

CADIXE (Projet Caderige : <http://caderige.imag.fr/>)

Manuel d’utilisation (Editeur version 2.0a1 du 23/11/04)

Contact : Gilles.Bisson@imag.fr, Hassiba.Zidelkheir@imag.fr

Objectifs de l’éditeur

Cet éditeur permet d’annoter interactivement tout ou partie d’un document textuel à l’aide de balises XML décrites dans une DTD. Voici quelques sites présentant ces technologies :

- Présentation du XML : <http://www.brics.dk/~amoeller/XML/overview.html>
- Ecriture d’une DTD : http://www.w3schools.com/dtd/dtd_intro.asp

Contrairement à ce qui est couramment proposé dans ce type d’éditeur, l’utilisateur n’a pas à structurer directement le texte sous la forme d’un arbre XML. Il peut introduire les balises une à une dans l’ordre où il le souhaite en « surlignant » la partie du texte qui l’intéresse et en y appliquant la balise de son choix parmi celles disponibles. On reste ainsi dans une optique voisine de celle d’un « traitement de texte ». Les zones de textes balisés sont visualisées à l’aide d’une combinaison de polices et couleurs qui sont décrites dans des « feuilles de styles » au format CSS librement configurables par l’utilisateur.

Configuration minimale

Le logiciel étant écrite en JAVA, il est exécutable sur un large nombre de plateformes : Linux, MacOS X et Windows. La machine virtuelle doit être récente : version $\geq 1.4.2$ et l’on doit disposer d’une quantité de mémoire raisonnable (≥ 256 Mo).

Installation de l’application

L’application distribué est nommée CADIXE.jar (il ne faut pas modifier ce nom !) et **il n’est pas nécessaire** de la « décompacter ». Lors de l’exécution le logiciel construit automatiquement (si elle n’existe pas déjà) la hiérarchie des fichiers dont il a besoin dans le dossier courant pour travailler. Elle a la forme suivante :

- CADIXE.jar : exécutable
- /resources : configuration de l’éditeur
 - cadixe.properties : fichier de préférence par défaut
 - ... : autres de fichiers de configuration
- /formats
 - /models : modèles de documents
 - /dtd : DTD et règles de complétion
 - /styles : feuilles de styles au format CSS

Lancement de l’application

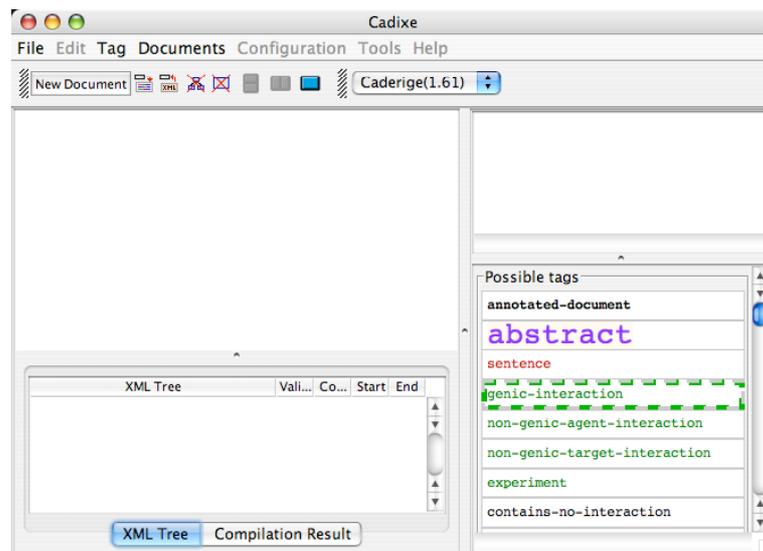
Le lancement de l’éditeur sous UNIX (ou Linux) s’effectue à partir d’un terminal (Xterm). Il faut se placer au niveau de l’application puis exécuter la ligne de commande suivante :

```
java -jar CADIXE.jar
```

Si l’application n’est pas au niveau du dossier actuel on peut aussi la lancer en indiquant le chemin d’accès. Il n’est plus nécessaire de mettre un argument *-root* comme dans la V1.16.

```
java -jar /Volumes/Memnos/editeur/version 2/CADIXE.jar
```

Sous d’autres systèmes (OS X notamment, Windows, ...) il est également possible de lancer l’application en double-cliquant simplement sur l’exécutable *SgmlJava.jar*. Après quelques instants, la fenêtre principale de l’éditeur apparaît ... L’application étant écrite avec Swing, l’aspect de la fenêtre change selon le systèmes (la copie ci-dessous est faite sur MacOS X). Dans certain cas les quatre parties peuvent être mal dimensionnées et il faut déplacer les séparateur pour obtenir l’écran ci-dessous ...



