

Aide à la formulation de requête : Un outil d'interrogation graphique de requêtes XQuery

L'interrogation de documents semi-structurés se base sur des langages d'interrogation comme XQL, XML-QL ou bien XQuery. Ces langages ont été conçus pour l'interrogation et la recherche d'information dans tout type de sources de données XML, dans lesquelles une interrogation est représentée comme une expression. Cependant, l'interrogation demeure difficile à réaliser pour la majorité des utilisateurs. Une solution d'interrogation graphique peut donc alléger la tâche d'interrogation.

Dans ce document, nous allons détailler notre outil d'aide à la formulation graphique de requête XQuery. Nous présentons les différents étapes d'interrogation en détaillant l'interface utilisée. Cette interface permet d'utiliser les expressions de XPath, de représenter les fonctions définies par l'utilisateur et d'utiliser les opérateurs booléens (or, not, and). Elle combine les avantages du texte et du graphique.

La Figure (1) montre les composants de cette interface.

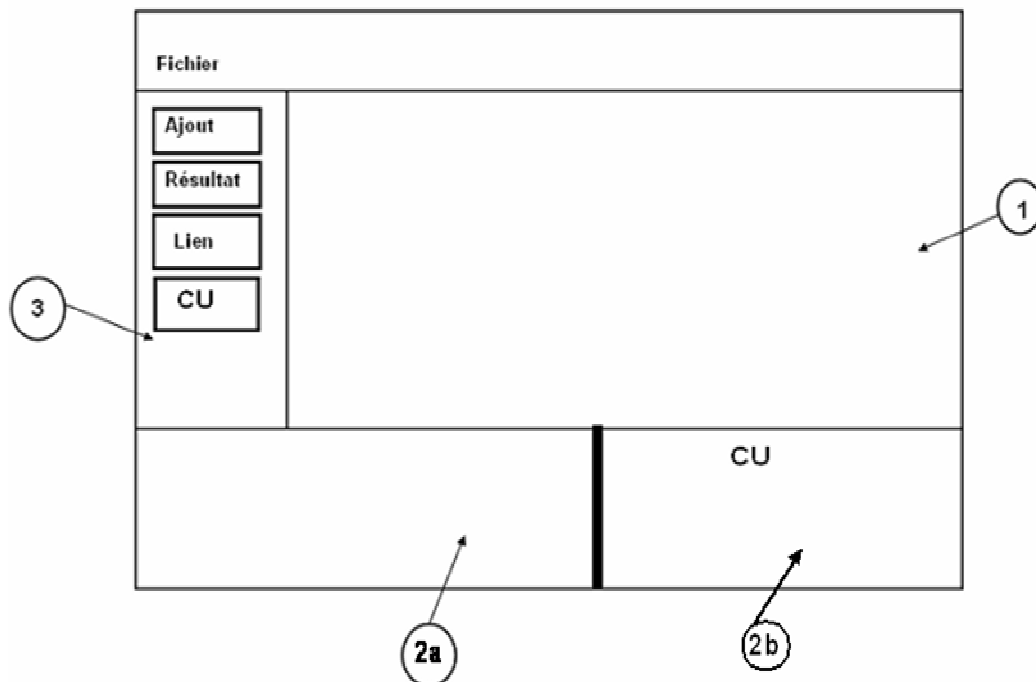


Figure (1) : L'interface graphique

Elle est composée de trois parties :

- ① : La zone graphique. Dans cette zone se construit le graphe qui représente la requête XQuery.
- ② : La zone textuelle. Cette zone se compose de deux parties :
 - ②a) contient le texte de la requête qui est représentée dans la zone graphique,

②b CU (Conditions Utilisateur) : contient le texte des fonctions (définies par l'utilisateur, logique, ...) s'appliquant sur la requête.

③ : La zone de boutons. Cette zone contient quatre boutons :

Le bouton Ajout permet de choisir le document source,

Le bouton Résultat déclenche l'évaluation de la requête et l'affichage du résultat de la requête sous forme graphique,

Le bouton Lien désigne le lien entre les éléments source et les éléments résultat pour montrer la correspondance entre les deux,

Le bouton CU permet de préciser les fonctions définies par l'utilisateur afin de les utiliser dans la CU (Conditions Utilisateur).


Dans la barre de menu, nous trouvons le menu Fichier qui contient les commandes :

Nouveau : pour choisir le document source,


Nouvelle requête : pour définir une nouvelle zone graphique et une nouvelle zone textuelle pour formuler la nouvelle requête.


