

CSC 102 • HOW THE INTERNET WORKS  
FINAL TAKE-HOME EXAM  
DUE 2:00 PM ON TUESDAY, 21 DECEMBER 2010  
PLEASE SUBMIT YOUR WORK ELECTRONICALLY VIA MOODLE

*You may not consult any person other than the professor when completing this exam, but you may use references online and in print. Wherever you use external sources, leave me a URL or citation. Don't forget to use quotation marks when copying verbatim from a source. There will be a certain amount of Web searching necessary to respond to several of the questions. Wikipedia will be accepted as a source of answers. You may also have to use Google or its equivalent intelligently, and some of the other tools we have studied.*

**1. Data Transfer (12 points)**

---

The Sloan Digital Sky Survey is a detailed digital map of the heavens used by astronomers. The images and associated data released to the public totals 40.5 TB (note: 1TB = 1000 GB = 1,000,000 MB). The astronomy department at Stardust College want to study the survey data, and are trying to figure out a way to get a full copy.

- a.) They are connected to the Internet with a very fast ethernet cable that has maximum capacity of 100 Mbps. In the unlikely case that they were able to devote the entire connection bandwidth to downloading the data set, how long would this take?
- b.) A graduate student in the department suggests that since Bigbucks U. has a copy of the full data set on BluRay disks, she could drive there in her car and pick them up. How many disks would be required to record the entire data set, if each has a capacity of 50 GB?
- c.) Assuming that the trip from Bigbucks U. to Stardust College takes two and a quarter hours, what is the effective data transfer rate of this operation?

**2. Communications Protocols (16 points)**

---

A home network is built around a wireless router that uses DHCP with network address translation. The router assigns local IP addresses in the reserved 192.168.x.x range; it takes the 192.168.0.1 address and gives successively higher addresses to each local connection. The global IP address provided by the household's internet service provider is 192.76.85.245.

- a.) The desktop computer connected to the router is at 192.168.0.2 on the local area network. Suppose that it sends out a request for a web page to www.smith.edu. As the packet travels (i) from the desktop to the router and (ii) from the router to www.smith.edu, what would it show as the sender IP address and as the destination IP address?

- b.) When www.smith.edu responds to the web page request, it sends packets back to the desktop computer. As the packet travels (i) from www.smith.edu to the router and (ii) from the router to the desktop computer, what would it show as the sender IP address and as the destination IP address?
- c.) Suppose that a laptop connected to the same router at 192.168.0.4 on the local area network also requests a web page from www.smith.edu at the same time. Packets are sent to the router intended for both recipients. How does the router know where to send them?
- d.) Give at least one advantage of using network address translation for small home networks like this one.

### 3. HTML (18 points)

---

Identify the outright errors in the web page source below, as well as instances where it departs from the "best practices" taught in class.

```
<HTML>
<TITLE>Exam Question</TITLE>
<BODY bgcolor="blue">
<H2>HTML Exam Question</H2>

This is part of the <B><EM>CSC 102 Fall 2010</B></EM> final exam.
<P>
You should try to find all the errors &ampersand bad habits in this page.
<FONT color="gray">It is very sneaky, with lots of hidden mistakes.</FONT>

<HR>~~~~~</HR>

<UL list-style-type="decimal">
<LI style="list-style-position: inside">One item</LI>
<li>Another item</li>
A third item
</UL list-style-type="decimal">

<PRE>
A heart: <3
</PRE>

<BLOCKQUOTE><BLOCKQUOTE>
I love indenting!
</BLOCKQUOTE></BLOCKQUOTE>

<CENTER>Middle!</CENTER>

<TABLE width=100%><TR><TD width="90%"></TD><TD>Right!</TD></TR></TABLE>

</BODY>
Last modified: 11 December 2010<BR>
</HTML>
```

#### 4. Images (6 points)

---

What three image file formats are recommended for web use? Suppose that you are designing a web site for a professional photographer, which will include a logo and samples of work. In this context, specify for each file type which parts of the site you would most appropriately use it for. If you would not use one of the three in this context, explain why.

#### 5. HTML Forms (8 points)

---

Devise a form that would produce the URL result shown below if submitted using its default values. (In other words, the form is submitted without making any changes and it produces this result.) The four results shown below should come from a checkbox, text input, popup menu, and button, respectively.

