Serveurs de grammaires et de dérivations

Eric de la Clergerie

Projet ATOLL – INRIA

26 Septembre 2001

Réunion ARC RLT

Serveurs

Mise en place de deux serveurs (prototypes)

Serveur de grammaires (Abdelaziz Khajour)

Accès aux grammaires (format XML DTD TAG)

BD MySQL + JServ + Apache (+ Cocoon)

Serveur de grammaires (Vartika Bhandari)

Accès à une banque de forêts de dérivation (format XML DTD Forest)

BD Postgres + Apache + Cocoon

Chacun des deux serveurs s'appuie sur un langage de requêtes qui n'est pas exactement le même.

Langage de requêtes

Forme générale des requêtes

```
[<variable declaration>]
select <path_expression_list>
[where <condition> ]
```

Notion centrale de `path_expression` :

1. `<type>(.<methode>)*` : `forest.op.deriv.anchor`
2. `<var>(.<methode>)*` : `F.op.deriv.anchor`

Condition : `forest.op.deriv.anchor = 'aime'`

Différences

La plus grande différence porte sur la sémantique des chemins :

```
select op.deriv.tree, op.deriv.anchor
```

Grammaire les deux occurrences de `op.deriv` ne dénotent pas le même objet

Dérivation les deux occurrences de `op.deriv` dénote pas le même objet

La sémantique « Dérivation » permet une écriture plus compacte mais demande la déclaration de variable pour des objets distincts :

```
op.deriv D1, D2;
select D1.tree, D2.tree
```

Extensions

Le langage du serveur de dérivation introduit quelques extensions :

- « `select distinct ...` » pour obtenir un ensemble distinct d'objets
- « `select xmlfrag(...)` » pour accèder au contenu XML d'un objet
- « `select count(...)` » pour compter le nombre d'objets vérifiant une propriété
- auto groupement « `select deriv.tree, count(deriv)` » : compte le nombre de dérivation pour chaque arbre apparaissant dans une dérivation.
- tri « `select deriv.tree,count(deriv) order by count(deriv) desc` »

Evolution

– Basculer « Grammaire » vers « Dérivation »

– Corriger les bugs

– Faire une vraie distribution
(pas simple ! Interactions Java, XML, Apache, Postgres, Cocoon)

– Mécanisme de définition de nouvelles méthodes

– Ajout d'un mode édition avec sécurité
(login avec mot de passe + éventuellement backup)