

UCLA CS 35L midterm, fall 2022

100 minutes total, open book, open notes, closed computer.

1 minute = 1 point. Write answers on exam.

Name: _____ Student ID: _____

1a (6 minutes). Explain why the shell command `ln . dot` fails while `ln -s . dot` succeeds.

1b (8 minutes). Explain why a hard link H to a symbolic link S can refer to a different file than S does, whereas a hard link H to another hard link J must reference the same file that J references.

2. Suppose the file 'mystery' is mode 755 and contains the following contents, derived from a simplified version of Internet RFC 3986, "Uniform Resource Identifier (URI): Generic Syntax":

```
#!/bin/sh
alpha='a-zA-Z'; digit='0-9'; unreserved="[$alpha$digit._~-]"
scheme="([$alpha][$alpha$digit+.-]*)"
hexdig="[A-Fa-f$digit]"; pct_encoded="(%%$hexdig$hexdig)"
sub_delims='[!$&'\''()*+;=]'
userinfo="(($unreserved|$pct_encoded|$sub_delims|:)*)"
deco="([$digit]|[1-9][$digit]|1[$digit][$digit]|2[0-4][$digit]|25[0-5])"
IPv4address="($deco\\. $deco\\. $deco\\. $deco)"
reg_name="(($unreserved|$pct_encoded|$sub_delims)*)"
host="($IPv4address|$reg_name)"; port="([$digit]*)"
authority="(($userinfo@)?$host(:$port)?)"
pchar="($unreserved|$pct_encoded|$sub_delims|:|@)"
segment="($pchar*)"; segment_nz="($pchar+)"
path_abempty="(/$segment)*"
path_absolute="(/($segment_nz(/$segment)*)?)"
path_rootless="($segment_nz(/$segment)*)"
hierpart="((//$authority$path_abempty|$path_absolute|$path_rootless)?)"
query="(($pchar|/|\\?)*)"
fragment="($query)"
grep -E "$scheme:$hierpart(\\?$query)?(#$fragment)?"
```

2a (5 minutes). What is the output of the following shell command? Briefly explain.

```
./mystery </dev/null
```

2b (5 minutes). Suppose 'mystery' were mode 644 instead of 755. How would that affect the output of the shell command in (a)?

2c (5 minutes). Briefly explain why the `'./'` is needed before `'mystery'` in the shell command in (a). What would happen if `'./'` were omitted? Might the command still work anyway?

2d (5 minutes). What is the output of the shell command `"echo '' | ./mystery"`? Briefly explain.

2e (6 minutes): What is a shortest nonempty string `F00` such that the shell command `"echo 'F00' | ./mystery"` outputs something? Explain.

2f (8 minutes). Almost every occurrence of `'\'` in `'mystery'` is doubled, so it appears as `'\\'`. Explain why this doubling is needed. Explain what would happen if every instance of `'\\'` were replaced by a single `'\'`, for example, in the `IPv4address` line.

3. You're annoyed that the following Emacs keystrokes:

`C-u 2 M-<`

put you 20% (i.e., 2/10) of the way through the current buffer. You'd rather have that command put you on line 2 of the current buffer instead. You type the following

`C-h k M-<`

and see the following on your screen:

`M-<` runs the command `beginning-of-buffer` (found in `global-map`), which is an interactive byte-compiled Lisp function in `'simple.el'`.

It is bound to `<begin>`, `C-<home>` and `M-<`. It can also be invoked from the menu: `Edit` → `Go To` → `Goto Beginning of Buffer`

`(beginning-of-buffer &optional ARG)`

Move point to the beginning of the buffer.

With numeric arg `N`, put point `N/10` of the way from the beginning. If the buffer is narrowed, this command uses the beginning of the accessible part of the buffer.

Push mark at previous position, unless either a `C-u` prefix is supplied, or `Transient Mark` mode is enabled and the mark is active.

This documentation bothers you in a different way, because you think this is too complicated, and that the last sentence quoted above should be removed and the resulting code fixed accordingly.

[continued on next page]

You click on 'simple.el' in the above text, and see the following function definition:

```

1 (defun beginning-of-buffer (&optional arg)
2   "Move point to the beginning of the buffer.
3   With numeric arg N, put point N/10 of the way from the beginning.
4   If the buffer is narrowed, this command uses the beginning of the
5   accessible part of the buffer.
6
7   Push mark at previous position, unless either a \\[universal-argument
] prefix
8   is supplied, or Transient Mark mode is enabled and the mark is active
."
9   (declare (interactive-only "use `(goto-char (point-min))' instead."
))
10  (interactive "^P")
11  (or (consp arg)
12      (region-active-p)
13      (push-mark))
14  (let ((size (- (point-max) (point-min))))
15    (goto-char (if (and arg (not (consp arg)))
16                  (+ (point-min) 1
17                    (/ (* size (prefix-numeric-value arg)) 10))
18                  (point-min))))
19  (if (and arg (not (consp arg))) (forward-line 1)))

```

3a (6 minutes). Change the definition of beginning-of-buffer so that it no longer implements the bothersome last sentence of the documentation. Use the line numbers of the source code in your description of the changes you'd make.

