COMP5323 – WebDB Assignment 1 (Spring 2012)

Question 1
*** It depends how the XML schema is designed. Below are just examples.

http://www.w3.org/TR/xquery/
http://www.w3.org/TR/xpath/
http://www.w3schools.com/xpath/xpath_functions.asp

a) for $x in db2-fn:xmlcolumn('xml_table.xml_column')/Form/Project/Leader
WHERE $x/Department = "COMP" AND $x/ModeOfStudy = "Full time"
RETURN $x/Name;

.//Leader[ModeOfStudy="Full time"][Department="COMP"]/Name

b) let $x := db2-fn:xmlcolumn('xml_table.xml_column')/Form/Project/Leader
for $i in db2-fn:xmlcolumn('xml_table.xml_column')/Form/Project/Member
WHERE $x/Name = $i/Name RETURN COUNT($x/Name);

No xpath as nested expression is not supported.

c) for $x in db2-fn:xmlcolumn('xml_table.xml_column')/Form/Project
WHERE $x/Activity/ActivityDate = date("2012-06-20")
RETURN $x/ProjectName;

.//Project[Activity/ActivityDate = "2012-06-20"]/NameOfProject

d) let $x := db2-fn:xmlcolumn
('xml_table.xml_column')/Form/Project[Activity/ActivityDate = date("2012-06-20")]
for $i in db2-fn:xmlcolumn
('xml_table.xml_column')/Form/Project[Activity/ActivityDate = date("2012-06-20")]
WHERE $x/Activity/Place = $i/Activity/Place
RETURN <Projec> {($x/NameOfProject), ($i/NameOfProject)} </Project>;

No xpath as nested expression is not supported.

e) It depends whether or not the "Total_Expenditure" is stored. If yes, it is simply to check the amount whether or not it is over HK$30,000 and the number of activities is les less than 3. If not, it is first to check the sum of the "Amount_Applied" whether it is over HK$30,000 or not, and then to check the number of activities is less than 3.

for $x in db2-fn:xmlcolumn('xml_table.xml_column')/Form/Project[Budget/
Total_Expenditure > 30000]
WHERE count($x/Activity/ActivityItem) > 3
RETURN <Project> {($x/NameOfProject)} </Project>
<NumberOfMember> {(count($x/Member)) </NumberOfMember>;

It is possible to have xpath for "Total_Expenditure" is used.

.//Project[Budget/Total_Expenditure > 30000][count(Activity/ActivityItem) > 3]
/NameOfProject/count(Member)

f) let $i := db2-fn:xmlcolumn('xml_table.xml_column')/Form/Project
RETURN
<Group_Below_100>
{ for $x1 in $i
WHERE $x1/Benefits/ServiceUsers <= 100
RETURN <Project> {($x1/NameOfProject), ($x1/Budget/Total_Expenditure div
$x1/Beneifits/ServiceUsers)} </Project>;
}
</Group_Below_100>
<Group_101_to_1000>
{ for $x2 in $i
WHERE ($x2/Benefits/ServiceUsers > 100) AND ($x2/Benefits/ServiceUsers <= 1000)
RETURN <Project> {($x2)} </Project>;
}
</Group_101_to_1000>
<Group_1000_More>
{ for $x3 in $i
WHERE $x3/Benefits/ServiceUsers > 1000
RETURN <Project> {($x3)} </Project>;
}
</Group_1000_More>

No xpath as nested expression is not supported.

Question 2

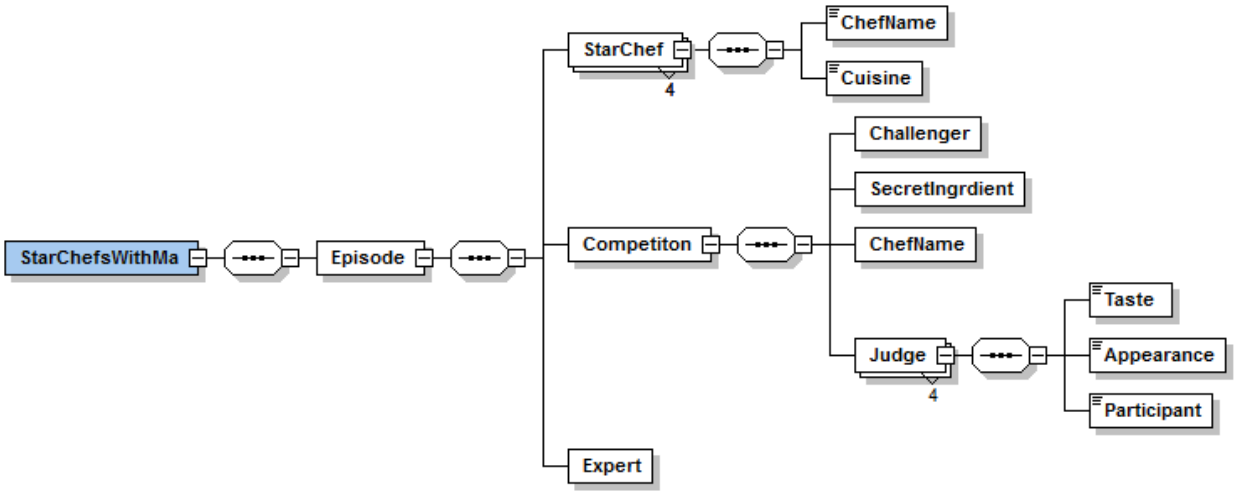
***Below is just an example for reference.

a)
```
<!ELEMENT DAG (NODE+, EDGE+)>
    <!ELEMENT NODE (#PCDATA)>
        <!ATTLIST NODE
            NodeID ID #REQUIRED
        >
    <!ELEMENT EDGE (HEAD, TAIL)>
        <!ELEMENT HEAD (#PCDATA)>
            <!ATTLIST HEAD
                HeadNode IDREF #REQUIRED
            >
        <!ELEMENT TAIL (#PCDATA)>
            <!ATTLIST TAIL
                TailNode IDREF #REQUIRED
            >
```

b) It seems it is difficult to determine that the graph contains cycle or not from such a DTD schema. For example, there is no restriction that how many edges are to be used to form a cycle. Also, there is no checking that the HEAD and TAIL nodes are to be the same NODE.

Question 3

a) Just a reference.



b) It is first to find out the list of distinct name of the "Star Chefs". For example,

```
for $i in db2-fn:xmlcolumn('xml_table.xml_column')/StarChefsWithMa/Episode
/StarChef[distinct-value(ChefName)] RETURN $i;
```

Then, count the appearance of the his/her name in each episode. Get the highest one.

```
let $x := db2-fn:xmlcolumn('xml_table.xml_column')/StarChefsWithMa/Episode
/StarChef[distinct-value(ChefName)]
for $i in db2-fn:xmlcolumn('xml_table.xml_column')/StarChefsWithMa/Episode
/StarChef
WHERE $x/ChefName = $i/ChefName
RETURN max(count($i/ChefName));
```

c) It would be better to use existing to explain or illustrate the M:N relationships and discuss about their pros and cons. For example: An episode has more than one "Star Chef". One "Star Chef" can attend more than one episode. Also, it is how the "key" and "keyref" are being used in the schema.

http://www.comp.polyu.edu.hk/~csamak/Ref/xml_m2n.pdf The article discusses the use of "IDREF" and "IDREFS" to eliminate the many-to-many relationship.

Question 4

It would be better to provide examples to show the differences between the two products. How product A is better than product B, for example, product A (100,000 transactions per minute) has a better processing power than product B (10,000 transactions per minute). You should identify which product could have a better performance rather than just to list out all the functions.