Final Exam
Compsci 82, December 9, 2008
Prof. Owen Astrachan

Name: _____

Login: _____

Honor code acknowledgment (signature) _____

1. (A) (B) (C) (D)   16. (A) (B) (C) (D)

2. (A) (B) (C) (D)   17. (A) (B) (C) (D)

3. (A) (B) (C) (D)   18. (A) (B) (C) (D)

4. (A) (B) (C) (D)   19. (A) (B) (C) (D)

5. (A) (B) (C) (D)   20. (A) (B) (C) (D)

6. (A) (B) (C) (D)   21. (A) (B) (C) (D)

7. (A) (B) (C) (D)   22. (A) (B) (C) (D)

8. (A) (B) (C) (D)   23. (A) (B) (C) (D)

9. (A) (B) (C) (D)   24. (A) (B) (C) (D)

10. (A) (B) (C) (D)   25. (A) (B) (C) (D)

11. (A) (B) (C) (D)   26. (A) (B) (C) (D)

12. (A) (B) (C) (D)   27. (A) (B) (C) (D)

13. (A) (B) (C) (D)   28. (A) (B) (C) (D)

14. (A) (B) (C) (D)   29. (A) (B) (C) (D)

15. (A) (B) (C) (D)   30. (A) (B) (C) (D)

                              31. (A) (B) (C) (D)

Questions 32 and 33 are answered on the back.

Question 32, Part A (2 points)

Question 32, Part B (1 point)

Question 32, Part C (2 points)

Question 33 (4 points)

PROBLEM 1:

Which of the following is most likely the reason for the term "copyleft" rather than "copyright" as used in the Gnu Public License from `www.gnu.org`, the organization started by Richard Stallman?

A. Proprietary software developers use copyright to take away freedom, so we use *copyleft* to guarantee freedom.

B. Proprietary software is sold by businesses whose politics are more likely "to the right"; our politics are "to the left" so we use *copyleft*.

C. We are working to change copyright law, we use *copyleft* to emphasize that our work is distributed outside of the US copyright system—we do not abide by US copyright law.

D. Software can be a liberating force in the world—it can literally change how people think, act, and work. To emphasize these opportunities we use a term different than "copyright" and *copyleft* makes sense in this regard.

PROBLEM 2:

In class an excerpt from a movie was shown to illustrate the liberating potential for free software and the free and open source software movement. What movie was the source of the excerpt?

A. Born Free

B. Live Free or Die Hard

C. Braveheart

D. Hackers

PROBLEM 3:

Which of the following is/are features of all open source licenses for software?

I. The source code must be available either free or at minimal cost.

II. The software can be sold or given away, no restrictions on how it is distributed can be made.

III. The software cannot discriminate against people or groups or fields of endeavor.

A. I only

B. III only

C. I, and II only

D. I, II, and III

PROBLEM 4:

Which of the following is **LEAST** likely as a reason that Al Yankovic's *White and Nerdy*, a parody of Chamillionaire and Krayzie Bone's *Dirty*, is acceptable under US copyright law?

  A. He paid the writers/singers a licensing fee to be able to record the song.

  B. US Copyright law specificies parody as a legitimate fair use.

  C. Al Yankovic has a history of writing creative parodies and precedent is important in considering legal cases.

  D. The *White and Nerdy* version likely did not detract from sales of *Dirty* and may have increased the sales.

PROBLEM 5:

Which of the following is **NOT** attributable to the website `4chan.org`?

  A. The first hosting of images of Sarah Palin's hacked Yahoo email account.

  B. The origin of *rickrolling* or a link to Rick Astley's song "Never Gonna Give You Up" instead of the link someone thinks they'll get when they click.

  C. The origin of the first *lolcat* images such as the one below.

  D. The origin of the phrase "on the Internet nobody knows you're a dog" associated with a cartoon related to anonymity and online personas.

PROBLEM 6:

The IETF (Internet Engineering Task Force) develops and creats the technical standards that drive the Internet. Which of the following would you be **LEAST** likely to see at an IETF meeting?

