

INF 551 – Spring 2018

Quiz 4: File format (10 points), 15 minutes

1. [6 points] The Unicode code point for the math symbol '∈' (meaning "is an element of") is U+2208. Derive its UTF-8 encoding in **both** binary and hexadecimal formats.

$(2208)_{\text{hex}} = (0010\ 0010\ 0000\ 1000)_{\text{binary}}$

Binary: 11100010 10001000 10001000

Hexadecimal: E2 88 88

2. [4 points] Consider the following XML document shown in class. Write an XPath for each of the following questions.

```
<?xml version="1.0"?>
<bib>
  <cd>abc</cd>
  <book>
    <publisher>Addison-Wesley</publisher>
    <author>Serge Abiteboul</author>
    <author>
      <first-name>Rick</first-name>
      <last-name>Hull</last-name>
    </author>
    <author age="20">Victor Vianu</author>
    <title>Foundations of Databases</title>
    <year>1995</year>
    <price>38.8</price>
  </book>
  <book price="55">
    <publisher>Freeman</publisher>
    <author>Jeffrey D. Ullman</author>
    <title>Principles of Database and Knowledge Base Systems</title>
    <year>1998</year>
  </book>
  <book>
    <title>xyz</title>
    <author/>
  </book>
</bib>
```

- a. [2 points] Find the titles of the books published after 1995.

`/bib/book[year > 1995]/title`

- b. [2 points] Find the titles of the books written by someone at the age of 20.

`/bib/book[author/@age = 20]/title`