Homework 4

This assignment must be submitted **electronically** using the Yale Classes server by **5pm EST, Tuesday, April 1, 2003**. It covers readings and lectures through Thursday, March 27, 2003. Late homeworks will not be accepted.

Please be sure to type your name and e-mail address at the top of your homework assignment. In addition, please include your NetID and "hw4" in the name of your file, *e.g.*, **`vr48-hw4.doc`**. Detailed instructions for submitting your homework are available on the course webpage at http://zoo.cs.yale.edu/classes/cs155/spr03/submit.html.

1. WEB SEARCHING AND GOOGLE (25 points)

A (5 points). Search engines make extensive use of *crawlers* (or *spiders*) to find and index available web pages. Give one reason that a web-page owner may not want a crawler to access his or her site.

B (5 points). Google makes money from advertising (*e.g.*, sponsored links) and from customizable search tools for large organizational web sites. Give an example of another company covered in class whose current business model is similar.

C (10 points). Recall that Google's "PageRank" technology is closely related to the "hub-and-authority" framework put forth by Kleinberg. What does it mean to say that a web page is a "hub," what does it mean to say that one is an "authority," and why are the concepts of hub and authority "mutually reinforcing?"

D (5 points). Why is Google able to deliver a web page in response to a query even if that page is currently inaccessible (*e.g.*, because a server or a link is down)?

2. WEB SERVICES I (20 points)

Recall that the promise of the web-service approach to Internet commerce is that "every data exchange is potentially a revenue opportunity."

Give an example of something you think might be a useful and profitable web service. Do *not* use a stock ticker as the underlying source of publicly available data.

Your answer should briefly address the following questions about the service you are proposing: (1) What is the underlying source of data? (2) What is the value proposition to the customer? (3) What is the revenue model?

3. WEB SERVICES II (30 points)

A (5 points). Which of the following security and privacy problems might arise in web services?

(i) identity theft
(ii) misuse of transaction data
(iii) denial of service
(iv) all of the above

B (5 points). Recall that the four layers of the Microsoft .NET framework are:

Devices (*e.g.*, PCs and cell phones)
User experiences (*e.g.*, AOL and MSN)
Web services (*e.g.*, authentication and data delivery)
Servers (*e.g.*, data repositories)

These are in fact four sub-layers of one of four basic Internet layers. Which one?

C (5 points). What is "screen scrapping," and why is it insufficient as a basis for web-service development?

D (5 points). Which Internet standard for information exchange that we covered in class provides the technical foundation for other web-services protocol standards?

E (10 points). Recall that the "Passport identity service" is one of the "Foundational Services" in the .NET framework. Briefly explain one potential advantage and one potential disadvantage for individual users and for Internet commerce generally of Microsoft's providing an "identity service."

4. XrML LICENSES (25 points)

Recall that XrML 2.0 is a rights-management language that defines a standard for expressing the conditions and rights for using resources. XrML defines a hierarchical set of objects (corresponding to XML elements) that represent the different components in a description of rights.

The high-level picture is as follows. A content publisher creates several resources, *e.g.*, eBooks or music files. To maintain control over the distribution and use of the content, the content publisher distributes a *license.* This is a document that *grants* specific *rights* for the use of *resources*. A license contains one or more grants, and each grant indicates the resource, the right granted, and the person to whom the right is granted. Here is a pictorial description of the components of a license.

license contains:
- **title** (a human-readable description of the document)
- **issuer** (a description of the license issuer)
- one or more **grant**s, each of which contains:
  - **principal** (the person to whom the right is granted)
  - **right** (the right granted)
  - **resource** (the relevant resource)
  - **condition** (the conditions under which the right applies)

Consider the following basic example of a license issued by a recording studio for a digital music file. The license describes that a consumer, Mallory, has the right to play the file “music.mp3” under the condition that she uses the studio’s own software program, “OMSRCplayer.exe,” to play the file. The license would look like this:

```
license =
    title = “Only-my-software Recording Company License”
    issuer = “Only-my-software Recording Company”
    grant =
        principal = “Mallory”
        right = play
        resource = music.mp3
        condition = “Helper program = OMSRCplayer.exe”
```

You can see how the different components of the license express its intent. In reality, these components are written in XML and the boldface words represent XML elements. To avoid the complexities of the actual XrML specification in this assignment, we will just use the form above for licenses. When components contain sub-components, those sub-components will be indented (*e.g.*, **title** is a sub-component of **license**).

Consider the following more complicated example of a license. Suppose a movie studio releases a DVD with a movie and some special features. The movie can always be played, but the special features can be accessed: (1) free, 5 times over 3 months; or (2) an unlimited number of times, after paying a fee. The license might look like this:

```
license =
    title = “Oscar-Winning DVD License”
    issuer = “Hollywood Movie Studio”
    grant =
        principal = anyone
        right = play
        resource = movie
        condition = none
    grant =
        principal = anyone
        right = play
        resource = special features
        condition = fee paid
    grant =
        principal = anyone
        right = play
        resource = special features
        condition = allConditions =
            condition = “valid not after” (purchase date + 3 months)
            condition = not viewed more than five times
```

In the third grant, the “valid not after” construction looks awkward, but it parallels the XML construction for the condition where a grant is valid for a specific time interval. (There is also a “valid not before” construction.) Also note that conditions can be “grouped together” using the **allConditions** element. For the **allConditions** condition to be true and the grant to be valid, all of the sub-conditions must be satisfied at the same time.