  A. A higher percentage of laptops in use that run the Linux operating system than is reflected in most technical businesses.

  B. Many men and few women reflecting the demographics of the technical group that makes up the IETF.

  C. Representatives from IBM, Microsoft, Google, Cisco and other technical, software, and web companies.

  D. Representatives from governments helping oversee the development of Internet standards.

PROBLEM 7:

Which of the following is **NOT** true with respect to videos on YouTube according to the DMCA and other copyright laws?

A. If you watch a video on YouTube that is subsequently removed because it violates copyright law then you are safe from prosecution and will not be charged with violation of copyright law, even under vicarious infringement guidelines in the laws.

B. If you appear in a video hosted on YouTube, but that was recorded by a classmate, you can have the video removed since YouTube must adhere to takedown notices from people that appear in videos hosted on the site.

C. If you make a video for a class, and your professor uploads it to YouTube, you can cause it to be removed because YouTube must adhere to takedown notices from copyright holders, even if someone else uploads to the site.

D. YouTube has a policy that will remove pornography that is tagged/flagged by YouTube staff according to YouTube's guidelines for "sexually suggestive" material.

PROBLEM 8:

Which of the following is/are true about anonymity on the Internet?

I. Some browsers now support "private browsing" in which no history, cookies, or other data is tracked during a browsing session, though this isn't always effective given the number of websites that require cookies.

II. Anonymizing Networks like TOR (The Onion Router) facilitate anonymous surfing even when IP addresses are stored by sites being browsed.

III. Sites that encourage anonymity such as `4chan` or `juicycampus` actually know who you are because of cookies that are tracked by the sites and which store your personal information even when you try to hide it.

A. I only

B. III only

C. I and II only

D. I, II, and III

PROBLEM 9:

Which of the following is/are true regarding encryption as used on the Internet today?

   I. Encryption is based on mathematics, and in particular on primality and factoring numbers.

  II. Good encryption such as PGP cannot be broken using brute force techniques (trying all possible keys) in our lifetime.

 III. Federal agents try to circumvent encryption technology and programs by logging the keystrokes that could include passwords/keys similarly to a wiretap or eavesdropping on phone calls.

   A. I only

   B. III only

   C. I and III only

   D. I, II, and III

PROBLEM 10:

The excerpt below appeared in February 2007.

> WASHINGTON: Hackers briefly overwhelmed at least three of the 13 computers that help manage global computer traffic Tuesday in one of the most significant attacks against the Internet since 2002.
>
> Experts said the unusually powerful attacks lasted as long as 12 hours but passed largely unnoticed by most computer users, a testament to the resiliency of the Internet. Behind the scenes, computer scientists worldwide raced to cope with enormous volumes of data that threatened to saturate some of the Internet's most vital pipelines.

Which reason is the **LEAST** likely reason the attack went unnoticed.

   A. DNS information is cached in computers near Internet users, queries don't typically make it to the root servers.

   B. The 13 root servers are replicated, most in many places, so even if one goes down due to a denial of service attack, its functionality is replicated elsewhere.

   C. Even if the DNS system slows down, it won't stop. Internet users are used to slow-downs and the DNS slow down wouldn't really have a discernible affect.

   D. A significant number of users use OpenDNS and other alternative DNS systems — these alternate systems weren't targeted and thus most people didn't notice when the conventional DNS system stopped working well.

PROBLEM 11:

I can purchase the movie Gattaca via iTunes or Amazon video on demand for $9.99. I can watch it for free via `hulu.com`. I can also download it using a Bittorrent client from: `http://www.mininova.org/tor/576963` (this site has not been verified in terms of availability or quality, but the site indicates Gattaca is available).

Which of the following claims is **NOT** true about watching Gattaca?

  A. Watching via hulu.com is possible in part because advertisements are shown before and during the movie, these ads can't be skipped.

  B. If I pay for the movie and watch it in iTunes than I can also download the movie using Bittorrent as a backup without violating copyright law.

