

MAS DSE 250 Beyond Relational Data Models

Homework 2

Due Thu Nov 2, 2017

This homework is a programming assignment.

Please turn it in *electronically in form of a single text file containing your queries. Stay tuned for directions on how to submit via github.*

In this assignment, we will develop some XQuery queries over the XML-annotated version of Shakespeare's play "Julius Caesar".

The document can be found at

https://mas-dse.github.io/DSE250/xml/j_caesar.xml.

It conforms to the dtd available at

<https://mas-dse.github.io/DSE250/xml/play.dtd>.

We will use an open-sourced XML engine called BaseX, which includes a W3C compatible XQuery processor. It is available for download from <http://basex.org>.

1.

Find out which person, and in what act, uttered the line "Et tu, Brute! Then fall, Caesar."

Your query should return an XML fragment conforming to the following DTD:

```
<!DOCTYPE result [  
    <!ELEMENT result    answer*>  
    <!ELEMENT answer    (who, when)>  
    <!ELEMENT who        #PCDATA>  
    <!ELEMENT when        #PCDATA>  
>
```

Intuitively, the text contents of the 'who' element should contain the text of the name of the person (not a nested element!), and the text contents of the 'when' element should contain the text contents of the act title.

2.

Find all speaking characters (those who have at least one line in the play).

The result should conform to the DTD

```
<!DOCTYPE speakers [  
    <!ELEMENT speakers    character*>  
    <!ELEMENT character    #PCDATA>  
>
```

3.

For each speaking character in the play, find all acts the character speaks in.

The result should have the following structure:

```
<!DOCTYPE result [  
    <!ELEMENT result    answer*>  
    <!ELEMENT answer    (who, when*)>  
    <!ELEMENT who        #PCDATA>  
    <!ELEMENT when        #PCDATA>  
>
```

4.

Find the characters who speak in every act.

Your query should not hardcode knowledge of the number of acts. That is, the query should work correctly on any XML-ized play conforming to the DTD play.dtd, not just for “Julius Caesar”. The result should have the following structure:

```
<!DOCTYPE result [  
    <!ELEMENT result    character*>  
    <!ELEMENT character    #PCDATA>  
>
```