Task documentation XTD: XML2DDL in english for IPP 2011/2012
Name and surname: Martin Riša
Login: xrisam00

- **Use of script**

Python script xtd.pl(further only script) transforms xml document into Data Definition Language. Command line argument(further only argument) --*input==file* is if is not set input is standard, same for argument --o*utput==file*. Argument -*a* causes from processing are omitted arguments of every node of input xml. Argument -*b* causes that multiple subnodes of same name will be treated as if there was only one. Argument –*heaher=head* prints head sa first line of output preceded with double *"-"*. Argument *etc=n* works by project settings and is exclusive with argument -*b*.

- **Implementation specification**

1. **internal data representation**

    Xml file is parsed using tool xml.dom.minidom this representation is key data storer and List of Objects of Tables are only summary information gatherer. Object Table constist of TableName and list of triples where first element represents attribute/reference/element name secon data type and third is determines whether element is text node , attribte or reference. Method print2DDL prints object to given output in DDL format.
    Input representation is initalised firstly from list of table names which is get from function *GetAllElementsFromNode* called on root node. Then for each table name is created objed from constructor and is stored in list *ListOfTableObjects.* Afterfwards for each text node, attribute and reference element are used to fill list *ListOfElementsAndTypes* especially with their name and data type

2. **argument -b,**

    causes that multiple elements with same name are treated as if there was only one with highest data type of them. This is done by checking list of child nodes of each node in xml except root. There are also used conversion functions *Get(Attribute/Element)*Type to get actual data type of table column.

3. **argument -a and –header**

    Argument -*a* causes that attribute of xml nodes are omitted from output DDL file. This is done in function which initializes internal data representation done by skipping every attribute subelement of xml node.
    Argument header causes to print commented header string which is value of that argument quoted with single or double quotes.

4. **argument -etc=n**

    causes that if in table A is more than n columns of same table B then table B implicitly contains relation to table A and A loses that realition as well. This is done by annalyzing that number for each node in xml input file. This argument is exclusive with argument -b but he requires to count number of same named subnodes of every node in xml file.