Les symboles

Pour la formulation graphique de la requête selon notre interface, nous utilisons les symboles suivants.

 représente un élément de documents XML. Désigner un rectangle dans la zone graphique, revient à choisir cet élément avec ses sous éléments.

Cercle : Nous proposons deux types de cercles:

 représente le contenu de l'élément (#PCDATA).

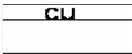
 représente un attribut de l'élément du document XML.

/ : Nous proposons deux types de liens qui ont la même forme :

Le lien (a) : Ce type de lien représente la relation directe entre les éléments. L'élément Y est le descendant direct de l'élément X. Il relie les éléments de document source ou les éléments du résultat.

Le lien (b) : Ce type de lien représente la relation de jointure entre les éléments du document source et les éléments du résultat.

// représente la relation indirecte entre les éléments. L'élément Y est un descendant indirect de l'élément X.

 : Cette fenêtre sert à écrire des expressions logiques, des conditions (if...then...else) et des fonctions définies par l'utilisateur.

 : Pour écrire les fonctions d'agrégations comme avg, sum, ...

Nous allons dans ce qui suit présenter des exemples d'interrogation graphique de requête XQuery en utilisons notre interface.

Pour illustrer nos exemples dans les sections suivantes, nous utilisons la collection de documents XML de la plateforme PRETI (<http://www.irit.fr/PRETI/accueil.fr.php>) comme document source.

Requête simple

" Pour chaque Gite, afficher son numéro et le nombre des chambres " :

```
for $a in document("gite.XML")//GITE
return
  <resultat>
    <GITE>
      {$a/NUMERO}
      {$a // NBCHAMBR}
    </GITE>
  </resutat>
```

Pour formuler cette requête sous forme graphique via notre interface nous suivons les étapes suivantes :

1- Nous sélectionnons le document de source (gite.XML), en activant dans la fenêtre principale le bouton Ajout ou l'ordre Nouveau de la barre de menu.

La figure (2) montre la fenêtre qui permet de choisir le document source :

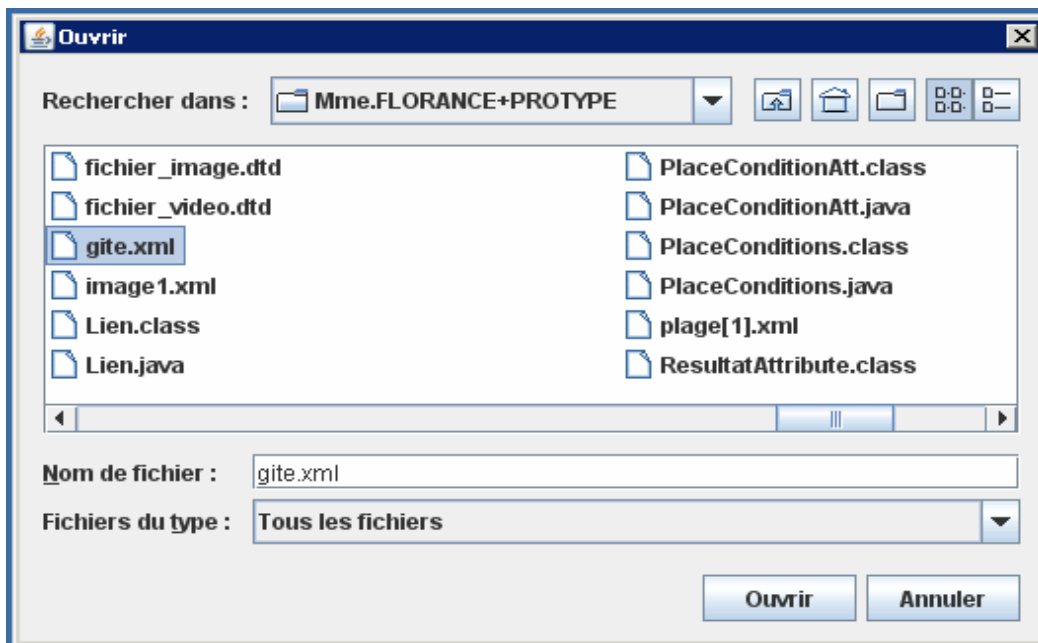


Figure (2) : La fenêtre de choix du document source

Après la sélection du document source, la racine du document portant le nom logique du document dans la zone graphique est représentée par un rectangle. Le nom logique du document est composé de mots "fichier" et le nom de document, comme il apparaît dans la figure (3).

