Chapter 12.2

Question 1

Which of the following Python data structures is most similar to the value returned in this line of Python:

```
x = urllib.request.urlopen('http://data.pr4e.org/romeo.txt')
```

- (•) file handle
- ( ) dictionary
- ( ) socket
- ( ) list
- ( ) regular expression

Question 2

In this Python code, which line actually reads the data?

```
import socket

mysock = socket.socket(socket.AF_INET, socket.SOCK_STREAM)
mysock.connect(('data.pr4e.org', 80))
cmd = 'GET http://data.pr4e.org/romeo.txt HTTP/1.0\n\n'.encode()
mysock.send(cmd)

while True:
    data = mysock.recv(512)
    if (len(data) < 1):
        break
    print(data.decode())
mysock.close()
```

- (•) mysock.recv()
- ( ) socket.socket()
- ( ) mysock.close()
- ( ) mysock.connect()
- ( ) mysock.send()

Question 3

Which of the following regular expressions would extract the URL from this line of HTML:

```
<p>Please click <a href="http://www.dr-chuck.com">here</a></p>
```

- (•) href="(.+)"

bash-hike

- href=".+"
- http://.*
- <.*>

Question 4

In this Python code, which line is most like the open() call to read a file:

```
import socket

mysock = socket.socket(socket.AF_INET, socket.SOCK_STREAM)
mysock.connect(('data.pr4e.org', 80))
cmd = 'GET http://data.pr4e.org/romeo.txt HTTP/1.0\n\n'.encode()
mysock.send(cmd)

while True:
    data = mysock.recv(512)
    if (len(data) < 1):
        break
    print(data.decode())
mysock.close()
```

- mysock.connect()
- import socket
- mysock.recv()
- mysock.send()
- socket.socket()

Question 5

Which HTTP header tells the browser the kind of document that is being returned?

- Content-Type:
- HTML-Document:
- Metadata:
- ETag:
- Document-Type:

Question 6

What should you check before scraping a web site?

- That the web site only has links within the same site
- That the web site allows scraping
- That the web site supports the HTTP GET command
- That the web site returns HTML for all pages

Question 7

What is the purpose of the BeautifulSoup Python library?

- ○ It animates web operations to make them more attractive
- ○ It optimizes files that are retrieved many times
- ○ It builds word clouds from web pages
- ● It repairs and parses HTML to make it easier for a program to understand
- ○ It allows a web site to choose an attractive skin

Question 8

What ends up in the "x" variable in the following code:

```
html = urllib.request.urlopen(url).read()
soup = BeautifulSoup(html, 'html.parser')
x = soup('a')
```

- ● "A list of all the anchor tags (<a..) in the HTML from the URL
- ○ True if there were any anchor tags in the HTML from the URL
- ○ All of the externally linked CSS files in the HTML from the URL
- ○ All of the paragraphs of the HTML from the URL

Question 9

What is the most common Unicode encoding when moving data between systems?

- ○ UTF-128
- ○ UTF-64
- ○ UTF-32
- ● UTF-8
- ○ UTF-16

Question 10

What is the ASCII character that is associated with the decimal value 42?

- ○ ^
- ○ /
- ○ !
- ● *
- ○ +

---

bash-hike

Question 11

What word does the following sequence of numbers represent in ASCII:

108, 105, 110, 101

- ○ lost
- ○ ping
- ● line
- ○ tree
- ○ func

Question 12

How are strings stored internally in Python 3?

- ○ ASCII
- ● Unicode
- ○ Byte Code
- ○ UTF-8
- ○ EBCDIC

Question 13

When reading data across the network (i.e. from a URL) in Python 3, what method must be used to convert it to the internal format used by strings?

- ○ find()
- ○ encode()
- ○ trim()
- ○ upper()
- ● decode()
