange them and produce the final questionnaire. He managed to come up with a robust code that did his portion of the job in time  $2\log n$  in the worst case. The two programs were sequenced to produce the exam. The worst-case complexity of the two programs in sequence is:

- A.  $O(\log n)$
- B.  $O(n^2 \log n)$
- C.  $O(6n^2)$  but not  $O(n^2)$
- D.  $O(n^2)$  but not  $O(\log n)$

Florian just got a job at Google. The following 6 questions are about his first day at job.

33. His first task was to store the names and phone numbers of Google's 2 billion customers. Assume every name is a unique string of 10 characters (denoted as name1, name 2 and so on) and each number is 10 digits long (denoted as num1 num2 and so on). Which of the following Matlab definitions will consume minimum memory in Matlab without losing any data? (Assume 'abcdefghij' represents the name of one of the employees and 9123456780 represents a 10 digit phone number)

- A. `name = 'abcdefghij'; number = 9123456780;`
- B. `name = ['abcdefghij']; number = int32(9123456780);`
- C. `name = int8('abcdefghij'); number = int32(9123456780);`
- D. `name = int8('abcdefghij'); number = int64(9123456780)`

34. He wanted to try storing the data as a cell array. He first considered the following three possible options:

- 1. `{{'name1' 'name2' 'name3'.....} {num1, num2, num3....}}`
- 2. `{{'name1' num1}}{'name2' num2} {'name3' num3}.....}`
- 3. `{{'name1' 'name2' 'name3'.....} [num1, num2, num3....]}`

Which of the following is the right relation of the memory consumption of the three options?

- A.  $1 > 2 > 3$
- B.  $1 = 2 > 3$
- C.  $3 > 1 > 2$
- D.  $2 > 1 > 3$



35. After considering all possible options, he decided on using the following structure for storing the entire data in cells: {'name1' num1 'name2' num2 'name3' num3.....}  
 The minimum amount of contiguous memory space required in Matlab for storing this is approximately:
- A. 112 GB
  - B. 28 GB
  - C. 448 GB
  - D. 224 GB
36. Next, he had to work on data retrieval. So, he wrote a constant time perfect hash function that gives the index of a name in the cell array (each customer name hashes to a unique index). What is the time complexity of retrieving a name and phone number pair from his new cell array database, created as per Problem 35, using this hash function?
- A.  $O(n)$
  - B.  $O(n \cdot \log n)$
  - C.  $O(1)$
  - D.  $O(\log n)$
37. Florian went to his supervisor with this idea and was instead asked to use linked lists. Luckily, he remembered one of his E7 assignments and decided to store the entire data as linked lists with each node storing data on a single customer. The definition of each node is as follows:

```
classdef myLNode < handle
    properties
        name = [];
        number = [];
        next = myLNode.empty;
    end
end
```

- The minimum amount of contiguous memory space approximately required in Matlab for storing data using this method falls in what range? (Assume each property definition consumes 112 bytes in addition to the data and a pointer consumes 64 bytes)
- A. 2-3 GB
  - B. 1000-1200 kilobytes
  - C. 500-600 bytes
  - D. 1000-1200 MB
38. Finally, he wants to retrieve data. What is the worst-case time complexity of retrieving a name number pair with this new linked list structure? (Assume there are no hash functions involved)
- A.  $O(1)$
  - B.  $O(n \cdot \log n)$
  - C.  $O(n)$
  - D.  $O(\log n)$