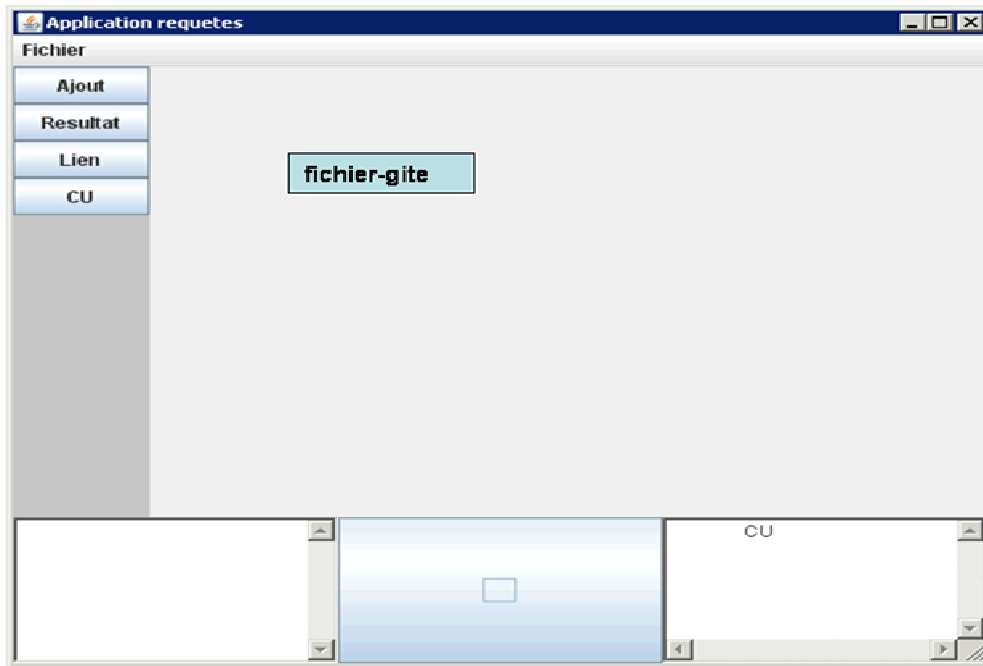


Figure (3) : L'interface graphique

- 2- Nous sélectionnons les éléments de la requête, en faisant un click droit sur le rectangle qui porte le nom logique du document source. Un menu apparaît dans la zone graphique :
- ✓ AjouterElement permet de choisir des éléments du document source.
 - ✓ Conditionnement permet de préciser les conditions de la requête sur les éléments et les attributs

Pour sélectionner les éléments, nous activons l'ordre AjouterElement du menu précédent pour faire apparaître la fenêtre suivante :



Figure (4) : La fenêtre qui permet le choix des éléments

Cette fenêtre contient tous les éléments de document source, nous choisissons les éléments demandés dans la requête puis nous cliquons sur le bouton (Ajouter Element) qui permet de dessiner dans la zone graphique l'élément choisi.

3- Pour choisir les éléments du résultat, nous activons le bouton Resultat situé dans la fenêtre principale.

La figure (5) montre la fenêtre qui permet de choisir les éléments du résultat

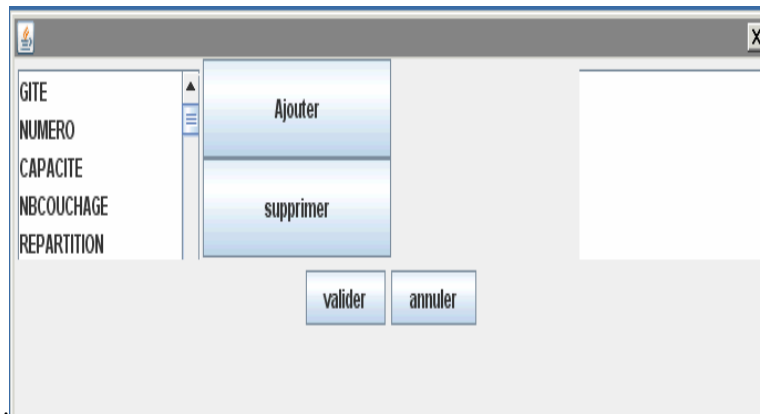


Figure (5) : La fenêtre d'aide au choix des éléments de résultat

3- Pour afficher le lien de jointure entre les éléments de document source et les éléments du résultat, nous activons le bouton Lien situé dans la fenêtre principale.