[page 6]

3b (8 minutes). Further change the definition of beginning-of-buffer so that with numeric arg N it puts you on line N rather than N/10 of the way through the buffer.

Consider the following source code adapted and simplified from random.py:

```

239     def _randbelow(self, n):
240         "Return a random int in the range [0,n).  Return 0 if n==0."
241         if not n:
242             return 0
243         getrandbits = self.getrandbits
244         k = n.bit_length() # don't use (n-1) here, as n can be 1
245         r = getrandbits(k) # 0 <= r < 2**k
246         while r >= n:
247             r = getrandbits(k)
248         return r
...
375     def choice(self, seq):
376         """Choose a random element from a non-empty sequence."""
377         return seq[self._randbelow(len(seq))]
...
826     def getrandbits(self, k):
827         "getrandbits(k) -> x.  Generates an int with k random bits."
...

```

and consider the following python3 transcript on SEASnet:

```

1  >>> random.choice([1,2,3])
2  Traceback (most recent call last):
3    File "<stdin>", line 1, in <module>
4  NameError; name 'random' is not defined
5  >>> import random
6  >>> random.choice([1,2,3])
7  2
8  >>> random.choice("abcdef")
9  'a'
10 >>> random.choice("")
11 Traceback (most recent call last):
12   File "<stdin>", line 1, in <module>
13   File "/usr/local/cs/Python-3.10.7/lib/python3.10/random.py", line 3
77, in choice
14     return seq[self._randbelow(len(seq))]
15  IndexError: string index out of range

```

4a (6 minutes). Explain Python's output in lines 2-4, 7, 9, and 11-15.

4b (8 minutes). `n.bit_length()` returns the number of bits needed to represent `n` in binary, excluding the sign and leading zeros. Assuming that `getrandbits` and `bit_length` each finish in less than one microsecond, that all the other code takes zero time, and that 'choice' is called on a sequence of length 1,000,000, how much CPU time will the 'choice' method take, in the worst case? Also, how would your answer change if 'choice' is called on a sequence of length 1,048,576 (2^{20}) instead? Briefly justify your answer.

[page 9]

5 (8 minutes). HTTP/3 is often touted as a major performance win for web applications. Explain why your solution to Assignment 3 would not materially benefit from switching to HTTP/3.

6 (6 minutes). Suppose you were supposed to use Python as much as possible when doing assignment 3, thus minimizing the use of JavaScript. Explain how this would affect how you'd develop your solution to Assignment 3. In particular, how would you debug your program, compared to debugging your program with Node.js as it is?

7 (10 minutes). Diagram the DOM tree for the following HTML5 document. (You need not diagram the first line.) Assume that elements lacking closing tags are void.

```

<!DOCTYPE html>
<html lang='en'>
<head>
<title>Assignments for UCLA Computer Science 35L, fall 2022</title>
<meta charset='UTF-8'>
<link rel='author' href='mailto:eggert@cs.ucla.edu'>
<link rel='license' href='copyright.html'>
</head>
<body>
<section>
<h1>Assignments for UCLA Computer Science 35L, fall 2022</h1>
<p><small>[<a href='index.html'>35L home</a>]</small></p>
<p>Most assignments are divided into two parts: laboratory exercises and
  homeworks. Laboratory exercises are expected to be done with more
  assistance from the TAs and LAs.</p>
</section>
<footer>
<hr>
<small>
  &copy; 1999, 2003&ndash;2022 <a href='mail-eggert.html'>Paul Eggert</a>.
  See <a href='copyright.html'>copying rules</a>.<br>
  $Id: assign.html,v 1.186 2022/10/13 01:13:20 eggert Exp $
</small>
</footer>
</body>
</html>

```

In your diagram, circle the element nodes. Represent each text string by a capital "T" under the element node, at the appropriate position. (For example, write a "T" to represent "Assignments for UCLA Computer Science 35L, fall 2022".) Write "A='V'" to the right of each node that has an attribute A with value V (using the actual attribute name and value, not a literal capital "A" or capital "V").

[Put the diagram on the next page.]

[page 11]

[Diagram the DOM tree described in the previous page. Neatness counts.]