Solution Set for Exam #2
CPSC 155b, Spring 2003

Please note that some of the following solutions contain more information than you needed to provide to receive full credit. Others only provide examples, rather than an exhaustive list, of correct answers.

1. (a) (i) Internet
      (ii) Telephone network
      (iii) Both
   (b) False
   (c) Browsers ↔ Application layer
      Packet routing ↔ IP
      Reliable byte streams ↔ TCP
      Nonreliable byte streams ↔ UDP
   (d) iii
   (e) ii

2. (a)
   • Advertiser-supported, basic services for individual users
   • Tools, customized portals, and services for corporate customers
   • Fees for "extra" or "premium" services and features for individual users (*e.g.*, extra storage for Yahoo! Mail customers)
   (b) i
   (c) iv
   (d) True
   (e) (The number of points that you get for your answer will depend on how well you justify your opinion.)
       The standard explanation of why portals are something fundamentally new is that the combination of personalization and access to a very broad range of services, from news to email to shopping to entertainment, really isn't available from any single business in the offline world.
       The standard answer to "what is the offline analog of a portal" is "a shopping mall." It's a flawed analogy, because shopping malls aren't personalizable, and portals are used for more than shopping, but there are some basic similarities.

3. (a) "Peer production" is production by users who are peers (*i.e.*, no one is above another in a formal organizational hierarchy) and who self-identify and self-organize by communicating directly with each other. Benkler distinguishes it from two other well-studied ways in which production occurs, namely markets (in which communication is facilitated by prices) and hierarchies (in which communication flows along a chain of command). He claims that peer production is superior to other production modes for Internet-scale information products, because "there are increasing returns to scale of the set of agents permitted to

work with a set of resources in pursuit of projects, and to the set of resources agents are allowed to work with. The unbounded sets of both human capital and information inputs that can be used in peer production capture these economies of scale more effectively than can firms, and to a lesser extent markets, both of which rely on securing access to bounded sets of agents and information inputs to reduce uncertainty about the success of projects."

(b) i

(c) Examples from Benkler's lecture notes include:
- academic research
- the web
- content (*e.g.*, Mars clickworkers)
- Accreditation (*e.g.*, slashdot)
- Distribution (*e.g.*, P2P networks)

Other valid examples will also be given full credit.

(d) Copyleft is a condition often included in an open-source license that requires the licensee to give users all of the same rights the licensee has when redistributing copies. For example, the licensee must allow those to whom he redistributes copies to have access to the source code.

(e) True (For example, IBM makes use of Linux in some of its products and services, and it also sells a large number of traditional software products under traditional licenses.)

(f) iv

4. (a) The "abundance problem" occurs when search-engine users submit "broad-topic" queries (*e.g.*, "cars," "computers," or "Afghanistan"); the crux of it is that the number of web pages that could reasonably be considered relevant is far too large for a human user to digest. Its opposite is the "scarcity problem." This occurs when a user submits a very specific query (*e.g.*, "Netscape support for the JDK 1.1 code-signing API"); there are very few relevant web pages, and it is difficult for search engines to find them. (The scarcity problem is not part of the exam question, and no definition of it is needed for full credit. The definition is included here just for your information.)

(b) False

(c) iii

(d) Web form, Index Server, Doc Server, Page Rank

(e) i

(f) True

5. (a) False

(b) The TCB ("trusted computing base") is the part of the computing environment that has to be trustworthy—if there is a successful attack on the TCB, then no guarantees about the correctness or security of the rest of the computing environment can be made. An "attestable TCB" is one that can prove (*e.g.*, by supplying digitally signed statements) to a remote machine that it is running on the local machine and that the required properties of the local computing environment hold.

(c) Correct answers include:
- Board seat(s)
- Sale, acquisition, or merger approval
- Budget approval
- Approval of changes to strategy or business plan
- Approval of executive removal or appointment
- Right of first refusal on sale of shares
- "Tag-along rights" (follow founder sale on a *pro rata* basis)
- "Drag-along rights" (force sale of company)
- Liquidation preference
- Non-compete conditions on founders
- Right to participate in subsequent rounds
- Anti-dilution protection

(d) Foundation documents are the government-issued basic identity documents (including birth certificates, passports, and social security cards) that are relied upon to bootstrap identity and authentication systems in diverse contexts. They are problematic, because they are issued by many government agencies, *e.g.*, all 50 states, and no clear, consistent, and effective set of definitions, rules, and information-security practices governs their issuance and maintenance. For example, drivers' licenses and passports may depend on each other for reissuance and replacement. Lack of well understood, secure handling of foundation documents creates risk of fraud and identity theft.

(e) XrML

6. (a) iv
   (b) iii
   (c) layering
   (d) advertising (Note that there was a typo in this question: The word "sources" was supposed to be "source." You will get full credit as long as your answer includes "advertising.")
   (e) DRM technology has been developed because distributors of copyrighted works want to maintain control over their digital content after it has been distributed to "untrusted" users. In web services, we have an analogous situation: Users would like to maintain control over their personal data after it has been transferred to "untrusted" servers.

7. (a) An open mail relay is an SMTP server that sends mail when the sender (or the machine that makes the request to send mail) and the recipient are both outside the domain of the SMTP server. For example, if a mail server at Yale accepted a request from a machine at Harvard to send mail to a recipient at Princeton, Yale's mail server would be an open mail relay.
   (b) A CAPTCHA is an automated test to tell humans and computers apart, *i.e.*, to distinguish whether the test-taker is a computer acting alone or an actual person. CAPTCHAs often require the test-taker to perform some task that humans can do easily but computers find difficult, such as image processing.

(c) Companies like Yahoo attempt to use CAPTCHAs to reduce e-mail Spam by preventing the automated creation of webmail accounts that can be used to send massive amounts of e-mail. Using these accounts for spam somewhat hides the identity of the originator and offloads the cost of sending mail onto the webmail service. Yahoo includes a CAPTCHA as part of the account registration process so that webmail accounts cannot be created without human effort.

(d) iii

(e)   i) e

       ii) d

       iii) b