La figure (6) montre la fenêtre qui permet d'afficher le lien de jointure

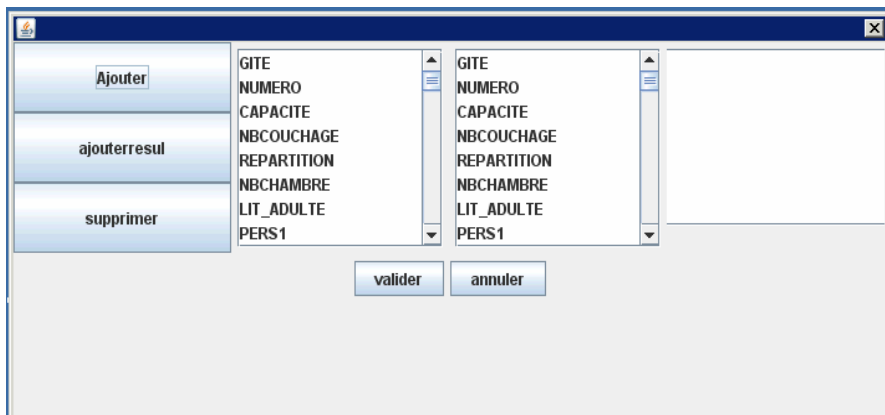


Figure (6) : La fenêtre d'affichage du lien

Cette fenêtre contient les éléments de document source et les éléments de résultat. Nous pouvons préciser les deux éléments qui sont liés.

La figure (7) montre la représentation graphique de cette requête :

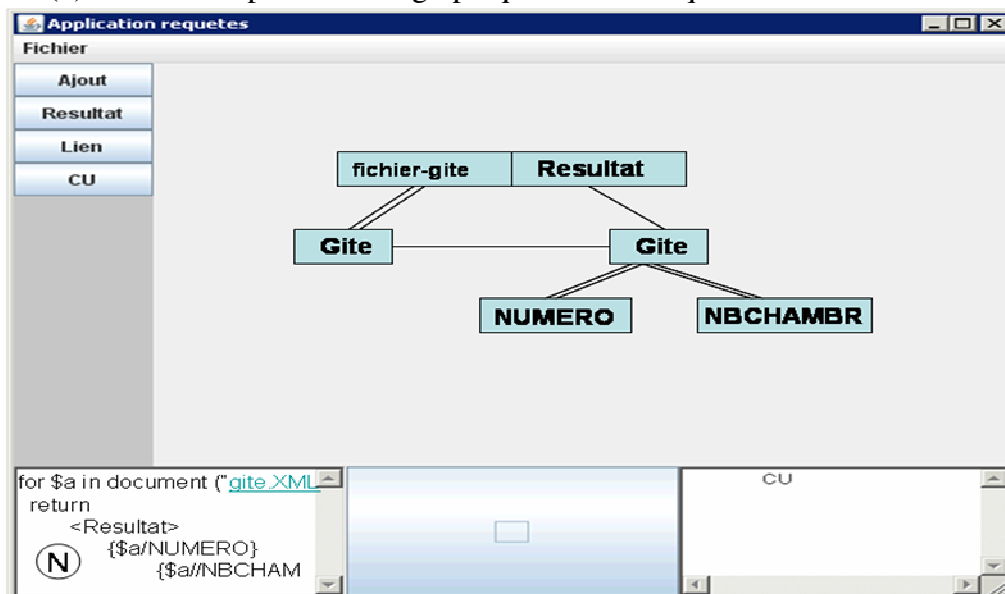


Figure (7) : La présentation graphique de l'exemple de la requête simple

Ainsi, la racine du graphe de notre interface qui présente la requête, est composée de deux parties :

- ✓ La première (filtre) porte le nom "fichier + le nom du document de source". Sous cette partie, nous trouvons des éléments et des attributs de document source qui servent la requête.
- ✓ La deuxième (résultat) porte le nom "Résultat". Sous cette partie nous trouvons les résultats de la requête c'est-à-dire les éléments qui réalisent les conditions de la requête.

Le texte de la requête est généré en XQuery dans la zone textuelle (N) via une écriture progressive qui correspond aux étapes de dessin du graphe de la requête.

L'exemple de requête présenté ici est un exemple simple car aucune condition sur les éléments ou les attributs n'y figure.