*`http://example.com/form.html?new=Yes&name=Anonymous&side=left&ok=ok`*

#### 6. Email (14 points)

---

Consider the email message source below. (Some irrelevant portions of the email have been redacted.) Please answer the questions that follow.

```
Return-path: <nicholas.r.howe@gmail.com>
Received: from mscreen3.smith.edu (mscreen.smith.edu [131.229.64.72])
    by gwsmtpl1.smith.edu with ESMTP; Sat, 11 Dec 2010 22:37:08 -0500
Received: from scmapp1.smith.edu (scmapp1.smith.edu [131.229.64.81])
    by mscreen3.smith.edu (8.14.3/8.14.3) with SMTP id oBC3b9IY007482
    for <nhowe@smith.edu>; Sat, 11 Dec 2010 21:37:09 -0600
Received: from (unknown [209.85.161.177]) by scmapp1.smith.edu with smtp
    id 5a4f_5c4e_187d6dd8_05a1_11e0_bd57_0014221cc49d;
    Sat, 11 Dec 2010 22:37:08 -0500
Received: by gxk27 with SMTP id 27so2968162gxk.36
    for <nhowe@smith.edu>; Sat, 11 Dec 2010 19:37:07 -0800 (PST)
MIME-Version: 1.0
Received: by 10.236.95.41 with SMTP id o29mr5588573yhf.40.1292125027865; Sat,
    11 Dec 2010 19:37:07 -0800 (PST)
Received: by 10.236.108.129 with HTTP; Sat, 11 Dec 2010 19:37:07 -0800 (PST)
Date: Sat, 11 Dec 2010 22:37:07 -0500
Message-ID: <AANLkTimcX54STYqBDWubNdceJpiZsaOmGxZ+EVnThapk@mail.gmail.com>
Subject: Test Email
From: Nicholas Howe <nicholas.r.howe@gmail.com>
To: nhowe@smith.edu
Content-Type: multipart/mixed; boundary=00235448f2594a956304972e4d2c

--00235448f2594a956304972e4d2c
Content-Type: multipart/alternative; boundary=00235448f2594a955804972e4d2a
```

--00235448f2594a955804972e4d2a  
Content-Type: text/plain; charset=ISO-8859-1

This is an email for the CSC 102 fall 2010 final exam.

--00235448f2594a955804972e4d2a  
Content-Type: text/html; charset=ISO-8859-1

This is an email for the CSC 102 fall 2010 final exam.<br>

--00235448f2594a955804972e4d2a--  
--00235448f2594a956304972e4d2c  
Content-Type: image/png; name="sun.png"  
Content-Disposition: attachment; filename="sun.png"  
Content-Transfer-Encoding: base64  
X-Attachment-Id: f\_ghldqb2h0

iVBORw0KGgoAAAANSUHEUgAAABAAAAQCAIAAACQkWg2AAAAAXNSR0IArs4c6QAAAAARnQU1BAACx  
jwv8YQUAAAgY0hSTQAAeiYAAICEAAD6AAAAgOgAAHUwAADqYAAAOpgAABdwnLpRPAAAAjTJREFU  
OE99UkkSwCAM8mk+zZ9bzQoa2+nBScgCoc36a3P0MtMySogV99SOJwwKdtQbK8haQBxLZVTk4JGT  
d3ed4Gmbrvv4H0GplISd9YY2YM0j1qNilwtEzrFGmMNFag8kNjowAF3MlkKJsDshwmcKEjrYlzp  
+gQCoE9ZpVfKijdGIU0l6ojW4EuHL9iFVwH6h8xyH+50c+3wD/S8XKesH5YgAAAAEIFTkSuQmCC

--00235448f2594a956304972e4d2c--

- a.) How many tags make up the email header in the message source as shown? (Count the lines, ignoring indented lines.)
- b.) What is the text of the message? Why does it apparently repeat itself in the message source?
- c.) There is an attachment to this message. What type of content does it contain? Could opening this attachment harm your computer, according to the guidelines discussed in class?
- d.) Suppose that you were suspicious of this message, and wanted to do everything you could to confirm the source. Describe what you can deduce or uncover about where this message came from.

## 7. Cryptography (6 points)

---

Imagine a world where strong cryptography does not exist, and compare it to a world where strong cryptography is widely available and widely used. Discuss at least three positive or negative aspects of each scenario. (You may count a positive for one as a negative for the other; e.g., you need to discuss at least three areas of significance, and describe why each would be advantageous/disadvantageous in one scenario or the other.)

## 8. Personal Safety Online (6 points)

---

Why do Microsoft and other software companies release security patches for their software, knowing that hackers will examine these closely to learn about vulnerabilities of unpatched systems? Discuss this topic, including its implications for individual computer users like you.

## 9. Web Search (16 points)

---

For each of the queries below, indicate which pages would be retrieved and their rank order, if a search engine uses a bag-of-words model with query augmentation, and ranks results according to Google's PageRank model. (Note: you are not expected to carry out the full PageRank computation, which is mathematically complex. Just rank according to the principles and example used in class. If you're not sure, include a justification of your choice.)

a.) Query: Flowers

b.) Query: Running

c.) Query: Bank runs

d.) Query: World