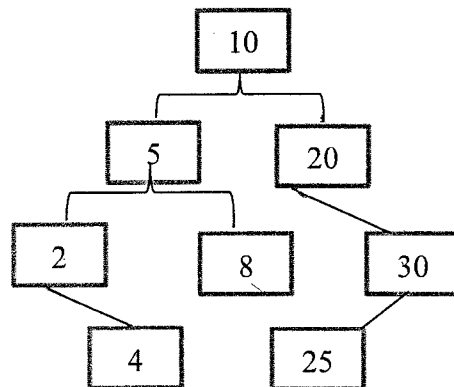
39. In order to have a FIFO structure, we should always:
- A. Enqueue at the head and dequeue at the head
  - B. Dequeue at the tail and enqueue at the tail
  - C. Neither A or B
  - D. Either A nor B
40. The time taken to enqueue or dequeue in a FIFO queue is:
- A. at least linear in the length of the data
  - B. constant time w.r.t. the length of the data
  - C. at least logarithmic in the length of the data
  - D. None of the above
41. Which of the following statements is NOT true about stacks?
- A. You can only insert an element after the last inserted element.
  - B. The 'Push' function allows you to insert an element at the bottom of the stack.
  - C. You can only insert an element at the top of the stack.
  - D. All of the above
42. You are given the following hash function that takes a string and its length and produces the index at which it will be stored in a hash table.

```
function index = hashfunc(str,length)
index = str(length/2)+str(1);
end
```

Which of the statements given below is true about this hash function?

- A. It is a perfect hash function
  - B. It has a worst case complexity of  $O(1)$
  - C. It is a minimum time perfect hash function
  - D. None of the above
43. Given a constant time perfect hash function, the best time complexity to hash and store  $n$  data items in an array is:
- A.  $O(1)$
  - B.  $O(n)$
  - C.  $O(\log n)$
  - D. None of the Above
44. Suppose we used a constant time imperfect hash function instead, what is the best time complexity to hash and store  $n$  data items in an array now?
- A.  $O(1)$
  - B.  $O(n)$
  - C.  $O(\log n)$
  - D. None of the above

45. Given  $n$  lists with upto  $m$  elements in each, the search time for elements that are located in the first two to three positions of each list would be:
- None of the below
  - $O(m)$
  - $O(m*n)$
  - Independent of the number of lists.
46. A list contains 255 names to be placed in a binary search tree lexicographically. What is the minimum and maximum height of that tree respectively?  
(note: height of the tree = number of layers in the tree - 1)
- 7, 254
  - 6, 254
  - 254, 255
  - 10, 255
47. Suppose you have an unordered linked list, with every node in the list having a 'value' and a 'next' pointer property pointing to the next node in the list. The list also has head and tail pointers. Which of the following operations can be implemented with a constant time complexity?
- I: Insertion at Head   II: Insertion at Tail   III: Data Retrieval   IV: Deletion at Tail
- I and II
  - III and IV
  - I, II, and III
  - I, II, III, and IV
48. A four layer binary tree resulted from data being inserted in the order in which it arrived. Which of the following represents a possible data arrival order?



- 10, 5, 20, 8, 30, 25, 2, 4;
- 10, 5, 20, 2, 8, 4, 25, 30;
- 4, 25, 2, 8, 30, 5, 20, 10;
- 4, 2, 5, 8, 25, 30, 20, 10;

49. Bassel is the supervisor mentioned in Problem 37. He himself deals with storing Google employee data. His next task is to store a list of all employees and their favorite songs. Since, the total number of employees may keep changing over time, he decides to use a hash table of size 10. Each employee's name is hashed to an index with a constant time hash function and then his/her information is stored as a node in a linked list at that index. Each node in the linked list has the following definition:

```

classdef myLNode < handle
    properties
        key =[];
        value = [];
        next = myLNode.empty;
    end
end

```

where, `key` stores the employee name and `value` is an array of favorite songs.

- What is the worst-case time complexity of searching for a given song in the database. (i.e. how does the worst-case time taken change in terms of the number of names( $n$ ) and number of songs (also  $n$ )). Assume that no two employees have the same favorite song (i.e. – Each song exists at only one place in the entire database) ?
- $O(n^2)$
  - $O(n \cdot \log n)$
  - $O(n)$
  - $O(1)$
50. Suppose the only motive of the above database is to search a specific song and which employee likes it, which of the following will improve the worst-case search complexity most? (Assume everything except what is mentioned in the option stays the same)
- Increase the hash table size to 20
  - Hash the employee name and use `key` to store the array of songs and `value` to store employee name in each node.
  - Store all data in a single linked list and don't use any hash table at all.
  - Last answer to finish E7. You are definitely tricking me. All have the same worst-case search complexity. Aha!