  C. If I download the movie using Bittorrent and show it to you in my office where you don't know how I obtained the movie we watch then you have not violated copyright law.

  D. I can back up the version I download from iTunes as part of my daily backup plan even if I store all my backed-up data on Amazon's S3 data-storage servers.

PROBLEM 12:

Which of the following is/are true about the Akamai overlay network?

  I. Akamai's technology is covered by the patent system so there are no other Content Distribution Networks.

  II. Akamai's technology was developed primarily by computer scientists rather than business entrepreneurs.

  III. Akamai's technology is based in part on storing replicated copies of content throughout the world.

  A. II only

  B. III only

  C. II and II only

  D. I, II, and III

PROBLEM 13:

Which of the following is/are true regarding p2p networks on college campuses.

I. The Higher Education Reauthorization Act, passed recently, requires colleges to offer alternatives to illegal downloading "to the extent practible".

II. Universities and colleges average spending more than $200,000/year each to detect and deal with p2p traffic according to a report published in the last two years.

III. Several colleges use deep packet inspection on network traffic to look for illegal filesharing via p2p networks in addition to shaping traffic by slowing down p2p packets.

A. I only

B. II only

C. I and II only

D. I, II, and III

PROBLEM 14:

In the paper *Tussle in Cyberspace* from the first week of the course the abstract says:

> As the Internet has moved from a research curiosity to a recognized component of mainstream society, new requirements have emerged that suggest new design principles, and perhaps suggest that we revisit some old ones.

Which of the following is/are likely to be one of the design principles that is being revisited and re-examined as outlined in that paper and in our studies this semester.

I. In the original Internet, users trusted each other, that's not the case today, the model of trust on the Internet (between users and between machines) should change to reflect changes in the user population.

II. Internet Service Providers have incentives to look at the packets broadcast on the Internet for reasons of service, e.g., to provide better service to some packets and for reasons of censorship or legality, e.g., copyright violations. The original Internet used an end-to-end design principle, the network was stupid. As packets are inspected the network becomes 'smarter' and this changes how things work.

III. The original purpose of DNS was to associate machines with names so that a name was independent of the machine's location. Today names are used for trademark and advertising, there are far more domain names, and so the entire DNS system has more than one purpose as it did orginally and so changes may be needed.

A. I only

B. II only

C. II and III only

D. I, II, and III

PROBLEM 15:

Which of the following is **LEAST** likely to be part of the sayings, writings, and philosophy of Free Software advocate Richard Stallman.

A. The idea of copyright did not exist in ancient times, when authors frequently copied other authors at length in works of non-fiction. This practice was useful, and is the only way many authors' works have survived even in part.

B. If programmers deserve to be rewarded for creating innovative programs, by the same token they deserve to be punished if they restrict the use of these programs.

C. Fighting patents one by one will never eliminate the danger of software patents, any more than swatting mosquitoes will eliminate malaria.

D. I could make money writing code, but since I openly advocate free software I give away all my code or auction it for charity: you must practice what you preach and software should be given away to those that need it, selling software is despicable.

PROBLEM 16:

In one of the readings from the first week of the course Jonathan Zittrain writes:

"generative systems are powerful and valuable, not only because they foster the production of useful things like Web browsers, auction sites, and free encyclopedias, but also because they can allow an extraordinary number of people to express themselves in speech, art, or code and to work with other people in ways previously not possible."

Which of the following is **NOT** an example of the kind of generative system Zittrain describes.

A. Flickr, a photo-sharing site, e.g., `http://en.wikipedia.org/wiki/Flickr`

B. Wikipedia, see `http://en.wikipedia.org`

C. Firefox, an open-source browser

D. The online version of the New York Times, see `http://www.nytimes.com`

PROBLEM 17:

The band that Steve Marks, General Counsel for the RIAA saw at the first concert he attended is:

A. Green Day

B. Cheap Trick

C. Duran Duran

D. Joe Jackson

E. Blondie

PROBLEM 18:

The IETF (Internet Engineering Task Force) has created the technical standards that drive the Internet. Which one of the following is **LEAST** likely to be (and in fact is not) a standard created by the IETF?

