

Données du Web : XML Relationnel

TD en binômes à rendre le 23 octobre.

1) Stockage Monet

1. Implémenter sous Oracle le schéma Monet permettant de stocker un document XML pour les bâtiments que vous avez proposé pour le TP 1.
2. Exprimez en SQL trois requêtes XPath de votre choix sur ce document.
3. Comment pourrait on étendre le schéma Monet pour gérer l'axe de navigation ancestor ?

2) Stockage schema-aware

1. À partir de la DTD proposée pour la presse (voir TP 1), définir le schéma de stockage relationnel associé.
2. Remplissez les tables avec des valeurs correspondants à l'un des documents XML que vous avez proposés lors du TP 1.
3. Traduisez en SQL trois requêtes XPath de votre choix.

3) Interval-encoding avec SAX

1. Donnez l'encodage begin/end de l'un des documents XML pour les bâtiments que vous avez proposé lors du TP 1, puis enregistrez-le dans la table NODE.
2. À l'aide de la classe SaxParser (.java sur Moodle), programmez l'encodage par intervalles begin/end. Cette classe permet de parser un document XML en streaming (noeud après noeud). Modifiez les méthodes startElement et endElement qui sont déclenchées à l'ouverture et fermeture d'un noeud de type element, ainsi que la méthode characters qui est déclenchée lorsque on rencontre un noeud de texte. L'output du programme sera une liste de mises à jour pour la table NODE.

```
INSERT INTO NODE (begin,end,parent,tag,nodtyp) VALUES(1,16,null,'racine','element')
```

3. Testez le programme sur le document XML concernant les Tweets, et reformulez sur la table NODE trois requêtes de votre choix.
4. Testez le programme avec le fichier XML <http://www.ins.cwi.nl/projects/xmark/Assets/standard.gz> et reportez le temps d'exécution.
5. Modifier le programme afin d'implémenter l'encodage pre/post du document.
6. Modifier le programme afin d'implémenter l'encodage Dewey du document.
7. Enfin, modifiez votre programme pour qu'il puisse répondre à des requêtes XPath de la forme `// α` avec α une balise quelconque.
8. Indiquer comment vous devriez étendre votre programme pour supporter n'importe quelle combinaison des axes child et descendant.