Text Processing

Outline
• Announcements:
  – Homework I: due Today. by 5, by e-mail
  – Discuss on Friday.
  – Homework II: on web
• HW I: question 7
• Finish functions
• Text
• Matlab path
• Survey

HW I: question 7
• Lots of people having trouble
  – $dt=0.1$, $t=(0:dt:21*dt)$;
  – $N$ is up to you, so pick something easy
This isn’t English!

- Why do we need text?
  - Comments from functions
  - File names
  - Label plots
  - Interact with users
  - Record-based I/O

Hello World!

- Create strings with single quote (‘)
  - a=’Hello World!’
- Believe it or not, characters are not doubles
  - a is an array of char
- Can display text nicely with disp(str)
  - if str is a matrix, each row is a new line

Working with text

- Concatenation--same as with other vectors
  - a=’Hello’; b=’World!’;
  - greetings=[a,’ ’;b];
  - Will greetings=[a;b] work?
**Number-to-String Conversions**

- `int2str` & `num2str` convert numbers to text
  - `int2str(2)` returns ‘2’
- `str2num` converts to numbers
  - `str2num('3')*2` returns 6

**Searching for strings**

- can search for single characters with `find`
  - `str='Scripts are evil!'; I=find(str=='');` % `I=[8 12]`
- search for substring `ss` in `str` with
  - `I=findstr(ss,str)`
  - `I=find('evil',str);` % `I=[13];`
  - `findstr(str,ss)` is the same
    - `findstr` always searches for small string in big string

**Working with ASCII**

- `double(str)` returns an array with ASCII codes
  - `str='012ABCabc'`
  - `num=double(str)=[48 49 50 65 66 67 97 98 99]`
- `char(num)` converts ASCII codes to char
  - `char(num)` returns ‘012ABCabc’
**Misc. Text Functions**

- `R=input(QuestionStr)`  
  - asks user for input, returned as R  
  - DO NOT USE IN THIS CLASS!!!  
  - For entertainment purposes only. Function arguments are the best way to get info into your functions

- `xlabel`, `ylabel`, `title` -- label plots  

- `text(x,y,str)` -- places string at x,y on plot

- `S=sprintf(str, val1, val2, ...)` -- C-like string creation  
  - `S=sprintf('Integer %d
Double %f
', 5, -pi);`  
  - S is 1-by-27 array of char

**String Summary**

- Matlab stores strings in arrays of char (ASCII)  
  - convert to ASCII values with double, to ASCII text with `text`  
  - Convert numbers to strings with `int2str`, `num2str`, `sprintf`  
  - Convert strings to numbers with `str2num`

- Search strings with `find` (single character) or `findstr` (substring)

**Matlab Path**

- Matlab maintains a list of directories where it searches for files  
  - m-files, data files  
  - Type "path" to see  
  - Can add directories using `addpath` or through GUI  
  - Ex: `addpath('D:\Andy\mfiles')`
• **startup.m**

  - startup.m is a special script (the only good one!) that (if it exists) is executed as Matlab starts
    - Not installed-You must create it
  - On UNIX/Mac, startup.m is in ~/matlab
  - Windows: %MATLABROOT%	oolbox\local
  - Windows NT/2000: in matlab in Profiles directory
    - Ex: C:\WINNT\Profiles\andy\matlab
    - Can find out where profiles are found by typing `getenv('USERPROFILE')`

• **startup.m**

  - Uses of startup.m
    - Personalize path--place addpath statements
    - Customize matlab
    - Set default directory
    - Set default graphics output (see 402)

• **Personal Opinion**

  - Create your own m-files directory, & put m-files there
    - group m-files into subdirectories by topics
  - Place addpaths in startup.m so you can always use your functions
  - CD into data directories & work there

*Above is the instructor’s opinion and does not necessarily reflect that of CIS or Cornell University*
You now know the basics of Matlab.
- The rest of the course will be spent extending and reinforcing that knowledge
- More Matlab or more applications?

<table>
<thead>
<tr>
<th>Survey</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Matlab</td>
</tr>
<tr>
<td></td>
<td>Polymorphic functions</td>
</tr>
<tr>
<td></td>
<td>Objects beyond arrays</td>
</tr>
<tr>
<td></td>
<td>Improving performance</td>
</tr>
<tr>
<td></td>
<td>Applications</td>
</tr>
<tr>
<td></td>
<td>File I/O (binary &amp; text)</td>
</tr>
<tr>
<td></td>
<td>Linear Systems</td>
</tr>
<tr>
<td></td>
<td>Diff. Equations</td>
</tr>
<tr>
<td></td>
<td>Statistics</td>
</tr>
<tr>
<td></td>
<td>Graphics</td>
</tr>
<tr>
<td></td>
<td>Polynomials &amp; splines</td>
</tr>
<tr>
<td></td>
<td>Signal processing (FFT)</td>
</tr>
<tr>
<td></td>
<td>Optimization</td>
</tr>
</tbody>
</table>