Requête avec conditions

" Quels sont les gites contenant 2chambres et dont le prix de la réservation pendant le mois de juin est inférieures de 1200 ? " :

```
for $a in document("gite.XML")/Gite
for $u in $a//NBCHAMBRE
for $t in $a//PRIXP
where
    contains ($u, "2")
and $t [@PERIODE_ID = "juin"]
and contains ($t, "<1200")
return
    <resultat>
        { $a }
</resultat>
```

Nous précisons les conditions des éléments et des attributs dans notre interface, en faisant un click droit sur l'élément pour lequel nous voulons affecter des conditions. Une nouvelle fenêtre s'affiche qui contient les opérations possibles pour les conditions sur les éléments ou sur les attributs.

La figure (8) montre la fenêtre qui permet de définir ces conditions :

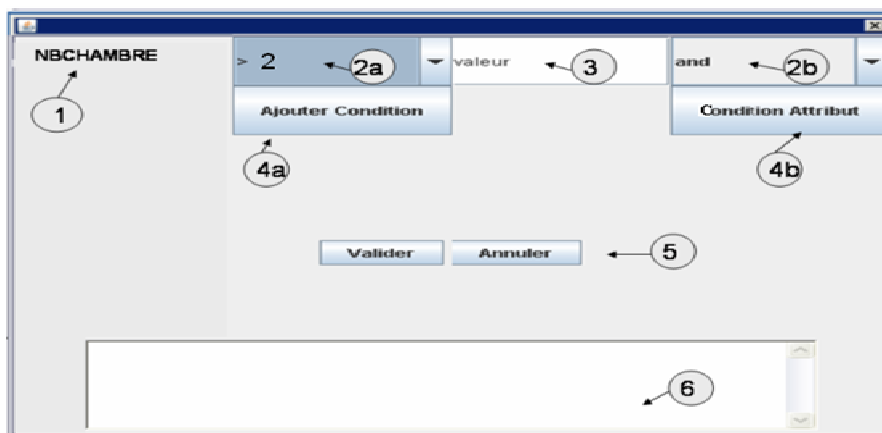


Figure (8) : La fenêtre d'aide pour les conditions sur les éléments

Cette fenêtre est composée de trois parties.

La première partie contient le nom d'élément ^①, deux listes ^{②a} ^{②b}, une zone de texte ^③ et deux boutons ^{④a} ^{④b}.

La liste ^{②a} contient les opérateurs que nous pouvons appliquer sur les éléments (ex : and,...).

La zone de texte ^③ permet de formuler le texte des conditions.

La liste ^{②b} contient les opérations booléennes (and,...), s'il y a plusieurs conditions sur l'élément.

Le bouton AjouterCondition ^{④a}, pour écrire le texte des conditions dans la deuxième partie

Le bouton ConditionAttribut ^{④b}, pour ajouter des conditions sur les attributs de cet élément.

La deuxième partie ^⑤ contient les boutons Valider et Annuler.

La troisième partie ^⑥, est une zone de texte, qui présente le texte des conditions.

Pour ajouter des conditions sur les attributs des éléments, nous devons activer le bouton (Condition Attribut) dans la Figure (8).

La figure (9) montre la représentation graphique de cette requête :