This license contains three grants. The first grant is for the *movie* resource. The second and third grants are both for the *play* right to the *special features* resource, but they have different conditions. This is how the license can implement an "or" between conditions. When a license contains multiple grants, any one of those grants can be used to give rights to the specified principal under the conditions of the grant. The second grant is only valid after a fee has been paid; the third grant is valid within a certain time interval. The principal can *play* the *special features* if he is granted that right by one of these two grants, and he can be granted that right only if the conditions of that grant are valid.

A (5 points). Using the form in the examples above, construct a license for the following scenario: Bob obtains a software package from XYZ, Inc. XYZ licenses Bob to *install* the *software* only once, *use* the *software* freely for 30 days, *use* the *software* after paying a small fee, and *view* the *tutorial* files freely. (Used the verbs *use* and *install* as rights, the nouns *software* and *tutorial* as resources.)

One special type of right is *issue*, which allows the principal to issue licenses to other people. This is implemented in the form above by setting the **resource** entry in an issue grant to be the new **grant** that can be given to someone else. That **grant** follows the same form as the **grant** components above (and, in particular, its **right** component can be *issue*, allowing that person to issue a license to yet another person).

For example, suppose that an eBook publisher gives a license to a distributor. The distributor is allowed to give eBooks to consumers so that they are permitted to (1) read the eBook and (2) have the eBook read aloud by a computer, under the condition that a royalty fee is paid to the publisher by the distributor when he grants the license to the user. The publisher's license might look like this:

```
license =
        title = "eBook Distribution License"
        issuer = "eBook Publishers, Inc."
        grant =
                principal = Distributor
                right = issue
                resource = grantGroup =
                        principal =  anyone
                        grant =
                                right = view
                                resource = "ebook.ebk"
                                condition = none
                        grant =
                                right = read aloud
                                resource = "ebook.ebk"
                                condition = uses approved eBook-to-Speech software
                condition = pay royalty fee
```

Just as conditions can be grouped together using **allConditions**, grants can be grouped together using **grantGroup**. When a **grantGroup** is specified as the **resource** for an *issue* grant, the license issued must contain all the sub-grants contained in the **grantGroup** (*i.e.*, the principal cannot just issue one grant in the set). The **principal** component of the **grantGroup** applies to all of the sub-grants; the individual **grant** components do not need the **principal** repeated.

Notice that the “pay royalty fee” condition is tied to the *issue* right, not the **grantGroup**. That is because paying a royalty fee is a condition that must occur for the *issue* to take place; the condition applies to the distributor (note the **principal** component of the grant with the “pay royalty fee” **condition**).

With the above license, it’s possible for Alice to obtain a license like the following from the distributor, once the distributor pays a royalty to the publisher:

```
license =
        title = “use an eBook”
        issuer = Distributor
        grantGroup =
                principal = Alice
                grant =
                        right = view
                        resource = “ebook.ebk”
                        condition = none
                grant =
                        right = read aloud
                        resource = “ebook.ebk”
                        condition = uses approved eBook-to-Speech software
```

Here, a **grantGroup** has been substituted for **grant** (which is legal). Note that the **grantGroup** matches exactly the **resource** in the publisher’s license to the distributor.

B (5 points). Write a license for the following scenario. A computer game manufacturer wants to give its *game* to distributor. The game manufacturer wants the distributor to be able to *play* the *game* freely and issue two types of licenses to consumers. “Basic” consumer licenses can be issued until August 1 for a $20 fee to the manufacturer and after August 1 for a $15 fee. “Basic” consumer licenses permit users to *play* the *game* freely. “Advanced” consumer licenses can be issued until August 1 for a $50 fee to the manufacturer and after August 1 for a $30 fee. “Advanced” licenses permit users to *play* the *game* freely, *modify* the *game*, and *print* the *maps* included with the game. (Use *play, modify,* and *print* as rights; use *game* and *maps* as resources.)

C (15 points). Consider the license on the following page.

(i) *(5 points)* The license gives the Distributor an *issue* right. Construct a license that the Distributor can give to another principal based on this right.

(ii) *(5 points)* The license you constructed in part (i) also contains an *issue* right. Construct a license that could be issued by someone that holds the license in part (i).

(iii) *(5 points)* The licenses you wrote in part (i) and (ii) both permit viewing the encyclopedia. Informally describe the differences between the licenses: From whom can they obtained, under what conditions are they issued, and under what conditions can the encyclopedia be viewed?

```
license =
    title = “Superdistribution license”
    issuer = “Digital encyclopedia publisher”
    grant =
        principal = Distributor
        right = issue
        resource = grantGroup =
            principal = anyone
            grant =
                right = view
                resource = encyclopedia
                condition = none
            grant =
                right = issue
                resource = grant =
                    principal = anyone
                    right = view
                    resource  = encyclopedia
                    condition = “valid not after” 12/31/2003
                condition = pay a $5 fee to publisher
        condition = none
```