  A. The protocol used for sending electronic mail.

  B. The protocol used for constructing IP packets.

  C. The Border Gateway Protocol (BGP) by which large systems share routing tables on the Internet.

  D. The opensocial protocol used for social networking sites to share information.

PROBLEM 19:

WWhich one of the following people did we **NOT** mention or discuss during the semester?

  A. Jon Postel who kept the original DNS system in a text file on his own computer.

  B. Bram Cohen who invented and developed Bittorrent.

  C. Linus Torvalds the inventor of Linux with whom Richard Stallman doesn't get along.

  D. Edsger Dijkstra who invented the algorithm used in many routers to help determine the optimal route that packets take across the Internet.

PROBLEM 20:

Which of the following is **NOT** true regarding Zephyr Teachout, who was Howard Dean's Internet Organizing efforts, and her visit to class.

  A. Prof. Teachout was able to do some coding and web design because she taught herself with help rather than receiving formal instruction in a school/technical setting.

  B. Prof. Teachout had almost no experience with the Internet and associated Internet campaign issues before taking on her role with the Dean campaign in 2004.

  C. Prof. Teachout attributes much of the successes that were part of the Dean campaign to the hierarchical organization and structure of the campaign.

  D. Prof. Teachout asked the students in class to close their laptops during her presentation and the subsequent/coincident question and answer session.

PROBLEM 21:

Suppose Duke gives every undergraduate an emergency cell-phone, but half the phones can only receive calls and not make them while the other half can call two student numbers that are hard-wired into the phones when they're distributed to students. No student's number is stored in more than one phone, and every student's number is stored in some phone.

Suppose a student Sam is called at noon, that Sam can also make calls on their phone, that Sam hears a message that takes one minute to hear, that Sam immediately calls the two numbers and repeats the message, and that everyone that gets called immediately picks up, listens, and immediately calls two numbers if they can (remember, half the students can't make any calls.)

Which of the following is/are true regarding this scenario.

  I. It's possible that fewer than 10 people ever hear the message because of the order in which calls are made, e.g., some people that can't make calls could be called early.

 II. It's possible that everyone will have heard the message before 1:00 pm if the phone tree described works perfectly, e.g., the people that can't make calls are called "last".

III. It's possible that the refusal of only one person who can make calls to make the calls (everyone else called who can make calls does so) will cause more than 90% of the students at Duke to not get the message

  A. I only

  B. II only

  C. I and II only

  D. I, II, and III

PROBLEM 22:

Which of the following is **NOT** a property of the p2p protocol Bittorrent and the p2p system based on Bittorrent.

  A. After a file is first introduced into the Bittorrent system subsequent users don't obtain a complete file from one place, rather the file is obtained in pieces from several places.

  B. The Bittorrent protocol was designed so that clients had to upload parts of files in order to be able to download files, e.g., sharing was encouraged.

  C. Some universities and other ISPs look for Bittorrent traffic by examining IP headers and may take action, e.g., slow down such packets.

  D. Bittorrent cannot handle large files (e.g., HD/Blu-ray movies) on its p2p network because re-assembling the pieces of a large file takes too long for typical computers in 2008.

PROBLEM 23:

There are an estimated 6 billion in the world today. The value of 2^{32} is roughly 4 billion. Which reason below is the **worst** explanation/reason of why we'll run out of IPv4 Internet addresses in the next ten years?

A. IP addresses cannot be re-used, once a machine is connected to the Internet the IP address used can be used only for that machine.

B. More devices are being connected to the Internet so addresses are being used at a greater rate.

C. IPv6 is not being deployed as rapidly as many would like.

D. Even with CIDR (classless inter-domain routing) address-allocation, IP addresses aren't allocated efficiently so many addresses are not in regular use.

