Computer Science 4411b/9538b

Assignment 2

## Due: Mar. 13, 2017

Please use on-line submission through owl.

Note: The data may be altered/corrected/added to before Feb. 14th.

**Purpose:**

The purpose of this assignment is to gain some experience with the XML languages XPath and XQuery.

**Rules/Background:**   
  
You are to work alone on this assignment.

Note that there is a tutorial on XPath and XQuery available in the resources on owl.

Galax is available on obelix in /usr/local/Galax/bin. Galax information is also found on <http://galax.sourceforge.net>. Galax is installed on the gaul network, and also on the linux machines: lxg-001.gaul.csd.uwo.ca through lxg-024.gaul.csd.uwo.ca, hack.gaul.csd.uwo.ca, joust.gaul.csd.uwo.ca, pong.gaul.csd.uwo.ca, rogue.gaul.csd.uwo.ca, tron.gaul.csd.uwo.ca, and zork.gaul.csd.uwo.ca. Students connecting from home should SSH to obelix.gaul.csd.uwo.ca. You can either work directly on obelix, or you can SSH from there to one of the MC 10 systems, if you prefer to work on Linux.

You can also download Altova xmlspy (for 30 days) from [www.altova.com](http://www.altova.com), or some other XQuery engine (listed on <http://www.w3.org/XML/Query/> scroll down to the bottom), in consultation with Dr. Osborn.

**Part 0, Schema and Data Files**:  
I have prepared a file of authors and papers and a context file for Galax.  The context file has a .txt extension so that IE will display it properly.  Change it to .xq when you store it on Gaul. It should be stored in the same directory as the two .xml files and where you are executing the queries from. These questions also make use of a factbook about countries of the world from around 2002. The data files are all available from the owl site.

# Part 1, XPath

For these questions, you need to save the XPath and a copy of the answer in files.

1. Give an XPath query and its answer to show the names of countries in the factbook which contain a city named “Victoria”.
2. Give an XPath query and its answer to list the names of the countries in the factbook where Seventh-Day Adventist is one of the religions, with a percentage >=5%.
3. Give an XPath query and its answer to count the number of countries in the factbook.
4. Give an XPath query and its answer to list the titles of papers which were published after 2014.
5. Give an XPath query and its answer to show the titles of papers which contain the word “Processing” (with either upper or lower case “p”).
6. Give an XPath query and its answer to show the locations of authors of papers which were published after 2014.

**Part 2, XQuery**

Using FLWOR queries, show XQuery expressions and their answers for the following:

1. Give a FLWOR query and its answer to list the countries in the factbook, and their population, ordered by country name, for countries where Seventh-Day Adventist is one of the religions with any percentage.
2. Give a FLWOR query and its answer to list the names of countries in Africa, and, nested within that, their ethnic groups as subelements, ordered by country name.
3. Give a FLWOR query and its answer to list names of countries in Europe, with a population less than 100,000. Give your answer as pairs with the country name and population. Make your answer a well-formed document.
4. Give a FLWOR query and its answer to show the First and Last names of authors in the authors file with a last name starting with “C”. Order the answer by Last name.
5. Give a FLWOR query and its answer to show the year and number of publications for that year, ordered by year of publication.
6. Give a FLWOR query and its answer to show the author names and locations of all the authors of papers whose title contains the word "processing" or “Processing”. Order your answer by author’s last names.
7. Give a FLWOR query and its answer to show the titles of publications in which one of the authors is at Facebook, ordered by year of publication.
8. Give a FLWOR query to give the titles of papers for which ALL of the authors are at the same location.
9. Give a FLWOR query and its answer which produces a well-formed XML document, containing each of the religions in the factbook file, and then for each religion, a sublist of all the countries (just their names) which contain that religions element.
10. Is the XQuery system you are using closed? Show why or why not with an example where an intermediate result which is not well formed is saved to a file and then try processing it with a subsequent query.

**What to hand in:**

* For each question, save the query as a file and the answer as a file, suitably named, e.g. path1query, path1ans or something like that.
* Prepare a Readme file which explains which XQuery engine you used to complete the assignment, and any other comments you think are necessary.
* Zip all of these files and submit them to the owl site, as an attachment. Students in 9538 should submit them to the 9538 owl site.