Lors du lancement de l’éditeur on peut spécifier dans la ligne de commande l’options suivante permettant de modifier le comportement par défaut de l’éditeur. Dans la prochaine version les autres arguments présents dans la V1.16 seront implémentés également :

-prefpath : indique le chemin d’accès aux préférences. Par défaut le chemin est l’endroit où se trouve l’application. En modifiant ce chemin on peut indiquer par exemple que les préférences sont stockées par rapport à l’utilisateur courant. Par exemple, sous Unix on pourra mettre dans le script de lancement le chemin « \$HOME ». Si le fichier n’existe pas (c’est le cas lors du premier appel) le système effectue une copie du fichier par défaut stocké dans le dossier “/Resources” de l’application ce qui permet facilement de « distribuer » une même configuration initiale aux utilisateurs (qu’ils pourront ensuite la personnaliser). Par exemple :

```
java -jar /CADIXE.jar -pref $HOME/Cadix
```

Historique des versions

La version 2.0a1 est une version « alpha » il manque donc encore beaucoup de fonctions, néanmoins outre le fait qu'elle permet d'annoter complètement des documents elle apporte d'ors et déjà de nombreuses nouveautés par rapport à la version précédente.

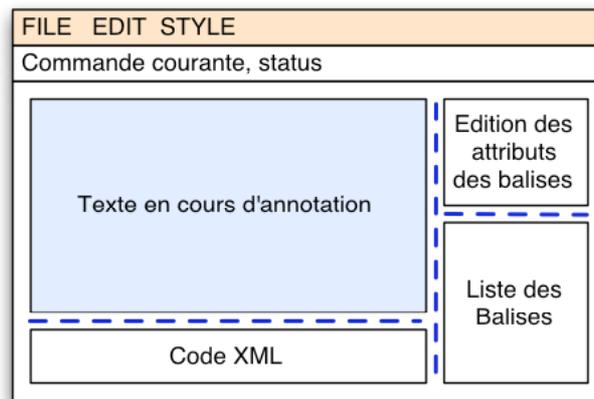
Nouveautés de la version 2.0a1 par rapport à la version 1.16

- Nouvelle implémentation
 - L'application a été intégralement réécrite ce qui a permis d'avoir un gain majeur en terme de vitesse et de stabilité. Les formats de fichiers ont été revus de manière à être plus simples (par exemple, les feuilles de styles sont au format CSS à 98%)
- Gestion multi-documents
 - L'éditeur permet maintenant de charger simultanément plusieurs documents annotés en mémoire, chaque document ayant sa propre DTD et ses propres styles, et d'en visualiser plusieurs simultanément ce qui facilite la comparaison des annotations.
- Copier/coller de textes et de balises
 - Le copier/coller prenant en compte les annotations est (enfin) implémenté, il permet également d'importer ou d'exporter du code XML. Par exemple si on copie une partie annotée et qu'on la colle dans un éditeur de texte (emacs,...) on voit apparaître le code XML correspondant. L'inverse étant également possible on peut donc importer du code écrit à la main, ou « à terme » généré par un script externe.
- Meilleur contrôle de l'introduction des balises
 - Dans la version précédente on ne pouvait insérer que des balise plus spécifiques que la balise courante, dans la V2.0 l'éditeur contrôle en outre que la balise que l'on applique sur un bloc est bien plus générale que toutes les balises incluses. Ainsi, n'est plus possible d'appliquer plusieurs fois une balise sur une même zone de texte.
- Nouvelles annotations graphiques
 - En plus de jouer sur les couleurs et les polices dans les feuilles de styles il est maintenant possible de dessiner des rectangles autour des annotations ce qui enrichie la palette visuelle (un exemple est fourni avec la distribution).
- Affichage personnalisable
 - Pour chacun des documents en mémoire l'utilisateur peut sélectionner le coefficient de zoom de l'affichage qu'il souhaite employer : 50%, 75%, 100%, ... Il n'est donc plus nécessaire de modifier les styles pour améliorer la lisibilité en outre l'utilisateur peut modifier la taille de l'interligne.
- Personnalisation des configurations
 - Via les paramètres passés dans la ligne de commande lors du lancement de l'éditeur il est possible de régler le comportement de l'éditeur de manière à décider quelles ressources (DTDs, Styles,, ...) sont locales ou globales aux utilisateurs. Cela facilite l'installation de CADIXE dans des environnements multi-utilisateurs. Par ailleurs l'éditeur "auto-installe" les dossiers dont il a besoin lors de la première exécution.