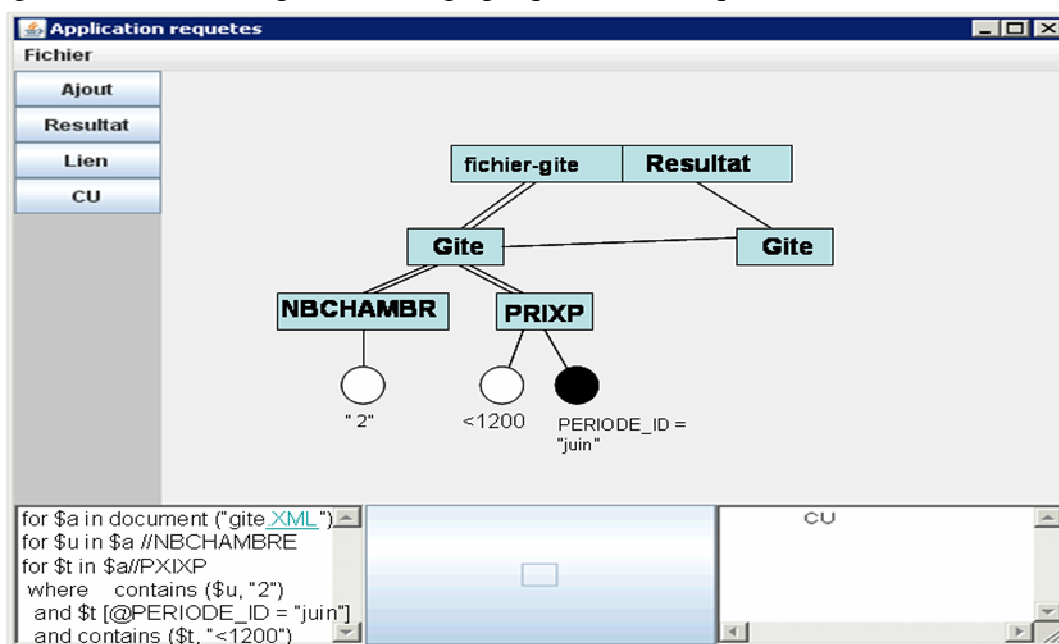


Figure (9) : Représentation graphique de la requête avec conditions

Requête avec les conditions utilisateur

Soit le document :

```
<!DOCTYPE meta_document SYSTEM "meta_document.dtd">
<Meta nom="Meta.XML">
  <fichier_video nom="pres_sig" langue="français" taille="80
    Mo" type="avi">
```

```

<sequence seq_debut="00:00:00" seq_fin="00:01:33">
  <son_cle_video scv="activité.wav"/>
  <son_cle_video scv="équipe.wav"/>
  <scene sce_debut="00:00:00" sce_fin="00:05:00"/>
</sequence>
<sequence seq_debut="00:01:34" seq_fin="00:05:00">
  <mot_cle_video mcv="équipe SIG"/>
  <image_cle_video icv="organisation.gif"/>
  <image_cle_video icv="preti.gif"/>
  <scene sce_debut="00:01:34" sce_fin="00:02:50"/>
  <scene sce_debut="00:02:51" sce_fin="00:05:00"/>
</sequence>

</fichier_video>
< fichier_audio nom="pres_sig" langue="Français" taille="300" type="wav">
  ...
</ fichier_audio >
  ...
</Meta>

```

Document (1) : Exemple de structure de description de document multimédia

"Restituer les séquences de la vidéo "pres-sig.avi" dans lesquelles on entend la voix de "Anis"
":

```

let $au := (for $s1 in doc (Meta.XML)// fichier_audio
  where $s1/ @nom = "pres_sig"
  return $s1)
let $loc:= ( for $g in $au/ segment
  where $q//locuteur//@ nom_loc= "Anis"
  return $q)
let $vi:= ( for $s2 in doc ( Meta.XML)//fichier_vidéo
  where $s2 /@nom= "pres_sig"
  return $s2)
for $segd in $loc, $seqd in $vi//sequence
let $x:= time ( $seqd/@seq_debut)
let $y:= time ( $seqd/@seq_fin)
let $a:= time ( $segd/@seq_debut)
let $b:= time ( $segd/@seq_fin)
let $res as xs: Boolean:=(local: CHEVAUCHE ($a, $b, $x, $y)
  or PENDANT ($a, $b, $x, $y) )
where $res
return
  <resultat>
  {$seqd}
</resultat>

```

Nous remarquons que cette requête contient deux fonctions définies par l'utilisateur CHEVAUCHE et PENDANT permettant respectivement de dire si deux intervalles temporels se chevauchent ou si l'un se déroule pendant l'autre, avec la prise en compte de l'inverse de chacune d'elle. Ces deux fonctions reçoivent en entrée 4 variables (\$a, \$b, \$x, \$y) représentant respectivement les instants de début et de fin du premier et deuxième intervalle. Elles retournent à la requête un résultat de type booléen indiquant 'vrai' si les conditions de la fonction sont vérifiées et 'faux' dans le cas contraire.

Pour présenter ces fonctions dans notre interface, nous activons le bouton CU (Conditions Utilisateur) (1) situé dans la zone de boutons.