PROBLEM 24:

Which of the following was **NOT** discussed by Luis Villa during his talk and visit to the class.

A. How Lawrence Lessig's book *Code and Other Laws of Cyberspace* helped motivate him to be interested in open source software and ultimately attend law school.

B. How he was a Cameron Crazie who wore a tie to games for more than three years.

C. How he was an adequate programmer/coder, but not one of the really best undergraduate programmers during his years at Duke.

D. How exciting it was to lecture/talk with a class meeting in B101 of the LSRC because he had not attended class in that room during his time at Duke.

PROBLEM 25:

Which of the following is **NOT** a case of Internet censorship?

A. Turkey asks YouTube to remove videos that are insulting to Ataturk.

B. China stores some Skype messages on servers in China when the messages are tagged as being in violation of Chinese law/policy.

C. Groups in France ask that sites selling Nazi memorabilia in the US via Yahoo! be inaccessible in France.

D. Students in Compsci 82 request that their pictures not be part of the pick-a-student program displaying student photos in class.

PROBLEM 26:

How many more IP addresses can be represented using IPv6 than using IPv4?

A. 2^{32}

B. 2^{64}

C. 2^{96}

D. 2^{128}

PROBLEM 27:

The first Internet worm that spread from machine to machine and slowed down the Internet appeared in 1988. In the Saturday, December 06, 2008 NY Times an article appeared saying

> "Security researchers concede that their efforts are largely an exercise in a game of whack-a-mole because botnets that distribute malware like worms, the programs that can move from computer to computer, are still relatively invisible to commercial antivirus software.
>
> [Gene Spafford, a computer scientist says] 'In many respects, we are probably worse off than we were 20 years ago because all of the money has been devoted to patching the current problem rather than investing in the redesign of our infrastructure.' "

Which of the following is **LEAST** likely as a reason for why more malware and worms are on the Internet today than were 20 years ago.

A. The Internet and the Web host much more e-commerce than was the case 20 years ago so there is more incentive for criminals to launch worms, botnets, and malware today.

B. According to Moore's Law computers double in speed roughly every two years so today they are more than 1,000 times faster than they were 20 years ago.

C. Far more people use the Internet on a regular basis than was the case twenty years ago so there's much more of an incentive for hackers/criminals to spread malware and worms to take over machines.

D. Apple computers were far more widepsread 20 years ago, as a percentage of total-computers online, than they are today and Apple computers are less susceptible to viruses, worms, and malware, so it makes sense that the number of infected computers is increasing rapidly.

PROBLEM 28:

Which of the following is **NOT** true regarding the files known as *cookies* which browsers save on a computer?

A. Some of the data stored in a cookie identifies when the cookie expires – i.e., when the browser will remove the cookie if the user doesn't remove it before the expiration time.

B. Visiting some sites causes so-called "third party" cookies to be saved, these are cookies that come from a site other than the one being visited.

C. Very few sites cause more than one or two cookies to be saved when the site is visited. Most sites only save one cookie that can be used to see when the user was last at the site.

D. Browsers on public computers, e.g., those in Perkins, Bostock, or the Durham public libraries should be set to remove stored cookies when the user logs off to assist with privacy concerns.

PROBLEM 29:

Which of the following is **NOT** a court case related to digital copyright issues?

A. Perfect10 v Amazon and Google for storing thumbnail images of porn stars.

B. MGM v Grokster for violation of copyright over p2p networks.

C. Capital Records v Thomas for sharing audio files over a p2p network.

D. Duke v Elliott Wolf for distribution of copyrighted materials over a campus network.

PROBLEM 30:

Which of the following is **NOT** a characteristic of Moot, aka Chris Poole, the founder of the online bulletin board `4chan.org`?

A. Moot started 4chan to make money and though he hasn't made lots of money, he has a plan to keep it going.

B. Moot didn't tell his parents about 4chan for the first several years of its operation.

C. Moot doesn't allow anonymous posts to 4chan that come via TOR (an anonymous network that keeps online users truly anonymous/private).