Organisation générale de l'interface

Outre la barre de menu et la ligne d'outils contenant les commandes usuelles, l'indicateur de position du curseur et les commandes de validation, la fenêtre est organisée en quatre zones qui peuvent être redimensionnées en faisant glisser les barres de séparations :

- TXT : Le (ou les) document en cours d'annotation (haut à gauche)
- XML : Les étiquettes XML qui ont été introduites dans le texte (bas à gauche)
- ATT : La valeur des attributs de la balise courante (haut à droite)
- BAL : La liste des balises utilisables (bas à droite) dans le contexte courant



Principes élémentaire de l'annotation d'un texte

Le principe d'utilisation de l'éditeur repose sur deux étapes : 1) la sélection à l'aide de la souris de la partie du document (mot, paragraphe, ...) que l'on souhaite annoter puis 2) sur l'application d'une balise à cette partie. Pour se faire, il suffit de cliquer sur l'étiquette voulue dans la liste BAL des balises. La zone ainsi distinguée apparaît alors dans le « format d'affichage » qui est associé à la balise et une nouvelle balise apparaît dans la zone XML.

A un instant donné, en fonction de la position du curseur, seules les balises de la DTD qui sont *plus spécifiques* que la balise courante peuvent être insérées. De manière symétrique lorsque l'utilisateur sélectionne tout une zone, seules les balises qui sont *plus générales* que chacune des balises apparaissant dans cette zone peuvent être insérées.

Selon la DTD utilisée, chaque balise peut avoir un certain nombre d'attributs qui servent à préciser sa signification. Ces attributs et leurs valeurs apparaissent dans la partie ATT de l'éditeur et peuvent être modifiés par l'utilisateur. Si les valeurs possibles appartiennent à un type énuméré (liste), ces valeurs sont sélectionnables à partir d'un menu, dans le cas contraire une simple zone d'édition permet de saisir le texte associé à l'attribut.

A tout instant la position du curseur détermine l'emplacement de la balise courante, dès lors pour examiner une annotation il suffit de placer le curseur dans celle-ci. Le nom et la liste des attributs de la balise courante s'affichent dans la partie ATT de l'éditeur.

La zone inférieure gauche de l'éditeur affiche les balises XML insérées dans le texte ainsi que la position de début et fin de chaque balise (exprimée en nombre de caractères). Elle permet de contrôler si l'annotation en cours est satisfaisante (notamment en ce qui concerne l'imbrication des balises). La navigation entre les deux fenêtres TXT et XML est

« synchronisée », en d'autres termes, lorsqu'on déplace le curseur dans l'une de ces zones l'autre est automatiquement décalée de manière à visualiser la même information.

Modification d'un texte et de son balisage

Comme dans tout éditeur, l'utilisateur est libre d'insérer, copier/coller, supprimer du texte dans le document. Pour supprimer une annotation il suffit de placer le curseur dans celle-ci puis de cliquer sur (ou de choisir dans le menu « Tag ») la commande "Delete current tag" pour effacer l'étiquette courante ou sur "Delete tag group" pour supprimer l'ensemble des étiquettes de cette zone c'est à dire l'étiquette courante et toutes celles qu'elle contenait éventuellement (le sous-arbre XML).

Il peut toutefois arriver que plusieurs balises soient appliquées sur une même zone de texte (même position de début et de fin). Dans ce cas pour indiquer à l'éditeur le niveau sur lequel on veut travailler, il suffit de cliquer dans la zone XML sur la balise concernée.

Gestion des fichiers (Menu FILE)

Les commandes de chargement de fichiers se trouvent dans le menu FILE et sont (pour une partie d'entre elles au moins) directement accessibles sous la forme d'icônes en haut de l'interface, sous la barre de menus. **IMPORTA NT** : L'éditeur CADIXE est désormais capable de travailler sur plusieurs documents en mémoire !

- **New Document** : l'éditeur construit un nouveau document en prenant comme modèle le fichier modèle par défaut qui est appelé pour l'instant « default.model » et qui se trouve dans le dossier « /formats/models/ ». On peut modifier les noms de la DTD et des feuilles de styles à charger en modifiant le modèle dans un éditeur de texte (évidemment les version ultérieure permettront d'effectuer cela dans l'interface).
- **Load Document** : Charge un document déjà annoté soit par l'éditeur actuel soit la version antérieure. Lorsqu'il charge une version antérieure le système charge automatiquement une feuille de style ayant le même nom que la DTD associée au document sinon il la créé. Par exemple, si le document pointait sur une DTD « caderige.dtd » il va automatiquement charger (ou créer) une feuille de styles « caderige.css » dans le dossier « /formats/styles ».