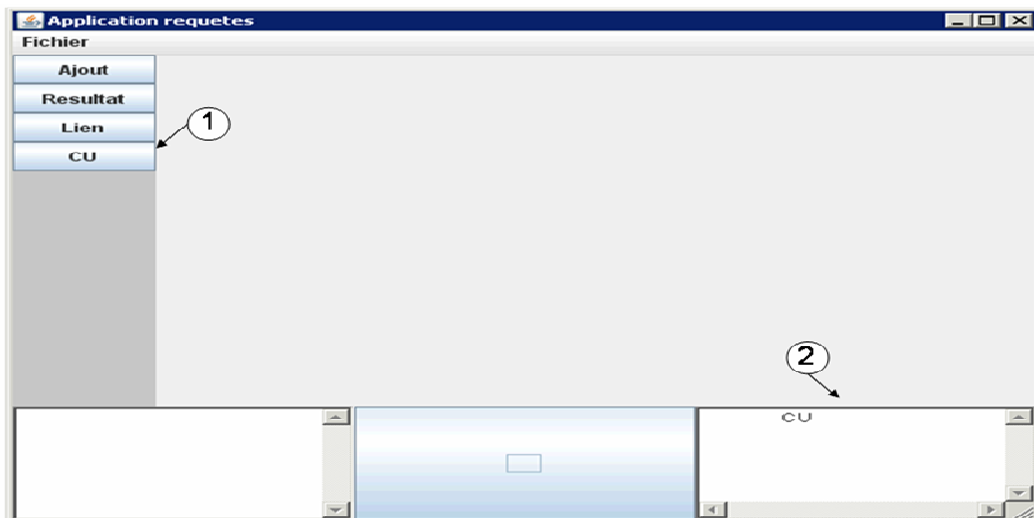


Figure (10) : L'interface graphique

La figure (11) montre la fenêtre qui permet de préciser les fonctions définies par l'utilisateur.

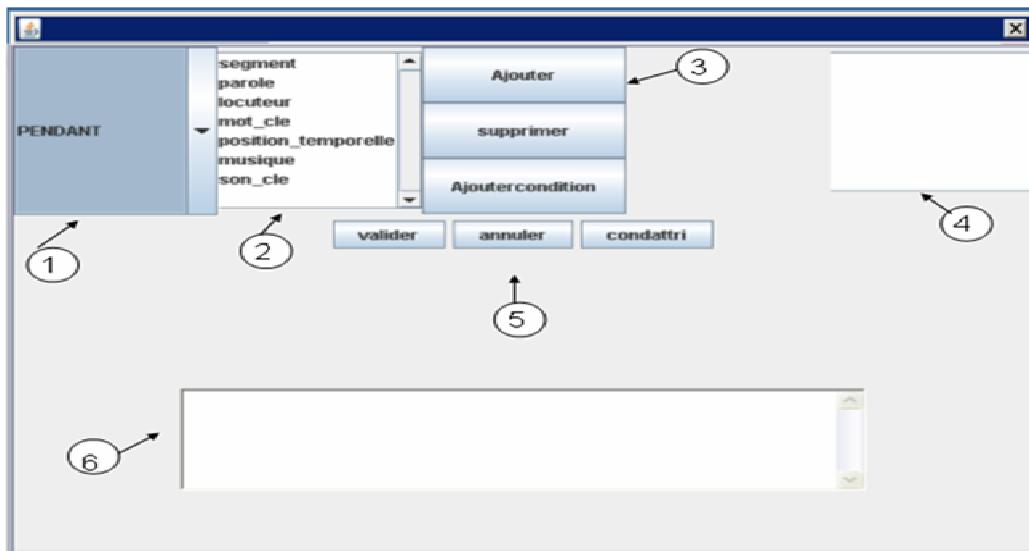


Figure (11) : La fenêtre d'aide à la formulation des conditions définies par l'utilisateur

Cette fenêtre est composée de cinq parties :

- ① : contient les fonctions définies par l'utilisateur.
- ② : contient les noms des éléments de document.
- ③ : contient trois boutons.
 1. Le bouton Ajouter : permet de choisir les éléments qui forment les variables de la fonction et de les ajouter dans ④.
 2. Le bouton supprimer : permet de supprimer des éléments ajoutés à ④.
 3. Le bouton Ajoutercondition : ajoute le nom de fonction et les noms des éléments sélectionnés dans ⑥.
- ④ : Est une zone de texte contenant les noms des éléments où nous appliquons la fonction.
- ⑤ : contient les boutons valider, annuler et condattri. Nous pouvons ajouter les conditions définies par l'utilisateur sur les attributs en activant ce dernier bouton.

⑥ : Est une zone de texte qui contient le texte des conditions.

Quand nous activons le bouton condadtri pour ajouter des conditions sur les attributs, la figure (12) apparaît.