D. Moot is not a hardcore programmer/hacker, he's a young adult who started 4chan because he liked Anime and other Japanese things.
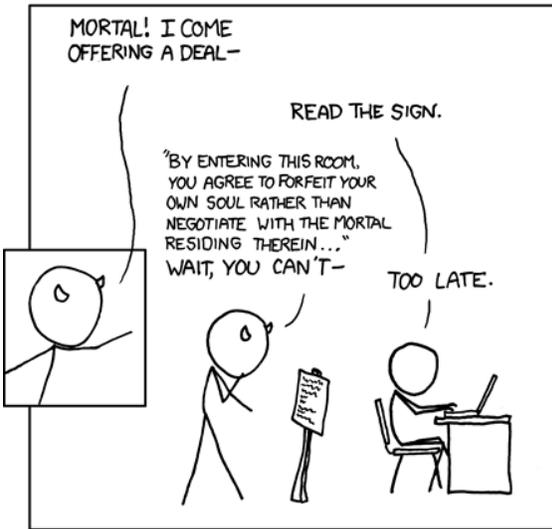

PROBLEM 31:

In class we saw an iPhone application called Shazam that can identify songs being played in a room and *listened to* by the iPhone. We also discussed a company named Audible Magic that offers software and hardware solutions to help identify illegal p2p traffic using deep packet inspection and comparing traffic against a database of copyrighted audio/songs.

Which of the following is **LEAST** likely to be true regarding these applications?

A. They both compare bits/data with a database of stored, copyrighted works – they transmit the bits/data and then receive feedback indicating what (if any) copyrighted work is identified by the bits/data.

B. The technique of identifying a work by a small digital fingerprint is shared by both applications, but the computer code doing the identification is different in each application.

C. If one million more songs/tracks are added to the databases used by these applications each year, the time it takes to identify a song will increase significantly – to the point that they'll likely have to be rewritten in the next few years.

D. If one of the companies holds a patent on the technique of identifying a song or audio work using a digital fingerprint it's possible that the other company has licensed the patent to be able to offer its service.

This question is about the xkcd comic below on the left. On the right is a picture of *Mephistopheles* from Wikimedia — he's a representative of the devil who according to German legend made a deal with a human named *Faust* who exchanged his soul for knowledge.

Faust 2.0

MORTAL! I COME OFFERING A DEAL—

READ THE SIGN.

"BY ENTERING THIS ROOM, YOU AGREE TO FORFEIT YOUR OWN SOUL RATHER THAN NEGOTIATE WITH THE MORTAL RESIDING THEREIN..."

WAIT, YOU CAN'T—

TOO LATE.

MEPHISTOPHELES ENCOUNTERS THE E.U.L.A.

PROBLEM 32:

Write answers on the answer sheet

Part A (2 points)

Provide a one sentence explanation for the title of the comic: *Faust 2.0*.

Part B (1 point)

On a scale of 0-4, where 4 is hysterically funny and 0 is not at all funny, what is your rating of the comic? You do not need to supply a reason, simply a number between 0 and 4. You must spell the number to receive credit, numerals receive no credit. Correctness of spelling does not count.

Part C (2 points)

Provide a one or two sentence explanation of the caption "Mephistopheles Encounters the E.U.L.A".

PROBLEM 33:

Richard Stallman said the following in 1990—before the Web, before p2p software:

> I believe that all generally useful information should be free. By 'free' I am not referring to price, but rather to the freedom to copy the information and to adapt it to one's own uses... When information is generally useful, redistributing it makes humanity wealthier no matter who is distributing and no matter who is receiving.

Stallman is also known for advocating free software, but "free as in speech, not free as in beer."

Many people will pay to purchase a book that can be downloaded for free—in some cases the book's free download increases sales of the hard-copy form of the book. At the same time, given a free (as in dollars), but illegal download for music, many people will not buy the music, but will download it.

Question: 4 points

In two to five sentences explain why all information is not created equal. Write your answer on the answer sheet.