ATTENTION : pour que le système soit capable de retrouver la DTD, si vous voulez mettre les DTD dans le nouvel emplacement « /formats/DTD », il faut aussi éditer manuellement l'entête des fichiers annotés de manière à ce qu'ils pointent bien sur ce nouveau chemin. L'autre solution est de remettre dans le dossier de l'éditeur, le dossier qui contenait dans V1.16 votre DTD

- **Save Document** : Sauve le fichier annoté, contrairement à la version précédente, il n'y a plus de fichier de style associé. Le nom des styles à charger (et le chemin d'accès) est directement inclus dans les fichiers XML ce qui est plus simple.

Edition du texte (Menu EDIT)

Le copier/coller fonctionne maintenant sur les zones annotées. Le texte annoté est copié dans le « presse-papier » sous la forme d’une chaîne de caractères contenant les balises XML, on peut donc facilement échanger des annotations avec d’autres logiciels.

- Cut : coupe la partie sélectionnée et les annotations
- Copy : copie la partie sélectionnée et les annotations
- Paste : colle le contenu du « presse-papier ». On ne peut coller le contenu que si le point d’insertion correspond à une balise plus générale ou identique (dans cec as on fusionne) à celle incluse dans le presse-papier.
- Paste as text : colle le texte « brut » sans les balises ...

Gestion des balises (Menu TAG)

- Delete current Tag: Suppression de la balise courante (désignée par la position du curseur). Cette commande ne modifie pas les balises qui seraient incluses dans celle-ci. Par exemple si l’on supprime la balise ayant la feuille de style rouge :

Avant : `contrast, both |the E213G and E213A Spo0A variants showed`

Après : `contrast, both the E213G and E213A Spo0A variants showed`

- Delete Tree Tags : Suppression récursive de l’ensemble des balises dans la balise courante (désignée par la position du curseur). Cette commande affiche une boîte de confirmation car elle est susceptible de faire perdre la totalité des annotations effectuées dans un document. Par exemple, dans le cas ci-dessus, on aurait :

Avant : `contrast, both |the E213G and E213A Spo0A variants showed`

Après : `contrast, both the E213G and E213A Spo0A variants showed`

Gestion des documents (Menu DOCUMENTS)

- Zoom : niveau de zomm du document courant (par défaut 100%)
- Spacing : espacement entre les lignes du document courant
- Window organisation : permet de décider le nombre de documentst visibles dans la partie « annotation » de l’interface. L’utilisateur peut voir soit un seul document soit deux documents organisés verticalement ou horizontalement.
- Close document : ferme le document courant
- *Liste des documents* : le menu se termine par la liste des documents chargés dans l’éditeur, pour changer le document courant il suffit d’en choisir un dans cette liste.

Gestion des styles

Actuellement il n'y a pas (encore) de boîtes de dialogues pour configurer les différents fichiers utilisés par CADIXE. Il faut donc les modifier via un éditeur de texte et c'est notamment le cas pour les feuilles de styles associées à une DTD. Dans cette nouvelle version les feuilles de styles sont codées sous forme de CSS comme dans l'HTML, ils se trouvent par défaut dans le dossier « formats/styles ». Voici la liste des attributs qui est gérée par CADIXE, les autres attributs des CSS sont pour l'instant ignorés.

Attributs standards :

- background-color : *color* ;
- color : *color* ;
- font-family : *font name* ;
- font-style: italic ;
- font-size : *number* ;
- font-weight : bold;
- text-decoration : {underline, line-through} ;

Au sujet des "boîtes" :

- border-color : *color* ;
- border-width : *number* px ;
- border-style : {solid, dotted, dashed} ;

Les attributs spécifiques à CADIXE (absent du standard CSS 2.0):

- border-shape : {rectangle} ;

Contact

Ce système est développé dans le cadre du projet CADERIGE (<http://caderige.imag.fr/>) qui est financé dans le cadre d'une ACI Bioinformatique. Pour toutes questions, remarques, propositions d'amélioration de l'éditeur, il faut contacter : gilles.bisson@imag.fr

Par ailleurs, si vous constatez un problème ou anomalie quelconque dans l'éditeur, afin de faciliter sa correction, il faut nous transmettre les fichiers (document, DTD, style) sur lesquels vous étiez en train de travailler au moment du problème ainsi qu'une description précise du problème que vous avez rencontré.