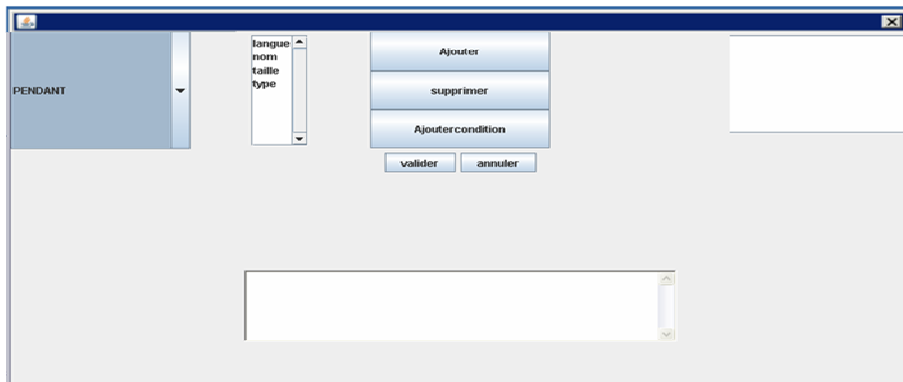


Figure 12 : La fenêtre d'aide à la formulation des conditions définies par l'utilisateur sur les attributs

Nous remarquons que cette fenêtre a les mêmes composants que de la fenêtre précédente (11) pour les attributs.

Après avoir précisé les conditions sur les éléments et les attributs, le texte des conditions apparaît dans la zone de texte (CU) ② (cf. figure (10)).

La représentation graphique de cette requête selon notre interface est :

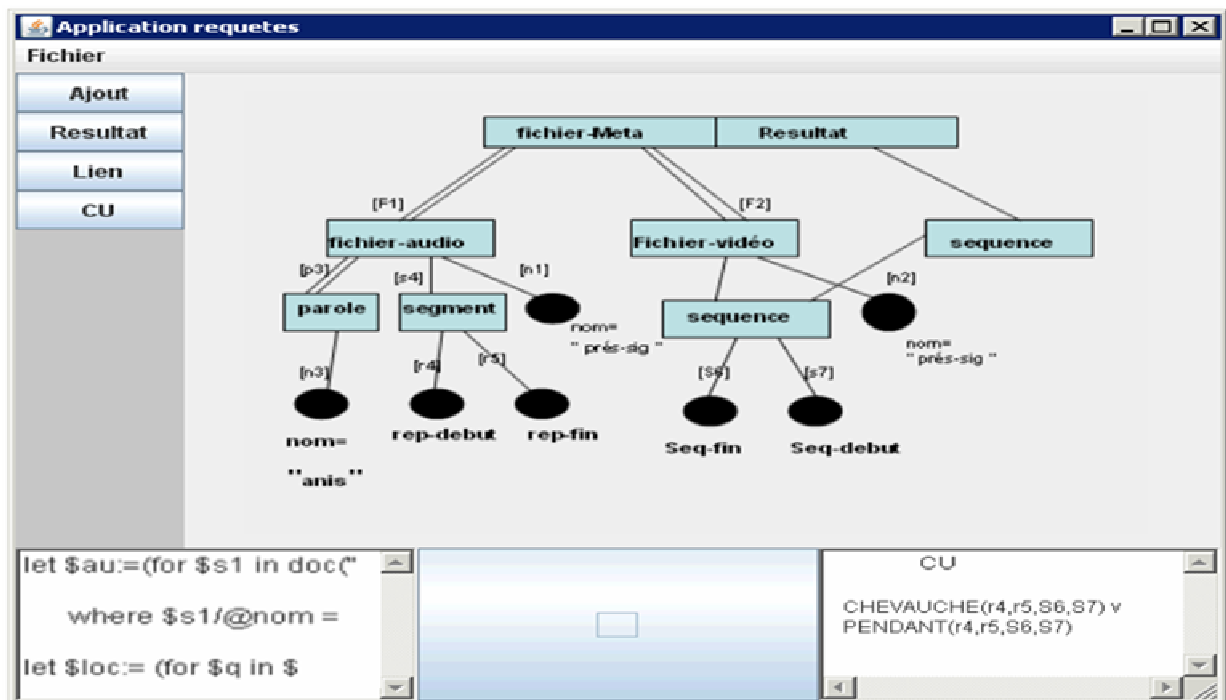


Figure (13) : Représentation graphique de la requête avec les conditions utilisateur

Pour présenter les variables de fonctions dans le graphe de la requête, nous proposons d'ajouter un identifiant à chacun des éléments dans la partie de filtre. Puis dans le texte des fonctions nous précisons les identifiants.